Dietary factors and all-cause and cardiovascular mortality in Eastern European cohorts

Dénes Stefler

Thesis submitted for the degree of Doctor of Philosophy

University College London

2015

DECLARATION

I, Denes Stefler, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ABSTRACT

Background. Unhealthy diet, particularly low fruit and vegetable consumption, has been proposed as an important reason for the high cardiovascular disease (CVD) mortality in Central and Eastern Europe (CEE) and the former Soviet Union (FSU). However, individual-level food and nutrient intake data in these regions and direct comparisons with Western European populations are sparse, and estimates of their health effects are not available.

Aims. The aim of this thesis was to compare dietary intake habits between adults who live in Eastern and Western European countries, and to assess the relationships between selected dietary habits and all-cause and cause-specific mortality in Eastern Europeans.

Methods. Data collected from the Czech, Polish and Russian participants of the Health, Alcohol and Psychosocial Factors in Eastern Europe (HAPIEE) prospective cohort study (n=28,947) were used. The comparison of food and nutrient intakes with British participants in the UK Whitehall II study was carried out using quantile regression analysis after dietary data harmonization. The associations between dietary habits and mortality outcomes in the Eastern European cohorts were assessed by Cox regression models. Missing data was imputed using multiple random imputation procedures.

Results. Compared to the British participants, fruit and vegetable intakes were significantly lower in the pooled Eastern European sample but not in all country cohorts. In the pooled HAPIEE sample, the healthy diet indicator score and the

Mediterranean diet score were significantly and inversely associated with CVD mortality even after multivariable adjustments. Regarding fruit and vegetable intake, the inverse association appeared to be the strongest with stroke mortality and especially among smokers.

Discussion. The findings of this thesis support the hypothesis that unhealthy diet has played a role in the high CVD mortality in Eastern Europe. Public health interventions which target fruit and vegetable consumption and/or other dietary factors should be considered in this region.

ACKNOWLEDGEMENTS

First of all, I would like to thank for the support and guidance throughout this PhD to my primary supervisor, Martin Bobak. His sensible approach to work and life made these years really productive and balanced for me. I am also grateful for Eric Brunner, Hynek Pikhart and Anne Peasey whose advice at the project`s critical points were truly invaluable. I also owe many thanks to the colleagues in Wageningen, Edith Feskens and Nicole Jankovic, who helped me to take the first crucial steps in the field of nutritional epidemiology.

I am grateful for the funding by the British Heart Foundation, which made it possible for me to undertake this work.

My special thanks go to my parents and brother, Péter, for their unconditional support, without which I would not be where I am now. Nagyon köszönöm.

Finally, and most importantly, Vidarshi and Maisey. Thank you for your patience, encouragement and happiness you have given me every single day. This work would have never been completed without your support.

THESIS OUTPUTS

Published articles (see also in appendix VII.)

Stefler D, Pajak A, Malyutina S, Kubinova R, Bobak M, Brunner EJ.: Comparison of food and nutrient intakes between cohorts of the HAPIEE and Whitehall II studies. *Eur J Public Health*. **2015** Dec 4. pii: ckv216.

Stefler D, Malyutina S, Kubinova R, Pajak A, Peasey A, Pikhart H, Brunner E, Bobak M.: Mediterranean diet score and total and cardiovascular mortality in Eastern Europe: the HAPIEE study. *Eur J Nutr*. **2015** Nov 17.

Stefler D, Bobak M. Does the consumption of fruits and vegetables differ between Eastern and Western European populations? Systematic review of cross-national studies. *Arch Public Health.* **2015**;73(1):29.

Stefler D, Pikhart H, Kubinova R, Pajak A, Stepaniak U, Malyutina S, Simonova G, Peasey A, Marmot MG, Bobak M.: Fruit and vegetable consumption and mortality in Eastern Europe: longitudinal results from the HAPIEE study. *Eur J Prev Cardiol.* **2015** Apr 22. pii: 2047487315582320.

Stefler D, Pikhart H, Jankovic J, Kubinova R, Pajak A, Malyutina S, Simonova G, Feskens E, Peasey A, Bobak M.: Healthy diet indicator and mortality in Eastern European populations: prospective evidence from the HAPIEE cohort. *Eur J Clin Nutr*. **2014**;68:1346-52.

Conference presentations

Stefler D, Pikhart H, Kubinova R, Pajak A, Malyutina S, Peasey A, Bobak M.: Mediterranean diet and mortality in Eastern Europeans: results from the HAPIEE study. European Congress of Epidemiology - Healthy Living; Maastricht, Netherlands; 25-27 June 2015.

Stefler D, Pajak A, Malyutina S, Kubinova R, Bobak M.: Fruit and vegetable consumption and mortality in Eastern Europe. 3rd World Congress of Public Health Nutrition; Las Palmas de Gran Canaria, Spain; 9-12 November 2014.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BMI	Body mass index
BMR	Basal metabolic rate
CEE	Central and Eastern Europe
CHD	Coronary heart disease
CI	Confidence interval
CINDI	Countrywide Integrated Non-communicable Diseases Intervention
CIS	Commonwealth of Independent States
CVD	Cardiovascular disease
DAFNE	Data Food Networking
DBP	Diastolic blood pressure
ECG	Electrocardiogram
EFSA	European Food Safety Authority
EHIS	European Health Interview Survey
EI	Energy intake
EPIC	European Prospective Investigation into Cancer and Nutrition study
EU	European Union
F&V	Fruit and vegetable
FAO	Food and Agriculture Organization of the United Nations
FBS	Food balance sheet
FFQ	Food frequency questionnaire
FSU	Former Soviet Union
GBD	Global burden of disease
GFQ	Graduate frequency questionnaire
GHO	Global Health Observatory
GINA	Global Database on the Implementation of Nutrition Action
HAPIEE	Health Alcohol and Psychosocial Factors in Eastern Europe study
HBS	Household budgetary survey
HDI	Healthy diet indicator
HR	Hazard ratio
ICD	International classification of diseases

IHD	Ischemic heart disease
IQR	Interquartile range
MAP	Mean arterial pressure
MAR	Missing at random
MCAR	Missing completely at random
MDS	Mediterranean diet score
MET	Metabolic equivalent
MNR	Missing not at random
NHS	National Health Service
NSP	Non-starch polysaccharides
OR	Odds ratio
PARF	Population attributable risk fraction
PP	Preventable proportion
PUFA	Polyunsaturated fatty acid
RCT	Randomised controlled trial
RR	Relative risk
SBP	Systolic blood pressure
SD	Standard deviation
SHR	Subhazard ratio
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
UK	United Kingdom
USDA	United States Department of Agriculture
WCRF	World Cancer Research Fund
WE	Western Europe
WHO	World Health Organization

INTRODUCTION

Unhealthy diet is the leading risk factor for morbidity and mortality worldwide (Lim *et al.* 2012). It plays a role in the development of the most common chronic non-communicable diseases, such as cardiovascular disease (CVD), diabetes and cancer. In fact, according to the WHO, unhealthy diet together with physical inactivity was responsible for 57% of cardiovascular and 19% of overall global mortality in 2004 (WHO 2009). However, although the relationship between nutrition and chronic diseases is one of the most intensively studied area in epidemiology, there are still important gaps in the literature which need to be filled.

While the research into the diet-disease relationships has a global relevance, it is particularly important in Eastern Europe. There are two main reasons for this. Firstly, previous analyses suggested that the disease burden due to unhealthy diet is greater in Eastern European countries than in any other regions of the world (WHO 2009; Lim *et al.* 2012). And secondly, despite the noteworthy findings of the aforementioned analyses, available individual-level dietary data in Eastern Europe are still sparse, and virtually no previous studies with prospective cohort design examined the association of diet with health in this region.

Investigating the link between dietary habits and health outcomes in Eastern European individuals can also help to explain the reasons for the large health gap which exist between Eastern and Western European populations. Data shows that the overall and cardiovascular mortality rates are significantly higher in Eastern European countries compared to Western European states (WHO Regional Office for Europe 2014). Although the role of several socio-economic and lifestyle factors in this East-West health gap have been intensively investigated (Bobak and Marmot 1996; Gilmore *et al.* 2004; Leon *et al.* 2007), our knowledge regarding the contribution of dietary habits is still limited.

Consumption of fruits and vegetables is one component of our diet where a health protective effect is supported by relatively strong epidemiological evidence (Dauchet *et al.* 2009; Wang *et al.* 2014). Fruit and vegetable intake in Eastern Europe has been often suggested to be inadequate, which might have contributed to the high CVD rates of the populations (Ginter 1998; Pomerleau, McKee, *et al.* 2003; Zatonski 2011). Analysing fruit and vegetable intake of Eastern European individuals in relation to Western European subjects and mortality outcomes has the potential to test the validity of these hypotheses and strengthen the respective evidence.

Examining dietary patterns, as opposed to specific foods or nutrients, offers the possibility to understand the health effects of the diet as a whole. The application of this holistic approach in nutritional epidemiological studies has gradually become more common over the recent years (Hu 2002; Kant 2004). In addition, some eating patterns, such as the Mediterranean diet, are amongst those few diet-related exposures which have been proved to be protective against chronic diseases and mortality not only in observational studies but interventional trials as well (Estruch *et al.* 2013; Sofi *et al.* 2014). This means that the research of dietary patterns in relation to mortality outcomes in Eastern European populations might answer some questions regarding overall eating habits in this region, and it corresponds well with the current trends of this scientific field.

Due to the modifiable nature of diet, nutritional epidemiological research has high importance from the public health point of view. Evidence from epidemiologic studies can help to design effective public health policies by offering clear targets for dietary intervention campaigns. Considering the poor health status of Eastern European populations, even small improvements in the preventative strategies can result in large benefits in terms of population health.

This PhD thesis is organised in six main chapters: background, aims and objectives, methods, results, discussion, and finally, conclusions and implications. Chapter 1 provides the background and context of the work, and it focuses on three specific topics. First, it describes the health status of Eastern European populations with particular attention to CVD, then the current knowledge on Eastern European dietary habits is detailed, and finally, an overview on the available evidence regarding the relationships between selected dietary habits and CVD is given. Chapter 2 outlines the aims and objectives of the work. In chapter 3, the applied methods are presented. In order to achieve the four main objectives, four distinct analyses were carried out: (1) comparison of dietary intakes between the HAPIEE and the Whitehall II cohorts; analysis of the relationships of (2) fruit, vegetable intake, (3) the healthy diet indicator and (4) the Mediterranean diet score with mortality in the HAPIEE study. While many methodological details are relevant for all parts of the work, some are applicable only for the specific analyses. These distinctions are made clear in the methods chapter. Chapter 4 shows the results of the work, presented separately for the four analyses as discussed above. In chapter 5, analysis-specific and overarching discussions of the findings are provided, including limitations and strengths of the work and the meaning of results in light of the existing literature. Overall conclusions and implications are considered in the final chapter.

- CHAPTER 1 -

BACKGROUND

This chapter describes the theoretical background of the thesis and presents the context of the research. Since the analytical part of the work focuses on selected dietary habits of Eastern European population samples in relation to Western European dietary data and CVD outcomes, the background is divided into three main parts. (1) CVD in Eastern European populations: the first part gives an overview of the differences in CVD mortality rates between Eastern and Western European countries, as well as examining the possible underlying reasons for this health gap (section 1.1). (2) Dietary habits of Eastern European populations: the second part (including sections 1.2, 1.3 and 1.4) describes the most important dietary habits which have been hypothesised by previous authors to contribute to the high CVD risk in Eastern European countries. A systematic literature review focusing on the available evidence regarding fruit and vegetable intake data, and a section dedicated to the three Eastern European countries from which the participants of the HAPIEE study are recruited from are also included in this part. (3) Relationship between selected dietary habits and CVD: in the final part (section 1.5), the available evidence for the associations of fruit and vegetable intake, healthy diet indicator and Mediterranean diet score with CVD risk is presented.

1.1 CVD in Eastern and Western Europe

This section starts by defining the key geographical terms which are used throughout the thesis. Subsequently, the differences in life expectancy at birth and CVD mortality and morbidity rates between Eastern and Western European countries are described. Finally, the possible reasons for the health gap are summarised, including upstream (social) and downstream (behavioural and metabolic) risk factors.

1.1.1 The East-West division of Europe

In health research, Europe is usually defined as the WHO European region which consists of 53 countries, including 11 which fully or partly belong to Asia in a geographical sense. The rich history and cultural heritage, the diverse climatic conditions and the large differences in economic performance between countries within a relatively small area make Europe an ideal region to study the determinants of population health.

The historical events of the 20th century, especially the east-west division of the continent during the Cold War era, are amongst the most important factors that influence the health of Europeans today. Although the Iron Curtain collapsed in 1991, the health gap between the former Eastern Bloc and Western European countries persists.

Although the terminology is not strictly defined, the term Central and Eastern Europe (CEE) usually refers to the group of countries which were members of the Eastern Bloc, but were not incorporated into the Soviet Union (Mackenbach *et al.* 2013). As in 2015, 13 independent states belong to this group, of which seven are members of the European Union (EU). The 15 countries of the former Soviet Union (FSU) gained their independence in 1991. Today, three of them (Latvia, Lithuania and Estonia) are EU members, and the others, with the exception of Georgia, belong to the Commonwealth of Independent States (CIS). In this thesis, the term Eastern Europe refers to the region which includes both CEE and FSU countries.

The term Western Europe can also be used in geographical context, but its more important political meaning was developed during the Cold War. It referred to countries on the other side of the Iron Curtain, and it includes, by convention, the 15 EU states which were members of the organization before 2004, as well as Iceland, Norway, Switzerland and the microstates within their territory (Mackenbach *et al.* 2013), see figure 1.1.

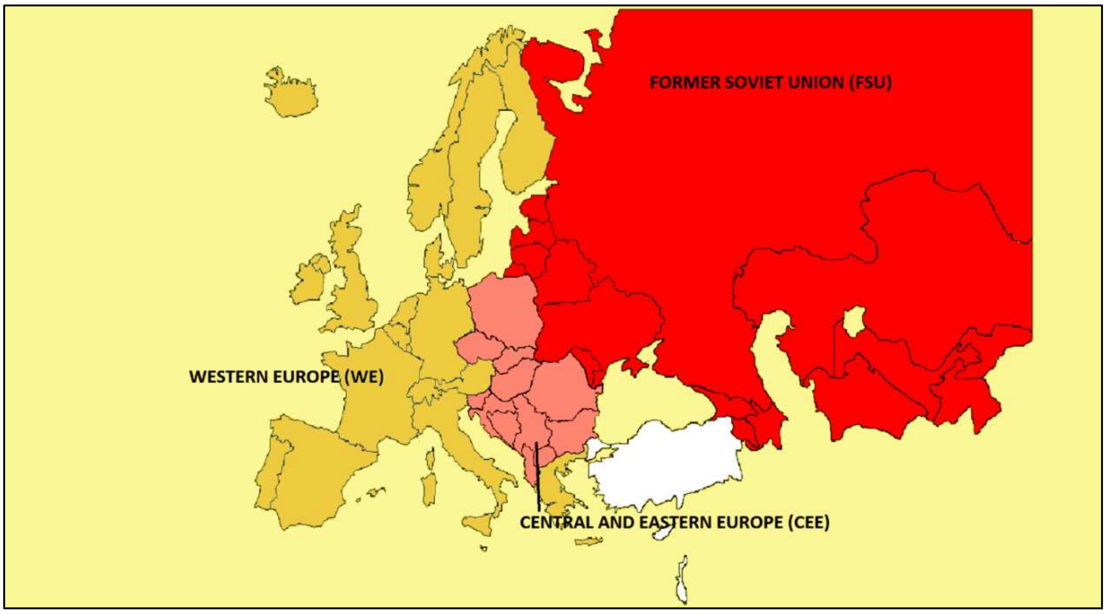


Figure 1.1: East-West division of Europe

1.1.2 Life expectancy at birth in European countries

As a result of improvement in personal and public hygiene and new discoveries in medical treatments, life expectancy at birth sharply increased in developed countries during the first half of the 20th century. Although the epidemiologic transition was interrupted by the two World Wars, this favourable trend was fairly consistent in most European countries until the 1960s (Kinsella 1992; Gelbard *et al.* 1999).

However, between 1970 and 1990, the health consequences of the different political systems in Eastern and Western Europe became apparent. While life expectancy at birth continued to grow further in the West, it stagnated or even declined in most Eastern Bloc countries over this period (Uemura and Pisa 1988) (figure 1.2). If the comparisons were made for adults only, the differences would be probably even more significant, because the rising death rates of adults in Eastern Europe were compensated by the improvement in child mortality (Chenet *et al.* 1996). In the 1990s, after the collapse of the communist regimes, CEE and FSU countries went through profound political, economic and social changes which affected the lives and health of the populations significantly (Bobak and Marmot 2009). Although life expectancy at birth started increasing during the early or mid-1990s in CEE, the level of disturbance was more remarkable in the FSU. In Russia and some Baltic states, for example, unprecedented fluctuations of death rates signalled a serious mortality crisis (Shkolnikov *et al.* 2001; Karanikolos *et al.* 2012). Steady improvement in life expectancy can be seen only from the mid-2000s in most of these countries. The overall trend in life expectancy between 1960 and 2010 shows converging pattern for Western European states but divergence for the countries of CEE and FSU, suggesting that the differences between countries became smaller in the former but bigger in the latter regions (Mackenbach *et al.* 2013).

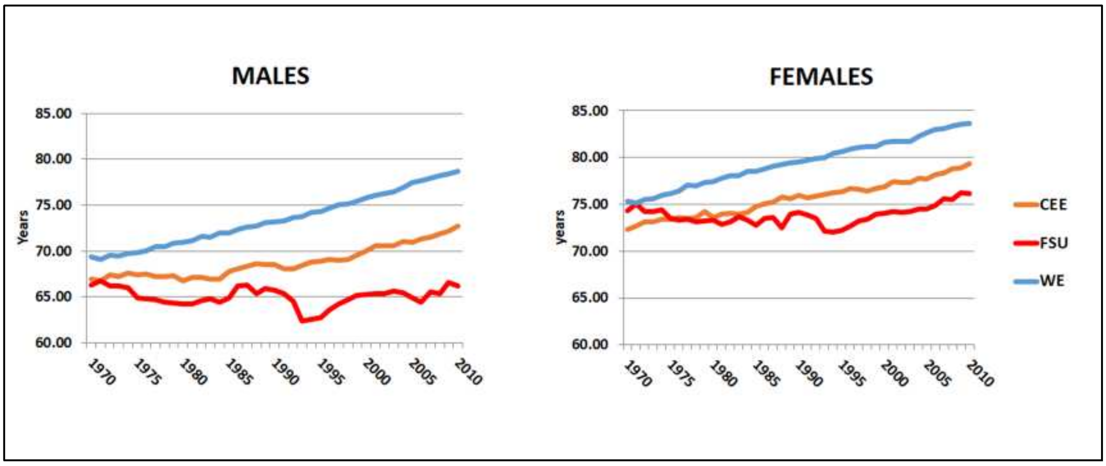
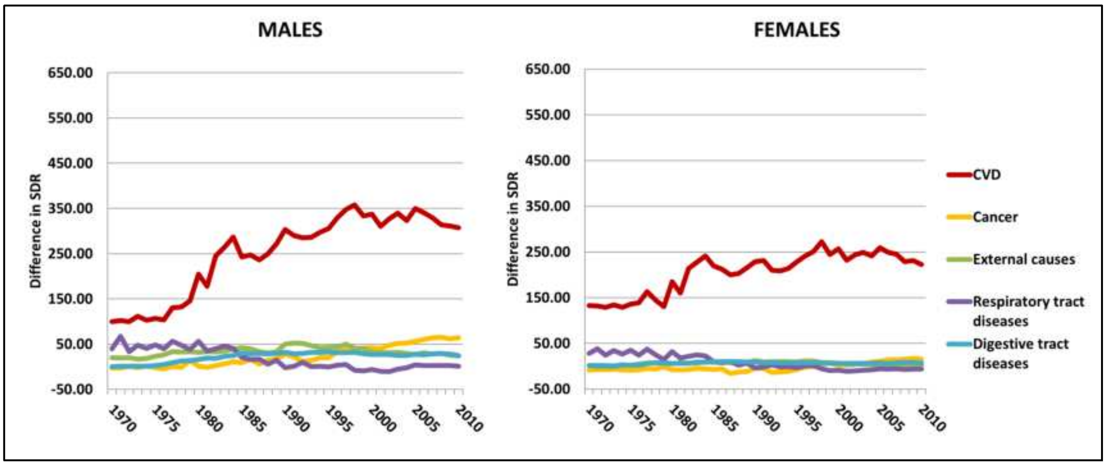


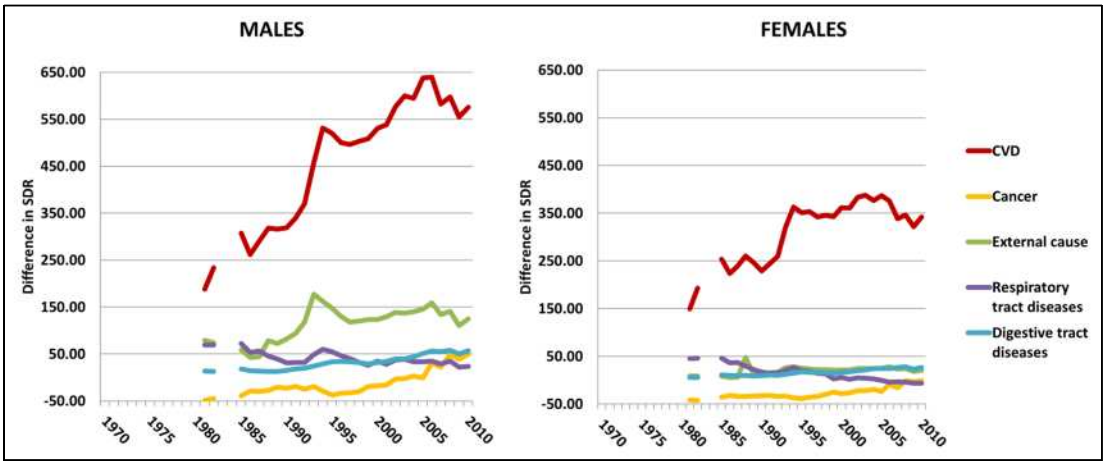
Figure 1.2: Average life expectancy at birth of males and females in Central and Eastern Europe (CEE), Former Soviet Union (FSU) and Western Europe (WE) between 1970 and 2010 *(Data source: WHO European Health for All Database)*

1.1.3 Differences in CVD mortality and morbidity rates between Eastern and Western Europe

Figure 1.3 and figure 1.4 show the differences in aggregate age-standardized cause-specific mortality rates between FSU and Western Europe, and between CEE and Western Europe from 1970 to 2010. These data demonstrate that mortality from CVD has been the most important difference between Eastern and Western European countries over the last four decades. The difference is more pronounced in FSU and in males. Previous analysis by the WHO showed that CVD was responsible for 54% of the mortality gap between Western Europe and CEE/FSU in 1992. The WHO analysis also indicated that the widest gap occurred in the 35-64 years age group and in males (Bobak and Marmot 1996). Although CVD death rates have been on a decline since the mid-1990s in most Eastern European countries, due to the consistent improvement in the West, the gap has hardly changed. Today, CVD accounts for approximately half of all deaths in the East, compared to one third in the West (WHO 2013).



<u>Figure 1.3:</u> Differences in average age-standardized cause specific mortality rates (SDR) between Central and Eastern European (CEE) and Western European (WE) countries between 1970 and 2010 ***(Data source: WHO European Health for All Database)***



<u>Figure 1.4:</u> Differences in average age-standardized cause specific mortality rates (SDR) between Former Soviet Union (FSU) and Western European (WE) countries between 1970 and 2010 ***(Data source: WHO European Health for All Database)***

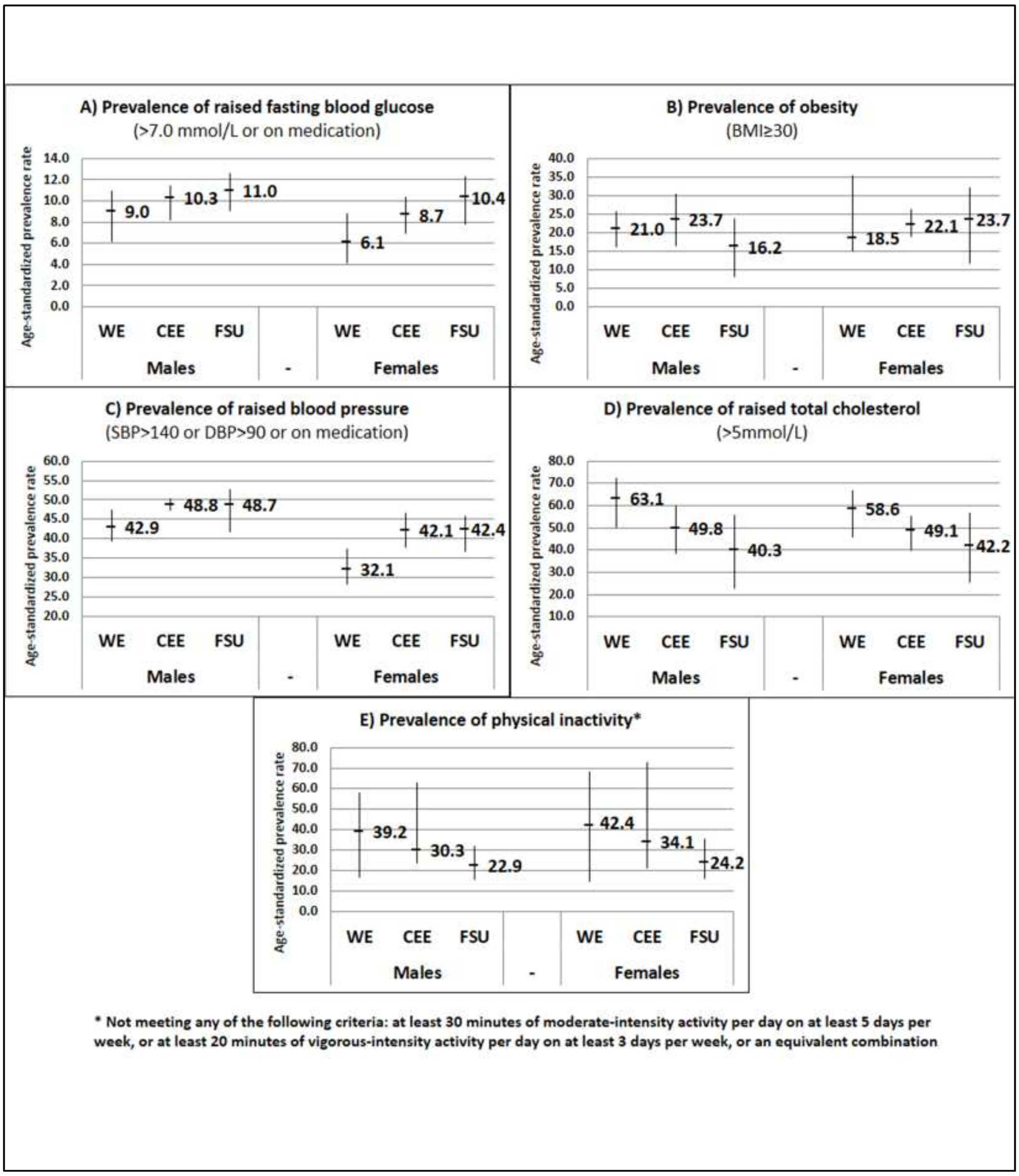
The WHO Global Burden of Disease project estimated that in 2010, Eastern European and Central Asian countries, majority of which were FSU states, had the highest ischemic heart disease (IHD) mortality rates not just compared to other parts of Europe but in the global context as well (Forouzanfar *et al.* 2012). Figures regarding Central Europe, covering CEE countries, were also amongst the highest globally. This study also found that IHD morbidity rates, calculated by statistical modelling using data from population-based surveys, followed the same global pattern as the death rates, which suggests that the primary problem is not the elevated fatality rates of IHD but its high incidence and prevalence in these countries.

1.1.4 Possible reasons for the health gap

Several possible explanations for the health gap between Eastern and Western Europe have been suggested. Amongst the upstream factors, one obvious reason could be the significant difference in economic performance between the two regions; on average, GDP is twice as high in Western European states compared to CEE or FSU countries (European Commission 2013). However, the picture is more complex, and other social and societal factors also need to be taken into account (Feachem 1994; Bobak *et al.* 2007). For example, the authoritarian, over-medicalized health care system and the lack of emphasis on non-communicable disease prevention in public health together with the easy availability of tobacco and alcohol products all contributed to the widespread occurrence of unhealthy lifestyle habits in CEE and FSU countries (McKee 2007; Zatonski 2011; Rechel *et al.* 2013).

From the proximal (lifestyle and metabolic) CVD risk factors, the role of alcohol consumption and tobacco smoking have been extensively investigated and confirmed (Leon *et al.* 1997; McKee *et al.* 1998; Pudule *et al.* 1999; Britton and McKee 2000; Gilmore *et al.* 2004; Nicholson *et al.* 2005; Leon *et al.* 2007; Tomkins *et al.* 2012; Lim *et al.* 2012).

Estimated prevalence rates of other behavioural (physical inactivity) and metabolic (hypertension, hypercholesterolemia, obesity, hyperglycaemia) CVD risk factors, published by the WHO Global Health Observatory (GHO), are presented in figure 1.5 (WHO 2013). Results suggest that in 2008 the average prevalence rate of hypertension and hyperglycaemia was higher in CEE and FSU states compared to Western Europe, while hypercholesterolemia and physical inactivity seems to have been less common in the East, and the picture regarding obesity was not clear. The figures also show that there were large differences between countries within a region. Consequently, comparisons of individual countries across regions could give significantly different results from the aggregate findings.



<u>Figure 1.5:</u> Average of estimated cardiovascular risk factor prevalence rates in Central and Eastern European (CEE), Former Soviet Union (FSU) and Western European (WE) countries, and the range of country-specific results within a region *(Data source: WHO Global Health Observatory)*

Socioeconomic deprivation can lead to poor health and increased risk of CVD via unhealthy lifestyle habits or directly, through the psychosocial pathway (Brunner and Marmot 2006). Studies have shown that the psychosocial stress due to effort-reward imbalance at work, low perceived control, job insecurity and low social support is higher in most CEE and FSU countries compared to Western Europe (Bobak, Pikhart, *et al.* 1998; Kopp *et al.* 2006; Steptoe *et al.* 2007; Lundberg *et al.* 2007; Laszlo *et al.* 2010; Salavecz *et al.* 2010). These findings suggest that the difference in psychosocial stress may also be an important reason for the health gap between East and West.

1.2 Dietary habits in Eastern and Western Europe

The role of unhealthy diet, as a possible contributing lifestyle factor, is more difficult to estimate, due to the complexity of dietary exposure. The evidence is summarized in this section.

First, the available data sources which provide information on food and nutrient supply and intake in European countries are discussed, including their strengths and limitations. In the second part of this section, I describe the two main dietary habits which have been previously suggested as contributing factors to the high CVD risk in Eastern European countries.

1.2.1 Sources of food and nutrient availability and intake data in Europe

International comparison of dietary intakes of foods and nutrients can be based on three data sources: (1) food balance sheet (FBS), (2) household budget survey (HBS), (3) individual level dietary survey.

FBSs are produced by the Food and Agriculture Organization of the United Nations (FAO). Data are collected from all member states annually and published on the organization`s website (FAO 2015). FBS do not give information about the actual consumption of the examined food items, only their availability on a country level. It is calculated by adding up the total quantity of foodstuff produced in and imported by a specific country, then subtracting the quantity which is exported, fed to livestock, used for non-food purposes or wasted during storage and transport. However, the amount lost in the households (i.e.: during meal preparation, plate-waste, given to pets, etc.) is not taken into account (Joffe and Robertson 2001). In addition, FBS data do not account for foods which are produced by individuals for self-supply, usually in small household gardens, allotments. For example, fruits and vegetables produced by such ways can contribute to the actual intake substantially. This contribution is probably larger in countries with weaker economy and extensive home-growing traditions, like Russia or other Eastern European countries. FBS data usually overestimate the intake levels of various food items. For example, it has been estimated that the discrepancy between FBS and dietary survey data regarding fruit and vegetable intake were approximately 30-39%, and the differences between countries were substantial (Joffe and Robertson 2001; Pomerleau, Lock, *et al.* 2003). More recent data from the WHO GBD project calculated even greater gap, suggesting 78.4% and 74.5% over-reporting for fruits and vegetables, respectively (Del Gobbo *et al.* 2015).

HBS has been used for dietary data collection by the Data Food Networking (DAFNE) project, and data is currently available for 24 European countries (National and Kapodistrian University of Athens 2005). This method provides food availability information on the household level collected from nationally representative population samples (Trichopoulou 1992).

Since both FBS and HBS data are ecological (characterizing populations, rather than individuals), they are suitable for hypothesis generation but are not ideal for testing causal associations. On the other hand, their advantage is that both of these methods are highly standardized, well-comparable between countries and readily available.

Nationally representative, individual level nutritional surveys are conducted regularly in most European countries in order to monitor the population`s dietary habits. Although they provide good evidence for public health recommendations in the specific countries, their applicability for international comparison is limited. The reason for this is that most surveys use different methods for data collection, different food classification and coding systems to categorize the items into food groups, different portion sizes to calculate g/day intakes and different food composition tables to calculate nutrient intake values (Charrondiere *et al.* 2002; Ireland *et al.* 2002; de Boer *et al.* 2011). In 2011, the European Food Safety Authority (EFSA) published the Comprehensive European Food Consumption Database which contains food intake data for most EU member states collected by national dietary surveys (EFSA 2011a). However, the authors emphasized that due to the differences in methods of data collection and analysis, the presented intake levels are not suitable for international comparison (EFSA 2011c).

1.2.2 Animal fat and vegetable oil intake

The unhealthy dietary habits which have been the most often proposed as major contributors to the high CVD rates in CEE and FSU countries are the high intake of animal fat and low intake of fruits and vegetables.

Regarding animal fat intake, FBS food availability data indicate that the aggregate vegetable oil vs. animal fat ratio was considerably lower in CEE and FSU countries compared to Western Europe before 1990. However, from the mid-90s, steady increase in the ratio in both Eastern regions suggests that animal fat was gradually replaced with vegetable oils in these countries (figure 1.6) (FAO 2015). Zatonski and others proposed that this change in diet was one of the main reasons for the sharp decline in CVD mortality rates in Poland and other CEE countries after 1991 (Zatonski *et al.* 1998; Waskiewicz *et al.* 2006; Zatonski *et al.* 2008). Similar trends in animal fat and vegetable oil intakes were observed in the Baltic States and the Czech Republic, which was also connected to the improvement in CVD mortality in these countries during the 1990s (Poledne and Skodova 2000; Puska *et al.* 2003; Kesteloot *et al.* 2006; Ramazauskiene *et al.* 2011). Estimations from the WHO GBD project also confirmed these trends suggesting a steady decline in saturated fat but increase in polyunsaturated fat intake in most Eastern European countries between 1990 and 2010 (Micha *et al.* 2014).

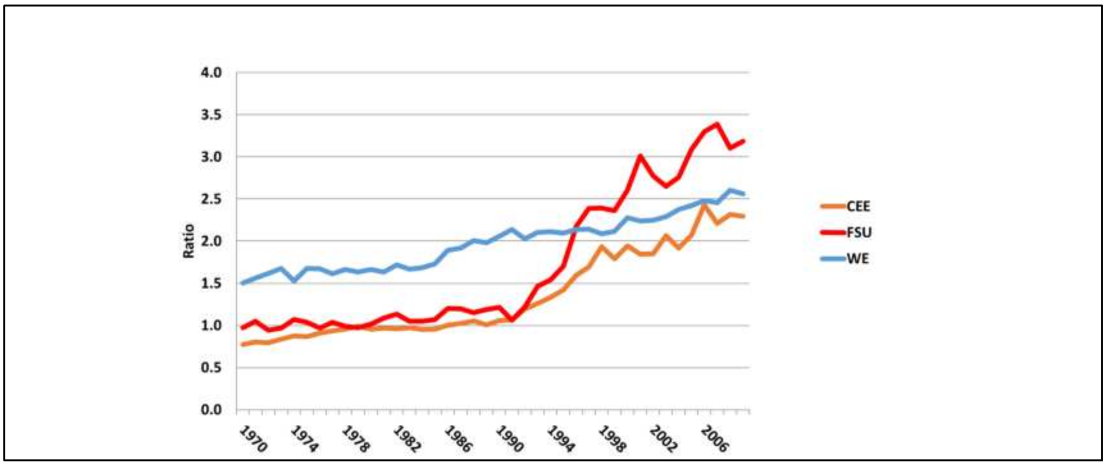
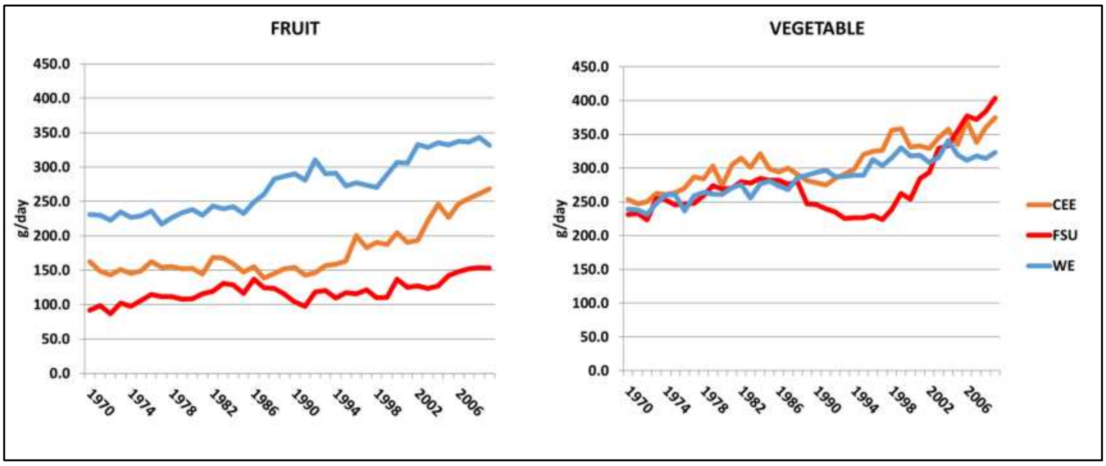


Figure 1.6: Ratio of average vegetable oil vs. animal fat availability in Central and Eastern Europe (CEE), Former Soviet Union (FSU) and Western Europe (WE) between 1970 and 2009 (Food balance sheet data) *(Data source: FAOSTAT)*

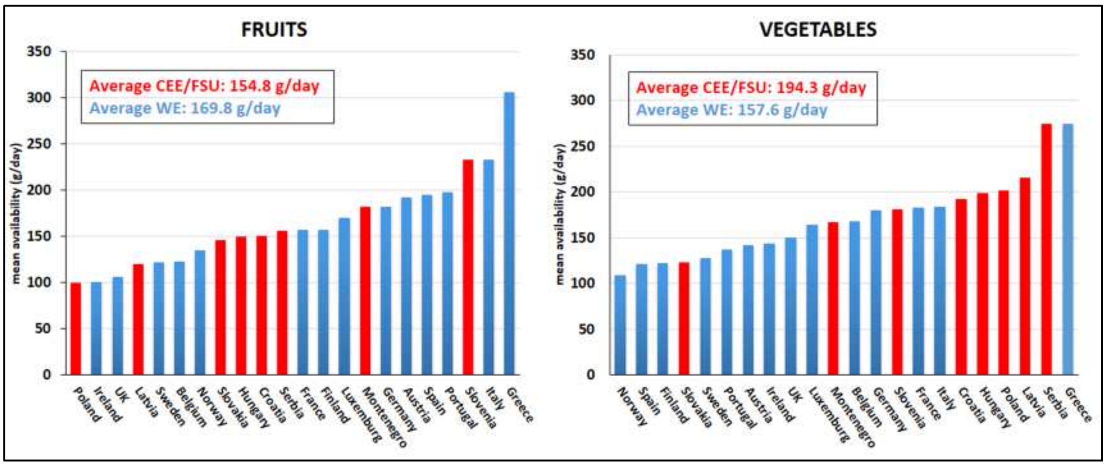
1.2.3 Fruit and vegetable intake

Inadequate consumption of fruits and the consequent low intake of antioxidant vitamins, as important reasons for the poor cardiovascular health in Eastern Europe, was first proposed by Ginter (Ginter 1995; Ginter 1998). Zatonski also suggested that increased fruit and vegetable supply, together with the reduced animal fat and increased vegetable oil intake, was responsible for the favorable trends in CVD mortality in Poland during the 1990s (Zatonski *et al.* 1998). The original hypothesis, similarly to the animal fat theory, was mainly supported by ecologic data from FAO`s FBSs. The average availability of fruits and vegetables in Western Europe, CEE and FSU countries between 1970 and 2010 calculated from the FAOSTAT database is presented in figure 1.7. The figure shows clearly higher fruit supply in Western Europe compared to CEE and FSU over the four decades, however, no differences between the regions can be seen for the availability of vegetables.



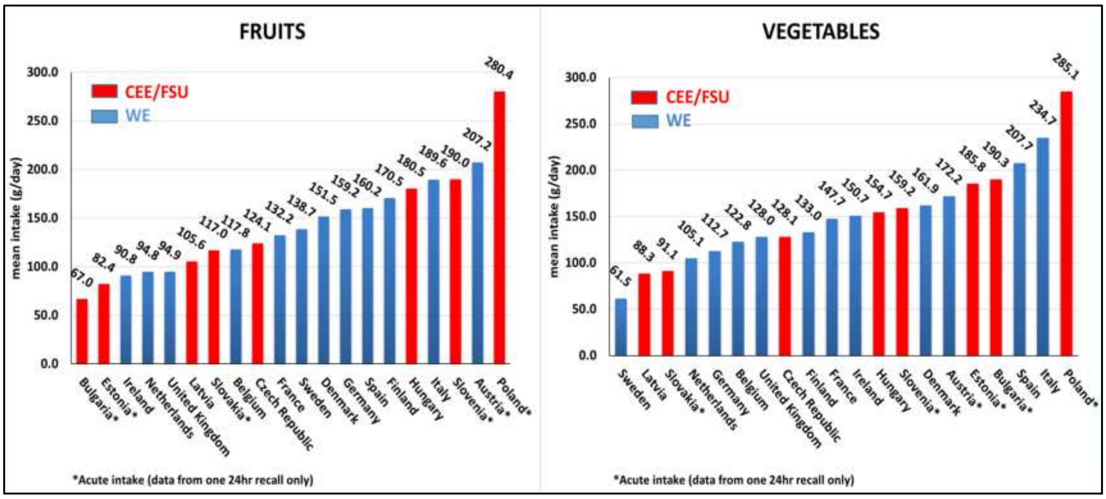
<u>Figure 1.7:</u> Average availability of fruits and vegetables in Central and Eastern Europe (CEE), Former Soviet Union (FSU) and Western Europe (WE) between 1970 and 2009 (Food balance sheet data) *(Data source: FAOSTAT)*

HBS data from DAFNE database collected around the year 2000 confirm the FAO results (figure 1.8). Average availability of fruits in households was higher in Western Europe than in the East, but vegetable supply shows opposite results.



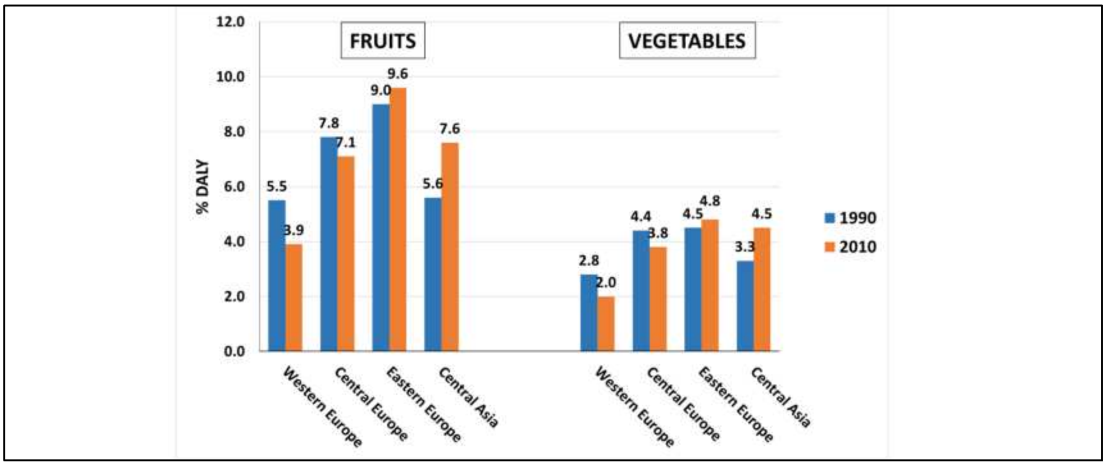
<u>Figure 1.8:</u> Availability of fruits and vegetables on the household level in European countries (Household budget survey data) *(Data source: DAFNE databank)*

The intake values of fruits and vegetables obtained from the EFSA`s Comprehensive European Food Consumption Database are presented in figure 1.9 (EFSA 2011a), although the international comparability of data from national dietary surveys is limited, as described earlier.



<u>Figure 1.9:</u> Intake of fruits and vegetables in European countries measured by national dietary surveys (Individual-level data) *(Data source: EFSA)*

Despite the limitations, data from nationally representative dietary surveys (supplemented by FAO statistics if survey data was not available) was used to calculate the disease burden attributable to low fruit and vegetable consumption in the WHO Global Burden of Disease projects (Lock *et al.* 2005; WHO 2009). The results of the most current version are presented in figure 1.10 showing disease burden estimates in 1990 and 2010 in Western Europe, Central Europe (CEE countries), Eastern Europe (FSU countries in Europe), and Central Asia (FSU countries in Asia + Mongolia) (Lim *et al.* 2012).



<u>Figure 1.10:</u> Percentage of disability-adjusted life years (DALY) due to low fruit and vegetable consumption in Western Europe, Central Europe, Eastern Europe and Central Asia. *(Data source: Lim et al 2012)*

The results suggest that the disease burden is significantly higher in CEE and FSU compared to Western Europe, and low fruit intake is a bigger problem than low vegetable consumption. The estimated disease burden reduced in Western and Central Europe between 1990 and 2010, but it got worse in Eastern Europe and Central Asia. The declining trend in fruit and vegetable intakes in the latter regions over the last decade has been also confirmed by a recent analysis (Abe *et al.* 2013). In comparison with other regions of the world, GBD estimates show that CEE and FSU countries have the highest disease burden due to low levels of fruit and vegetable consumption, not only in Europe but globally as well (WHO 2009; Lim *et al.* 2012).

Apart from the EFSA database, systematic reviews also compared fruit, vegetable and micronutrient intake levels and status between CEE/FSU and Western Europe countries using data from studies which had been separately conducted in the two regions (Lesser *et al.* 2008; Novakovic *et al.* 2013). They found that the methodological differences between studies seriously limited the interpretation of the results; the lack of comparable data is especially salient in CEE/FSU countries. On the other hand, cross-national studies which include participants from both CEE/FSU and Western Europe countries, and use identical methods for data collection and analysis in the two samples are more suitable designs to compare food consumption levels directly between the two regions. I systematically reviewed the literature for such cross-national studies which reported data on consumption of fruits, vegetables, or their surrogate indicators, vitamin C or carotenoids. The methods, results and detailed discussion of the systematic review are presented in the following section.

1.3 Systematic literature review of cross-national studies

The aim of this systematic review was to collect and summarize the results of all cross-national studies which reported data on consumption of fruits, vegetables, or their surrogate indicators, such as vitamin C and carotenoids, of participants from CEE/FSU and Western European countries using identical methods for data collection and analysis in the two samples.

1.3.1 Methods

1.3.1.1 Search strategy

MEDLINE, EMBASE and Web of Science databases were searched from inception to September 2014, using search terms described in figure 1.11. References and citation lists of selected papers were studied for additional papers, and hand search of key journals (*Public Health Nutrition, European Journal of Clinical Nutrition, European Journal of Public Health*) was also performed. No restriction on language was applied.

exp Europe, Eastern/ OR exp USSR/ OR exp Czechoslovakia/ OR exp Germany, East/ OR exp Yugoslavia/ OR exp Transcaucasia/ OR exp Asia, Central/ OR central europe*.mp. OR eastern europe*.mp. OR alban*.mp. OR armen*.mp. OR azerbajan*.mp. OR belarus*.mp. OR bosnia*.mp. OR hercegovina*.mp. OR bulgar*.mp. OR croat*.mp. OR czechslovak*.mp. OR czech*.mp. OR east german*.mp. OR eston*.mp. OR georgia*.mp. OR hungar*.mp. OR kazakh*.mp. OR kyrgiz*.mp. OR latvia*.mp. OR lithuan*.mp. OR montenegro*.mp. OR poland*.mp. OR polish*.mp. OR moldova*.mp. OR roman*.mp. OR russia*.mp. OR serb*.mp. OR slovak*.mp. OR sloven*.mp. OR tajik*.mp. OR macedon*.mp. OR turkmen*.mp. OR ukrain*.mp. OR soviet*.mp. OR uzbeg*.mp. OR ussr*.mp. OR yugoslav*.mp.

AND exp Nutritional physiological phenomena/ OR exp Vegetables/ OR exp Fruit/ OR exp Carotenoids/ OR exp Ascorbic acid/ OR vegetable*.mp. OR fruit*.mp. OR caroten*.mp. OR lycopen*.mp. OR ascorbic*.mp.

AND exp Epidemiologic Methods/ OR exp multicenter study/ OR exp comparative study/

limit to humans

mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier, text word

exp …/ = Explode MeSH term

Figure 1.11: Search terms used for MEDLINE search

1.3.1.2 Inclusion and exclusion criteria

Original, quantitative, observational epidemiological studies which described fruit, vegetable, antioxidant intakes or antioxidant status of adult participants who live in CEE or FSU countries and provided comparison populations from Western Europe were included in the review. Based on the data collection methods and reported dietary data, the following studies were considered for inclusion: (1) Dietary surveys:

studies which reported data on fruit and vegetable intake levels using established nutritional assessment methods such as food frequency questionnaire (FFQ), diet history, dietary record and 24-hour diet recall. (2) Health behavioural surveys: reporting data on fruit and vegetable intakes using lifestyle questionnaires with questions regarding fruit or vegetable consumption habits. (3) Antioxidant studies: reporting data on average vitamin C or carotenoid intakes or status (including plasma, serum and adipose tissue concentrations).

Studies were excluded if data collection methods or the inclusion criteria of participants differed substantially between the two regions. Studies which compared dietary habits between the former East and West Germany were used only if their data collection took place before 1991, because food consumption patterns of East Germans seem to have changed rapidly after the reunification (Winkler *et al.* 1998).

To avoid bias towards studies which reported more than one exposure of interest from the same participants, only one set of data from these studies was included in the review: data on carotenoid and vitamin C intake or status were included only if no data on fruit or vegetable consumption were available. If both antioxidant intake and status were reported, only intake data was used, and if data on more than one type of carotenoid concentration were available, only beta-carotene was extracted.

1.3.1.3 Quality assessment

Quality of the included studies was assessed by a modified version of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement (Vandenbroucke *et al.* 2007). Modification of the checklist was necessary because several studies described only the nutritional characteristics of the subjects and the analysis of the relationship with disease outcomes was not reported. Therefore, four items of the statement, which refer to the variables and outcome results of an analytic study (item nos. 7, 11, 15 and 16), were omitted and the assessment was carried out using the remaining 18 items (table I-1 in appendix).

1.3.1.4 Data analysis

Most studies described dietary data of participants from more than one country within a certain region. For these studies, the average values for CEE/FSU and Western Europe were calculated and reported in the review.

To take into account the well-documented difference in fruit and vegetable consumption between Northern and Southern European countries (Agudo *et al.* 2002; Trichopoulou *et al.* 2002), both CEE/FSU and Western European regions were divided into “south” and “north” sub-regions (table I-2 in appendix). If a study reported g/day intake levels of fruits or vegetables of participants from opposite sub-regions, north/south weighting was applied: the intake figure of the “south” country was multiplied with a weighting factor calculated from FAO data (FAO 2015) by dividing the average fruit or vegetable supply of all northern countries of that region between 1970 and 2009 by the specific country`s average supply over the same time period. For studies reporting data on the percentages of participants eating daily fruits or vegetables, or antioxidant data, no such weighting was carried out because appropriate weighting factors were not available.

If data were collected in winter or spring months in one region and during summer or autumn in the other, seasonal weighting of the CEE/FSU data was applied: the intake figures were multiplied with a weighting factor which was calculated from the

Health Alcohol and Psychosocial Factors in Eastern Europe (HAPIEE) study, which is the largest study in CEE/FSU with dietary data (Peasey *et al.* 2006). The weighting factor was determined as the ratio of the energy standardized mean intake level between participants who completed the questionnaire in the summer/autumn months and those who completed it during the winter or spring months. Weighting for seasonal variation was applied only in CEE/FSU because seasonal differences in this region are more substantial than in Western Europe (Powles *et al.* 1996; Capita and Alonso-Calleja 2005; Zatonski 2011).

Most reviewed studies did not report statistical significance of the differences between CEE/FSU and Western Europe. In order to assess whether the reported differences were statistically significant, power calculation was applied. If a study had more than 80% power to show the described difference as statistically significant on the 0.05 significance level, the reported difference was considered statistically significant. If the power was between 20% and 80%, than the observed difference was considered non-significant but the trend was worth noting, and if the power was lower than 20%, the difference was considered negligible. Power calculations were carried out using STATA 13.1 statistical software (StataCorp Texas, USA).

If standard deviation (SD) value was required for power calculation but it was not available from the specific study, the average SD of fruit, vegetable, vitamin C and beta-carotene intake and concentration levels reported in the European Prospective Investigation into Cancer and Nutrition (EPIC) study cohorts was assumed (Agudo *et al.* 2002; Al-Delaimy *et al.* 2004). This assumption was considered appropriate because EPIC is the largest international study with such data available and its results suggest that SD values vary in a narrow range irrespectively of study size and mean intake level. In the study which measured adipose tissue beta-carotene concentration (Kardinaal *et al.* 1993) the SD reported on a subsample of the same study participants were used (Su *et al.* 1998). In studies where south/north or seasonal weighting was applied, SDs were multiplied with the same figures as the mean values.

1.3.2 Results

1.3.2.1 Characteristics of the reviewed studies

Twenty-two studies met the inclusion criteria: ten dietary surveys (Kromhout *et al.* 1989; Winkler *et al.* 1992; Schroll *et al.* 1996; Karamanos *et al.* 2002; Serra-Majem *et al.* 2003; Petkeviciene *et al.* 2009; Lixandru *et al.* 2010; Paalanen *et al.* 2011; Crispim *et al.* 2011; El Ansari *et al.* 2012), six health behavioural surveys (Wardle *et al.* 1997; Prättälä *et al.* 2007; Prattala *et al.* 2009; Hall *et al.* 2009; European Commission 2013; Burisch *et al.* 2014) and six antioxidant studies (Kardinaal *et al.* 1993; Kristenson *et al.* 1997; Bobak, Brunner, *et al.* 1998; Bobak *et al.* 1999; Miere *et al.* 2007; Woodside *et al.* 2013). Figure 1.12 shows the study selection process and table 1.1 describes the main features of the included studies. Most studies were cross-sectional in design or reported cross-sectional data from cohort studies. In two studies (Kardinaal *et al.* 1993; Lixandru *et al.* 2010), data were extracted from case-control setting. Participants from 18 CEE/FSU countries and 18 Western European states were included in the comparisons and most countries were covered by more than one study. The earliest study reported data from the early 1960s, while the latest data collection took place in 2010. Sample sizes ranged from 30 to 85,921 per region. Five studies recruited only males but the majority gave dietary data for both genders.

More than half of the studies applied random sampling method at recruitment and eight used the general population as the sampling frame.

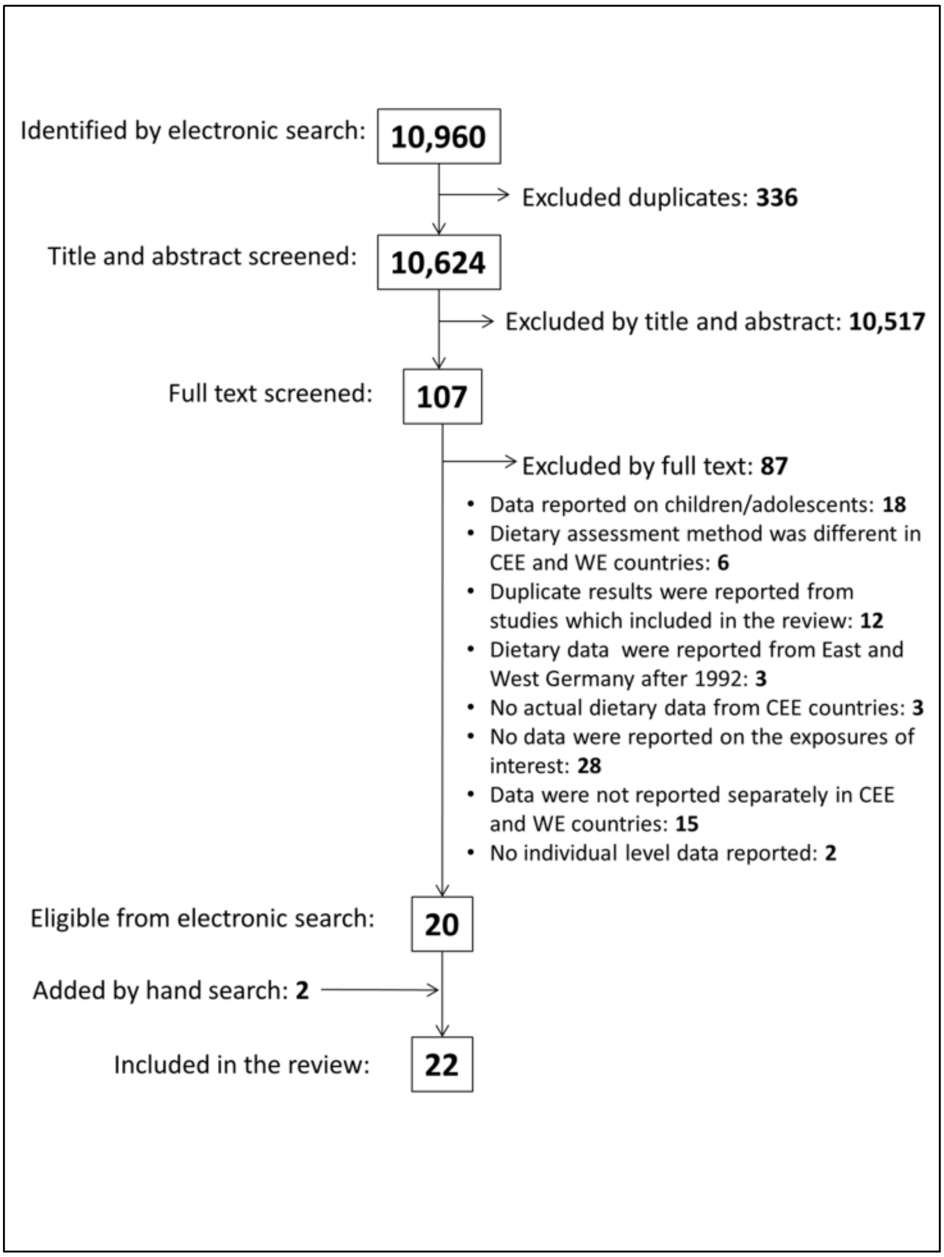


Figure 1.12: Flow diagram of the study selection process

Table 1.1: Characteristics of included studies

1st author, year of publication	Name of study	Examined food or antioxidant	Dietary assessment	Participants` country of origin	Year of data collection	Month of data collection	Sample size	Response rate (%)	Females (%)	Age range or mean (years)	Sampling method	Basis of sample	Quality score[1] (max:18)
1. DIETARY SURVEYS													
Kromhout 1989	Seven Countries Study	Fruits, vegetables	7d record	**CEE:** Yugoslavia	1960-64	Jan-May, Sep	150	nd	0	40-59	random	farm/factory workers, academics	9
				WE: Finland, Italy, Greece Netherlands	1959-65	Feb-Sep	286	nd	0	40-59	random	village inhabitants, railroad workers	
Winkler 1992		Fruits, vegetables	3d record	**CEE:** GDR	1987	Oct-Dec	132	73	0	45-64	random	urban inhabitants	11
				WE: FDR	1984-85	Oct-May	424	70	0	45-64	cluster	urban inhabitants	
Schroll 1996	SENECA	Fruits, vegetables	Diet history	**CEE:** Poland	1993	Jan-Jun	120	51	61	74-79	random	urban inhabitants	13
				WE: Belgium, Denmark, France, Italy, Netherlands, Portugal, Spain, UK, Switzerland	1993	Jan-Jun	1237	51	51	74-79	random	urban inhabitants	
Karamanos 2002		Fruits, vegetables	Diet history	**CEE:** Bulgaria	Nd	nd	288	nd	50	35-60	random	urban inhabitants	14
				WE: Italy, Greece	Nd	nd	1058	nd	54	35-60	random	urban and rural inhabitants	
Serra-Majem 2003	WHO-CINDI	Fruits, vegetables	24hr recall	**CEE:** Poland	1991-94	nd	4440	nd	50	20-65	random	factory workers	14
				WE: Spain	1992	nd	2757	69	nd	6-75	random	general population	
Petkeviciene 2009	NORBAGREEN	Fruits, vegetables	FFQ	**CEE:** Lithuania	2002	Apr	99	68	57	19-75	random	general population	15
				WE: Finland	2002	Jan-May	125	91	nd	25-64	random	general population	
Lixandru 2010		Fruits, vegetables	FFQ	**CEE:** Romania	2005	Apr-Nov	40	nd	30	63	convenience	diabetic patients	12
				WE: Belgium	2005	Apr-Nov	30	nd	20	62	convenience	diabetic patients	
Paalanen 2011		Fruits, vegetables	FFQ	**CEE:** Russia	1992-07	Mar-May	2672	45-92	57	25-64	random	general population	16
				WE: Finland	1992-02	Mar-May	4365	67-81	53	25-64	random	general population	

1st author, year of publication	Name of study	Examined food or antioxidant	Dietary assessment	Participants` country of origin	Year of data collection	Month of data collection	Sample size	Response rate (%)	Females (%)	Age range or mean (years)	Sampling method	Basis of sample	Quality score[1] (max:18)
Crispim 2011	EFCOVAL	Fruits, vegetables	24hr recall	**CEE:** Czech Republic	2007-08	Oct-Apr	118	nd.	51	45-65	convenience	healthy individuals	16
				WE: Belgium, France, Norway Netherlands,	2007-08	Apr-Jul, Oct-Apr	482	nd.	50	45-65	convenience	healthy individuals	
El Ansari 2012	CNSHS	Fruits, vegetables	FFQ	**CEE:** Bulgaria, Poland	2005	nd	1143	95	70	21	convenience	university students	14
				WE: Denmark, Germany	2005	nd	1236	85-92	53	21	convenience	university students	
2. HEALTH BEHAVOIUR SURVEYS													
Wardle 1997	EHBS	Fruits	na	**CEE:** Poland, Hungary, GDR	1989-92	nd	2293	90-100	51	22	convenience	university students	13
				WE: Austria, Belgium, FDR, UK Denmark, Finland, Spain, France, Greece, Iceland, Ireland, Italy, Sweden, Netherlands, Norway, Portugal, Switzerland	1989-92	nd	14,192	90-100	56	21	convenience	university students	
Prattala 2007	Finbalt Health Monitor project	Fruits	na	**CEE:** Estonia, Latvia, Lithuania	1998-02	Apr-May	15,740	62-80	57	20-64	random	general population	16
				WE: Finland	1998-02	Apr-May	9354	65-70	53	20-64	random	general population	
Prattala 2009	EUROTHIENE	Vegetables	na	**CEE:** Estonia, Latvia, Lithuania	2000-04	nd	14,219	60-73	58	20-64	random	general population	15
				WE: Finland, Denmark, Spain, Germany, France, Italy	1998-04	nd	86,924	61-87	51	20-64	random	general population	
Hall 2009	WHS	Fruits, vegetables	na	**CEE:** Bosnia and Herzegovina, Croatia, Czech Republic, Estonia, Georgia, Hungary, Kazakhstan, Latvia, Russia, Slovakia, Slovenia, Ukraine	2002-03	nd	22,475	69-100	53	18-99	random	general population	15
				WE: Spain	2002-03	nd	5448	86	60	18-99	random	general population	

1st author, year of publication	Name of study	Examined food or antioxidant	Dietary assessment	Participants` country of origin	Year of data collection	Month of data collection	Sample size	Response rate (%)	Females (%)	Age range or mean (years)	Sampling method	Basis of sample	Quality score[1] (max:18)
European Commission 2013	EHIS	Fruits, vegetables	na	**CEE:** Bulgaria, Czech Republic, Estonia, Latvia, Hungary, Poland, Romania, Slovakia, Slovenia	2006-09	nd	85,921	56-89	53	15-99	random	general population	na
				WE: Belgium, Greece, Spain, France	2006-09	nd	62,700	60-96	55	15-99	random	general population	
Burisch 2014	ECCO-EpiCom	Fruits, Vegetables	na.	**CEE:** Croatia, Czech Rep, Estonia, Hungary, Lithuania, Moldova, Romania, Russia	2010	Jan-Dec	249	76	42	15+	Convenience	IBD patients (at diagnosis)	16
				WE: Cyprus, Denmark, Finland, Greece, Iceland, Ireland, Israel, Italy, Portugal, Spain, Sweden, UK	2010	Jan-Dec	933	76	46	15+	Convenience	IBD patients (at diagnosis)	
3. ANTIOXIDANT STUDIES													
Kardinaal 1993	EURAMIC	Beta-carotene in adipose tissue	na	**CEE:** Russia	1991-92	nd	200	79-97	0	51	convenience	hospital patients, healthy controls	16
				WE: Finland, Germany, Netherlands, Norway, UK, Spain, Switzerland	1991-92	nd	1180	50-98	0	54	convenience	hospital patients, healthy controls	
Kristenson 1997	LiVicordia	Beta-carotene in plasma	na	**CEE:** Lithuania	1993-94	Oct-Jun	100	83	0	50	random	urban inhabitants	14
				WE: Sweden	1993-94	Oct-Jun	95	83	0	50	random	urban inhabitants	
Bobak 1998		Beta-carotene in plasma	na	**CEE:** Czech Republic	1992	Sep-Nov	136	70	49	40-59	random	urban inhabitants	14
				WE: UK	1991-93	nd	358	73	31	40-59	random	civil servants	

1st author, year of publication	Name of study	Examined food or antioxidant	Dietary assessment	Participants` country of origin	Year of data collection	Month of data collection	Sample size	Response rate (%)	Females (%)	Age range or mean (years)	Sampling method	Basis of sample	Quality score[1] (max:18)
Bobak 1999		Beta-carotene in plasma	na	**CEE:** Czech Republic	1995	Apr-Jun	188	70	0	45-64	random	general population	17
				WE: Germany	1995	Apr-Jun	153	70	0	45-64	random	general population	
Miere 2007		Vitamin C intake	24h recall	**CEE:** Romania	nd	nd	312	nd	87	21	convenience	university students	8
				WE: Spain	nd	nd	918	nd	58	22	convenience	university students	
Woodside 2013	EUREYE	Vitamin C and Beta-carotene in plasma	na	**CEE:** Estonia	2000-03	nd	833	59	66	65+	random	general population	15
				WE: Norway, UK, France, Italy, Greece, Spain	2000-03	nd	3300	36-56	52	65+	random	general population	

WHO-CINDI, World Health Organization Countrywide Integrated Non-communicable Disease Intervention; NORBAGREE, Consumption of vegetables and fruits and other dietary health indicator foods in the Nordic and Baltic countries; EFCOVAL, European Food Consumption Validation; CNSHS, Cross National Student Health Survey; EHBS, European Health and Behaviour Survey; WHS, World Health Survey; EHIS, European Health Interview Survey; EURAMIC, European Community Multicentre Study on Antioxidants, Myocardial Infarction and Breast Cancer; LiVicordia, Linkoping-Vilnius Coronary Disease Risk Assessment Study; ECCO-EpiCom, European Crohn`s and Colitis Organization`s Epidemiological Committee study; FDR, Federal Republic of Germany; GDR, German Democratic Republic; CEE: Central and Eastern Europe (or Former Soviet Union); WE, Western Europe; FFQ, Food frequency questionnaire; na, not applicable; nd, no data available; IBD, Inflamatory bowel disease

1 Based on evaluation using a modified STROBE checklist;

Overall, the quality of the reviewed studies was good. Fifteen studies scored 14 or more points on the 18 point scale and only two scored less than ten points. While most studies gave clear descriptions regarding their design, setting and participants, almost all of them failed to report how the analytical sample size was arrived at (only one study out of 21 met this criterion). Further weaknesses included the lack of detailed discussion of study limitations (10/21) and the lack of description of how potential sources of bias were addressed (10/21) (table I-1 in appendix). Quality of one study (European Commission 2013) was not assessed because it was published as an online database, with no peer-reviewed research paper available.

1.3.2.2 Findings of the reviewed studies

Table 1.2 shows the average intake, percentage and concentration values of CEE/FSU and Western European participants regarding fruit, vegetable and antioxidants reported by the reviewed studies. The directions of the observed differences and the extent of their significance, determined by power calculation, are also summarised.

Most studies reported their results separately for fruits and vegetables and for males and females. Majority of dietary surveys gave average fruit or vegetable consumption values as mean gram per day intakes, and most of the health behavioural surveys as the percentage of the sample who eat these foods at least once a day.

Regarding fruit intake, both dietary and health behavioural surveys showed consistently lower intakes in CEE/FSU compared to Western Europe. Although six out of nine dietary survey comparisons with adequate power found higher vegetable intake in CEE/FSU countries, the estimates were consistently lower in health behavioural surveys. All antioxidant studies indicated lower concentration of beta-carotene in CEE/FSU subjects, but the results for vitamin C were not consistent. No consistent difference was found between males and females.

Table 1.2: Results of the reviewed studies

1st author, year of publication	Unit of measurement	Sex	CEE countries: Average intake, cc. or %	CEE countries: Range[1]	CEE countries: SD	WE countries: Average intake, cc. or %	WE countries: Range[1]	WE countries: SD	Power	SUMMARY: CEE compared to WE[2]
1. DIETARY SURVEYS										
					FRUITS					
Kromhout 1989	g/day intake	M	**58.6**	1.0-153.6	207.3[3]	**132.1**	21.3-310.9	178.3[3]	0.96	**LOWER**
Winkler 1992	g/day intake	M	**98.0**		145.3	**101.0**		164.3	0.05	no difference
Schroll 1996	g/day intake	M	**186.0**		239.1[3]	**234.0**	120.0-532.5	230.2[3]	0.26	lower-ns
		F	**162.0**		210.2[3]	**208.0**	135.0-399.6	202.4[3]	0.43	lower-ns
Karamanos 2002	g/day intake	M	**293.0**		239.1[3]	**315.0**	236.0-355.0	239.1[3]	0.16	no difference
		F	**303.0**		210.2[3]	**325.7**	234.0-377.0	210.2[3]	0.21	lower-ns
Serra-Majem 2003	g/day intake	M+F	**137.0**		224.7[3]	**290.0**		218.0[3]	1.00	**LOWER**
Petkeviciene 2009	p/month intake	M+F	**20.8**		84.3[3]	**29.4**		84.3[3]	0.12	no difference
Lixandru 2010	% eat daily	M	**100.0**		na	**89.5**		na	0.34	higher-ns
		F	**100.0**		na	**100.0**		na	na	no difference
Paalanen 2011	% eat daily	M	**14.0**	2.0-31.0	na	**52.3**	43.0-61.0	na	1.00	**LOWER**
		F	**26.0**	4.0-50.0	na	**73.3**	66.0-82.0	na	1.00	**LOWER**
Crispim 2011	g/day intake	M	**207.0**		176.7	**197.0**	163.0-228.0	175.1	0.07	no difference
		F	**226.0**		155.7	**230.5**	194.0-265.0	151.1	0.05	no difference
El Ansari 2012	% eat daily	M	**31.6**	23.8-39.4	na	**30.4**	28.6-32.1	na	0.05	no difference
		F	**46.8**	39.5-54.1	na	**51.6**	47.8-55.4	na	0.42	lower-ns

1st author, year of publication	Unit of measurement	Sex	CEE countries			WE countries			Power	SUMMARY: CEE compared to WE[2]
			Average intake, cc. or %	Range[1]	SD	Average intake, cc. or %	Range[1]	SD		
VEGETABLES										
Kromhout 1989	g/day intake	M	**240.0**	159.0-276.0	198.2[3]	**102.6**	57.3-227	88.1[3]	1.00	**HIGHER**
Winkler 1992	g/day intake	M	**126.0**		154.8	**124.0**		154.8	0.05	no difference
Schroll 1996	g/day intake	M	**341.0**		154.8[3]	**288.0**	82.4-461.0	128.1[3]	0.63	higher-ns
		F	**297.0**		143.9[3]	**238.0**	77.0-383.0	121.0[3]	0.92	**HIGHER**
Karamanos 2002	g/day intake	M	**243.0**		154.8[3]	**189.0**	168.0-214.0	154.8[3]	0.96	**HIGHER**
		F	**291.0**		143.9[3]	**197.3**	178.0-222.0	143.9[3]	1.00	**HIGHER**
Serra-Majem 2003	g/day intake	M+F	**288.0**		149.4[3]	**97.1**		68.7[3]	1.00	**HIGHER**
Petkeviciene 2009	p/month intake	M+F	**29.9**		56.0[3]	**29.1**		56.0[3]	0.05	no difference
Lixandru 2010	g/day intake	M	**287.0**		189.4	**269.9**		108.1	0.07	no difference
		F	**258.3**		157.9	**283.3**		125.2	0.06	no difference
Paalanen 2011	% eat daily	M	**15.0**	10.0-24.0	na	**48.7**	44.0-54.0	na	1.00	**LOWER**
		F	**22.3**	11.0-35.0	na	**70.7**	69.0-72.0	na	1.00	**LOWER**
Crispim 2011	g/day intake	M	**162.0**		121.1	**201.0**	168.0-222.0	112.8	0.60	lower-ns
		F	**157.0**		99.1	**202.3**	166.0-254.0	108.5	0.87	**LOWER**
El Ansari 2012	% eat daily	M	**37.8**	23.9-51.6	na	**24.4**	23.3-25.4	na	0.99	**HIGHER**
		F	**44.9**	28.0-61.8	na	**42.0**	37.5-46.4	na	0.18	no difference

1st author, year of publication	Unit of measurement	Sex	CEE countries: Average intake, cc. or %	CEE countries: Range[1]	CEE countries: SD	WE countries: Average intake, cc. or %	WE countries: Range[1]	WE countries: SD	Power	SUMMARY: CEE compared to WE[2]
2. HEALTH BEHAVIOURAL SURVEYS										
FRUITS										
Wardle 1997	% eat daily	M	**40.0**	36.0-45.0	na	**42.9**	23.0-78.0	na	0.43	lower-ns
		F	**65.0**	59.0-74.0	na	**61.1**	36.2-86.0	na	0.72	higher-ns
Prattala 2007	% eat daily	M	**11.0**	10.0-12.0	na	**18.0**		na	1.00	**LOWER**
		F	**20.3**	17.0-25.0	na	**36.0**		na	1.00	**LOWER**
EHIS 2013	% eat daily	M	**52.8**	39.4-66.8	na	**60.6**	57.9-66.0	na	1.00	**LOWER**
		F	**67.0**	49.2-82.3	na	**69.1**	62.3-74.5	na	1.00	**LOWER**
Burisch 2014	% eat daily	M+F	**43.4**		na	**54.3**		na	0.87	**LOWER**
VEGETABLES										
Prattala 2009	% eat daily	M	**22.5**	16.1-27.5	na	**32.1**	24.7-39.1	na	1.00	**LOWER**
		F	**30.4**	25.0-33.4	na	**45.9**	36.9-59.1	na	1.00	**LOWER**
EHIS 2013	% eat daily	M	**54.8**	44.2-71.3	na	**68.6**	56.0-82.7	na	1.00	**LOWER**
		F	**62.5**	55.0-78.6	na	**74.2**	65.3-87.4	na	1.00	**LOWER**
Burisch 2014	% eat daily	M+F	**49.0**		na	**60.1**		na	0.88	**LOWER**
FRUITS and VEGETABLES										
Hall 2009	% eat >=5 p/day	M	**18.1**	8.0-44.5	na	**22.0**		na	0.98	**LOWER**
		F	**23.5**	9.4-49.7	na	**24.9**		na	0.38	lower-ns

1st author, year of publication	Unit of measurement	Sex	CEE countries			WE countries			Power	SUMMARY: CEE compared to WE[2]
			Average intake, cc. or %	Range[1]	SD	Average intake, cc. or %	Range[1]	SD		
3. ANTIOXIDANT STUDIES										
				BETA CAROTENE						
Kardinaal 1993	ug/g fatty acid	M	**0.51**	0.45-0.56	0.80	**0.42**	0.18-0.59	0.80	0.31	higher-ns
Kristenson 1997	umol/l cc.	M	**0.38**		0.20	**0.51**		0.32	0.92	**LOWER**
Bobak 1998	umol/l cc.	M	**0.39**		0.26[3]	**0.77**		0.26[3]	1.00	**LOWER**
		F	**0.52**		0.40[3]	**0.97**		0.40[3]	1.00	**LOWER**
Bobak 1999	umol/l cc.	M	**0.11**		0.08	**0.20**		0.21	1.00	**LOWER**
Woodside 2013	umol/l cc	M	**0.25**		0.26	**0.34**	0.19-0.48	0.31	1.00	**LOWER**
		F	**0.36**		0.34	**0.44**	0.30-0.67	0.37	1.00	**LOWER**
				VITAMIN C						
Miere 2007	mg/day intake	M	**80.3**		54.8	**106.2**		83.4	0.77	lower-ns
		F	**88.8**		67.9	**124.4**		94.8	1.00	**LOWER**
Woodside 2013	umol/l cc	M	**42.0**		23.8	**38.0**	32.7-44.4	23.1	0.74	higher-ns
		F	**54.5**		27.7	**48.5**	43.5-52.4	23.4	1.00	**HIGHER**

na, not applicable; cc, concentration;

[1]Range of intake levels, percentages or concentrations if data was reported from more than one country or site; [2]**LOWER**: Intake level, percentage or concentration significantly lower in CEE/FSU countries compared to data from WE, (power>0.80); **HIGHER:** Intake level, percentage or concentration significantly higher in CEE/FSU countries compared to data from WE, (power>0.80); **lower-ns**: Intake level, percentage or concentration lower in CEE/FSU but difference not significant (power<0.80 and >0.20); **higher-ns**: Intake level, percentage or concentration higher in CEE/FSU but difference not significant (power<0.80 and >0.20); **no difference:** power<0.20; [3]SD assumed from EPIC study

1.3.3 Discussion

This systematic review of cross-national studies on fruit and vegetable intake found consistently lower fruit intake figures in CEE/FSU populations compared to Western Europe, but no consistent difference for vegetable intake between the two regions.

These results are congruent with ecological dietary data of food availability based on FBS and HBS. Comparison of average fruit and vegetable supply in CEE/FSU and Western Europe countries between 1970 and 2009 suggests clear difference only for fruits but not for vegetables (FAO 2015). Similarly, comparison of HBS data from DAFNE database indicates that, on average, the availability of fruits is lower but vegetables is higher in CEE/FSU countries (National and Kapodistrian University of Athens 2005).

The inconsistency of the findings regarding vegetable intake can be due to the lack of north/south weighting of health behavioural survey results. For example, in the European Health Interview Survey (EHIS), the largest health behavioural survey included in the review, most participants came from southern countries of Western Europe and northern part of CEE/FSU. If, as a sensitivity analysis, I applied the weighting factors calculated from FAO database for the EHIS results, the comparison showed that the proportion of individuals who consumed vegetables at least once a day was higher in CEE/FSU countries, which is similar to most dietary surveys.

On the other hand, most health behaviour surveys had larger sample size than the dietary surveys, and they are also less prone to measurement error. Furthermore, since the main food sources of beta-carotene are vegetables (Jenab, Salvini, *et al.*

2009), the findings of the antioxidant studies are also in support of the health behavioural survey results and the lower vegetable intake in Eastern Europe.

On the whole, it is not possible to exclude that the reason for the inconsistent results regarding vegetable consumption is that there is no actual difference in intake between CEE/FSU and WE populations.

This systematic review has several limitations. Firstly, it is possible that further published or non-published studies exist which were not identified during the search. However, cross-national studies tend to require substantial funding, logistics and international cooperation between institutions, which often go hand in hand with the endeavour to publish the work in internationally reputable journals which can be found in the electronic databases we searched. In addition, as no language restriction was applied in the electronic search, the possibility of finding studies from non-English speaking countries was increased.

Secondly, the data analysis involved several assumptions. The weighting factors from FAO database and HAPIEE study were the best options currently available for these purposes, and the SD values brought over from EPIC study did not influence the direction of the results, it only helped to decide whether the studies were sufficiently large to draw meaningful conclusions of their findings.

Although the reviewed studies included participants from a large number of CEE/FSU and Western European countries, some of them providing nationally representative food consumption data, specific comparisons were representative only for a small proportion of the whole CEE/FSU and Western European populations. Because large differences exist in fruit and vegetable intakes within the regions, the reported comparisons can only be seen as pixels of a much larger picture. The complete picture will emerge only when nationally representative, comparable dietary data is available for most European countries; in fact, this is the main aim of EFSA`s on-going “EU Menu” project (EFSA 2010).

1.3.4 Conclusion

This systematic review supports previous data that people in CEE/FSU countries consume less fruit than Western Europeans, and that the difference in vegetable intake is probably less clear-cut.

1.4 Czech Republic, Poland and the Russian Federation: CVD mortality and the characteristics of diet

This section focuses on three Eastern European countries which are represented in the multi-centre cohort which is the basis of this thesis. It is therefore important to describe existing information on nutrition in these countries in more detail. Poland and the Czech Republic are two countries in CEE with the first and third largest population in this region, respectively. The Russian Federation is the largest country in the FSU encompassing half of the FSU`s entire population (WHO Regional Office for Europe 2014).

Life expectancy at birth and CVD mortality rates have been very similar in the Czech Republic and Poland over the last 40 years (figure 1.13). Although they have also followed closely the average CEE trend throughout this period, the decline in CVD mortality rates since the early 90s seems to be more pronounced in these two countries than in the CEE as a whole. CVD mortality rates in the Russian Federation have been somewhat higher than the average FSU figures. In fact, in the mid-90s, this country has gone through such a severe mortality crisis which has never been seen before, and the extreme fluctuation in mortality rates persisted even in the 2000s (Shkolnikov *et al.* 2001; Leon *et al.* 2009).

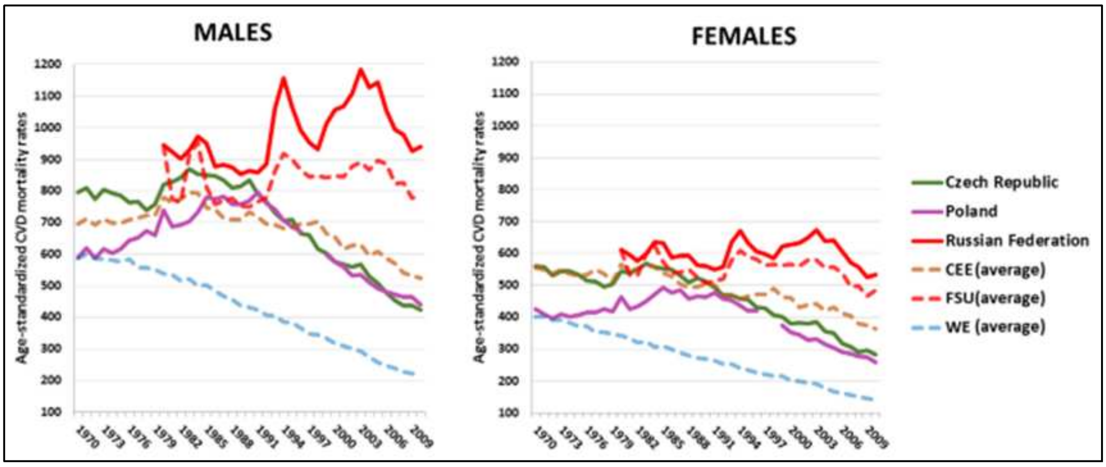
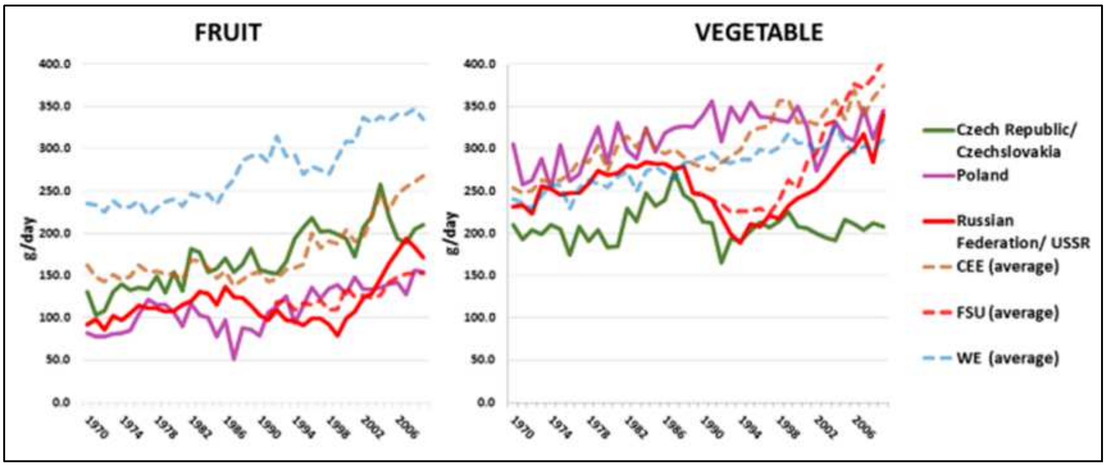


Figure 1.13: Age-standardized CVD mortality rates in the Czech Republic, Poland, Russia, Central and Eastern Europe (CEE), Former Soviet Union (FSU) and Western Europe (WE) between 1970 and 2010 *(Data source: WHO European Health for All Database)*

Availability of fruits and vegetables in Poland, the Czech Republic and Russia, together with the average CEE, FSU and Western European figures between 1970 and 2010 are shown in figure 1.14. Over the 40 years, fruit supply in the Czech Republic and Russia has been more or less typical to the average CEE and FSU trends, respectively. However, the figures suggest that the availability of fruits in Poland have been consistently lower than the CEE average. Regarding vegetables, the availability seems to be the lowest in the Czech Republic and the highest in Poland. Overall, similar to the average regional trends, both fruit and vegetable availability has increased considerably over the last decades in all three countries.



<u>Figure 1.14:</u> Availability of fruits and vegetables in the Czech Republic, Poland, Russia, Central and Eastern Europe (CEE), Former Soviet Union (FSU) and Western Europe (WE) between 1970 and 2010 *(Data source: FAOSTAT)*

Food supply data also suggest that, over the last two decades, the availability of several other food items in the Czech Republic (for example, animal fats and nuts) and Poland (for example, vegetable oil, milk and egg) was similar to the CEE average (table 1.3) (FAO 2015). Similarly, Russian food availability data was close to the FSU average regarding a number of food products (i.e.: animal fat, vegetable oil and meat) during the same period. On the other hand, availability of many food groups (i.e.: vegetable oil and meat for the Czech Republic; animal fat, pulses and nuts for Poland; fish for Russia) differed substantially between the three examined countries and the CEE and FSU regions as a whole.

Table 1.3: Average g/day/capita availability of the different food groups in the Czech Republic, Poland, Russia and the two Eastern European regions between 1991 and 2011 ***(Data source: FAOSTAT)***

Food groups	Czech Republic	Poland	CEE[1]	Russia	FSU[2]
Animal fats	29.7	44.2	27.4	21.9	16.7
Vegetable Oils	48.7	32.0	32.1	27.1	22.1
Meat	226.4	199.8	165.6	148.3	111.0
Fish, seafood	25.5	27.5	16.3	52.4	25.4
Offal	12.6	7.0	9.5	14.0	9.7
Milk	540.4	524.4	493.4	404.6	442.5
Eggs	40.5	28.8	28.3	35.6	22.1
Cereals	331.2	415.4	385.8	413.8	455.1
Pulses	6.1	5.7	10.3	4.5	2.9
Starchy Roots	206.3	357.6	181.6	315.9	238.6
Tree nuts	4.9	2.3	5.1	1.5	3.7
Sugars, sweeteners	119.9	119.3	91.5	118.4	78.1
Alcoholic beverages	18.6	13.6	14.9	25.7	16.3

[1] Average of all Central and Eastern European countries
[2] Average of all Former Soviet Union countries

Nationally representative dietary surveys were carried out in 2003-2004 in the Czech Republic, in 2000 in Poland and annually in Russia as part of the Russian Longitudinal Monitoring Survey (EFSA 2011a; National Research University Higher School of Economics *et al.* 2013). The method used for data collection was repeated 24-hour recall in the Czech Republic and single 24-hour recall in Poland and Russia. The comparability of these survey data with each-other or with intake data collected from the same populations using different dietary assessment methods is limited, as described earlier. For example, the very high fruit and vegetable consumption values in Poland (see figure 1.9) were probably due to the fact that the survey was conducted during the peak season of their intake (Elmadfa 2009).

Eating habits of individuals are often strongly influenced by the traditional dishes and customs of a specific country or region (Shepherd 2005; Abbott *et al.* 2006). As a result of similarities in their history and cultural heritage, the traditional Czech, Polish and Russian cuisines are also similar to each other (and to other Eastern European countries as well). For example, according to a monograph on ethnic foods (Zibart 2001), the central ingredient of many traditional dishes in these countries is meat of any origins, including processed meat (i.e. sausages, salami) and offal (i.e. brain, kidney). Foods are traditionally cooked in animal fat rather than vegetable oil, and large amounts of salt added during cooking or/and after serving is not unusual. Fruits and vegetables are often consumed in a preserved form (i.e. sauerkraut, kompot, picked gherkin). Pastries (i.e. pierogi, pelmeni, groats), high fat dairy products (i.e. sour cream, kefir, cottage cheese) and sweet desserts (i.e. mazurek, babovka) are also popular.

1.5 Selected dietary habits and CVD

In order to contribute to the East-West difference in CVD mortality, fruits, vegetables and other dietary factors need to show different intake levels between regions, and also need to be causally associated with CVD.

1.5.1 Overview

Existing evidence on the relationship between diet and CVD and established CVD risk factors has been summarized by several traditional (Hu and Willett 2002; Bhupathiraju and Tucker 2011) and systematic reviews (Mente *et al.* 2009; Zhao *et al.* 2011; Huang *et al.* 2011; USDA 2014). National (Food Standards Agency 2006; Australian Government 2013; USDA 2015) and international (WHO 2003a) dietary guidelines, which give recommendations on the intake of foods, nutrients and dietary patterns with sufficient evidence for their beneficial or harmful health effects, can also be seen as reviews of the most recent scientific knowledge. For example, the latest dietary guidelines published by the US Department of Agriculture (USDA) recommends eating more fruits, vegetables, whole grains, seafood, legumes and nuts, while the consumption of red and processed meat, refined grains and sugar-sweetened foods and drinks should be restricted. It also encourages moderate consumption of low- and non-fat dairy products and alcohol for adults. Regarding sodium, saturated fat and added sugars, the reduced intake is recommended to be achieved not in isolation but as part of a balanced healthy eating pattern (table 1.4) (USDA 2015).

Table 1.4: Key recommendations of the 2015 US Department of Agriculture dietary guidelines

Groups	Foods/Nutrients
Increased intake advised *(protective factors)*	Fruits Vegetables Legumes Nuts Whole grains Seafood
Reduced intake advised *(risk factors)*	Red and processed meat Sugar-sweetened foods and drinks Refined grains Sodium Saturated fat
Moderate intake advised	Low- or non-fat dairy products Alcohol (for adults)

These reviews and guidelines are based on the results of a large number of observational studies, trials and systematic reviews which investigated the health effects of individual foods, nutrients and dietary patterns.

While summarizing the findings of previous research is useful and necessary, it has been suggested that in nutritional epidemiology no literature review can be completely unbiased (Willett 2013d). This is because although most observational studies collect data on and analyse the health effects of a large number of foods and nutrients, only few of the associations, usually those which are in line with prior expectations, are reported in peer-reviewed journals.

1.5.2 Fruits, vegetables and their relationship with CVD

1.5.2.1 Definition, classification

Although most people have fairly good idea what fruits and vegetables are, their scientific definition is less straightforward. Botanically, fruit is a part of a flowering plant that derives from the flower, ovaries and accessory tissues (Lewis 2002). Vegetable has no botanical meaning. In everyday life, and also in nutritional science, the culinary definitions of fruits and vegetables are used: both are edible parts of a plant, and while fruits are sweet and often eaten raw, vegetables are savoury in taste and often cooked before consumption.

Despite the large variety of fruits and vegetables, their chemical compositions and nutritional properties are surprisingly similar. They are important sources of fibre, vitamin C, carotenes and minerals like calcium and iron (Passmore and Eastwood 1986). However, there are some vegetables which belong to this group by definition, but differ considerably in composition. For example, potatoes and other starchy roots contain large amounts of carbohydrates, and legumes are rich sources of protein. These foods are included amongst fruits and vegetables in some nutritional epidemiological studies but excluded in others. Similarly, the literature regarding the inclusion of fruit and vegetable juices, condiments (ketchup) or jams is inconsistent (Roark and Niederhauser 2013).

These discrepancies have a significant impact on the comparability of the results of any study and emphasize the need of a standard food classification system. The FoodEx2 food classification and description system, developed by the European Food Safety Authority, is a good example of such standardized tools, which helps categorising fruits, vegetables and other food items (EFSA 2015). It consists of a comprehensive list of food products which are aggregated into food groups and larger food categories in a hierarchical way.

1.5.2.2 Relationship with CVD

Between 2005 and 2007, two research groups independently published four systematic reviews with meta-analyses of observational studies examining the association of fruit and vegetable intake with CHD and stroke. Higher fruit and vegetable intakes were significantly related to decreased risk of both disease outcomes in the pooled analyses (Dauchet *et al.* 2005; Dauchet *et al.* 2006; He *et al.* 2006; He *et al.* 2007). Since the publication of these reviews, a number of large scale prospective studies have been published, and majority showed similar results (Nagura *et al.* 2009; Zhang *et al.* 2011; Leenders *et al.* 2013). However, some of them found no significant associations (Bendinelli *et al.* 2011). More recently, Wang and colleagues published another systematic review and meta-analysis which summarised the dose-response effect size of fruit and vegetable intakes on total and cause-specific mortality outcomes (Wang *et al.* 2014). They found that 1 serving per day increase in fruit and vegetable intake was related to 5% and 4% decrease in the risk of total and CVD death, respectively, and they found similar values when fruit and vegetable intakes were assessed separately. Their results also indicated a threshold of 5 portion/day intake, above which no further protective effect seems to exist.

Although the results of observational studies are consistent, no clinical or population-based experimental trials with CVD endpoints using solely fruit and/or vegetable intake as intervention have been conducted. This presents a clear gap in the literature and an important weakness of the evidence (Dauchet *et al.* 2009; Hartley *et al.* 2013).

Other than its link with CVD outcomes, the effects of fruit and vegetable consumption on the occurrence of established CVD risk factors are also important components of the complete picture. The strongest evidence supports the blood pressure lowering effect of fruit and vegetable intake. Both observational studies and clinical trials showed clearly significant anti-hypertensive effect (Appel *et al.* 1997; John *et al.* 2002). Results of large prospective cohort studies also suggest that the consumption of certain types of fruits and vegetables, such as green leafy vegetables, blueberries, grapes or apples, decrease the risk of diabetes mellitus (DM), however no protective effect was shown for fruit and vegetable intake as a whole (Hamer and Chida 2007; Carter *et al.* 2010; Muraki *et al.* 2013). Consumption of fruits and vegetables also seem to have a weight stabilizing effect, although their relationship with BMI and weight loss is not clear (Rolls *et al.* 2004; Dauchet *et al.* 2009; Boeing *et al.* 2012). Majority of the intervention studies which examined the link with lipid levels showed no association (Dauchet *et al.* 2009).

There are several bioactive components of fruits and vegetables which can be responsible for the beneficial effects, including fibres, antioxidant vitamins (vitamin

C and carotenoids), potassium, magnesium and polyphenols (i.e.: flavonoids) (Bhupathiraju and Tucker 2011). Vitamin C, carotenoids and other antioxidant vitamins were especially in the focus of research during the 1990s when several observational studies found their significant associations with reduced risk of CVD, CHD, stroke and cancer (Voutilainen *et al.* 2006; Bhupathiraju and Tucker 2011). However, subsequent randomized controlled trials (RCTs) showed no associations with most disease outcomes, and some even found increased risk (Vivekananthan *et al.* 2003; Bjelakovic *et al.* 2007; Druesne-Pecollo *et al.* 2010). Although a number of hypotheses have been proposed to explain this discrepancy, the most likely explanation is that the negative relationship described in observational studies is a result of residual confounding, and that RCTs reflect the real association between these antioxidants and the examined diseases (Vivekananthan *et al.* 2003; Lawlor *et al.* 2004). Concentration of vitamin C and carotenoids in various body tissues are good indicators of fruit and vegetable consumption of individuals and populations. Consequently, they are often used to validate dietary data regarding fruit and vegetable intakes in epidemiological studies (Jenab, Slimani, *et al.* 2009).

Overall, the evidence for the protective effect of fruit and vegetable consumption against CVD seems to be strong, but further studies with experimental design and CVD clinical endpoints would be necessary to confirm the findings of previous research.

Most studies of fruit and vegetable intake and health outcomes have been carried out in Western European or North American population samples. Despite the fact that the WHO Global Burden of Disease project estimated that the disease burden due to inadequate fruit and vegetable consumption is higher in CEE/FSU than any other parts of the world (WHO 2009; Lim *et al.* 2012), reliable individual-level dietary data in CEE and FSU countries are scarce and, to date, no well-powered studies of fruit and vegetable intakes in relation to CVD have been reported in the region.

1.5.3 Dietary patterns and CVD

In order to emphasise the importance of the diet as a whole in the development of chronic diseases, the focus of nutritional epidemiology has shifted over the last two decades from single foods and nutrients towards dietary patterns (Hu 2002; Kant 2004; Kant 2010; Tucker 2010; Bhupathiraju and Tucker 2011). Studies which investigate the intakes of individual foods or nutrients in relation to disease outcomes have a number of inherent limitations (Slattery 2010; Willett 2013d; USDA 2014). (1) Foods or nutrients are rarely consumed on their own, most often they are eaten in combination with each other. Consequently, studies that focus on one food or nutrient do not reflect real life circumstances and usually do not allow inferences to the overall diet. (2) Foods and nutrients in our diet can interact with each other or confound each other`s health effect. These inter-relationships between the components of diet cannot be taken into account by these studies. (3) Health effects of some foods or nutrients can be small and may remain undetected even if large sample size is applied. (4) When modifying a person`s diet, substitution-effect can occur: increased intake of one dietary component might result in the reduction of the other. Studying dietary patterns allows us to overcome many of these limitations.

On the other hand, the methods of dietary pattern analysis have important disadvantages as well (Kant 2004; Newby and Tucker 2004; Waijers *et al.* 2007; Tucker 2010; Willett 2013d). (1) The construction of predefined diet quality scores (“a priori” method), as well as the statistical techniques applied in the “a posteriori” method (see below) require the researchers to make many arbitrary decisions without clear standardized guidelines. This subjectivity has been often in the centre of criticism of dietary pattern analysis (Martínez *et al.* 1998; Jacques and Tucker 2001; Newby and Tucker 2004). (2) Public health interventions cannot be easily designed based on dietary pattern analyses because they do not provide sufficiently specific information on which area of the diet needs special attention. (3) If the effect of a dietary pattern on a disease outcome is mediated through a specific food or nutrient, which is not carefully separated from the overall pattern, this effect could be easily overlooked.

On the whole, dietary patterns reflect a comprehensive picture of diet and provide a holistic approach to study the relationship between diet and health. However, because of the methodological pitfalls, care is needed when the methods of dietary pattern analysis are applied. Dietary pattern analysis does not necessarily represent higher quality research than the reductionist studies which are focused on individual foods or nutrients. It should rather be seen as a complementary strategy (Willett 2013d; USDA 2014).

Two main approaches have been used for dietary pattern analysis. The “a priori” method uses predefined diet quality scores, and ranks individuals based on how closely they follow healthy eating patterns or dietary guidelines. On the other hand, the “a posteriori” or “data-driven” method applies statistical techniques (most often principal component analysis or cluster analysis) to determine the inherent nutritional characteristics of the study population. Finally, reduced rank regression is often referred to as a hybrid between “a priori” and “a posteriori” methods. It identifies the combination of dietary intakes that best explains the variance in a set of intermediate markers of a disease (Hoffmann *et al.* 2004; Kant 2010; Tucker 2010).

Predefined diet quality scores can assess the adherence to (1) healthy diet patterns, or (2) national or international dietary guidelines (Waijers *et al.* 2007). Mediterranean diet score (MDS) is probably the best well-known example for the former group, and the healthy diet indicator (HDI) belongs to the second category.

1.5.3.1 Mediterranean diet score (MDS)

Mediterranean diet is the traditional eating pattern of populations around the Mediterranean Sea in Southern Europe (Keys 1980; Trichopoulou *et al.* 2014). It is usually characterised by high consumption of fruits, vegetables, legumes, cereals, fish and olive oil, low consumption of milk and meat and moderate intake of alcohol (Trichopoulou *et al.* 2005; Bach *et al.* 2006). Mediterranean diet score (MDS), the indicator of someone`s adherence to the Mediterranean diet, was first introduced by Trichopoulou in 1995 (Trichopoulou *et al.* 1995). It consisted of eight components and applied dichotomous scoring system (table 1.5). Sex-specific median intake levels were used as cut-off values between those who scored zero and one points for the various components. Alcohol intake was an exception: those with moderate consumption scored one point while low and high consumers scored zero. The overall MDS was calculated by adding up the individual component scores. Although several modified versions of the original indicator has been developed since then, its association with chronic diseases, including CVD, shows largely consistent beneficial results across studies (Sofi *et al.* 2008). The most recent systematic review and meta-analysis of observational studies found that 2-point increase in the MDS was related to 8% decrease in total and 10% decrease in CVD mortality risk (Sofi *et al.* 2014). What makes this dietary pattern unique in nutritional epidemiology is the fact that it has shown to be significantly protective against CVD not just in observational studies but in primary and secondary prevention trials as well (De Lorgeril *et al.* 1996; Estruch *et al.* 2013). In addition, consistent inverse associations of the MDS have been shown with diabetes mellitus, obesity, metabolic syndrome, depression and some other chronic conditions (Kastorini *et al.* 2010; Kastorini *et al.* 2011; Rees *et al.* 2013; Psaltopoulou *et al.* 2013; Koloverou *et al.* 2014; Chiva-Blanch *et al.* 2014).

Table 1.5: Components and scoring criteria of the original Mediterranean diet score (MDS) by Trichopoulou and colleagues (Trichopoulou *et al.* 1995)

MDS components	MDS component scores	
	0 point	**1 point**
Fruits and nuts	Below median	Above median
Vegetables	Below median	Above median
Legumes	Below median	Above median
Monounsaturated vs. saturated fatty acid ratio	Below median	Above median
Cereals	Below median	Above median
Meat and meat products	Above median	Below median
Milk and dairy products	Above median	Below median
Alcohol	M: 10-50g/d; F: 5-25g/d	M: <10 or >50g/d F: <5 or >25g/d

median= sex-specific median

A major disadvantage of the various MDSs is that their component scores are given based on sample-specific cut-off values (usually sex-specific medians) which can differ greatly between studies and not necessarily reflect the threshold between healthy and unhealthy intake levels (Waijers *et al.* 2007). In addition, relative cut-off points do not allow comparison of MDSs between populations. More recently, Sofi and colleagues developed a scoring system that applies absolute cut-off values which were determined by systematic literature review of previous MDS studies using data from more than 4 million individuals and 35 prospective cohorts (Sofi *et al.* 2014). In order to determine the cut-off values between the component scores for each food groups, the authors first calculated the mean values of the weighted medians published in previous MDS studies. Then, in the second step, the actual absolute cut-offs were determined by using the ±1SD values around this mean. Although this newly developed MDS has never been tested in relation to disease outcomes to date, it has the potential to overcome the above mentioned limitations.

The associations between MDS and mortality outcomes have been investigated primarily in Southern European population samples, and relatively few studies examined the link in non-Mediterranean individuals. I found one study which was carried out in Eastern European participants, but the restricted age range (75-80 years) and the small sample size (n=411) seriously limits the generalizability of its findings (Frackiewicz *et al.* 2010).

1.5.3.2 Healthy diet indicator (HDI)

Diet quality scores which measure adherence to dietary guidelines have been also linked with CVD risk in observational studies, however, the strengths of the associations were usually modest and the overall results were less consistent (Kant 2004; Kant 2010; Waijers *et al.* 2007; Fransen and Ocke 2008). The healthy diet indicator (HDI) was originally developed in 1997, reflecting the WHO`s 1990 dietary recommendations for the prevention of chronic diseases (WHO 1990; Huijbregts *et al.* 1997). The indicator consisted of nine components, and, similarly to the original MDS, dichotomous scoring system was applied: participants scored one point for each specific component for which their dietary intake was within the recommended range, and no points were given if the intake level was outside this range (table 1.6). The overall HDI score was the sum of the individual component scores. Being based on international guidelines, its application is not restricted to a specific country or region, thus often used in cross-cultural settings. It has been shown to be associated with overall and cardiovascular disease (CVD) mortality (Huijbregts *et al.* 1997; Knoops *et al.* 2006); an inverse but not statistically significant association was observed in a recent Swedish study using an adapted score (Sjogren *et al.* 2010).

Table 1.6: Components and scoring criteria of the healthy diet indicator (HDI) by Huijbregts and colleagues (Huijbregts *et al.* 1997)

HDI components	**HDI component scores**	
	0 point	**1 point**
Saturated fatty acids, *energy%*	>10	0-10
Polyunsaturated fatty acids, *energy%*	<3 or >7	3-7
Complex carbohydrates, *energy%*	<50 or >70	50-70
Mono- and disaccharides, *energy%*	>10	0-10
Protein, *energy%*	<10 or >15	10-15
Cholesterol, *mg/day*	>300	0-300
Fruits/vegetables, *g/day*	<400	≥400
Pulses/nuts/seeds, *g/day*	<30	≥30
Dietary fibre, *g/day*	<400	≥400

energy % - Percentage of alcohol-free energy intake

1.6 Summary

As a result of the different political systems, wide health gap between Eastern and Western Europe developed over the second half of the 20th century, which was primarily due to the high CVD mortality rates in the East. While alcohol consumption and smoking have been shown to be important lifestyle factors in this context, the evidence regarding dietary habits is inconclusive and largely based on ecological data. Comparable individual-level dietary data in CEE and FSU countries are scarce, and there are few studies which compared fruit and vegetable intakes directly between Eastern and Western European population samples.

Strong body of evidence suggests that increased fruit and vegetable intake can reduce the risk of CVD. Similarly, high adherence to healthy dietary patterns, such as the Mediterranean-style diet or the diet that follows the WHO nutritional recommendations, has been shown to be related to lower CVD risk. Despite the large number of observational epidemiological studies carried out worldwide in this topic, virtually no studies examined these relationships in large Eastern European population samples.

- CHAPTER 2 -

AIMS AND OBJECTIVES

This chapter outlines the aims and objectives of the thesis.

2.1 Aims

The overall aim of the thesis is to assess and compare dietary intake habits of adult participants of large Eastern and Western European population-based cohorts, and to investigate the relationship between selected dietary habits and all-cause and cause-specific mortality in the Eastern European cohorts.

Achieving these aims is important for a number of reasons. First of all, as described in the background chapter, the hypothesis that unhealthy diet contributes to the high CVD morbidity and mortality rates in Eastern European countries is mainly supported by ecological data but individual-level evidence is limited. This PhD work seeks to fill this gap in the literature and help to better understand the role of diet in the poor health of Eastern European populations. Secondly, the results will contribute to the general discussion on the relationship between diet and health. Replication of previously established analyses in population samples with different covariate structure can make the existing evidence more robust. Diet is associated with many other factors which may confound the link between diet and health, but the association of diet with confounders (e.g. socioeconomic status) is likely to differ between populations. If the associations between the examined dietary factors and health outcomes in Eastern Europe are consistent with previous studies, the overall evidence for these relationships will become more robust. Finally, the findings can be used to support public health intervention campaigns in the Czech Republic, Poland and Russia, and possibly other Eastern European countries. The results will provide some guidance as to whether dietary interventions have the potential to reduce CVD burden in these countries; by focusing on specific food groups (i.e.: fruits and vegetables), the thesis can provide some evidence for targeted dietary campaigns.

2.2 Objectives

In order to achieve these aims, specific objectives are identified. These are:

Objective 1: To assess the consumption of foods and nutrients, estimated using food frequency questionnaires (FFQ), of Czech, Polish and Russian participants of the HAPIEE study and compare it with British individuals from the Whitehall II cohort. The contribution of fruit and vegetable intakes to the mortality differences between cohorts will be also assessed.

As detailed in the background chapter, very few studies have compared dietary intakes directly between Eastern and Western European individuals. The systematic literature review showed that no previous dietary surveys have carried out such comparison regarding individual-level fruit and vegetable intakes in large sample size such as the HAPIEE and Whitehall II studies. While keeping the main limitation in mind, which is due to the fact that neither study populations are fully representative to their respective countries, let alone the entire Western and Eastern European regions, the results of this analysis will indicate whether there are any differences in eating habits which worth investigating further, or whether there are any food groups/nutrients which might be candidates for potential targeted public health interventions in CEE/FSU countries. Estimation of the extent by which fruit and vegetable consumption contribute to the mortality differences between cohorts will indicate the possible importance of these foods for the East-West health divide.

Objective 2: To investigate the association between fruit and vegetable intakes and all-cause and CVD (including CHD and stroke) mortality in participants of the HAPIEE study. The mediating effect of blood pressure between fruit/vegetable intake and mortality, as well as the proportion of deaths which could be prevented if the individuals` fruit and vegetable intakes were increased will be also estimated.

Although inadequate fruit and vegetable intake has been often suggested as important reason for the high CVD mortality and morbidity rates in CEE and FSU countries, the relationship between these dietary habits and health outcomes has not been examined empirically in large Eastern European population cohorts. In addition to filling this gap in the literature, the results will also help estimation of the health benefits which would be realized if fruit and vegetable consumption increased in the populations. Consequently, this will provide a guidance about the potential public health value of dietary interventions. This analysis will also contribute to the scientific debate whether the health protective effect of fruit and vegetable intake is mediated through the lowering of blood pressure.

Objective 3: To estimate the association between the healthy diet indicator (HDI) score and all-cause, CVD, CHD, stroke, cancer and other cause (non-CVD-non-cancer) of death in participants of the HAPIEE study, using a newly developed version of the HDI which is based on adherence to the WHO dietary guidelines published in 2003 and applies continuous scoring system to determine component scores.

As explained in the background chapter, “a priori” diet quality scores have been developed to characterize an individual`s overall diet. HDI is primarily a nutrient based diet quality score which is often used in international settings but never before tested in large Eastern European populations. The findings of this analysis will show how well the HDI, and the underlying adherence to the WHO dietary guidelines, predicts mortality outcomes in Eastern European settings.

Objective 4: To examine the association between the Mediterranean diet score (MDS) and all-cause, CVD, CHD and stroke mortality in participants of the HAPIEE study, using a recently proposed scoring system which applies absolute cut-off values to determine component scores.

Mediterranean diet score (MDS) is primarily a food based diet quality index which, similarly to HDI, has not been applied in large Eastern European populations studies before. In fact, MDS with the recently proposed absolute scoring system has not yet been tested in any populations. Consequently, the results will provide evidence whether MDS is a suitable indicator of healthy diet in Eastern Europeans, and also, whether the new version of the MDS performs as well as previous ones in predicting mortality outcomes. Due to the food based nature of this score and the fact that the absolute scoring method clearly indicates recommended intake levels of the various food groups, the results of this analysis will be relatively easy to translate into public health interventions.

- CHAPTER 3 -

METHODS

This chapter describes the dataset and analytical methods used in this thesis. The characteristics of the HAPIEE study population and the methods which were used for collecting data on dietary habits, mortality follow up and covariates in this study (forming the core dataset for my analytical work) are detailed in the first part of the chapter (sections 3.1, 3.2). Sections 3.3, 3.4, 3.5 and 3.6 give further information on how missing data was dealt with and how the analytical samples were selected, as well as some common characteristics of the statistical methods. In order to reach the four objectives set out in the previous chapter, I carried out four distinct epidemiological analyses. The analysis-specific methodological steps, including the construction of exposure variables and the application of statistical procedures, are described in the final part of the chapter (section 3.7).

3.1 Study populations

Most of the analytical work in this thesis has been carried out using data collected from the Czech, Polish and Russian participants of the Health Alcohol and Psychosocial Factors in Eastern Europe (HAPIEE) prospective cohort study. In addition, data from the Whitehall II cohort study of British civil servants was used when dietary intakes between the participants of the two studies were compared (Objective 1). The following sections give an overview of the HAPIEE study, while the characteristics of the Whitehall II cohort are described in the analysis-specific section (Section 3.7.1).

3.1.1 The HAPIEE study

The HAPIEE study is a recent and one of the largest studies in CEE and FSU with available data on dietary habits of general population samples. This multi-centre prospective cohort study was designed to investigate the relationship between traditional, non-conventional and psychosocial risk factors and chronic non-communicable diseases, particularly CVD, in middle-aged and older individuals in Eastern Europe. A detailed description of the study`s rationale, protocol and data collection procedures has been published previously by Peasey (Peasey *et al.* 2006). The baseline survey was carried out between February 2002 and July 2005, and it recruited population samples of men and women aged 45-69 years in Novosibirsk (Russia), Krakow (Poland) and six towns (Havířov/Karviná, Jihlava, Ústí nad Labem, Liberec, Hradec Králové, and Kroměříz) in the Czech Republic (figure 3.1). In 2006, further participants were recruited in Kaunas (Lithuania). As no dietary data was collected in the Lithuanian arm of the study, these individuals were not included in the thesis. The methodological description will therefore focus on the Czech, Polish and Russian arms of the HAPIEE study.

Figure 3.1: Geographic location of HAPIEE cohorts *(Data source: Googlemaps)*

Novosibirsk is the third largest city in the Russian Federation with a population of 1.5 million. It is the administrative, cultural and commercial capital of the Siberian Federal District. Although it is located in the Asian continent, the cityscape is not different from any European metropolis. Existing data suggest that, in terms of CVD mortality trends and some selected lifestyle factors, Novosibirsk is similar to other urban areas in Russia (Malyutina *et al.* 2001; Malyutina *et al.* 2002). Participants of the study were selected from two separate districts of the city, each with different socio-economic profiles.

Krakow is situated in Southern Poland. It has approximately 760,000 inhabitants (1.4 million including urban agglomeration) which makes it the second largest city of the country. Being a major economic, cultural and educational centre, Krakow and the surrounding Malopolskie region have lower unemployment rates and higher life expectancy than most other areas in Poland (Central Statistical Office of Poland 2015). For the purpose of the study, four city districts were selected which represented different levels of the socio-economic spectrum.

The six towns in the Czech Republic also have varying socio-economic profiles. They include former coal mining town with high unemployment rate (Havířov/Karviná) as well as a market town with relatively prosperous population (Hradec Králové). Their respective population ranges from 30,000 (Kroměříž) to 140,000 (Havířov/Karviná), giving the combined population of approximately 600,000.

Participants of the study were selected using stratified random sampling in all three countries. Eligible individuals were identified using national and regional population registers in the Czech Republic and Poland, respectively, and electoral list in Russia. The sampling frame consisted of all inhabitants who were between the age of 45 and 69 years on the 1st July 2002 and lived in the two selected districts of Novosibirsk, the four selected districts of Krakow or any of the six Czech towns. In all three centres, the eligible subjects were stratified by sex and five year age groups in order to make sure that equal number of individuals was invited to participate in all age groups and both sexes.

The target sample size was 10,000 participants per country cohort. Eventually, a total of 28,947 persons were recruited with an overall response rate of 59% (table 3.1) (Peasey *et al.* 2006). All participants signed informed consent form. The study protocols were approved by ethical committees at University College London and all participating centres.

Table 3.1: Number of participants and response rates in the three HAPIEE cohorts (Peasey *et al.* 2006)

Country cohort	No. participants	Response rate
Czech Republic	8856	55%
Poland	10,728	61%
Russia	9363	61%
Total	28,947	59%

3.2 Measurements

As part of the baseline survey, participants filled in an extensive questionnaire, underwent a short medical examination and provided blood samples. In the Czech Republic and Poland, the main questionnaire was completed in the participants` home during a visit by a research nurse. After the visit, participants completed the

FFQ which was then checked for completeness at a subsequent visit to a study clinic where the medical examination took place. In Russia, all questionnaires (including the FFQ) were completed with a nurse in a clinic at the same day as the medical examination. As a result of this difference in methodology, the proportion of subjects with complete data was nearly 100% in the Russian cohort, compared to 87% in Poland and 82% in the Czech Republic (Peasey *et al.* 2006).

The structured questionnaire covered a wide range of health-related topics. The questions aimed to collect information on the participants` health (both physical and mental health), health behaviour/lifestyle (including diet, smoking habits, alcohol intake, physical activity, etc.), past and present socio-economic characteristics, psychosocial factors and physical functioning. The original questionnaire was developed in English, which was then translated into Czech, Polish and Russian languages. In order to check accuracy and consistency, all three non-English versions were also translated back into English.

The medical examination included measurements of anthropometry (weight, height, waist circumference, etc.), blood pressure, lung function and cognitive function.

Blood samples were collected in Becton Dickinson SST II (1x10ml) and K_2-EDTA (1x10ml and 2x3ml) vacutainers. In order to separate plasma and serum samples, the 10ml vacutainers were centrifuged at 4000rpm for 15 minutes. Plasma samples were then divided into three, and serum samples into four aliquots. One 250µl aliquot of plasma was prepared in a way to make it suitable for measuring vitamin C concentration by adding 250µl of 10% metaphosphoric acid stabiliser. All blood, serum and plasma samples were subsequently stored at -80ºC (Peasey *et al.* 2006).

Concentrations of total cholesterol, HDL cholesterol and triglycerides of all individuals were determined in local laboratories in the Czech Republic, Poland and Russia. Concentrations of other selected compounds, including vitamin C and beta-carotene, were measured on a random subsample of 3000 participants (1000 per country cohort). In this case, the laboratory analysis was carried out in a central laboratory (Clinical Trial Service Unit, Oxford).

3.2.1 Dietary assessment

Dietary data were collected using a semi-quantitative FFQ which was based on the instrument developed by Willett and colleagues (Willett *et al.* 1985) and subsequently modified for the Whitehall II study (Brunner *et al.* 2001). Detailed description of the FFQ and the process how the dataset regarding nutrient intakes was compiled has been provided by Boylan (Boylan *et al.* 2009). Briefly, the list of foods and drinks on the FFQ consisted of 136, 147 and 148 items in the Czech Republic, Russia and Poland, respectively (see appendix II). The only differences in the country-specific FFQ versions were those due to country-specific food items, which were added by local dietitians. Participants indicated how frequently they consumed a given amount (usually medium serving or average size) of a particular food or drink item during the previous three months. The nine possible answers ranged from “never or less than once a month” to “six or more times a day”. As mentioned above, the FFQ was self-administered in the Czech and Polish cohorts and subsequently checked for completeness or unclear entries by a nurse in a clinic, while it was completed during an interview with a research nurse in Russia.

In order to estimate daily intakes of foods and drinks, FFQ answers were first converted into portion/day intakes (table 3.2). Gram/day intakes were then calculated by multiplying these values with the portion sizes determined by local dietitians.

Table 3.2: Conversion of FFQ answers to portion per day intakes

Consumption of a food or drink item indicated on the FFQ	Portion per day consumption
6+ per day	6.0
4-5 per day	4.5
2-3 per day	2.5
1 per day	1.0
5-6 per week	0.79
2-4 per week	0.43
1 per week	0.14
1-3 per month	0.07
<1 portion per month; no data	0.0

Daily intakes of energy and 41 nutrients were calculated by adding together the amounts consumed through the individual food and drink items. To do this, food consumption tables of energy and nutrient content of foods and drinks were required. Since the existing country-specific food composition tables were not comparable with each other, the McCance & Widdowson food composition table, the most comprehensive database available, was used to estimate nutrient content of most (92%) foods and drinks (McCance and Widdowson 2002). In case of some country specific foods (eight Polish and two Russian items), local food composition tables had to be used for this purpose. In addition, the United States Department of Agriculture Nutrient Database and manufacturer data were also used for one item each (Boylan *et al.* 2009). The amount of nutrient and energy consumed by each individual in the study was calculated with the Wfood 2002 nutrient analysis software which had been developed previously for the Whitehall II study.

3.2.2 Mortality follow up

Deaths in the three cohorts were ascertained using the city death register in Novosibirsk, the city and regional death register in Krakow, and the national death register in the Czech Republic. Linkage of study participants with data from these registers was possible through their national insurance number in Krakow and the Czech Republic, and by matching name and date of birth in Novosibirsk. Mortality data for the thesis were available until the 31st December 2010 in Russia and Poland and until the 31st December 2011 in the Czech Republic.

In addition to total mortality, data on the causes of death were also available. Coding of the cause of death was based on the 9th and 10th revision of the International Classification of Diseases (ICD) (WHO 2015b): CVD (ICD-9: 390-459; ICD-10: I00-I99), CHD (410-414; I20-I25), stroke (430-438; I60-I69), cancer (140-239, C00-D48). Deaths were categorized as other-cause (non-CVD-non-cancer) if the main underlying cause of death was coded with any other ICD codes.

There were 1183 (4.1%) participants in the study who did not provide consent with follow up or who did not have national insurance number. Mortality data from these participants were not available. Furthermore, there were 127 (0.4%) subjects who died during follow up but data on the cause of their death was not available. These individuals were included in the analysis of total mortality but excluded from analyses of cause-specific mortality. Finally, 198 (0.7%) participants were lost during follow up; their records were censored at the last date of contact through postal questionnaires sent out to participants in 2006 and 2009.

3.2.3 Covariates

In addition to outcomes and main exposures, several other variables were considered as possible confounders, effect modifiers or intermediate variables and were included in the statistical analyses. These covariates were largely selected on the basis of previous knowledge. This section gives a detailed description of how data on these selected covariates were collected and prepared for analysis. The covariates are categorised into three groups: (1) socio-demographic factors, (2) lifestyle factors and (3) anthropometric, biological and medical factors (table 3.3).

Table 3.3: List of covariates used in the statistical analyses (see detailed description of variables in text)

COVARIATE	TYPE	CATEGORIES
Socio-demographic factors		
Sex	Binary	Male Female
Age	Continuous	
Marital status	Binary	Married/cohabiting Single/divorced/widowed
Education	Ordinal	Primary or less Vocational Secondary University
Household amenities score	Ordinal	Low Medium High
Employment status	Nominal	Employed Retired Non-employed-non-retired
Lifestyle factors		
Smoking	Nominal	Current smoker Ex-smoker Never smoker
Alcohol intake	Ordinal	Abstainer Light to moderate drinker Heavy drinker
Physical activity	Ordinal	Low Moderately active High
Vitamin supplement usage	Ordinal	Non-user Irregular user Regular user
Anthropometric, biological and medical factors		
Body mass index (BMI)	Continuous	
Obesity	Binary	Obese Not obese
Mean arterial blood pressure (MAP)	Continuous	
Hypertension	Binary	Hypertensive Not hypertensive
Plasma cholesterol cc.	Continuous	
Hypercholesterolemia	Binary	Hypercholesterolemia Not hypercholesterolemia
Medical history (CVD, diabetes)	Binary	Positive Negative

3.2.3.1 Socio-demographic factors

Sex. The study population included both males and females. Information on the participants` sex was collected by the structured questionnaire.

Age. The exact age of participants was calculated as the time in years between the date of birth and the date when the questionnaire was completed. It was included in all analyses as continuous variable.

Marital status. Participants were asked about their marital status in the questionnaire and five possible answers were given: single, married, widowed, divorced or cohabiting. In all analyses, marital status was applied as a binary variable with two groups based on whether the participant lived with a companion (married or cohabiting) or not (single, divorced, widowed).

Education. Educational attainment was assessed as the highest completed level of education. According to the six possible answers, participants were categorised into four groups: (1) Primary, incomplete or no formal education, (2) vocational training, (3) secondary education and (4) university or college degree.

Household amenities score. Household amenities score was used as an indicator of the participants` socio-economic position. Subjects were asked how many of the following 12 household items they possessed: microwave, video recorder, colour TV, washing machine, dishwasher, car, freezer, holiday cottage, video camera, satellite/cable TV, telephone, mobile phone. The household amenities score was considered low if less than five items were indicated, moderate between five and seven, and high if eight or more items were answered.

Employment status. When the dietary habits between HAPIEE and Whitehall II participants were compared, household amenities score could not be used as the indicator of the participants` socio-economic position because no such questions were asked in the Whitehall II study. In order to include some kind of information on economic activity, data on employment status, which was comparable across cohorts, was used in these analyses. Participants were grouped in three categories (employed, retired, non-employed-non-retired).

3.2.3.2 Lifestyle factors

Smoking habits. Participants were asked if they smoked cigarettes. Based on the answer, subjects were grouped into never smokers, ex-smokers, and current smokers. In some sensitivity analyses regarding current smokers (see section 4.3.3), the number of cigarettes smoked per day and the number of years how long they had smoked was also taken into account as possible confounder.

Alcohol intake. Alcohol consumption of participants in the HAPIEE study was assessed by several methods. In this thesis, in order to take alcohol intake into account as a possible confounder, indices derived from the graduate frequency questionnaire (GFQ) were used. The GFQ measured how frequently the participants consumed specific amounts of alcohol over the previous year. The amounts of alcohol were expressed in local units and ranged from half to ten or more drinks across six categories. The frequency of intake could be indicated on a 9-point scale ranging from never to daily/almost daily. Based on the answers, the annual and daily alcohol consumption could be estimated. In all analyses, study participants were grouped into three categories according to their daily alcohol intake: (1) abstainers (0g/day alcohol intake); (2) light to moderate drinkers (<15g/day alcohol intake for women, and <20g/day intake for men); (3) heavy drinkers (≥15g/day alcohol intake for women, and ≥30g/day intake for men). The cut-off values were selected in line with current guidelines (USDA 2010).

Physical activity. In order to estimate physical activity of participants, leisure time and occupational activities were both taken into account. The sources of the information on leisure time activity were two questions in the questionnaire which asked how many hours per week the subjects spend with (1) sport activities and (2) household activities, such as gardening, housework or maintenance (i.e.: DIY). Information on occupational activity also came from two questions: (1) the current economic activity of the participant (employed, owner of a company, self-employed, housewife, farmer, pensioner-still employed, pensioner-not employed, unemployed); (2) the way how participants described their job (sedentary, standing, manual, physical).

Participants were categorised into overall physical activity groups through a number of steps. The applied method was similar to the procedures used previously in the European Prospective Investigation into Cancer and Nutrition (EPIC) studies which showed fairly good agreement with accelerometer data (Friedenreich *et al.* 2006; Cust *et al.* 2008). Firstly, using the Compendium of Physical Activities (Ainsworth *et al.* 2011), metabolic equivalent intensity values (MET, defined as the ratio of metabolic rate during a specific activity in relation to a standard resting rate of 1 (4184 kJ) kg^{-1} $hour^{-1}$) were assigned to each reported hour of sport (5 MET) and household (3 MET) activities. As a result, it was possible to express the participants` leisure time physical activity in MET-hours/day. Second, according to their economic activity and type of job they had, participants were categorised into five distinct occupational activity groups: sedentary, standing, physical, heavy manual and non-workers. The third step was to cross-tabulate the sex-specific quartiles of leisure time physical activity expressed in MET-hours/day with the five occupational activity categories, as shown in table 3.4 (Cust *et al.* 2008). Finally, as the mean age, sex, cohort and BMI adjusted energy intake/basal metabolic rate ratio did not differ between the inactive and moderately inactive categories, these two groups were combined into one “inactive” category.

Table 3.4: Cross-tabulation of occupational and leisure time physical activities (Cust *et al.* 2008)

Occupational Activity	**Leisure time activity** (MET-hours/day in sex-specific quartiles)			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Sedentary	Inactive	Inactive	Moderately inactive	Moderately active
Standing	Moderately inactive	Moderately inactive	Moderately active	Active
Manual	Moderately active	Moderately active	Active	Active
Heavy manual	Moderately active	Moderately active	Active	Active
Non-worker	Moderately inactive	Moderately inactive	Moderately active	Moderately active

When the dietary habits of HAPIEE and Whitehall II cohorts were compared, total physical activity of participants was not comparable between cohorts. In these analyses, therefore only leisure time physical activity, expressed as MET-hours/day and categorised in low (≤5 MET-hrs/day) moderate (5-15 MET-hrs/day) and high (>15 MET-hrs/day) groups, was used.

Vitamin supplement intake. Participants in all three HAPIEE cohorts were asked if they took any vitamin supplements. Three categories were created: (1) no vitamin supplement users, (2) irregular users who took supplements less than three times a week, and (3) regular users who took them at least three times a week.

3.2.3.3 Anthropometric, biological and medical factors

Body mass index (BMI). BMI (kg/m^2) of most study subjects was calculated based on measured height and weight using the standard formula (weight/height^2). Since the correlations between measured and self-reported values (height: r=0.97; weight: r=0.98) were high, if data on measured height and weight were not available, self-reported values were used for this purpose.

Other than the continuously treated BMI, participants were also categorised as obese and non-obese. In line with the WHO guidelines, those with BMI higher than 30kg/m^2 were classified as obese (WHO 2015c).

Blood pressure. Blood pressure measurement was taken by a trained nurse following a standard protocol using an Omron M5-I digital sphygmomanometer. Subjects were in a sitting position after five minutes of quiet rest. The measurements were taken three times with two minutes intervals between them. The means of the second and third measurement of systolic and diastolic blood pressure values were used in the analyses (Peasey *et al.* 2006; Pajak *et al.* 2013). Systolic (SBP) and diastolic (DBP) blood pressure could not be included simultaneously in the multiple imputation procedures and also in the Cox regression models due to multicollinearity (Slinker and Glantz 1985). In order to overcome this problem and still include both SBP and DBP data in the regression models, the mean arterial blood pressure (MAP) was calculated for all participants using the following formula: MAP=1/3(SBP)+2/3(DBP) (Sesso *et al.* 2000).

Participants were also categorised in two groups based on whether they had hypertension or not. All subjects whose MAP was higher than 110 mmHg or had been taking antihypertensive medication were considered hypertensive.

Total cholesterol level. Plasma concentration of total cholesterol was analysed enzymatically using autoanalyzers and conventional methods in local laboratories in the Czech Republic (IKEM, Prague), Poland (Jagellonian University, Krakow) and Russia (Institute of Internal and Preventive Medicine, Novosibirsk). Total cholesterol level was treated as continuous variable in the analyses and was also used to help categorising participants in two groups. All individuals with total cholesterol concentration higher than 5.2 mmol/l or took lipid lowering medication were considered hypercholesteraemic, as opposed to subjects with non-elevated cholesterol concentrations.

Medical history. When the dietary habits between HAPIEE and Whitehall II cohorts were compared the results were also adjusted for self-reported medical history. Participants who indicated that they had been diagnosed with diabetes or CVD (including heart attack, acute myocardial infarction, angina, ischemic heart disease or stroke) were classified as having history of CVD/diabetes.

In order to avoid reverse causation, participants with prevalent CVD were excluded from the analyses when the association between dietary habits and mortality was assessed. In contrast to CVD, there is no clear consensus in the international literature how participants with diabetes should be treated. In some similar studies diabetes was considered as confounder which needs to be adjusted for (Hung *et al.* 2004), while others, including most studies in the EPIC cohort (Leenders *et al.* 2013), excluded these individuals from the analysis. Since the presence of diabetes can be not just a confounding factor but an important mediator on the causal pathway between diet and mortality, the latter approach was considered more appropriate and these participants were excluded from the analytical samples.

3.3 Missing data

Missing data can cause loss of statistical power and information bias in any epidemiological analysis. In order to reduce the extent of these issues, multiple random imputation of missing covariate data was applied.

Data can be missing due to several reasons which can determine the pattern of missingness. According to the relationships between the missing and observed values, three main types of missing data can be distinguished (Sterne *et al.* 2009; He 2010). If data is missing completely at random (MCAR) than there are no systematic differences between the missing and observed values. In this case the probability that an observation is missing (missingness) is independent from any measured or unmeasured variables. If data is missing at random (MAR), than there are systematic differences between the missing and observed values, but these differences can be explained by the measured variables in the dataset. The missingness is independent from the unmeasured factors if the measured variables in the dataset are controlled for. Finally, data is missing not at random (MNAR) if the systematic differences between the missing and observed values cannot be explained by the measured variables in the dataset. Missingness is related to unmeasured factors even after all variables in the dataset are taken into account.

Multiple random imputation can be applied if missing data are MAR (Sterne *et al.* 2009). Although no statistical test is available to distinguish between MCAR, MAR or MNAR (Sterne *et al.* 2009; Bhaskaran and Smeeth 2014), there are a number of reasons which indicate that the MAR assumption can be justified in the current dataset. First of all, missingness was significantly related to several variables in the dataset. For example, males, ex- or current smokers, and those with lower education or higher household amenities score were more likely to have missing data (table 3.5). This suggests that the missing data was not likely to be MCAR. It is not possible to say whether the difference between the missing and observed values was related to any unmeasured factors or the available variables could fully account for it, in other words whether the missing data was MAR or MNAR. However, it is likely that the different data collection procedures in the three cohorts were responsible for most of the differences in missingness (questionnaires were nurse-administered and completed in the research clinic in Russia, but mainly self-administered (only supervised or checked by a nurse) and completed in the participants` home in the Czech Republic and Poland). The proportion of participants with missing data in the Czech, Polish and Russian cohorts were 31%, 21% and 1%, respectively.

Table 3.5: Associations between missingness and covariates

Variables	Categories	Missingness[1]	
		OR	**p-value**
Sex	Males (ref.)	1.00	
	Females	0.93	0.010
Age (years)		1.01	0.002
Marital status	Married/cohabiting (ref.)	1.00	
	Single/divorced/widowed	0.93	0.045
Education	Primary or less (ref.)	1.00	
	Vocational	1.00	0.939
	Secondary	0.81	<0.001
	University	0.66	<0.001
Amenities score	Low (ref.)	1.00	
	Medium	1.02	0.603
	High	1.22	<0.001
Smoking	No smokers (ref.)	1.00	
	Ex-smokers	1.37	<0.001
	Current smokers	1.38	<0.001
Alcohol intake	Abstainers (ref.)	1.00	
	Light to moderate drinkers	0.69	<0.001
	Heavy drinkers	0.89	0.051
Physical activity	Low (ref.)	1.00	
	Moderate	0.78	<0.001
	High	0.73	<0.001
Energy intake (MJ/day)		0.96	<0.001
Vitamin supplement usage	Non-users (ref.)	1.00	
	Irregular users	1.15	<0.001
	Regular users	1.28	<0.001
Mean arterial blood pressure (mmHg)		1.00	0.098
Body mass index (kg/m2)		0.97	<0.001
Serum cholesterol cc. (mmol/l)		0.90	<0.001
Follow up time (years)		1.28	<0.001
All-cause mortality	Alive (ref.)	1.00	
	Dead	1.35	<0.001

[1] Probability that a participant has missing data in any of the covariates. All ORs were calculated with logistic regression using missingness, coded as “1” or “0”, as the outcome and the covariates as the exposure variables.

The MAR assumption could not be justified for the missing FFQ data because missing answer for a particular FFQ item suggests no consumption rather than a random miss. Missing mortality outcome data was imputed than deleted as recommended by von Hippel (von Hippel 2007). Consequently, only missing covariate and olive oil usage (component of the Mediterranean diet score) data were imputed using multiple imputation procedures and subsequently included in the statistical models. There were 6564 participants (22.7%) in the full HAPIEE study population who had missing data in any of the variables listed in table 3.6. Multiple imputation was carried out using the "mi impute chained" command in STATA version 13.1 (van Buuren 2007; White *et al.* 2011). Ten imputed datasets were created and, other than the covariates with missing data, the following predictor variables were included in the procedure: age, sex, cohort, follow-up time and all-cause mortality.

Table 3.6: Number of participants with missing covariate data and the applied imputation methods

Variables with missing data	No. missing	Imputation method
Marital status	66	Simple logistic regression
Education	60	Ordered logistic regression
Household amenities score	758	Predictive mean matching
Smoking habits	152	Multinomial logistic regression
Alcohol intake	378	Predictive mean matching
Physical activity	1754	Ordered logistic regression
Vitamin supplement usage	157	Ordered logistic regression
Mean arterial blood pressure	3668	Predictive mean matching
Body mass index	52	Predictive mean matching
Serum cholesterol cc.	3415	Predictive mean matching
Olive oil usage	1451	Multinomial logistic regression

Multiple imputation was carried out in a separate procedure when dietary habits between the HAPIEE and Whitehall II cohorts were compared (section 4.2) because a different set of covariates were applied in this analysis. Missing data on marital status (number of participants with missing data in the combined HAPIEE/Whitehall II dataset = 77), education (654), smoking (129), employment status (103), leisure-time physical activity (897) and medical history of previous CVD or diabetes (313) were imputed using the same predictor variables as described above.

As a sensitivity analysis, the association of fruit and vegetable intake with mortality was also estimated using data from participants without missing data only. This complete case analysis showed largely similar results to the analysis of the imputed dataset, although, as a result of the smaller sample size, confidence intervals were somewhat wider (table III-1 in appendix).

3.4 Analytical samples

Not all individuals who were part of the HAPIEE study population were included in the actual statistical analyses. The selection of analytical samples was carried out in several steps. Although most of these steps were identical across the main analyses of the thesis, there were some important differences as well. As a result, the size of the analytical samples differed between specific analyses. The number of participants excluded from the HAPIEE study population in the different analyses due to the various exclusion/inclusion criteria is presented in table 3.7.

Table 3.7: Selection of analytical samples from the HAPIEE study population in the main analyses of the thesis (see details in text)

Criteria for exclusion	No. excluded participants	Main analyses of the thesis			
		HAPIEE vs. Whitehall II: descriptive dietary comparison	Fruit/vegetable intake vs. mortality	HDI vs. mortality	MDS vs. mortality
<90% completed FFQ	717	X	X	X	X
FFQ not representative of their diet	776	X	X	X	X
Extreme energy intake reporting	548	X	X	X	X
Missing mortality data	1048		X	X	X
Previously diagnosed CVD or diabetes	6525		X	X	X
Previously diagnosed cancer	774			X	
Analytical sample size		26,906	19,333	18,559	19,333

Firstly, all participants who answered less than 90% of the FFQ questions were excluded from all analyses. Secondly, those who stated that the FFQ was not representative of their diet were also omitted.

Energy misreporting was assessed using the energy intake (EI) to basal metabolic rate (BMR) ratio (Schofield 1985). Participants in the lowest and highest 1% of the EI/BMR distribution were excluded from the analyses, which criteria is often used in EPIC studies (Leenders *et al.* 2013). In addition, as a sensitivity analysis, I also estimated the association between the healthy diet indicator (HDI) and mortality rates after using different exclusion criteria for energy misreporting (i.e.: participants in the top and bottom 5% of the EI/BMR ratio, or those with above or below a specified reported energy intake level were excluded), or when the implausibility of reported dietary intake data was defined based on the reported number of FFQ items consumed a day (i.e.: participants who reported to consume more than 65 items or less than 5 items a day were excluded). Changes in exclusion criteria had only small impact on the hazard ratios (table V-1 in appendix).

In all prospective analyses when the associations with mortality outcomes were assessed, individuals whose mortality follow up data was not available (due to missing national ID number or refusal to be followed up) were excluded. In order to avoid reverse causation, those with previously diagnosed CVD or diabetes were also omitted in these analyses.

Since mortality data was not relevant in the descriptive dietary comparison, subjects with missing mortality follow up data or prevalent CVD/diabetes were excluded only when the contribution of fruit and vegetable intakes to the between-cohort mortality differences was assessed.

In the analysis when the relationship between HDI and mortality was estimated, deaths from cancer and other causes (non-CVD-non-cancer) were also included as additional outcomes. Consequently, to avoid reverse causation, all subjects with previously diagnosed cancer were excluded from this analysis.

3.5 Power calculations

Power calculations showed that the pooled sample size had a power of 80% to demonstrate HR=0.92-0.95 as statistically significant at the 0.05 significance level in analyses of dietary habits and all-cause mortality. For CVD mortality, the power of 80% would demonstrate as statistically significant HR between 0.87 and 0.91. However, for CHD and stroke, as a result of the lower number of deaths, the power was adequate to detect only relatively strong effects (table 3.8).

Table 3.8. Smallest detectable HRs of the analyses between dietary exposures and mortality outcomes

Mortality outcomes	Dietary exposures					
	Fruit and vegetable intake Per 100g/day increase (SD=1.45)		**HDI** Per 1SD increase (SD=1.0)		**MDS** Per 1SD increase (SD=1.0)	
	No. events	HR[1]	No. events	HR[1]	No. events	HR[1]
All-cause	1314	0.95	1209	0.92	1314	0.93
CVD	438	0.91	423	0.87	438	0.87
CHD	226	0.88	220	0.83	226	0.83
Stroke	109	0.83	105	0.76	109	0.76
Cancer			437	0.87		
Non-CVD-non-cancer			284	0.85		

[1] Smallest detectable HR if power=0.80 and alpha=0.05

3.6 Statistical software

Data preparation and all statistical analyses were carried out using the statistical software STATA versions 12.1 and 13.1 (StataCorp, Texas, US).

3.7 Methodology of specific analyses

This thesis presents four distinct epidemiological analyses. All of them used data from the HAPIEE study, as described in the previous sections. However, several important methodological procedures were different across the four analyses. These specific methodological steps are detailed below.

3.7.1 Objective 1: comparison of dietary intakes between the HAPIEE and Whitehall II cohorts

In order to reach the first objective of the thesis, dietary habits of the HAPIEE study participants were compared with individuals who took part in the London-based

Whitehall II prospective cohort study of civil servants. Details of the Whitehall II study, as well as the methodological steps of the dietary data harmonization process and the applied statistical techniques are explained in this section.

3.7.1.1 The Whitehall II study

Study population and measurements. The Whitehall II study is a prospective cohort study of civil servants set up in 1985-88 with the central aim to examine social inequalities in physical and mental health (Marmot *et al.* 1991; Marmot and Brunner 2005). At baseline, 10,308 participants (6895 men and 3413 women), aged between 35 and 55 years, were recruited from 20 civil service departments in London. The overall response rate of the baseline survey was 73%.

Participants were asked to complete a self-administered questionnaire and attended a short screening examination at baseline. The questionnaire included topics on socio-demographic factors, health status, work and other social environmental characteristics and health behaviours/lifestyle. During the examination, anthropometric characteristics were measured, blood pressure was taken, electrocardiogram (ECG) was recorded and blood samples were taken (Marmot *et al.* 1991). Every five years since the baseline survey participants have completed a similar questionnaire and undergone medical examination (waves 3, 5, 7 and 9). In addition, participants were asked to complete postal questionnaire (without examination) between the screening phases (forming waves 2, 4, 6 and 8).

The 7th wave of the study took place between 2002 and 2004, at approximately the same time as the baseline data collection of the HAPIEE study. In this phase 6967 participants (68% of baseline responders) took part with an age range of 50-74 years.

Dietary assessment. Dietary data were first collected in 1991-93 during the 3rd phase of the study using FFQ and 7-day diet diary. The FFQ was developed based on the questionnaire constructed by Willett and colleagues in the US Nurses Health study (Willett *et al.* 1985). The Whitehall II FFQ was also used as a template during the development of the HAPIEE study FFQs, which means that the FFQs used in the two studies are very similar. Since the 3rd wave, participants have been asked to complete the FFQ every second (5th, 7th, 9th) phase of the study. In the 7th wave, the FFQ consisted of 116 items (see appendix II). As in the HAPIEE study, a common unit or standard portion size was specified for each food item, and participants could indicate how frequently they consumed a particular item over the previous year using a 9-point scale ranging from “never, or less than once a month” to “more than 6-times a day” (Brunner *et al.* 2001). To calculate nutrient intake levels, similarly to HAPIEE study, the McCance and Widdowson Food Composition Database and the in-house Wfood 2002 nutrient analysis software was used.

An earlier analysis of the relative validity of the FFQ data in the 3rd wave of the study indicated good agreement with 7-day diet diary data and plasma biomarker concentrations (Brunner *et al.* 2001).

Mortality follow up. Data for mortality follow up is provided by the National Health Service (NHS) Central Registry which allows data linkage for nearly all individuals (n=10,297) who took part in the baseline survey. In the current analysis, mortality data registered until the 31st August 2012 was used. Similarly to the HAPIEE study, the cause of death was defined by the underlying cause indicated on the death certificate and coded according to ICD-9 and ICD-10: CVD (ICD-9: 390-459; ICD-10: I00-I99), CHD (410-414; I20-I25), stroke (430-438; I60-I69).

Analytical sample. Exclusion of participants from the full study population followed the same procedures as it was applied for the HAPIEE study. From the 6967 individuals who took part in the 7th wave of the study, participants who did not complete the FFQ or answered less than 90% of its questions (n=1363), indicated that the FFQ was not representative of their diet (n=61) or provided implausible dietary data (participants in the lowest and highest 1% of the EI/BMR ratio distribution) (n=110) were excluded from the analysis. Subjects with missing mortality follow up data (n=5) or prevalent CVD or diabetes (n=467) were excluded when the contribution of fruit and vegetable intake to the between-cohort mortality differences was assessed but not in the descriptive dietary comparison. Overall, the analytical sample of the Whitehall II study consisted of 5433 participants in the descriptive comparison and 4961 individuals when mortality differences between cohorts were taken into account.

3.7.1.2 Dietary data harmonization

The FFQs completed by the Czech, Polish, Russian and UK cohorts consisted of 136, 147, 142 and 116 food and drink items, respectively. There were two reasons for the discrepancies: (1) Some food products were combined into one FFQ item in one country, but asked separately in others. For example, apricots, peaches and plums were combined in one question in the UK but were included as three separate questions in the HAPIEE cohorts. (2) Certain items were not included in all FFQs, because some of them were country-specific foods (e.g. pirogi, borscht). However, the majority of these FFQ-specific items (77%, 66%, 67% and 59% in the Czech, Polish, Russian and British questionnaires, respectively) were consumed in all four countries (e.g. pineapple, aubergine, cucumber, lasagne).

The estimated intake of a given food group is likely to be proportional to the number of relevant items in the FFQ. Unless the differences between the FFQs represent country-specific differences in dietary habits (i.e. country-specific food items), which is not the case in the current comparison as described above, these discrepancies in the number of FFQ items may introduce reporting bias and need to be taken into account. Accordingly, I first excluded those items from the analysis which were not common in all four FFQs. Secondly, regarding food and drink items which were asked separately in one but in combination in other FFQs, the portion/day intake levels were summarized and the data on the combined intakes were used in all cohorts. Overall, dietary intake data from 81 food and drink items (including 9 combined items) were compared.

Participants had to estimate their intakes using an average portion or medium sized food or drink item in all four FFQs. In order to calculate g/day intake of a specific item, standard portion sizes, provided by local dietitians, were specified (Brunner *et al.* 2001; Boylan *et al.* 2009). These country-specific portion sizes were identical or similar for most items, however, for 29 (36%) of 81 items the difference was more than 50%. Although some of the small differences might reflect real regional differences, large discrepancies are likely due to arbitrary choices made by local dietitians during the construction of the FFQs. To avoid information bias due to different portion sizes, the g/day intake of each food and drink items were recalculated (i.e.: producing identical portion sizes in all cohorts, using the portion sizes published by the UK`s Food Standard Agency (Food Standards Agency 2002)). Alcoholic drink sizes were an exception, because the size of a standard drink clearly differs between countries and the questions on the FFQs were asked in line with the local habits. (i.e.: 1 beer is 1/2 pint=287ml in the UK but 1 glass=250ml in

CEE/FSU.) Mean energy intakes increased between 4% and 9% by cohort when standard portion sizes were used instead of cohort-specific portion sizes, which suggests that this change had only a small impact on the overall results.

In the HAPIEE cohorts, participants were asked to estimate their eating habits over the past three months. In contrast, the questions referred to the previous year in Whitehall II study, and regarding seasonal foods (i.e. fruits, vegetables), participants were asked to estimate their intakes in the time period when that particular item is in season. In order to eliminate the differences due to the different reference periods of the FFQs, weighted intake data for fresh fruits and vegetables were compared: for those participants of the HAPIEE cohorts who completed the FFQ during winter or spring, the intake of fresh fruits and vegetables were multiplied by the within-cohort summer-autumn vs. winter-spring ratio of median fresh fruit and vegetable intake.

3.7.1.3 Statistical analysis

The food and drink items listed in the FFQs were categorised into food/drink groups and subgroups according to the European Food Safety Authority`s Foodex2 food classification system (EFSA 2011b). The cross-cohort dietary comparisons were carried out on absolute intake values for food/drink groups and subgroups, and on energy standardized intake values (calculated by the residual method) for nutrients (Willett *et al.* 1997; Willett 2013c).

To take into account of possible information bias, food/drink groups and nutrients were categorised as fully comparable, partially comparable or not comparable between cohorts, according to the contribution of the 81 identical items to their total intake. Food/drink groups and nutrients were considered fully comparable if more than 80% of intake was provided by common items in all cohorts. If the contribution was 60-80% in one or more of the cohorts, they were considered partially comparable. If the contribution was less than 60% of intake in one or more of the cohorts then the food, drink or nutrient was not considered comparable and results were not shown.

In the multivariable adjusted models, quantile regression method was used because of the non-normal distribution of food, drink and nutrient intakes (Marrie *et al.* 2009; Koenker *et al.* 2013). All comparisons were adjusted for age, sex, energy intake, marital status, education, employment status, smoking, leisure time physical activity and medical history.

Differences in mortality rates between cohorts were assessed by Cox regression models using the British cohort as the reference category. Schoenfeld residuals indicated no violation of the proportionality assumption (Schoenfeld 1980). Hazard ratios (HR) of mortality differences between cohorts were calculated in four models. In model 1, HRs were adjusted for age and sex. In model 2, they were further adjusted for energy intake, smoking, leisure time physical activity, education, marital status and employment status. Finally, HRs were also adjusted for fruit intake in model 3, or vegetable intake in model 4. In order to assess the impact of conventional risk factors and fruit and vegetable intake on mortality rates, the percentage changes of HRs were also calculated in the different models. HR change in relation to the basic model (model 1) was calculated in model 2, and in relation to the multivariable adjusted model (model 2) in model 3 and 4. The following formula was used:

% change=(HR1-HR2)/(HR1-1)*100.

3.7.2 Objective 2: association between fruit, vegetable intake and mortality

In line with the second objective of the thesis, the association of fruit and vegetable intakes with total and CVD mortality was assessed using data collected from participants of the HAPIEE study.

3.7.2.1 Assessment of fruit and vegetable intake

The European Food Safety Authority`s FoodEx 2 food classification and description system was used to categorise food items into fruit and vegetable food groups (EFSA 2011b). All items which are listed in the group of “fresh fruits” [A04RK] or “vegetable and vegetable products” [A00FJ], with the exception of “vegetable products” [A00ZA], were considered as fruits and vegetables. Overall, 21 fruit and 24 vegetable items were included (table 3.9). Daily consumption of the different fruit and vegetable items were calculated by multiplying the number of portions per day by average portion sizes determined by local dieticians. A person’s daily overall fruit and vegetable consumption was calculated by adding up the intake values of the different items.

Table 3.9: Fruit and vegetable items included in the analysis

FOOD GROUPS	ITEMS
Fruits	apple, pear, peach, apricot, plum, cherry, strawberry, raspberry, red currant, black currant, gooseberry, blueberry, orange, mandarin, lemon, grapefruit, kiwi, melon, pineapple, banana, grape
Vegetables	broccoli, cauliflower, cabbage, Brussels sprouts, garlic, onion, leek, tomato, cucumber, pepper, aubergine, courgette/marrow, sweet maize, green salad (lettuce), spinach, beetroot, carrot, celeriac, turnip/swedes, parsnip, radish, green beans/runner beans, parsley, mushrooms

3.7.2.2 Assessment of fruit and vegetable intake data`s relative validity against biomarkers

As self-reported dietary intakes are often imprecise, the relative validity of fruit and vegetable intake data against plasma biomarker concentrations, measured in a random sub-sample of participants in all three cohorts and determined in a central laboratory, was assessed. In a previous analysis, the correlations between the intakes and plasma concentrations of antioxidant vitamins, as well as the correlations of fruit and vegetable intakes with plasma vitamin C and beta-carotene, were the lowest amongst those participants who took vitamin supplements regularly (Stefler 2011). Therefore, for the purpose of this thesis, the correlations between fruit, vegetable intakes and vitamin C and beta-carotene plasma concentrations were re-calculated including only those subjects in the analysis who took no regular vitamin supplements. From the 2327 and 2647 participants with available data on plasma vitamin C and beta-carotene concentrations, 1929 and 2180 were included, respectively. Data on both intakes and antioxidant plasma concentrations were log-transformed. Pearson`s partial correlation coefficients, adjusted for energy intake, country-cohort and sex, were calculated.

3.7.2.3 Statistical analysis

Cox proportional hazard model was applied to estimate the association of fruit and vegetable intake with all-cause and cause-specific mortality. Follow up time for each participant was calculated from the date of baseline questionnaire completion until the end of observational period (December 2011 for Czech and December 2010 for Russian and Polish participants) or the date of death, whichever happened first. For participants who were lost during follow up, the last date of contact was used as exit date. Proportionality assumption in all Cox models was checked using the Schoenfeld equations (Schoenfeld 1980). Sensitivity analyses using competing risk assessment models or excluding those who died during the first two years of follow up were also carried out (tables III-2 and III-3 in appendix).

Fruit and vegetable intake, categorised into cohort-specific quartiles, was used as the main exposure variable. Additionally, the HRs of mortality per one unit (100g/day) increase across six absolute intake categories (<100g/d, 1-200g/d, 2-300g/d, 3-400g/d, 4-500g/d, >500g/d) were also calculated. In model 1, the associations were adjusted for sex, age and cohorts. In model 2, the associations were further adjusted for education, household amenities score, marital status, alcohol intake, smoking, physical activity, vitamin supplement intake and diet quality (using the healthy diet indicator [HDI] without the fruit and vegetable component (see section 3.7.3)). Since the correlation between fruit and vegetable intake was moderate (Spearman`s rho=0.21), when I examined their association with mortality outcomes separately, the HRs were further adjusted for each other.

Assuming causal relationship between fruit, vegetable intake and mortality, preventable proportions (PP) of deaths which could be avoided if participants in the lowest three quartiles would shift their intake one quartile upward were calculated using the same formula as in previous studies (Wahrendorf 1987; Leenders *et al.* 2013):

$$PP = \frac{\sum_{i=0}^{K} p_i r_i - \sum_{i=0}^{K} p_i^* r_i}{\sum_{i=0}^{K} p_i r_i} \quad \text{(ref: Wahrendorf 1987)}$$

Where *p* and *p** are the proportion of participants in quartile *i* before and after the shift, and *r* is the corresponding hazard ratio. This approach models the effect of an overall positive shift in the exposure distribution, rather than assuming that all individuals increase their fruit and vegetable intake above a specific threshold (i.e. 400g/day). The shifting model is probably a more realistic description of what would happen if primary preventive measures implemented effectively in a population (Wahrendorf 1987).

Because of a statistically significant interaction between overall fruit and vegetable intake and smoking for all-cause mortality (p=0.008), I also report results separately by smoking groups. Although there was no significant interaction between fruit and vegetable intake and cohorts, data were also analysed separately by country cohorts.

In order to assess the mediating effect of blood pressure, the associations were further adjusted for mean arterial blood pressure (MAP) in the subsample of participants who were not taking antihypertensive medications (n=13,966).

The relationship between intakes of selected fruit and vegetable subgroups and mortality outcomes was also analysed. The examined subgroups included citrus fruits (orange, mandarin, grapefruit and lemon), berries (black currant, blueberry, gooseberry, red currant, strawberry and raspberry), green/leafy vegetables (broccoli, Brussel sprout, cabbage, cauliflower, lettuce and spinach) and processed fruits or vegetables (mixed frozen vegetables, pickled beet-root, pickled gherkin, sauerkraut, dried fruits and tinned/canned fruits). HRs of cohort-specific tertiles and per one unit increase across four absolute intake categories (>30g/d, 30-60g/d, 60-90g/d, >90g/d) were calculated.

3.7.3 Objective 3: healthy diet indicator and mortality

In accordance with the third objective of the thesis, the association between the healthy diet indicator (HDI) and total and cause-specific mortality was investigated in the HAPIEE study.

3.7.3.1 Construction of the healthy diet indicator (HDI)

The HDI was constructed to reflect the WHO`s dietary recommendations for the prevention of chronic diseases published in 2003 (WHO 2003a). From the 15 dietary items listed in the WHO guideline, nine were included in the score. Total fat, total polyunsaturated fatty acids, monounsaturated fatty acids and total carbohydrates were excluded to avoid overlap with other components of the score, and sodium was excluded because such information was unavailable. Since no data was available on fibre intake, the intake of non-starch polysaccharides (NSP) was used instead. As opposed to the dichotomised scoring method used in the original HDI study (Huijbregts *et al.* 1997), continuous scoring was used. This approach reflects the fact that the health effect of various nutritional factors does not follow definite cut-off points, but it rather changes on a continuous scale. In addition, the continuous scoring results in greater variation of scores between individuals, which improves the statistical power to detect associations with health outcomes.

Participants scored ten points for each component if their intake level met the WHO recommendation. No points were given if the intake level was above the 85% of the population distribution regarding the “moderation” components (saturated fat, trans fatty acids, mono- and disaccharides, cholesterol), or if the intake level was zero regarding the “adequacy” (fruits and vegetables, NSP) and “moderation range” (n3-

PUFA, n6-PUFA, protein) components. In case of "moderation range" components, the intake levels above which no point is given were chosen to reflect equal deviation from the ideal intake on both sides of the recommended range. Participants whose intake was between the ideal (10 points) and "no point" ranges scored between zero and ten points, proportionately to their deviation from the recommended intake. The total HDI score was calculated as the sum of individual component scores. The scoring criteria for the different components are shown in table 3.10.

Table 3.10: Scoring criteria of the HDI

HDI components	HDI component scores		
	0 point	**0 - 10 points**	**10 points**
Saturated fatty acids, *energy%*	>15	10-15	0-10
n3-Polyunsaturated fatty acids, *energy%*	>3	0-1 or 2-3	1-2
n6-Polyunsaturated fatty acids, *energy%*	>13	0-5 or 8-13	5-8
Trans fatty acids, *energy%*	>2	1-2	<1
Mono- and disaccharides, *energy%*	>30	10-30	0-10
Protein, *energy%*	>25	0-10 or 15-25	10-15
Cholesterol, *mg/day*	>400	300-400	0-300
Fruits/vegetables, *g/day*	0	0-400	>400
Non-starch polysaccharides, *g/day*	0	0-20	>20

energy% – percentage of daily alcohol-free energy intake

3.7.3.2 Statistical analysis

Simple, multinomial and ordered logistic regression was used to compare HDI scores between covariate categories, and p-values of the crude and age, sex, country-cohort and energy intake adjusted comparisons were reported.

Cox regression was used to investigate the association between the HDI score and all-cause and cause-specific mortality. The estimated HRs indicated the change in mortality risk by one standard deviation (SD) increase in HDI score. One SD was equal to 8.93 points in the HDI score.

Because no interactions between countries and HDI were detected, the Cox regression analysis was performed in the pooled sample, as well as separately in each country cohort. The analyses were conducted in two steps. First, HDI was adjusted for age, sex and cohort. Second, HDI was further adjusted for the highest level of education, household amenities score, marital status, alcohol, smoking, physical activity and energy intake and vitamin supplement intake. BMI was not included; as it could be on the causal pathway, controlling for BMI might lead to over-adjustment.

In order to illustrate the shape of the relationship between HDI and the mortality outcomes, participants were classified into four groups based on their HDI score`s distance from sample mean (Group1: HDI ≤ -1SD; Group2: HDI > -1SD and HDI ≤ mean; Group 3: HDI > mean and HDI ≤ +1SD; Group 4: HDI > +1SD) and HRs were also calculated across categorised HDI scores. Preventable proportions (PP) of deaths which could be avoided if participants in the lowest three HDI groups would shift their diet one group upward were calculated using the same formula as described previously (see section 3.7.2.3).

Finally, I investigated the extent to which differences in death rates between the three cohorts could be explained by the HDI. For this purpose, age and sex-adjusted hazard ratios of mortality differences between cohorts were first adjusted for potential lifestyle and socio-economic risk factors (model 2). Subsequently, the HRs were further adjusted for HDI in model 3. The Czech cohort (with the lowest mortality rate) was used as reference category in this analysis.

3.7.4 Objective 4: Mediterranean diet score and mortality

In line with the fourth objective of the thesis, the relationship between the Mediterranean diet score and total and CVD mortality was assessed amongst the participants of the HAPIEE study.

3.7.4.1 Construction of the Mediterranean diet score

The MDS applied in this analysis followed the recommendations of Sofi et al who defined absolute cut-off values for all MDS components based on comprehensive literature review, and applied a three-tier scoring system with zero, one or two points for each component (table 3.11) (Sofi *et al.* 2014). The component regarding olive oil usage had to be modified because the corresponding question in the FFQ did not allow distinction between occasional, frequent and regular users. One point was given for this component to those participants who stated that they used olive oil for cooking, and zero point to those who reported to cook with any other type of oil (vegetable oil, butter, margarine or lard). As a result, after adding up the individual component scores, overall MDS ranged from zero to 17.

Table 3.11: Scoring criteria of the MDS

MDS components	MDS component scores		
	0 point	1 point	2 points
Vegetables (g/day)	<100	100-250	>250
Fruits and nuts (g/day)	<150	150-300	>300
Legumes (g/week)	<70	70-140	>140
Cereals (g/day)	<130	130-195	>195
Fish (g/week)	<100	100-250	>250
Meat and meat products (g/day)	>120	80-120	<80
Dairy products (g/day)	>270	180-270	<180
Alcohol (g/day)	>24	<12	12-24
Olive oil usage	Not used for cooking	Used for cooking	-

3.7.4.2 Statistical analysis

Participants` adherence to the Mediterranean diet was classified as low (0-7 points), moderate (8-10 points) and high (11-17 points) according to their MDS. These categories reflect similar fraction of the maximum score as those applied by Trichopoulou et al in the most commonly used scoring system with the maximum of 9 points (Trichopoulou *et al.* 2005).

Crude and basic (cohort, sex, age and energy intake) adjusted logistic and linear regression models were used to estimate the relationships between covariates and MDS categories.

The associations between the MDS and mortality outcomes were assessed using Cox proportional hazard models with MDS as both a categorical and a continuous variable. In the latter case, the associations of mortality risk with 1 SD increase in the

MDS were calculated. One SD in the MDS was equal to 2.2 points in the pooled sample. Proportionality assumptions were tested with Schoenfeld residuals. In the multivariable models, the associations were adjusted for age, sex, cohort, education, household amenities score, marital status, smoking, physical activity, total energy intake and vitamin supplement intake.

The proportion of deaths which could be prevented if participants in the lowest two MDS categories increased their adherence to the Mediterranean diet one category upwards was calculated using a formula applied previously (see section 3.7.2.3) but modified for three exposure categories.

Since the dietary assessment methods in the three cohorts were very similar and there were no interactions between MDS and cohort, sex or smoking status, the associations were calculated in the pooled sample, but results are also presented by country cohorts.

In order to assess the impact of the individual components to the overall MDS, the associations between the MDS component scores and mortality outcomes were also calculated. Multivariable adjusted HR per 1-point increase in each component score is presented.

- CHAPTER 4 -

RESULTS

This chapter provides a detailed description of the thesis` findings. First, the descriptive characteristics of the HAPIEE study population is presented (section 4.1), which is then followed by the results of the four main epidemiological analyses separately: comparison of dietary habits between HAPIEE and Whitehall II cohorts (section 4.2); estimation of the association of fruit and vegetable intake (section 4.3), healthy diet indicator (section 4.4) and Mediterranean diet score (section 4.5) with total and cause-specific mortality in the HAPIEE study.

4.1 Descriptive characteristics of the HAPIEE study participants

Table 4.1 shows the demographic, socio-economic and lifestyle characteristics of the study participants in the whole HAPIEE sample and by country cohorts. There were more females than males in each study centre, and there was no substantial difference in age between centres and genders. Energy intake in Russia was higher than in the other two cohorts in both sexes but BMI was increased only in females, which is consistent with the relatively high proportion of Russian men who were physically active. Blood pressure, serum cholesterol level, as well as the prevalence of obesity, hypertension and hypercholesterolemia also seemed to be higher in Russian females. Although university degree was relatively infrequent among the Czech participants, their household amenities score was higher than the other two cohorts`. There was a large contrast in smoking prevalence between Russian men and women, and the proportion of heavy drinkers was the highest among Czech men.

Table 4.1: Baseline characteristics of the HAPIEE study population

Covariate	Category	CZECH		POLISH		RUSSIAN		TOTAL
		Males (n=4125)	Females (n=4731)	Males (n=5230)	Females (n=5498)	Males (n=4269)	Females (n=5094)	(n=28,947)
Mean age, years (SD)		58.6 (7.2)	57.9 (7.1)	58.0 (6.9)	57.4 (7.0)	58.3 (7.1)	57.4 (7.0)	58.0 (7.1)
Mean energy intake, MJ/day (SD)[1]		9.0 (4.0)	8.5 (3.9)	9.7 (4.0)	8.8 (3.8)	11.8 (3.8)	9.9 (3.1)	9.6 (3.9)
Mean body mass index, kg/m^2 (SD)		28.2 (4.0)	28.0 (5.1)	27.8 (4.0)	28.1 (5.1)	26.6 (4.4)	30.2 (5.7)	28.2 (4.9)
Mean arterial blood pressure, mmHg (SD)		108.5 (14.2)	102.9 (15.1)	106.0 (15.2)	101.1 (14.5)	107.8 (15.7)	107.5 (16.7)	105.4 (15.0)
Mean serum cholesterol cc., mmol/l (SD)		5.6 (1.2)	5.8 (1.1)	5.7 (1.2)	5.9 (1.1)	6.0 (1.2)	6.5 (1.3)	6.0 (1.2)
		%	%	%	%	%	%	%
Marital status	Single/divorced/widowed	15.8	31.9	13.5	33.6	12.2	40.6	25.2
	Married/cohabiting	84.2	68.1	86.5	66.4	87.8	59.4	74.8
Education	Primary or less	6.1	18.3	9.5	13.5	11.4	9.6	11.5
	Vocational	44.0	31.3	27.4	15.2	21.7	30.5	27.8
	Secondary	31.6	40.5	32.9	44.3	35.0	33.5	36.5
	University	18.2	9.9	30.2	27.0	31.9	26.3	24.1
Household amenities score	Low	14.2	19.9	16.9	25.0	26.0	37.7	23.5
	Medium	41.1	44.7	44.5	47.3	47.9	45.1	45.2
	High	44.6	35.4	38.7	27.7	26.1	17.2	31.3
Smoking habits	Never smoker	31.8	54.7	27.9	50.8	25.7	85.3	47.0
	Ex-smoker	38.7	21.5	36.1	20.8	24.8	4.4	23.9
	Current smoker	29.5	23.8	36.0	28.4	50.0	10.3	29.1
Alcohol intake	Abstainers	6.6	18.8	21.9	46.3	13.5	17.9	21.9
	Light to moderate drinkers	71.5	74.1	70.7	52.1	70.7	80.6	69.6
	Heavy drinkers	21.9	7.1	7.4	1.6	15.8	1.6	8.5

Covariate	Category	CZECH		POLISH		RUSSIAN		TOTAL
		Males (n=4125)	**Females** (n=4731)	**Males** (n=5230)	**Females** (n=5498)	**Males** (n=4269)	**Females** (n=5094)	(n=28,947)
Physical activity	Inactive	48.8	54.5	47.6	50.9	43.6	52.0	49.7
	Moderately active	39.9	38.7	43.3	42.5	40.5	40.4	41.0
	Active	11.2	6.8	9.0	6.6	16.0	7.6	9.3
Vitamin supplement usage	Non-users	58.0	39.1	59.9	45.1	78.0	59.3	56.0
	Irregular users	23.7	29.3	27.2	33.7	15.0	23.9	25.9
	Regular users	18.3	31.6	12.9	21.1	7.0	16.8	18.1
Obesity (BMI>30kg/m2)	Obese	28.4	30.7	26.3	32.9	20.7	47.0	31.4
	Not obese	71.6	69.3	73.7	67.1	79.3	53.0	68.6
Hypertension	Hypertensive	59.6	48.2	54.3	47.0	48.7	57.1	52.3
	Not hypertensive	40.4	51.8	45.7	53.0	51.3	42.9	47.7
Hypercholesterolaemia	Hypercholesterolaemic	69.7	76.7	72.3	79.5	75.2	86.2	76.9
	Not hypercholesterolaemic	30.3	23.3	27.7	20.5	24.8	13.8	23.1

[1] Not imputed. n=28230

The median follow up time of the study participants was 7.1 years, however, on average, it was shorter for Russians and approximately one year longer for the Czech cohort (table 4.2). During this follow up period, all-cause mortality rates were similar in the Czech and Polish cohorts but they were substantially higher amongst Russians, especially for males. High total mortality in the Russian sample was mainly due to their increased CVD death rates. Compared to the other two cohorts, CVD, CHD and stroke mortality rates of Russian men were higher by two-, three- and five-times, respectively. Although no large difference in cancer mortality rates were seen between cohorts, non-CVD-non-cancer deaths, which included mainly deaths due to injuries, were also the most common amongst Russians. Nearly all death rates were higher in males than in females.

<u>Table 4.2:</u> Mortality follow up of the HAPIEE study population

		CZECH		POLISH		RUSSIAN		TOTAL
		Males	**Females**	**Males**	**Females**	**Males**	**Females**	
Median follow up time, years (IQR)		8.1 (7.7-8.9)	8.2 (7.8-8.9)	7.1 (6.8-7.7)	7.1 (6.9-7.7)	6.2 (5.7-6.9)	6.7 (6.0-7.1)	7.1 (6.7-7.8)
No. deaths	All-cause	478 (15.4)	266 (7.3)	543 (16.2)	288 (7.9)	696 (27.5)	286 (8.8)	2557 (13.1)
(per 1000 person-years)	CVD	181 (5.8)	91 (2.5)	178 (5.3)	97 (2.7)	349 (13.9)	140 (4.3)	1036 (5.3)
	CHD	85 (2.7)	34 (0.9)	105 (3.1)	31 (0.9)	226 (9.0)	80 (2.5)	561 (2.9)
	Stroke	26 (0.8)	16 (0.4)	22 (0.7)	28 (0.8)	100 (4.0)	46 (1.4)	238 (1.2)
	Cancer	194 (6.2)	126 (3.5)	209 (6.2)	124 (3.4)	154 (6.1)	78 (2.4)	885 (4.5)
	Non-CVD-non-cancer	101 (3.3)	49 (1.3)	120 (3.6)	52 (1.4)	141 (5.6)	46 (1.4)	509 (2.6)

Not all participants of the HAPIEE study were included in the statistical analyses. As noted in the previous chapter, participants with missing, non-representative or implausible dietary data, those whose mortality follow-up data was not available or had previously diagnosed CVD or diabetes were omitted from the analyses. Majority of the excluded subjects belonged to the Polish cohort (44%), while the proportion of Czechs (30%) and Russians (26%) was smaller in this group.

Table 4.3 shows that included and excluded individuals differed in most baseline characteristics and mortality rates. Participants who were excluded from the analyses were older, had higher blood pressure, BMI and somewhat lower energy intake and serum cholesterol level. They were more likely to be males, and had lower education attainment and household amenities score. Consistent with the fact that majority of the excluded participants had previously diagnosed CVD or diabetes, relatively larger proportion of them were ex-smokers and alcohol abstainers, which may be the result of their conscious decision related to their medical conditions. Due to these pre-existing diseases, mortality rates were also significantly higher in this group.

Table 4.3: Comparison of HAPIEE study participants who were included in and excluded from the analyses

Covariate	Category	Included (n=19,333)	Excluded (n=9614)	p value[1]
Mean age, years (SD)		57.0 (7.0)	60.1 (6.7)	<0.001
Mean energy intake, MJ/day (SD)[2]		9.7 (3.1)	9.4 (5.3)	<0.001
Mean body mass index, kg/m^2 (SD)		27.8 (4.7)	29.1 (5.1)	<0.001
Mean arterial blood pressure, mmHg (SD)		104.8 (15.3)	106.7 (15.6)	<0.001
Mean serum cholesterol cc., mmol/l (SD)		6.0 (1.2)	5.9 (1.3)	<0.001
No. all-cause deaths (per 1000 person years)		1314 (9.6)	1243 (21.4)	<0.001
No. CVD deaths (per 1000 person years)		438 (3.2)	582 (10.3)	<0.001
		%	**%**	
Sex	Males	45.5	50.3	
	Females	54.5	49.7	<0.001
Marital status	Single/divorced/wid.	24.6	26.4	
	Married/cohabiting	75.4	73.6	0.001
Education	Primary or less	10.0	14.7	
	Vocational	27.0	29.4	
	Secondary	37.2	35.1	
	University	25.7	20.9	<0.001
Household amenities score	Low	21.3	28.1	
	Medium	45.1	45.3	
	High	33.6	26.7	<0.001
Smoking habits	Never smoker	48.1	44.6	ref.
	Ex-smoker	21.2	29.4	<0.001
	Current smoker	30.7	26.0	0.003
Alcohol intake	Abstainers	18.4	29.0	
	Light-moderate drink.	72.5	63.8	
	Heavy drinkers	9.2	7.2	<0.001
Physical activity	Inactive	49.0	51.3	
	Moderately active	40.4	42.2	
	Active	10.6	6.6	<0.001
Vitamin supplement usage	Non-users	55.6	56.7	
	Irregular users	26.5	24.7	
	Regular users	17.9	18.6	0.788

[1] All p values were calculated with logistic regression using inclusion/exclusion as outcome variable and the covariates as explanatory variables

[2] Not imputed. n=28230

4.2 Objective 1: comparison of dietary intakes between the HAPIEE and Whitehall II cohorts

To address objective 1, food and nutrient intakes were compared between the British participants of the Whitehall II study and the Czech, Polish and Russian subjects of the HAPIEE study. Results of this dietary comparison analysis, as well as the findings of the analysis which estimated the contribution of fruit and vegetable intake to the mortality gap between cohorts is described in this section.

4.2.1 Descriptive characteristics

Table 4.4 shows the basic socio-demographic, lifestyles characteristics and mortality rates of the British, Czech, Polish and Russian participants included in this analysis. The British sample included more males, older and higher educated individuals than the Eastern European cohorts. The Whitehall II study also included fewer smokers, physically inactive persons and individuals with previously diagnosed CVD or diabetes. Mortality rates were lower in the British sample compared to any of the other cohorts.

Table 4.4: Characteristics of the Whitehall II and HAPIEE cohorts

Covariate	Category	BRITISH (n=5433)	CZECH (n=7864)	POLISH (n=9900)	RUSSIAN (n=9142)
Mean age, years (SD)		61.2 (6.0)	58.1 (7.1)	57.7 (7.0)	58.2 (7.1)
Median follow-up time, years (IQR)[1]		9.0 (8.6-9.4)	8.2 (7.8-8.9)	7.1 (6.9-7.7)	6.5 (5.9-7.1)
No. all-cause deaths (per 1000 person-years)[1]		249 (5.7)	364 (7.5)	388 (8.4)	562 (13.1)
No. CVD deaths (per 1000 person-years)[1]		59 (1.3)	106 (2.2)	99 (2.1)	233 (5.4)
No. CHD deaths (per 1000 person-years)[1]		29 (0.7)	43 (0.9)	45 (1.0)	138 (3.2)
No. stroke deaths (per 1000 person-years)[1]		8 (0.2)	18 (0.4)	20 (0.4)	71 (1.7)
		%	%	%	%
Sex	Males	72.2	46.6	49.0	45.4
	Females	27.8	53.4	51.0	54.6
Marital status	Single/divorced/wid.	23.4	23.8	23.5	27.7
	Married/cohabiting	76.6	76.2	76.5	72.3
Education	Primary or less	9.8	11.9	11.8	10.4
	O-level/vocational	25.5	36.8	21.4	26.5
	A-level/secondary	29.3	37.1	38.4	34.1
	BA/BSc or higher	35.5	14.2	28.5	28.9
Employment status	Employed	49.2	53.0	43.4	53.5
	Retired	45.6	43.5	50.0	41.5
	Non-employed-non-retired	5.2	3.5	6.6	5.0

Covariate	Category	BRITISH (n=5433)	CZECH (n=7864)	POLISH (n=9900)	RUSSIAN (n=9142)
Smoking habits	Never smoker	49.5	44.0	39.8	58.4
	Ex-smoker	43.2	29.5	28.1	13.6
	Current smoker	7.3	26.5	32.1	28.0
Leisure time physical activity	Inactive	15.5	34.5	29.5	28.6
	Moderately active	44.2	49.6	52.4	56.9
	Active	40.3	15.9	18.1	14.5
Medical history (CVD, diabetes)	Negative	91.4	78.5	70.6	75.3
	Positive	8.6	21.5	29.4	24.7

[1] Without participants with missing follow-up data or previously diagnosed CVD or diabetes (British: n=4961; Czech: n=5967; Polish: n=6543; Russian: n=6823)

4.2.2 Comparison of dietary intakes

On average, approximately 75% of total food/drink and energy intakes were captured by the 81 identical FFQ items in each cohort (table 4.5 and table 4.6). However, this proportion varied widely across food/drink groups, nutrients and cohorts. For example, on average, 2.2% of vegetable oil intake was provided by the common item in the Russian sample, while nearly all (96.1%-100%) of the fresh meat intake came from identical items in all four cohorts (table 4.5).

Table 4.5: Comparison of the FFQs used in the British, Czech, Polish and Russian cohorts

Overall food and drink categories	**Food and drink groups and subgroups (FoodEx2)**	**No. items in FFQ**				**No. items identical across the 4 FFQs**	**Mean percentage of food and drink intakes from the identical items[1]**			
		UK	**CZE**	**POL**	**RUS**		**UK**	**CZE**	**POL**	**RUS**
Foods of animal origin	Meat and meat products	9	15	14	15	**8**	98.2	76.2	81.5	86.2
	Animal fresh meat / animal offals	5	6	6	7	**5**	100.0	96.2	98.9	98.9
	Processed meat products / sausages and comminuted meat	4	9	8	8	**3**	92.1	40.5	56.2	53.7
	Milk and dairy products	9	13	15	12	**6**	25.4	49.4	50.2	59.8
	Eggs and egg products	1	1	1	1	**1**	100.0	100.0	100.0	100.0
	Fish, seafood, amphibians, reptiles and invertebrates	5	5	7	7	**3**	75.6	37.0	54.2	36.3
Foods of plant origin	Grains and grain-based products	15	10	10	10	**7**	72.6	74.1	72.1	66.1
	Fruits and fruit products	11	23	22	23	**11**	100.0	86.7	85.4	86.8
	Fresh fruits	8	20	19	20	**8**	100.0	85.5	84.1	81.6
	Processed fruit products	3	3	3	3	**3**	100.0	100.0	100.0	100.0
	Vegetables and vegetable products	18	25	28	26	**16**	94.9	79.9	72.5	87.2
	Vegetables (all non-products)[2]	18	22	24	23	**16**	94.9	89.0	86.2	94.2
	Vegetable products	0	3	4	3	**0**	na.	0.0	0.0	0.0
	Legumes, nuts, oilseeds and spices	6	6	4	6	**4**	87.9	60.4	100.0	78.5
	Starchy roots or tubers and products	4	3	3	3	**3**	84.2	100.0	100.0	100.0
	Sugar, confectionery and water-based sweet desserts	3	4	5	4	**3**	100.0	94.5	96.3	98.1

Overall food and drink categories	Food and drink groups and subgroups (FoodEx2)	No. items in FFQ				No. items identical across the 4 FFQs	Mean percentage of food and drink intakes from the identical items[1]			
		UK	**CZE**	**POL**	**RUS**		**UK**	**CZE**	**POL**	**RUS**
Foods of mixed origin	Animal and vegetable fats and oils	5	7	9	7	**3**	38.7	60.4	58.3	32.7
	Animal fats and oils	1	4	4	4	**1**	100.0	78.9	86.5	95.2
	Vegetable fats and oils	2	2	2	2	**1**	8.3	31.9	23.8	2.2
	Fats and oils of mixed origin	2	1	3	1	**1**	11.8	100.0	48.7	100.0
	Seasoning, sauces and condiments	6	3	4	3	**3**	64.2	100.0	95.4	100.0
	Composite dishes	10	8	13	13	**3**	58.5	64.7	47.9	41.0
Drinks	Alcoholic beverages	5	5	5	5	**5**	100.0	100.0	100.0	100.0
	Water and water-based beverages	2	4	2	2	**2**	100.0	25.0	100.0	100.0
	Coffee, cocoa, tea and infusions	5	2	3	3	**2**	89.3	100.0	98.4	99.2
	Fruit and vegetable juices and nectars	2	2	2	2	**1**	80.1	65.8	66.2	88.7
TOTAL		**116**	**136**	**147**	**142**	**81[3]**	**80.4**	**68.3**	**79.1**	**78.6**

[1] Values were calculated for each participant (in g/day) as follows: Intake from the 81 identical FFQ items*100 / Intake from all items in the original FFQs, for each food/drink group and overall

[2] Including: brassica vegetables; bulb, stalk and stem vegetables; fruiting vegetables; leafy vegetables; legume greens, sprouts; non-starchy root and tuber vegetables; fungi; marine algae, aromatic herbs or flowers

[3] Including nine which included more than one items each (combined items)

na. - not applicable

Table 4.6: Mean percentage of nutrient and energy intake from the identical items compared to the original FFQs in the four cohorts[1]

Nutrients/energy	UK	CZE	POL	RUS
Total carbohydrate (g/day)	76.4	76.7	75.8	74.7
Sugar (g/day)	81.0	78.2	76.5	83.9
Protein (g/day)	75.1	75.3	74.2	72.1
Total fat (g/day)	73.4	70.9	69.5	63.3
Saturated fat (g/day)	74.8	76.9	75.3	71.0
Polyunsaturated fat (g/day)	65.5	65.2	64.9	60.7
Trans fat (g/day)	57.2	76.9	78.0	79.3
Cholesterol (mg/day)	83.7	84.2	81.6	77.1
Alcohol (g/day)	100.0	100.0	100.0	100.0
Non-starch polysaccharides (g/day)	78.6	79.0	73.5	76.8
Vitamin C (mg/day)	86.8	80.1	72.3	66.8
Beta-carotene (ug/day)	91.7	89.7	89.8	94.9
Total energy (kJ/day)	76.7	75.0	73.4	70.4

[1] Values were calculated for each participant as follows:
Intake from the 81 identical FFQ items*100 / Intake from all items in the original FFQs, for each nutrient and energy

Table 4.7 shows the median (IQR) g/day intakes of foods and drinks which were considered fully or partially comparable across cohorts. Multivariable adjusted cross-cohort comparisons, using the UK values as reference, are also shown. Average total and fresh fruit intake was significantly lower in Russian and Polish participants but higher in Czechs compared to the UK cohort. Russians had the lowest fresh fruit intakes, with average consumption less than half of any other cohort. In contrast, vegetable intake was significantly higher in Russians but lower in Poles and Czechs compared to the British sample. British participants reported higher consumption of starchy roots, alcohol, coffee, tea, legumes and fruit juices, but less meat products, sweets and animal fats than any of the Eastern European cohorts.

Table 4.7: Average intake of foods and drinks in the British, Czech, Polish, Russian cohorts and the pooled Eastern European sample

Food groups and subgroups (FoodEx2)	UK	CZE		POL		RUS		POOLED Czech, Polish and Russian sample	
	Median[1] (IQR)	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]
Fully comparable foods and drinks[3]									
Animal fresh meat / animal offals	**74.2** (49.0-102.0)	**76.8** (47.6-111.6)	<0.0001	**76.8** (58.8-103.2)	<0.0001	**117.2** (68.4-154.8)	<0.0001	**85.2** (57.4-120.0)	<0.0001
Eggs	**7.0** (3.5-21.5)	**7.0** (7.0-21.5)	1.0	**21.5** (7.0-21.5)	<0.0001	**21.5** (7.0-21.5)	<0.0001	**21.5** (7.0-21.5)	<0.0001
Fruits and fruit products	**257.4** (157.4-385.4)	**277.9** (153.9-479.4)	<0.0001	**211.6** (123.6-347.9)	<0.0001	**129.3** (69.6-219.1)	<0.0001	**189.8** (104.1-339.4)	<0.0001
Fresh fruits	**232.1** (137.1-353.7)	**257.6** (138.6-452.0)	<0.0001	**189.0** (112.3-325.6)	<0.0001	**91.4** (43.1-179.7)	<0.0001	**164.4** (79.2-311.0)	<0.0001
Processed fruit products	**16.5** (7.0-32.0)	**14.7** (7.7-25.2)	<0.0001	**9.5** (2.5-18.8)	<0.0001	**21.5** (7.7-48.5)	<0.0001	**14.7** (7.0-29.2)	<0.0001
Vegetables *(all non-products)*[4]	**247.2** (169.7-341.2)	**186.1** (114.6-295.7)	<0.0001	**196.8** (127.2-303.2)	<0.0001	**291.0** (224.7-380.4)	<0.0001	**233.6** (143.8-332.4)	<0.0001
Starchy roots or tubers	**98.3** (75.3-151.8)	**86.8** (75.3-101.2)	<0.0001	**86.8** (75.3-141.1)	<0.0001	**86.8** (64.5-146.2)	<0.0001	**86.8** (75.3-138.3)	<0.0001
Sugars, confectionery and water-based sweet dessert	**8.1** (3.5-24.9)	**8.8** (3.5-21.5)	<0.0001	**19.6** (7.0-35.1)	<0.0001	**31.1** (15.6-42.9)	<0.0001	**19.1** (7.0-36.0)	<0.0001
Alcoholic beverages (portion/day)	**1.0** (0.3-2.5)	**0.3** (0.1-1.0)	<0.0001	**0.1** (0.0-0.2)	<0.0001	**0.1** (0.0-0.5)	<0.0001	**0.1** (0.0-0.5)	<0.0001
Coffee, cocoa, tea and infusions	**855.0** (503.0-1055.0)	**581.7** (390.0-690.0)	<0.0001	**675.0** (503.0-975.0)	<0.0001	**561.0** (475.0-855.0)	<0.0001	**675.0** (475.0-883.0)	<0.0001

Food groups and subgroups (FoodEx2)	UK	CZE		POL		RUS		POOLED Czech, Polish and Russian sample	
	Median[1] (IQR)	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]
Partially comparable foods and drinks[5]									
All meat and meat products	**90.1** (59.8-122.6)	**91.8** (59.8-130.9)	<0.0001	**104.8** (79.6-136.1)	<0.0001	**135.5** (90.9-179.3)	<0.0001	**109.1** (75.4-149.9)	<0.0001
Grains and grain based products	**185.9** (125.7-265.3)	**162.6** (109.1-229.5)	0.6978	**190.7** (134.8-263.4)	<0.0001	**217.1** (135.6-295.3)	<0.0001	**189.3** (127.0-267.7)	<0.0001
Legumes, nuts, oilseeds, spices	**31.3** (16.1-49.7)	**11.2** (6.3-18.2)	<0.0001	**11.2** (6.3-18.2)	<0.0001	**8.4** (4.9-14.7)	<0.0001	**11.2** (4.9-17.5)	<0.0001
Animal fats and oils	**0.0** (0.0-4.3)	**1.4** (0.7-10.0)	<0.0001	**7.9** (0.0-25.0)	<0.0001	**4.3** (1.4-10.0)	<0.0001	**4.3** (0.7-10.0)	<0.0001
Seasoning, sauces, condiments	**10.8** (4.3-26.7)	**12.2** (7.8-28.1)	<0.0001	**8.7** (4.3-20.0)	0.0034	**14.7** (4.3-32.9)	<0.0001	**12.2** (5.7-28.7)	<0.0001
Fruit and vegetable juices and nectars	**86.0** (14.0-200.0)	**14.0** (0.0-28.0)	<0.0001	**28.0** (0.0-86.0)	<0.0001	**14.0** (0.0-86.0)	<0.0001	**14.0** (0.0-86.0)	<0.0001

[1] Values are g/day intakes except for alcoholic beverages where portion/day intake is shown

[2] All p-values were calculated with quantile regression using the intake values in the UK cohort as reference category, adjusted for sex, age, energy intake, smoking, education, employment status, marital status, leisure time physical activity, CVD/diabetes in medical history

[3] On average, more than 80% of their intake was provided by the common items (n=81) in all four cohorts

[4] Including: brassica vegetables; bulb, stalk and stem vegetables; fruiting vegetables; leafy vegetables; legume greens, sprouts; non-starchy root and tuber vegetables; fungi; marine algae, aromatic herbs or flowers

[5] On average, 60-80% of their intake was provided by the common items (n=81) in at least one of the cohorts, and more than 80% in the other cohorts

Table 4.8 shows the medians (IQR) of energy-standardised nutrient intakes in the four cohorts, as well as the results of the quantile regression analysis. Only alcohol and beta-carotene intakes were fully comparable across cohorts (i.e.: more than 80% of their intake was provided by the 81 included items in all four cohorts). There was higher intake of beta-carotenes but lower intake of vitamin C in Russians compared to the other cohorts which is in line with the high vegetable and low fruit intake in this sample. Total fat, saturated fat and cholesterol intake were significantly higher in all three Eastern European cohorts than in the British sample, consistent with the food intake data. Alcohol consumption of British participants was the highest of any cohort.

In order to take into account the fact that multiple statistical tests were carried out, the p-values, which are used to indicate the threshold of statistical significance, were also calculated with Bonferroni`s correction method (Bland and Altman 1995). This approach suggested that in case of 112 (4x28) statistical tests, the threshold p-value of statistical significance is 0.00045 instead of 0.05. Since almost all p-values were lower than 0.0001, the differences seem to be statistically significant even if we take into account the issue of multiple testing.

An important difference between the Whitehall II and HAPIEE study participants was that the British cohort was based on civil service office workers, while large proportions of the Eastern European cohorts were engaged in physical occupations. However, in a sensitivity analysis restricting the comparisons to office workers the results were substantially similar (tables IV-1 and IV-2 in appendix). Further, the results were similar when the analysis was carried out separately in males or females (tables IV-3, IV-4, IV-5 and IV-6 in appendix).

Table 4.8: Average intake of nutrients in the British, Czech, Polish, Russian cohorts and the pooled Eastern European sample

Nutrients	UK	CZE		POL		RUS		POOLED Czech, Polish and Russian sample	
	Median[1] (IQR)	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]
Fully comparable nutrients[3]									
Alcohol (g/day)	**10.5** (3.0-24.7)	**2.6** (0.6-9.7)	<0.0001	**0.0** (0.0-2.4)	<0.0001	**1.1** (0.0-4.8)	<0.0001	**1.1** (0.0-4.9)	<0.0001
Beta-carotene (mg/day)	**6.4** (3.7-8.8)	**5.1** (3.6-8.1)	<0.0001	**7.3** (4.5-10.3)	<0.0001	**11.5** (7.8-14.3)	<0.0001	**7.7** (4.6-12.0)	<0.0001
Partially comparable nutrients[4]									
Total carbohydrate (g/day)	**235.0** (205.8-261.8)	**220.9** (194.3-247.9)	<0.0001	**225.4** (200.9-249.2)	<0.0001	**225.6** (200.2-249.8)	<0.0001	**224.4** (198.6-249.0)	<0.0001
Sugar (g/day)	**116.6** (94.9-140.0)	**108.4** (83.5-137.0)	<0.0001	**103.5** (83.3-126.9)	<0.0001	**107.4** (86.9-129.1)	<0.0001	**106.2** (84.7-130.3)	<0.0001
Protein (g/day)	**72.4** (64.0-82.1)	**78.3** (68.2-88.1)	<0.0001	**81.5** (73.1-90.6)	<0.0001	**81.9** (71.2-93.0)	<0.0001	**80.7** (71.0-90.7)	<0.0001
Total fat (g/day)	**66.8** (58.3-76.1)	**76.0** (67.2-85.0)	<0.0001	**78.0** (68.4-87.5)	<0.0001	**76.4** (67.9-85.2)	<0.0001	**76.9** (67.9-86.0)	<0.0001
Saturated fat (g/day)	**25.3** (21.2-30.2)	**31.3** (26.9-36.2)	<0.0001	**32.5** (27.2-38.8)	<0.0001	**29.2** (25.0-33.7)	<0.0001	**30.9** (26.2-36.2)	<0.0001
Polyunsaturated fat (g/day)	**11.4** (9.5-14.2)	**11.2** (9.5-13.1)	<0.0001	**10.7** (9.0-12.7)	<0.0001	**13.9** (11.0-17.5)	<0.0001	**11.6** (9.7-14.3)	0.7074
Cholesterol (mg/day)	**218.3** (171.7-274.2)	**308.7** (255.2-370.1)	<0.0001	**348.1** (294.9-403.8)	<0.0001	**319.8** (263.0-386.8)	<0.0001	**327.6** (271.8-389.2)	<0.0001
Non-starch polysaccharides (g/day)	**16.7** (14.1-20.0)	**15.8** (12.7-19.9)	<0.0001	**14.9** (12.4-18.0)	<0.0001	**14.4** (12.4-16.8)	<0.0001	**14.9** (12.5-18.0)	<0.0001

Nutrients	UK	CZE		POL		RUS		POOLED Czech, Polish and Russian sample	
	Median[1] (IQR)	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**
Vitamin C (mg/day)	**144.8** (102.7-199.8)	**137.2** (90.4-221.0)	0.0003	**108.6** (73.2-163.6)	<0.0001	**81.8** (56.8-131.2)	<0.0001	**106.1** (69.6-168.5)	<0.0001
Total energy (MJ/day)	**7.3** (6.0-8.9)	**6.4** (5.1-8.1)	<0.0001	**6.9** (5.6-8.4)	0.0015	**7.7** (6.1-9.4)	<0.0001	**7.0** (5.6-8.7)	0.1504

[1] All values are energy standardized around 8MJ/day, except for alcohol and total energy intake for which absolute intakes are shown

[2] All p-values were calculated with quantile regression using the intake values in the UK cohort as reference category, adjusted for sex, age, energy intake, smoking, education, employment status, marital status, leisure time physical activity, CVD/diabetes in medical history

[3] On average, more than 80% of their intake was provided by the common items (n=81) in all four cohorts

[4] On average, 60-80% of their intake was provided by the common items (n=81) in at least one of the cohorts, and more than 80% in the other cohorts

4.2.3 Contribution of fruit and vegetable intakes to the mortality differences between cohorts

Table 4.9 shows all-cause, CVD, CHD and stroke mortality rates in the Czech, Polish and Russian cohorts in relation to the British sample. Changes in HRs after different levels of multivariable adjustment are also indicated. In the basic adjusted model, the mortality rates of the three Eastern European cohorts were significantly higher compared to the British sample in all outcomes. The excess mortality was especially remarkable in the Russian sample. HRs decreased considerably after social and lifestyle factors were adjusted for in model 2. Approximately half of the excess mortality was explained by these factors in the Czech and Polish cohorts and about one third amongst Russians. When the associations were further adjusted for fruit or vegetable intake in model 3 and 4, there were no further reductions in the HRs for all-cause mortality. On the other hand, after adjusting for fruit intake, HRs for CVD, CHD and stroke mortalities decreased by 10.2%, 5.6% and 13.5%, respectively, in the Russian cohort. There was also a notable reduction in HRs for stroke mortality in the Czech and Polish samples (7.9% and 7.3%, respectively) after vegetable intake was adjusted for.

Table 4.9: **Differences in all-cause, CVD, CHD and stroke mortality rates between cohorts, and the change in hazard ratios after different levels of multivariable adjustment (n=24,294)**

Cause of death	Cohort	Model 1	Model 2		Model 3		Model 4	
		HR (95% CI)	HR (95% CI)	% change vs. model 1[1]	HR (95% CI)	% change vs. model 2[2]	HR (95% CI)	% change vs. model 2[3]
All-cause	**UK**	1.00	1.00		1.00		1.00	
	Czech	2.16 (1.83-2.55)	1.60 (1.34-1.91)	-48.3	1.60 (1.34-1.91)	0	1.61 (1.35-1.92)	+1.6
	Polish	2.70 (2.28-3.20)	1.92 (1.61-2.29)	-45.9	1.92 (1.61-2.29)	0	1.93 (1.61-2.30)	+1.1
	Russian	4.19 (3.56-4.92)	3.29 (2.77-3.90)	-28.2	3.29 (2.76-3.92)	0	3.28 (2.76-3.89)	-0.4
CVD	**UK**	1.00	1.00		1.00		1.00	
	Czech	2.91 (2.10-4.03)	2.03 (1.44-2.86)	-46.1	2.12 (1.50-2.99)	+8.7	2.02 (1.43-2.85)	-1.0
	Polish	3.37 (2.39-4.74)	2.25 (1.57-3.20)	-47.3	2.23 (1.55-3.17)	-1.6	2.23 (1.56-3.19)	-1.6
	Russian	8.52 (6.25-11.61)	6.21 (4.47-8.62)	-30.7	5.68 (4.06-7.94)	-10.2	6.24 (4.49-8.67)	+0.6
CHD	**UK**	1.00	1.00		1.00		1.00	
	Czech	2.48 (1.53-4.01)	1.70 (1.03-2.82)	-52.7	1.74 (1.05-2.89)	+5.7	1.70 (1.02-2.81)	0
	Polish	2.98 (1.82-4.87)	1.93 (1.16-3.21)	-53.0	1.91 (1.15-3.19)	-2.2	1.92 (1.15-3.21)	-1.1
	Russian	10.07 (6.54-15.49)	6.92 (4.37-10.95)	-34.7	6.59 (4.12-10.55)	-5.6	6.94 (4.38-11.00)	+0.3
Stroke	**UK**	1.00	1.00		1.00		1.00	
	Czech	3.49 (1.50-8.14)	2.40 (1.00-5.75)	-43.8	2.55 (1.06-6.14)	+10.7	2.29 (0.95-5.49)	-7.9
	Polish	4.45 (1.90-10.39)	3.06 (1.28-7.33)	-40.3	2.99 (1.25-7.17)	-3.4	2.91 (1.21-6.99)	-7.3
	Russian	16.32 (7.58-35.14)	11.80 (5.30-26.26)	-29.5	10.34 (4.58-23.37)	-13.5	12.34 (5.54-27.47)	+5.0

Model 1: adjusted for age and sex
Model 2: adjusted for all variables in model 1 and energy intake, smoking, education, employment status, marital status, leisure time physical activity
Model 3: adjusted for all variables in model 2 and ***fruit intake***
Model 4: adjusted for all variables in model 2 and ***vegetable intake***
[1] %=(HR2-HR1)/(HR1-1)*100; [2] %=(HR3-HR2)/(HR2-1)*100; [3] %=(HR4-HR2)/(HR2-1)*100

4.3 Objective 2: association between fruit, vegetable intake and mortality

The second objective of this thesis was the assessment of the association between fruit and vegetable intake and mortality in the HAPIEE study. As part of this analysis, the correlations between fruit and vegetable intake and plasma biomarker concentrations were re-assessed. Furthermore, the role of blood pressure, as a possible mediator between fruit and vegetable intake and mortality, was also assessed.

4.3.1 Correlation between fruit, vegetable intakes and plasma biomarkers

Table 4.10 shows the correlations of fruit and vegetable intakes with plasma vitamin C and beta-carotene levels on a subsample of participants who provided blood samples and did not take vitamin supplements regularly. The correlations between intakes and plasma concentrations of vitamin C and beta-carotene are also shown. The correlation coefficients indicated low and moderate agreements. Fruit intake correlated better with vitamin C plasma concentration, while vegetable intake showed higher agreement with beta-carotene. Correlation coefficients seemed to be higher in the Russian cohort than for Czechs and Poles. The agreement between vegetable intake and antioxidant vitamins was especially low for Czech males and Polish females.

Table 4.10: Correlations between fruit, vegetable, vitamin C, beta-carotene intakes and vitamin C, beta-carotene plasma concentrations

			Intake[1]:	FRUIT				VEGETABLE				VITAMIN C		BETA-CAROTENE	
			Plasma concentration[1]:	Vitamin C		Beta-carotene		Vitamin C		Beta-Carotene		Vitamin C		Beta-carotene	
Cohort	Sex group	n1[2]	n2[3]	r[4]	95%CI	r[4]	95%CI	r[4]	95%CI	r[4]	95%CI	r[4]	95%CI	r[4]	95%CI
Czech	Males	231	268	0.33	(0.21-0.44)	0.09	(-0.03-0.21)	0.11	(-0.02-0.24)	0.06	(-0.04-0.16)	0.29	(0.17-0.40)	0.00	(-0.12-0.12)
	Females	218	257	0.21	(0.08-0.33)	0.07	(-0.05-0.19)	0.12	(-0.01-0.25)	0.07	(-0.05-0.19)	0.20	(0.07-0.32)	0.08	(-0.04-0.20)
	All	**449**	**525**	**0.27**	**(0.18-0.35)**	**0.08**	**(-0.01-0.16)**	**0.11**	**(0.02-0.20)**	**0.07**	**(-0.02-0.16)**	**0.24**	**(0.15-0.33)**	**0.05**	**(-0.04-0.14)**
Polish	Males	262	364	0.17	(0.05-0.29)	0.17	(0.07-0.27)	0.18	(0.06-0.30)	0.16	(0.06-0.26)	0.26	(0.14-0.37)	0.17	(0.07-0.27)
	Females	243	340	0.11	(-0.02-0.23)	0.07	(-0.04-0.18)	-0.04	(-0.17-0.09)	0.08	(-0.03-0.19)	0.10	(-0.03-0.22)	0.06	(-0.05-0.17)
	All	**505**	**704**	**0.15**	**(0.06-0.23)**	**0.12**	**(0.05-0.19)**	**0.09**	**(0.00-0.18)**	**0.11**	**(0.04-0.18)**	**0.19**	**(0.10-0.27)**	**0.11**	**(0.04-0.18)**
Russian	Males	613	600	0.26	(0.19-0.33)	0.22	(0.14-0.30)	0.19	(0.11-0.27)	0.27	(0.19-0.34)	0.34	(0.27-0.41)	0.21	(0.13-0.29)
	Females	362	351	0.24	(0.14-0.34)	0.21	(0.11-0.31)	0.24	(0.14-0.34)	0.26	(0.16-0.36)	0.32	(0.22-0.41)	0.12	(0.02-0.22)
	All	**975**	**951**	**0.25**	**(0.19-0.31)**	**0.21**	**(0.15-0.27)**	**0.20**	**(0.14-0.26)**	**0.27**	**(0.21-0.33)**	**0.33**	**(0.27-0.39)**	**0.17**	**(0.11-0.23)**
Pooled	Males	1106	1232	0.30	(0.25-0.35)	0.11	(0.05-0.17)	0.14	(0.08-0.20)	0.18	(0.13-0.23)	0.36	(0.31-0.41)	0.18	(0.13-0.23)
	Females	823	948	0.26	(0.20-0.32)	0.03	(-0.03-0.09)	0.09	(0.02-0.16)	0.16	(0.10-0.22)	0.28	(0.22-0.34)	0.14	(0.08-0.20)
	All	**1929**	**2180**	**0.29**	**(0.25-0.33)**	**0.07**	**(0.03-0.11)**	**0.12**	**(0.08-0.16)**	**0.17**	**(0.13-0.21)**	**0.32**	**(0.28-0.36)**	**0.16**	**(0.12-0.20)**

[1] All data on intake and plasma concentration are log-transformed

[2] Number of participants with available data on plasma vitamin C concentration

[3] Number of participants with available data on plasma beta-carotene concentration

[4] Cohort, sex and energy intake adjusted partial Pearson`s correlation coefficient (cohort and sex adjustment were omitted in case of cohort- and sex-specific results)

4.3.2 Bivariate analysis of fruit and vegetable intakes

Table 4.11 shows the distribution of participants' socio-demographic and lifestyle characteristics and CVD risk factors across cohort-specific quartiles of fruit and vegetable intakes. Being female, higher education and higher household amenities score were positively associated with fruit and vegetable consumption. Those who ate more fruits and vegetables also seem to have had better overall diet, and were less likely to be heavy drinkers, smokers, or physically inactive. Among the potential mediators, mean arterial blood pressure declined but BMI increased and serum cholesterol level did not change with increasing consumption, which suggests that blood pressure was a possible but BMI and cholesterol were unlikely mediators between fruit and vegetable intake and CVD.

Table 4.11: Distribution of sample characteristics across cohort-specific fruit and vegetable intake quartiles

		Cohort-specific fruit and vegetable intake quartiles							
		Q1		Q2		Q3		Q4	
F&V intake	Median fruit intake (IQR), *g/day*	**75.2**	(36.4-127.1)	**170.2**	(95.7-246.0)	**268.8**	(158.0-369.8)	**482.3**	(306.6-686.7)
	Median vegetable intake (IQR), *g/day*	**119.4**	(80.3-161.8)	**189.4**	(138.1-234.1)	**247.0**	(183.1-318.0)	**371.3**	(262.6-495.4)
	Median fruit and vegetable intake (IQR), *g/day*	**214.1**	(165.2-251.3)	**352.1**	(318.7-412.6)	**514.7**	(449.1-591.1)	**831.4**	(698.5-1067.4)
Socio-demographic characteristics	Mean age (SD), *years*	**57.1**	(7.1)	**57.0**	(7.1)	**57.1**	(7.0)	**56.7**	(6.8)
	Sex: Females, %	**42.7**		**51.2**		**58.5**		**65.8**	
	Marital status: Married, %	**72.2**		**76.6**		**76.2**		**76.5**	
	Education: Primary or less, %	**11.2**		**10.0**		**10.1**		**8.6**	
	Education: University, %	**23.3**		**25.0**		**25.3**		**29.3**	
	Household amenities score: Low, %	**27.4**		**21.5**		**19.5**		**16.7**	
	Household amenities score: High, %	**28.4**		**32.5**		**34.8**		**38.7**	
Lifestyle characteristics	Mean energy intake (SD), *MJ/day*	**8.4**	(2.6)	**9.2**	(2.8)	**9.8**	(2.9)	**11.2**	(3.3)
	Mean HDI score (without F&V component) (SD)	**45.3**	(8.8)	**45.6**	(8.4)	**46.4**	(8.6)	**46.2**	(8.4)
	Median alcohol intake (IQR), *g/day*	**1.9**	(0.2-11.0)	**1.7**	(0.2-8.5)	**1.2**	(0.2-6.7)	**1.0**	(0.1-5.7)
	Alcohol: Moderate to heavy drinkers, %	**12.0**		**9.7**		**7.6**		**7.0**	
	Smoking: Current smokers, %	**38.6**		**30.5**		**27.7**		**25.9**	
	Physical activity: Low, %	**50.1**		**49.0**		**48.3**		**47.6**	
	Vitamin supplement intake: regular, %	**13.3**		**15.4**		**19.9**		**22.9**	
Possible mediators	Mean BMI (SD), *kg/m^2*	**27.3**	(4.7)	**27.7**	(4.7)	**28.0**	(4.7)	**28.1**	(4.8)
	BMI >30kg/m^2 , %	**24.6**		**27.7**		**29.0**		**30.7**	
	Mean MAP (SD), *mmHg*	**105.5**	(15.4)	**105.2**	(15.2)	**104.7**	(15.2)	**103.8**	(14.9)
	Hypertension, %	**46.9**		**47.7**		**47.1**		**44.8**	
	Mean serum cholesterol level (SD), mmol/l	**6.0**	(1.2)	**6.0**	(1.2)	**6.0**	(1.2)	**6.0**	**(1.2)**
	Hypercholesterolemia, %	**75.4**		**76.6**		**77.1**		**77.5**	

4.3.3 Multivariable Cox regression analysis

The associations between fruit and vegetable intake and the mortality outcomes are presented in table 4.12. Although inverse associations were found for all four mortality outcomes, statistically significant lower mortality risk in the highest compared to the lowest combined fruit and vegetable intake quartiles was found only for stroke after multiple adjustment. The trends were borderline significant for CVD and stroke, and non-significant for all-cause and CHD mortality. The preventable proportion (PP%) of death estimates indicated that if there is causal relationship between fruit, vegetable intake and mortality, and the intake increased by one quartile across the population distribution, than the reduction in mortality would be the greatest for stroke, potentially preventing 16% (95%CI: 0.5-34%) of cerebrovascular deaths. When the effects of fruit and vegetable intakes were analysed separately, the multivariable adjusted results indicated inverse but mostly statistically non-significant associations.

In the subgroup analysis, statistically significant inverse associations were found between overall fruit and vegetable intake and total mortality in current smokers but not in ex- or never smokers (table 4.13). Significantly reduced CVD and stroke mortality risk in the highest vs. lowest intake quartiles was also found only for smokers. When the results were further adjusted for the number of cigarettes smoked per day and the number of years has smoked, the associations remained statistically significant in this subgroup (table III-4 in appendix). In cohort-specific analysis, similarly to the pooled sample, most associations were found to be inverse but statistically not significant (table 4.14).

Table 4.12: Results of Cox regression analysis on the pooled sample

Cause of death	Deaths/n	Model	Cohort-specific quartiles: Q1 HR		Q2 HR	Q2 (95%CI)	Q3 HR	Q3 (95%CI)	Q4 HR	Q4 (95%CI)	p-value (trend)	PP% (95%CI)[2]	Per 100g/day increase[1] HR	(95%CI)
FRUIT AND VEGETABLE INTAKE														
All-cause	1314/19,333	model1	**1.00**	ref.	**0.78**	(0.68-0.90)	**0.77**	(0.66-0.89)	**0.67**	(0.58-0.79)	<0.001	**10.1** (5.9-14.4)	**0.90**	(0.87-0.93)
		model2	**1.00**	ref.	**0.93**	(0.80-1.08)	**0.97**	(0.83-1.13)	**0.91**	(0.76-1.08)	0.356	**2.4** (-1.9-7.1)	**0.98**	(0.94-1.02)
CVD	438/19,263	model1	**1.00**	ref.	**0.66**	(0.51-0.84)	**0.65**	(0.51-0.84)	**0.54**	(0.41-0.72)	<0.001	**16.1** (8.2-24.3)	**0.87**	(0.81-0.93)
		model2	**1.00**	ref.	**0.80**	(0.62-1.03)	**0.83**	(0.64-1.09)	**0.74**	(0.54-1.01)	0.060	**7.7** (-0.2-16.4)	**0.95**	(0.89-1.02)
CHD	226/19,263	model1	**1.00**	ref.	**0.62**	(0.44-0.87)	**0.61**	(0.43-0.87)	**0.60**	(0.41-0.87)	0.003	**14.3** (3.6-25.9)	**0.87**	(0.80-0.96)
		model2	**1.00**	ref.	**0.79**	(0.55-1.13)	**0.85**	(0.59-1.25)	**0.92**	(0.60-1.39)	0.608	**2.4** (-8.2-14.6)	**0.99**	(0.89-1.09)
Stroke	109/19,263	model1	**1.00**	ref.	**0.62**	(0.37-1.02)	**0.69**	(0.42-1.13)	**0.50**	(0.28-0.88)	0.019	**17.9** (3.0-34.8)	**0.88**	(0.77-1.00)
		model2	**1.00**	ref.	**0.67**	(0.40-1.12)	**0.73**	(0.44-1.24)	**0.52**	(0.28-0.98)	0.056	**16.3** (0.5-34.0)	**0.91**	(0.78-1.05)
FRUIT INTAKE[3]														
All-cause	1314/19,333	model1	**1.00**	ref.	**0.75**	(0.65-0.86)	**0.72**	(0.62-0.84)	**0.68**	(0.58-0.79)	<0.001	**10.2** (6.0-14.7)	**0.92**	(0.88-0.96)
		model2	**1.00**	ref.	**0.95**	(0.82-1.10)	**0.97**	(0.83-1.13)	**0.99**	(0.83-1.18)	0.845	**0.3** (-4.1-4.9)	**1.00**	(0.96-1.04)
CVD	438/19,263	model1	**1.00**	ref.	**0.77**	(0.61-0.97)	**0.55**	(0.42-0.72)	**0.53**	(0.40-0.70)	<0.001	**16.6** (8.8-24.7)	**0.84**	(0.78-0.91)
		model2	**1.00**	ref.	**1.00**	(0.79-1.28)	**0.75**	(0.56-0.99)	**0.78**	(0.57-1.07)	0.034	**6.2**(-1.6-14.7)	**0.92**	(0.84-0.99)
CHD	226/19,263	model1	**1.00**	ref.	**0.67**	(0.48-0.94)	**0.51**	(0.35-0.74)	**0.54**	(0.36-0.80)	<0.001	**16.9** (5.7-29.2)	**0.85**	(0.76-0.95)
		model2	**1.00**	ref.	**0.91**	(0.65-1.28)	**0.73**	(0.49-1.08)	**0.86**	(0.55-1.33)	0.235	**4.0** (-7.0-16.7)	**0.95**	(0.85-1.07)
Stroke	109/19,263	model1	**1.00**	ref.	**0.90**	(0.57-1.43)	**0.62**	(0.37-1.05)	**0.51**	(0.28-0.93)	0.011	**16.0** (1.6-32.4)	**0.82**	(0.70-0.97)
		model2	**1.00**	ref.	**1.12**	(0.69-1.82)	**0.79**	(0.45-1.38)	**0.66**	(0.34-1.29)	0.164	**9.4** (-5.3-26.6)	**0.87**	(0.73-1.03)

Cause of death	Deaths/n	Model	Cohort-specific quartiles								p-value (trend)	PP% (95%CI)[2]	Per 100g/day increase[1]	
			Q1		Q2		Q3		Q4					
			HR		HR	(95%CI)	HR	(95%CI)	HR	(95%CI)			HR	(95%CI)
VEGETABLE INTAKE[3]														
All-cause	1314/19,333	model1	**1.00**	ref.	**0.76**	(0.65-0.88)	**0.75**	(0.65-0.87)	**0.72**	(0.62-0.83)	<0.001	**8.8** (4.7-13.0)	**0.93**	(0.89-0.97)
		model2	**1.00**	ref.	**0.82**	(0.70-0.95)	**0.84**	(0.72-0.98)	**0.85**	(0.72-1.00)	0.052	**4.4** (0.0-8.9)	**0.98**	(0.93-1.03)
CVD	438/19,263	model1	**1.00**	ref.	**0.83**	(0.65-1.06)	**0.69**	(0.53-0.89)	**0.67**	(0.51-0.88)	<0.001	**10.3** (3.1-18.2)	**0.90**	(0.84-0.98)
		model2	**1.00**	ref.	**0.93**	(0.72-1.20)	**0.81**	(0.62-1.07)	**0.88**	(0.66-1.19)	0.249	**3.2** (-4.3-11.3)	**0.99**	(0.90-1.07)
CHD	226/19,263	model1	**1.00**	ref.	**0.81**	(0.58-1.15)	**0.65**	(0.45-0.94)	**0.71**	(0.49-1.02)	0.027	**9.2** (-0.5-20.2)	**0.91**	(0.82-1.01)
		model2	**1.00**	ref.	**0.94**	(0.66-1.34)	**0.82**	(0.55-1.20)	**1.00**	(0.66-1.51)	0.745	**0.0** (-10.1-11.8)	**1.01**	(0.89-1.14)
Stroke	109/19,263	model1	**1.00**	ref.	**0.73**	(0.44-1.21)	**0.64**	(0.38-1.08)	**0.64**	(0.38-1.07)	0.066	**12.1** (-1.6-28.2)	**0.91**	(0.78-1.06)
		model2	**1.00**	ref.	**0.76**	(0.45-1.26)	**0.65**	(0.38-1.13)	**0.69**	(0.39-1.24)	0.157	**10.0** (-5.2-27.5)	**0.94**	(0.79-1.12)

Model 1: adjusted for sex, age, cohort

Model 2: adjusted for sex, age, cohort, alcohol intake, smoking, education, household amenities score, marital status, energy intake, physical activity, vitamin supplement Intake, HDI (without the fruit and vegetable component)

[1] Per one unit increase across six intake categories (<100g/d, 1-200g/d, 2-300g/d, 3-400g/d, 4-500g/d, >500g/d)

[2] Preventable proportion of death if participants in the lowest three quartiles increased their intake one quartile upward

[3] In model 2, fruit and vegetable intakes were mutually adjusted for each-other

Table 4.13: Results of Cox regression analysis by smoking groups

Cause of death	Subgroup	Deaths/n	Cohort-specific fruit and vegetable intake quartiles: Q1 HR		Q2 HR	Q2 (95%CI)	Q3 HR	Q3 (95%CI)	Q4 HR	Q4 (95%CI)	p-value (trend)	PP% (95%CI)[2]	Per 100g/day increase[1] HR	(95%CI)
All-cause	Current smokers	638/5905	**1.00**	ref.	**0.90**	(0.73-1.11)	**0.87**	(0.70-1.09)	**0.70**	(0.53-0.91)	0.011	**8.8** (2.2-15.9)	**0.93**	(0.87-0.98)
	Ex-smokers	300/4080	**1.00**	ref.	**0.94**	(0.69-1.28)	**0.93**	(0.67-1.29)	**1.09**	(0.76-1.57)	0.748	**-2.3** (-11.1-7.7)	**1.00**	(0.92-1.10)
	Never smokers	369/9272	**1.00**	ref.	**1.02**	(0.76-1.38)	**1.22**	(0.91-1.66)	**1.20**	(0.86-1.67)	0.168	**-4.5** (-11.7-4.0)	**1.05**	(0.97-1.14)
CVD	Current smokers	226/5871	**1.00**	ref.	**0.75**	(0.53-1.06)	**0.78**	(0.54-1.14)	**0.62**	(0.40-0.97)	0.037	**11.9** (0.7-24.3)	**0.94**	(0.85-1.04)
	Ex-smokers	94/4062	**1.00**	ref.	**0.93**	(0.55-1.57)	**0.71**	(0.39-1.32)	**1.06**	(0.55-2.03)	0.782	**-1.6** (-17.4-18.1)	**0.92**	(0.79-1.08)
	Never smokers	117/9254	**1.00**	ref.	**0.79**	(0.47-1.32)	**1.06**	(0.64-1.77)	**0.80**	(0.44-1.45)	0.747	**5.5** (-8.1-22.0)	**1.01**	(0.87-1.16)
CHD	Current smokers	125/5871	**1.00**	ref.	**0.72**	(0.44-1.16)	**0.82**	(0.50-1.37)	**0.76**	(0.43-1.35)	0.340	**7.3** (-7.2-24.1)	**0.98**	(0.86-1.12)
	Ex-smokers	49/4062	**1.00**	ref.	**1.07**	(0.52-2.20)	**0.52**	(0.20-1.37)	**1.48**	(0.63-3.47)	0.828	**11.9** (-30.7-15.7)	**0.91**	(0.73-1.13)
	Never smokers	51/9254	**1.00**	ref.	**0.76**	(0.34-1.71)	**1.32**	(0.62-2.85)	**0.97**	(0.39-2.40)	0.710	**0.8** (-17.6-26.0)	**1.10**	(0.88-1.37)
Stroke	Current smokers	50/5871	**1.00**	ref.	**0.76**	(0.37-1.56)	**0.66**	(0.30-1.46)	**0.30**	(0.10-0.94)	0.038	**25.6** (1.2-50.8)	**0.85**	(0.68-1.06)
	Ex-smokers	18/4062	**1.00**	ref.	**0.70**	(0.15-3.23)	**1.86**	(0.49-7.00)	**2.09**	(0.49-8.87)	0.172	**19.3** (-39.2-23.9)	**1.34**	(0.91-1.98)
	Never smokers	41/9254	**1.00**	ref.	**0.55**	(0.23-1.30)	**0.57**	(0.24-1.34)	**0.43**	(0.16-1.17)	0.110	**22.1** (-3.5-51.5)	**0.85**	(0.66-1.08)

All HRs are adjusted for sex, age, cohort, alcohol intake, education, household amenities score, marital status, energy intake, physical activity, vitamin supplement intake, HDI (without the fruit and vegetable component)

[1] Per one unit increase across six intake categories (<100g/d, 1-200g/d, 2-300g/d, 3-400g/d, 4-500g/d, >500g/d)

[2] Preventable proportion of death if participants in the lowest three quartiles increased their intake one quartile upward

<u>Table 4.14:</u> Results of Cox regression analysis by country cohorts

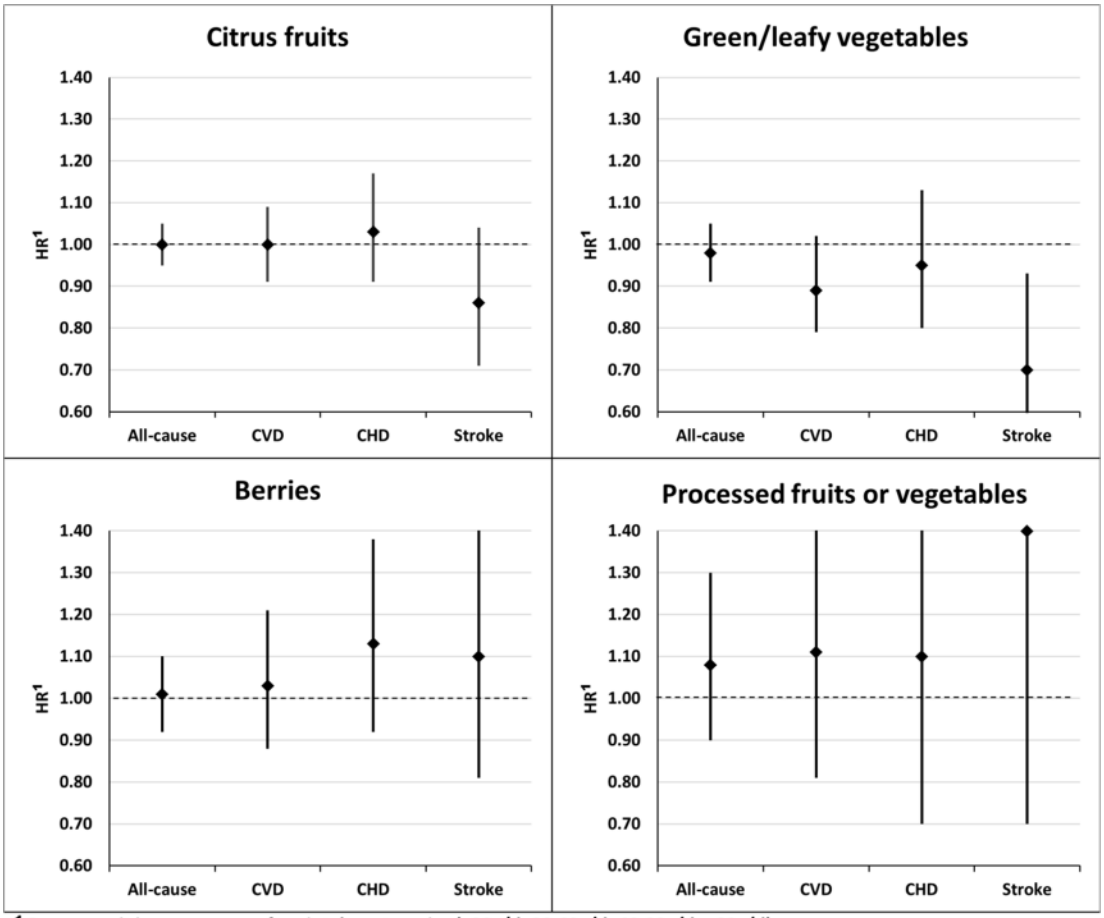
			Fruit and vegetable intake quartiles										Per 100g/day increase[1]	
			Q1		Q2		Q3		Q4					
Cause of death	**Subgroup**	**Deaths/n**	**HR**		**HR**	**(95%CI)**	**HR**	**(95%CI)**	**HR**	**(95%CI)**	**p-value (trend)**	**PP% (95%CI)[2]**	**HR**	(95%CI)
All-cause	Czech	364/5967	**1.00**	ref.	**0.82**	(0.62-1.10)	**0.92**	(0.68-1.24)	**0.94**	(0.67-1.32)	0.817	**1.6** (-6.9-11.1)	**0.97**	(0.90-1.05)
	Polish	388/6543	**1.00**	ref.	**1.00**	(0.76-1.31)	**1.04**	(0.78-1.39)	**1.05**	(0.75-1.47)	0.716	**-1.3** (-9.1-7.6)	**1.01**	(0.94-1.09)
	Russian	562/6823	**1.00**	ref.	**0.98**	(0.79-1.23)	**0.98**	(0.77-1.24)	**0.85**	(0.65-1.12)	0.314	**3.9** (-2.6-10.9)	**0.97**	(0.91-1.03)
CVD	Czech	106/5965	**1.00**	ref.	**0.59**	(0.35-1.00)	**0.64**	(0.37-1.11)	**0.69**	(0.37-1.30)	0.197	**10.5** (-6.8-30.1)	**0.90**	(0.78-1.03)
	Polish	99/6517	**1.00**	ref.	**1.01**	(0.59-1.73)	**1.29**	(0.75-2.21)	**0.91**	(0.45-1.85)	0.815	**2.1** (-12.5-19.7)	**1.06**	(0.91-1.24)
	Russian	233/6781	**1.00**	ref.	**0.84**	(0.59-1.18)	**0.80**	(0.55-1.16)	**0.77**	(0.51-1.16)	0.168	**6.8** (-3.6-18.5)	**0.95**	(0.86-1.05)
CHD	Czech	43/5965	**1.00**	ref.	**0.51**	(0.21-1.22)	**0.66**	(0.28-1.56)	**0.76**	(0.29-2.00)	0.564	**8.2** (-17.3-39.9)	**0.92**	(0.74-1.14)
	Polish	45/6517	**1.00**	ref.	**0.85**	(0.37-1.96)	**1.45**	(0.66-3.15)	**0.82**	(0.27-2.50)	0.798	**4.4** (-17.4-31.7)	**1.11**	(0.88-1.38)
	Russian	138/6781	**1.00**	ref.	**0.90**	(0.57-1.40)	**0.79**	(0.47-1.31)	**1.07**	(0.64-1.80)	0.938	**-1.9** (-14.5-13.4)	**1.00**	(0.87-1.14)
Stroke	Czech	18/5965	**1.00**	ref.	**0.23**	(0.05-1.16)	**0.54**	(0.14-2.05)	**0.59**	(0.13-2.68)	0.554	**17.2** (-24.4-65.9)	**0.97**	(0.68-1.39)
	Polish	20/6517	**1.00**	ref.	**0.91**	(0.30-2.81)	**0.62**	(0.17-2.24)	**0.44**	(0.09-2.06)	0.250	**18.9** (-13.1-58.3)	**0.89**	(0.64-1.25)
	Russian	71/6781	**1.00**	ref.	**0.72**	(0.38-1.35)	**0.84**	(0.44-1.59)	**0.52**	(0.23-1.14)	0.157	**15.8** (-2.8-37.6)	**0.89**	(0.74-1.08)

All HRs are adjusted for sex, age, alcohol intake, smoking, education, household amenities score, marital status, energy intake, physical activity, vitamin supplement intake, HDI (without F&V component)

[1] Per one unit increase across six intake categories (<100g/d, 1-200g/d, 2-300g/d, 3-400g/d, 4-500g/d, >500g/d)

[2] Preventable proportion of death if participants in the lowest three quartiles increased their intake one quartile upward

The associations with mortality outcomes were largely non-significant when fruit and vegetable subgroups were analysed separately. From the four examined subgroups, green leafy vegetables showed the most consistent inverse association across the four mortality outcomes, reaching statistical significance for stroke (figure 4.1 and table III-5 in appendix)



[1] Per one unit increase across four intake categories (<30g/d; 30-60g/d; 60-90g/d; >90g/d)

<u>Figure 4.1:</u> Multivariable adjusted hazard ratios of all-cause and cause-specific mortality outcomes per 30g/day increase in the intake of selected fruit and vegetable subgroups

4.3.4 Mediating effect of blood pressure

To assess the potential mediating role of blood pressure, the analysis was conducted with and without additional adjustment for mean arterial blood pressure (MAP) on a subsample of participants who took no antihypertensive medication at baseline (table 4.15). After adjusting for MAP, the associations with fruit and vegetable were attenuated for all four mortality outcomes. The reduction in the strength of the association was largest for CVD (the change in the HR between highest vs. lowest quartile was 37%).

Table 4.15: Results of Cox regression analysis before and after adjustment for blood pressure (MAP) on a subsample of participants who took no antihypertensive medications

			Cohort-specific fruit and vegetable intake quartiles								Per 100g/day increase[1]	
			Q1		Q2		Q3		Q4			
Cause of death	**Deaths/n**	**Model**	**HR**		**HR**	**(95%CI)**	**HR**	**(95%CI)**	**HR**	**(95%CI)**	**HR**	**(95%CI)**
All-cause	939/13,966	model1	**1.00**	ref.	**0.92**	(0.77-1.09)	**0.94**	(0.78-1.13)	**0.88**	(0.71-1.08)	**0.95**	(0.91-1.00)
		model2	**1.00**	ref.	**0.92**	(0.77-1.09)	**0.96**	(0.80-1.15)	**0.90**	(0.73-1.11)	**0.96**	(0.91-1.01)
		Percentage change[2] (%)			**0**		**33.3**		**16.7**		**20.0**	
CVD	305/13,915	model1	**1.00**	ref.	**0.79**	(0.58-1.07)	**0.76**	(0.55-1.06)	**0.81**	(0.56-1.17)	**0.94**	(0.87-1.03)
		model2	**1.00**	ref.	**0.81**	(0.60-1.10)	**0.80**	(0.58-1.12)	**0.88**	(0.61-1.28)	**0.96**	(0.88-1.05)
		Percentage change[2] (%)			**9.5**		**16.7**		**36.8**		**33.3**	
CHD	175/13,915	model1	**1.00**	ref.	**0.71**	(0.47-1.07)	**0.76**	(0.49-1.18)	**0.99**	(0.62-1.57)	**0.98**	(0.87-1.10)
		model2	**1.00**	ref.	**0.73**	(0.49-1.11)	**0.80**	(0.51-1.24)	**1.07**	(0.67-1.71)	**1.00**	(0.89-1.12)
		Percentage change[2] (%)			**6.9**		**16.7**		**na.**		**100.0**	
Stroke	65/13,915	model1	**1.00**	ref.	**0.81**	(0.43-1.53)	**0.66**	(0.32-1.34)	**0.53**	(0.23-1.22)	**0.86**	(0.71-1.05)
		model2	**1.00**	ref.	**0.85**	(0.45-1.60)	**0.72**	(0.35-1.48)	**0.62**	(0.26-1.44)	**0.89**	(0.73-1.08)
		Percentage change[2] (%)			**21.1**		**17.6**		**19.1**		**21.4**	

Model 1: adjusted for sex, age, cohort, alcohol intake, smoking, education, household amenities score, marital status, energy intake, physical activity, vitamin supplement Intake, HDI (without F&V component)

Model 2: adjusted for all covariates in model 1 + MAP

[1] per one unit increase across six intake categories (<100g/d, 1-200g/d, 2-300g/d, 3-400g/d, 4-500g/d, >500g/d)

[2] % change=(HR2-HR1)/(1-HR1)*100

na, not applicable

4.4 Objective 3: healthy diet indicator and mortality

The healthy diet indicator is a predefined diet quality index which was constructed according to the WHO dietary recommendations for the prevention of chronic diseases. Its relationship with mortality outcomes was assessed in the HAPIEE study and the results are presented below.

4.4.1 HDI components and bivariate analysis

The median (IQR) HDI component scores by cohort and sex are shown in table 4.16, and table 4.17 presents the mean (SD) overall HDI scores by covariate categories. The differences in the overall HDI score between country cohorts were due to differences in specific HDI components. In particular, the intakes of n-3 and n-6 polyunsaturated fatty acids and mono/disaccharides were further from the WHO recommendations amongst Polish participants compared to Czechs and Russians, which resulted in lower component scores, and consequently, lower overall HDI score in this cohort.

The HDI scores were higher in women, older participants and regular vitamin supplement users, but lower in heavy drinkers and current smokers. Surprisingly, the mean HDI score seemed lower in people with higher education and in subjects with higher household amenities score.

<u>Table 4.16:</u> HDI component scores by cohort and sex

Components of the HDI	Median scores (IQR)					
	CZECH		POLISH		RUSSIAN	
	Males	Females	Males	Females	Males	Females
Saturated fatty acids, *energy%*	2.3 (0.0-5.4)	3.3 (0.2-6.6)	0.0 (0.0-3.2)	1.0 (0.0-4.6)	0.3 (0.0-3.7)	1.5 (0.0-4.9)
n3-Polyunsaturated fatty acids, *energy%*	4.2 (3.3-5.4)	4.5 (3.5-5.6)	3.2 (2.4-4.3)	3.0 (2.2-4.0)	5.4 (4.2-7.3)	6.2 (4.9-8.6)
n6-Polyunsaturated fatty acids, *energy%*	4.7 (3.7-5.7)	4.5 (3.5-5.6)	3.4 (2.6-4.5)	3.1 (2.4-4.2)	6.6 (4.7-8.8)	7.5 (5.5-9.8)
Trans fatty acids, *energy%*	9.7 (7.8-10.0)	9.8 (7.7-10.0)	9.7 (7.4-10.0)	9.9 (7.8-10.0)	10.0 (9.2-10.0)	10.0 (10.0-10.0)
Mono- and disaccharides, *energy%*	4.4 (2.3-6.5)	2.5 (0.0-4.7)	4.4 (2.5-6.2)	2.8 (0.7-4.8)	6.2 (4.8-7.6)	5.1 (3.2-6.6)
Protein, *energy%*	6.9 (4.9-8.6)	7.6 (5.7-9.4)	6.8 (5.2-8.3)	7.1 (5.4-8.6)	7.4 (5.8-8.9)	7.8 (5.8-9.6)
Cholesterol, *mg/day*	10.0 (1.5-10)	10.0 (6.1-10.0)	0.3 (0.0-8.1)	6.5 (0.0-10.0)	0.0 (0.0-2.3)	2.2 (0.0-10.0)
Fruits/vegetables, *g/day*	10.0 (6.3-10.0)	10.0 (9.4-10.0)	10 (7.2-10.0)	10 (8.6-10.0)	8.2 (6.0-10.0)	9.5 (6.8-10.0)
Non-starch polysaccharides, *g/day*	7.7 (5.9-10.0)	8.8 (6.6-10.0)	9.0 (7.1-10.0)	9.1 (7.0-10.0)	8.7 (7.2-10.0)	8.4 (6.8-10.0)

energy% – percentage of daily alcohol-free energy intake

Table 4.17: Overall HDI scores by covariate categories

Covariate[1]	Category	Mean HDI score (SD)	p-value (crude)	p-value (adjusted)[2]
Cohort[3]	Czech	55.8 (8.0)	ref.	ref.
	Polish	49.8 (7.1)	<0.001	<0.001
	Russian	57.3 (9.2)	<0.001	<0.001
Sex[4]	Males	52.7 (8.5)		
	Females	55.7 (8.9)	<0.001	<0.001
Age groups[5]	<50 years	53.6 (8.4)		
	50-54 years	53.7 (8.6)		
	55-59 years	54.2 (8.8)		
	60-64 years	54.8 (9.0)		
	65+ years	55.7 (9.3)	<0.001	<0.001
Marital status[4]	Single/divorced/widowed	55.5 (9.4)		
	Married/cohabiting	53.9 (8.6)	<0.001	0.433
Education[5]	Incomplete/primary	54.8 (9.3)		
	Vocational	55.0 (8.8)		
	Secondary	54.2 (8.7)		
	University	53.6 (8.8)	<0.001	0.003
Household amenities score[5]	Low	55.8 (9.6)		
	Moderate	54.3 (8.8)		
	High	53.3 (8.2)	<0.001	0.006
Alcohol intake[5]	Abstainers	54.2 (9.1)		
	Moderate drinkers	54.4 (8.7)		
	Heavy drinkers	53.7 (8.5)	0.514	0.006
Smoking habits[3]	No smoker	55.6 (9.0)	ref.	ref.
	Ex-smoker	53.6 (8.4)	<0.001	0.772
	Current smoker	52.8 (8.6)	<0.001	<0.001
Physical activity[5]	Inactive	54.6 (9.1)		
	Moderately active	54.4 (8.7)		
	Active	53.8 (8.2)	0.006	0.810
Vitamin supplement usage[5]	Non-users	54.2 (9.0)		
	Irregular users	54.1 (8.4)		
	Regular users	54.9 (8.7)	0.017	<0.001
Energy intake[5]	Low (<8MJ/day)	55.7 (8.9)		
	Moderate (8-10MJ/day)	54.9 (9.6)		
	High (>10MJ/day)	52.9 (8.0)	<0.001	<0.001
BMI[5]	Low (<25kg/m^2)	53.8 (8.8)		
	Moderate (25-30kg/m^2)	54.2 (8.7)		
	High (>30kg/m^2)	55.1 (9.0)	<0.001	<0.001

Covariate[1]	Category	Mean HDI score (SD)	p-value (crude)	p-value (adjusted)[2]
Hypertension[4]	Hypertensive	55.0 (9.0)		
	Not hypertensive	54.0 (8.6)	<0.001	<0.001
Hyper-cholesterolemia[4]	Hypercholesterolaemic	54.6 (8.9)		
	Not hypercholesterol.	54.1 (8.8)	0.007	0.109

[1] Only participants with complete data were included; [2] cohort, sex, age and energy intake adjusted p-values; [3] p-values calculated with multinomial logistic regression; [4] p-values calculated with simple logistic regression; [5] p-values calculated with ordered logistic regression

ref. – reference category

4.4.2 Multivariable Cox-regression analysis

Table 4.18 shows the results of the Cox regression analysis for the association between HDI and mortality on the pooled sample and in each cohort. In the pooled sample, one SD increase in the HDI was inversely and statistically significantly associated with CVD and CHD mortality but not with deaths from other causes. As a result, there was an inverse but statistically not significant association with all-cause mortality. Most cohort-specific results were similar; there were statistically significant associations between HDI and both CVD and CHD mortality in the Russian cohort and with all-cause mortality in the Polish cohort. The adjustment for covariates (model 2) resulted in a small attenuation in the strengths of most associations but did not radically change the pattern of results.

When participants were classified into four categories based on their HDI score`s distance from the sample mean, the results indicated an approximately linear relationship between HDI and CVD and CHD mortality (figure 4.2 and table V-2 in appendix). Preventable proportion of deaths was also the highest for CVD and CHD outcomes.

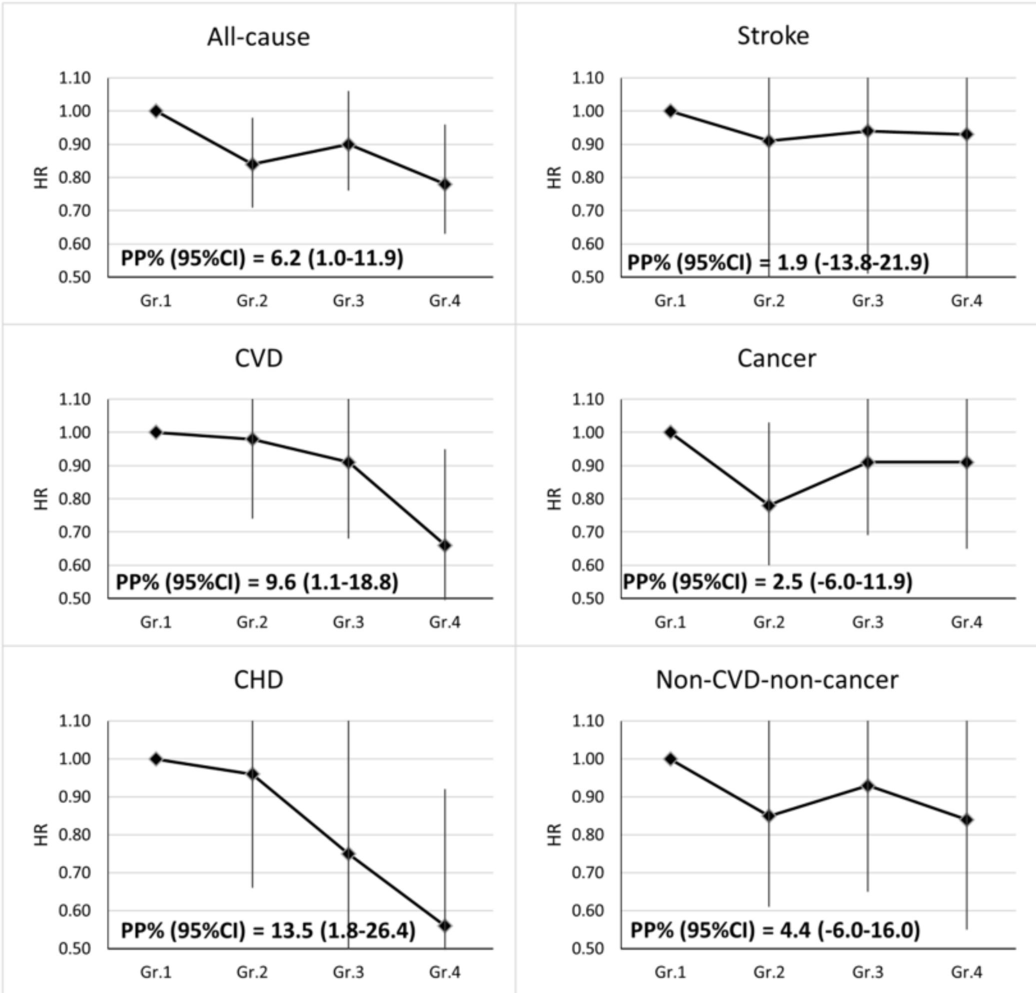
Table 4.18: **Results of Cox regression analysis for the association between HDI and mortality on the pooled and cohort-specific samples**

Cause of death	Sample	Dead/n	Model 1		Model 2	
			HR/SD (95%CI)[1]	p-value	HR/SD (95%CI)[1]	p-value
All-cause	**Pooled**	**1209/18,559**	**0.94 (0.89-1.00)**	**0.055**	**0.95 (0.89-1.00)**	**0.068**
	Czech	330/ 5632	0.96 (0.85-1.08)	0.512	0.97 (0.86-1.09)	0.611
	Polish	343/ 6278	0.83 (0.72-0.95)	0.007	0.86 (0.75-0.98)	0.027
	Russian	536/ 6649	0.99 (0.91-1.08)	0.879	0.98 (0.90-1.06)	0.506
CVD	**Pooled**	**423/18,494**	**0.89 (0.81-0.99)**	**0.030**	**0.90 (0.81-0.99)**	**0.030**
	Czech	102/ 5630	0.95 (0.77-1.18)	0.646	0.95 (0.77-1.17)	0.620
	Polish	92/ 6256	0.94 (0.72-1.22)	0.632	0.96 (0.74-1.25)	0.762
	Russian	229/ 6608	0.88 (0.77-1.00)	0.048	0.87 (0.77-0.99)	0.029
CHD	**Pooled**	**220/18,494**	**0.85 (0.74-0.97)**	**0.020**	**0.85 (0.74-0.97)**	**0.018**
	Czech	43/ 5630	0.94 (0.68-1.30)	0.698	0.98 (0.71-1.35)	0.907
	Polish	41/ 6256	0.77 (0.52-1.14)	0.197	0.84 (0.57-1.25)	0.400
	Russian	136/ 6608	0.84 (0.71-1.00)	0.044	0.83 (0.70-0.97)	0.020
Stroke	**Pooled**	**105/18,494**	**0.95 (0.78-1.16)**	**0.623**	**0.96 (0.79-1.16)**	**0.657**
	Czech	17/ 5630	0.89 (0.53-1.48)	0.644	0.87 (0.52-1.46)	0.600
	Polish	19/ 6256	1.22 (0.70-2.14)	0.485	1.20 (0.67-2.13)	0.540
	Russian	69/ 6608	0.95 (0.76-1.19)	0.653	0.95 (0.76-1.19)	0.657
Cancer	**Pooled**	**437/18,494**	**0.98 (0.88-1.08)**	**0.670**	**0.98 (0.89-1.09)**	**0.712**
	Czech	153/ 5630	0.96 (0.81-1.14)	0.654	0.97 (0.82-1.16)	0.760
	Polish	143/ 6256	0.84 (0.68-1.04)	0.102	0.86 (0.69-1.06)	0.151
	Russian	141/ 6608	1.10 (0.94-1.29)	0.223	1.08 (0.92-1.27)	0.345
Non-CVD-non-cancer	**Pooled**	**284/18,494**	**0.96 (0.84-1.09)**	**0.500**	**0.96 (0.84-1.08)**	**0.474**
	Czech	73/ 5630	0.97 (0.75-1.25)	0.795	0.98 (0.76-1.26)	0.881
	Polish	86/ 6256	0.71 (0.54-0.94)	0.030	0.76 (0.58-1.00)	0.053
	Russian	125/ 6608	1.08 (0.91-1.29)	0.379	1.03 (0.87-1.22)	0.702

Model 1: adjusted for age, sex, cohort

Model 2: adjusted for age, sex, cohort, education, household amenities score, marital status, smoking, alcohol intake, energy intake, physical activity, vitamin supplement intake

[1] effect of one standard deviation (SD) increase in the score



Groups are based on distance of HDI score from sample mean:

- **Gr.1.:** ≤-1SD
- **Gr.2.:** >-1SD and ≤ mean
- **Gr.3.:** > mean and ≤ 1 SD
- **Gr.4.:** > 1 SD

PP% – Preventable proportion of deaths if participants in the lowest three groups increased their HDI one group upward

Figure 4.2: Multivariable adjusted hazard ratios (95% CIs) of all-cause and cause-specific mortalities across categorical HDI groups (reference category: Gr. 1), and preventable proportions of deaths

When the analysis included subjects with prevalent diabetes, CVD or cancer (increasing the sample size to 25,858), no significant associations between HDI and CVD or CHD mortality was found but there was a suggestion of an inverse association with non-CVD-non-cancer mortality and with all-cause mortality (table V-3 in appendix). This finding supports the view that people who are diagnosed with chronic diseases are likely to change their diet as a result of their condition, and that this reverse causation can have significant impact on the associations observed.

I also assessed the effects on mortality of the original HDI score, based on the earlier dichotomous scoring method by Huijbregts and colleagues (Huijbregts *et al.* 1997). No association between this "original" HDI and mortality outcomes was found (table V-4 in appendix). This negative finding may be explained by the fact that the correlation between the "original" and newly constructed HDI scores was low (Pearson`s r = 0.25).

Age- and sex-adjusted mortality differences between the Czech and Polish cohorts were not statistically significant for most outcomes, which made it unfeasible to assess the contribution of the HDI in the mortality differences between these two cohorts (table 4.19). Although mortality rates in the Russian cohort were significantly higher compared to Czechs, diet quality measured by the HDI did not seem to explain any of these differences.

Table 4.19: **Differences in mortality rates between cohorts, and the change in hazard ratios after different levels of multivariable adjustment**

Cause of death	Cohort	Model 1	Model 2		Model3	
		HR (95% CI)	HR (95% CI)	Percentage change in HR[1]	HR (95% CI)	Percentage change in HR[2]
All-cause	Czech	1.0	1.0		1.0	
	Polish	1.18 (1.01-1.38)	1.12 (0.95-1.32)	-33.3	1.08 (0.91-1.30)	-33.3
	Russian	1.97 (1.70-2.27)	1.82 (1.55-2.14)	-15.5	1.85 (1.57-2.17)	+3.7
CVD	Czech	1.0			1.0	
	Polish	1.08 (0.81-1.45)	0.93 (0.69-1.27)	>-100	0.87 (0.64-1.20)	+85.7
	Russian	2.86 (2.23-3.67)	2.44 (1.84-3.22)	-22.6	2.51 (1.89-3.32)	+4.9
CHD	Czech	1.0	1.0		1.0	
	Polish	1.10 (0.71-1.72)	0.98 (0.62-1.55)	>-100	0.89 (0.55-1.42)	+450.0
	Russian	3.95 (2.74-5.70)	3.25 (2.17-4.88)	-23.7	3.39 (2.25-5.09)	+6.2
Stroke	Czech	1.0	1.0		1.0	
	Polish	1.26 (0.64-2.45)	1.09 (0.55-2.19)	-65.4	1.06 (0.52-2.15)	-33.3
	Russian	4.74 (2.73-8.25)	3.93 (2.14-7.20)	-21.7	3.97 (2.16-7.30)	+1.4
Cancer	Czech	1.0	1.0		1.0	
	Polish	1.08 (0.85-1.36)	1.07 (0.83-1.38)	-12.5	1.06 (0.82-1.38)	-14.3
	Russian	1.10 (0.87-1.40)	1.16 (0.89-1.51)	+60.0	1.17 (0.89-1.53)	+6.3
Non-CVD-non-cancer	Czech	1.0	1.0		1.0	
	Polish	1.26 (0.91-1.74)	1.32 (0.94-1.85)	+23.1	1.28 (0.91-1.82)	-12.5
	Russian	2.07 (1.53-2.80)	1.91 (1.37-2.68)	-15.0	1.93 (1.38-2.72)	+2.2

Model 1: adjusted for age, sex

Model 2: adjusted for all variables in model 1 and energy intake, marital status, education, household amenities score, smoking, alcohol intake, physical activity

Model 3: adjusted for all variables in model 2 ***and HDI***

[1] %=(HR2-HR1)/(HR1-1)*100; [2] %=(HR3-HR2)/(HR2-1)*100

4.5 Objective 4: Mediterranean diet score and mortality

To address the final objective, I examined the association of Mediterranean dietary pattern with total and cause-specific mortality in the HAPIEE study. A recently developed modified Mediterranean diet score which gives component scores based on absolute cut-off values was applied as indicator of the participant`s adherence to the Mediterranean diet.

4.5.1 MDS components

The proportions of participants in the three cohorts who scored the maximum points for the various MDS components are shown in table 4.20. While a high proportion of participants scored maximum points for cereal intake in all three country cohorts, less than 25% of all subject reached this “ideal intake” category regarding meat and alcohol intake and olive oil usage. Adequate intake of fruits and nuts and olive oil was especially rare amongst Russians. Although the proportion of participants with adequate vegetable, fruit and nut and meat consumption was higher in females than males, for all other MDS components, maximum score was more common in males.

<u>Table 4.20:</u> Percentage of participants with maximum MDS component scores

Components	**Percentage of participants with maximum[1] component scores**							
	Czech		**Polish**		**Russian**		**TOTAL**	
	Males (n=2648)	Females (n=3319)	Males (n=3083)	Females (n=3460)	Males (n=3056)	Females (n=3767)	Males (n=8787)	Females (n=10,546)
Vegetables (g/day)	21.3	35.9	29.5	32.8	39.0	41.4	30.4	36.9
Fruits and nuts (g/day)	37.9	59.1	33.0	45.5	8.2	15.0	25.9	38.9
Legumes (g/week)	60.1	58.4	42.9	38.0	29.6	29.6	29.7	27.1
Cereals (g/day)	67.9	59.1	80.3	75.7	87.9	73.2	79.2	69.6
Fish (g/week)	34.1	31.1	42.6	33.0	36.9	33.1	38.0	32.4
Meat and meat products (g/day)	15.1	30.1	9.5	22.6	9.7	21.0	11.3	24.4
Dairy products (g/day)	55.9	40.3	46.8	34.3	52.7	49.0	51.6	41.4
Alcohol (g/day)	16.7	5.9	10.4	1.5	19.9	2.0	15.6	3.0
Olive oil usage	5.6	6.0	41.7	40.0	0.3	0.3	16.4	15.1

[1] 1-point for olive oil usage and 2-points for all other components

4.5.2 Bivariate analysis

Table 4.21 shows the distribution of the sample characteristics across the three MDS categories. Overall, 25% of the participants had high (>10) MDS. The proportion of these high scorers was the largest in the Polish cohort and smallest amongst Russians.

Female sex, married status, high education, high household amenities score, high total energy intake and regular vitamin supplement intake were related to high MDS. The proportion of smokers was lower amongst those with high MDS, and not surprisingly, the mean HDI score increased sharply with increasing MDS. CVD risk factors were not significantly related to MDS categories after the differences in cohort, sex, age and energy intake were accounted for. However, there was a clear inverse trend of total and cause-specific mortality rates across MDS categories.

Table 4.21: Characteristics of the study sample by MDS categories

	MDS categories			p-value (trend)[1]	
	Low (0-7 points)	Moderate (8-10 points)	High (11-17 points)	crude	adj.[2]
Number of participants[3]:	4790	8941	4589		
Cohorts					
Czech, %	26.1	28.0	36.8	<0.001	<0.001
Polish, %	23.9	33.9	41.8	<0.001	<0.001
Russian, %	50.0	38.1	21.5	<0.001	<0.001
Socio-demographic characteristics					
Mean age (SD), *years*	56.9 (7.1)	57.1 (7.0)	56.7 (6.9)	0.216	0.002
Sex: Females, %	49.8	55.8	58.5	<0.001	<0.001
Marital status: Married, %	73.8	75.4	77.1	<0.001	<0.001
Education: Primary or less, %	10.3	10.1	9.1	0.045	0.002
Education: University, %	25.8	25.5	27.5	0.068	<0.001
Household amenities score: Low, %	24.3	21.7	17.1	<0.001	<0.001
Household amenities score: High, %	29.2	33.0	38.5	<0.001	<0.001
Lifestyle characteristics					
Mean energy intake (SD), *MJ/day*	9.3 (3.1)	9.7 (3.1)	10.0 (3.1)	<0.001	<0.001
Smoking: Current smokers, %	33.7	30.0	27.2	<0.001	<0.001
Physical activity: Low, %	49.8	48.8	48.4	0.188	0.201
Vitamin supplement intake: Regular, %	13.9	17.5	22.7	<0.001	<0.001
Mean healthy diet indicator score (SD)	50.9 (8.4)	54.9 (8.6)	57.4 (8.8)	<0.001	<0.001

	MDS categories			p-value (trend)[1]	
	Low (0-7 points)	Moderate (8-10 points)	High (11-17 points)	crude	adj.[2]
CVD risk factors					
Mean BMI (SD), *kg/m^2*	27.8 (4.8)	27.8 (4.7)	27.7 (4.6)	0.260	0.707
BMI >30kg/m^2, %	28.9	28.2	27.1	0.042	0.270
Mean MAP (SD), *mmHg*	105.3 (15.6)	104.8 (15.0)	104.2 (15.5)	0.001	0.081
Hypertension, %	46.6	47.1	45.1	0.148	0.492
Mean total cholesterol (SD), *mmol/l*	6.0 (1.2)	6.0 (1.2)	5.9 (1.2)	<0.001	0.811
Hypercholesterolemia, %	76.5	76.8	75.9	0.507	0.782
Mortality outcomes					
All-cause, *per 1000 person-years*	12.2	9.0	7.3	<0.001	<0.001
CVD, *per 1000 person-years*	4.3	3.3	1.9	<0.001	<0.001
CHD, *per 1000 person-years*	2.4	1.7	0.9	<0.001	<0.001
Stroke, *per 1000 person-years*	1.2	0.8	0.4	<0.001	<0.001

[1] p-values were calculated by logistic regression for categorical and linear regression for continuous variables

[2] Adjusted for cohort, sex, age and energy intake

[3] Including only participants with complete MDS data

4.5.3 Multivariable Cox regression analysis

Basic and multivariable adjusted associations of MDS with total and cause-specific mortality in the pooled sample are shown in table 4.22. In the multivariable adjusted models, 1 SD (=2.2 points) increase in the MDS was significantly associated with reduced risk of total and CVD deaths after potential confounders were taken into account. The association with CHD and stroke mortality were also inverse but statistically non-significant. The preventable proportion of deaths was the highest for stroke mortality.

Country-specific analyses revealed inverse but not statistically significant associations between MDS and most mortality outcomes in individual cohorts (table 4.23).

In addition to analyses using the modified MDS, the relationship between the most frequently used MDS based on sex-specific median cut-offs for component scores (Trichopoulou *et al.* 2005) and mortality outcomes was also examined (table 4.24). The agreement between the two Mediterranean diet scores was moderate: Spearman`s correlation coefficient was 0.69, and the linear weighted kappa between the three MDS categories in each score was 0.50. The results suggested somewhat weaker associations with mortality than the main analyses using the modified MDS.

<u>Table 4.22:</u> **Results of Cox regression analysis between MDS and mortality outcomes on the pooled sample**

Cause of death	**dead/n**	**model**	**MDS categories**						**Per 1SD[1] increase in MDS score**		
			Low (0-7p)	**Moderate (8-10p)**		**High (11-17p)**					
			HR	**HR**	**(95% CI)**	**HR**	**(95% CI)**	**PP% (95%CI)[2]**	**HR**	**(95%CI)**	**p-value**
Any-cause	1314/19,333	model1	**1.00**	**0.79**	(0.70-0.90)	**0.72**	(0.62-0.85)	**11.2** (5.5-16.4)	**0.87**	(0.82-0.92)	<0.001
		model2	**1.00**	**0.85**	(0.75-0.96)	**0.85**	(0.73-1.00)	**5.6** (0.0-10.9)	**0.93**	(0.88-0.98)	0.012
CVD	438/19,263	model1	**1.00**	**0.87**	(0.70-1.07)	**0.64**	(0.48-0.86)	**14.3** (4.8-23.9)	**0.83**	(0.75-0.92)	<0.001
		model2	**1.00**	**0.94**	(0.76-1.16)	**0.78**	(0.58-1.05)	**8.1** (-1.6-17.9)	**0.90**	(0.81-0.99)	0.036
CHD	226/19,263	model1	**1.00**	**0.88**	(0.66-1.17)	**0.64**	(0.42-0.96)	**14.3** (1.3-27.9)	**0.82**	(0.71-0.94)	0.004
		model2	**1.00**	**0.97**	(0.72-1.29)	**0.81**	(0.53-1.23)	**6.8** (-6.5-20.9)	**0.90**	(0.78-1.03)	0.132
Stroke	109/19,263	model1	**1.00**	**0.86**	(0.57-1.30)	**0.64**	(0.35-1.16)	**14.4** (-4.6-33.9)	**0.84**	(0.69-1.03)	0.093
		model2	**1.00**	**0.88**	(0.58-1.34)	**0.71**	(0.39-1.30)	**11.2** (-8.2-31.0)	**0.87**	(0.71-1.07)	0.201

Model 1: adjusted for sex, age, cohort

Model 2: adjusted for sex, age, cohort, smoking, education, household amenities score, marital status, energy intake, physical activity, vitamin supplement Intake

[1] 1SD=2.2 MDS points

[2] Preventable proportion of death if participants in the lowest two categories increased their adherence to Mediterranean diet one category upward

Table 4.23: Results of Cox regression analysis between MDS and mortality outcomes by country cohort

Cause of death	Cohort	Death/n	MDS categories						Per 1SD[1] increase in MDS score		
			Low	Moderate		High					
			HR	HR	(95% CI)	HR	(95% CI)	PP% (95%CI)[2]	HR	(95%CI)	p-value
Any-cause	Czech	364/5967	**1.00**	**0.72**	(0.56-0.93)	**0.82**	(0.62-1.09)	**7.1** (-3.0-17.4)	**0.94**	(0.85-1.05)	0.277
	Polish	388/6543	**1.00**	**0.83**	(0.64-1.06)	**0.78**	(0.59-1.05)	**8.4** (-1.6-18.4)	**0.89**	(0.81-0.99)	0.038
	Russian	562/6823	**1.00**	**0.94**	(0.79-1.12)	**0.98**	(0.75-1.29)	**0.7** (-8.5-9.8)	**0.97**	(0.88-1.05)	0.459
CVD	Czech	106/5965	**1.00**	**0.80**	(0.51-1.27)	**0.87**	(0.51-1.49)	**4.9** (-13.0-24.3)	**0.92**	(0.76-1.12)	0.424
	Polish	99/6517	**1.00**	**1.27**	(0.76-2.12)	**0.92**	(0.50-1.71)	**2.5** (-14.7-22.1)	**0.93**	(0.75-1.13)	0.458
	Russian	233/6781	**1.00**	**0.92**	(0.70-1.21)	**0.73**	(0.45-1.17)	**10.2** (-5.0-25.6)	**0.89**	(0.79-1.03)	0.117
CHD	Czech	43/5965	**1.00**	**1.28**	(0.60-2.73)	**1.10**	(0.44-2.72)	**-3.0** (-26.7-27.5)	**1.03**	(0.75-1.42)	0.841
	Polish	45/6517	**1.00**	**1.28**	(0.61-2.70)	**0.94**	(0.38-2.30)	**1.9** (-21.7-31.2)	**0.90**	(0.67-1.21)	0.481
	Russian	138/6781	**1.00**	**0.86**	(0.70-1.21)	**0.77**	(0.42-1.40)	**8.7** (-11.1-27.4)	**0.87**	(0.73-1.04)	0.120
Stroke	Czech	18/5965	**1.00**	**0.78**	(0.25-2.46)	**0.91**	(0.25-3.32)	**3.3** (-34.2-50.0)	**0.83**	(0.51-1.35)	0.450
	Polish	20/6517	**1.00**	**0.65**	(0.23-1.82)	**0.45**	(0.12-1.65)	**26.2** (-14.5-56.2)	**0.82**	(0.52-1.28)	0.378
	Russian	71/6781	**1.00**	**0.94**	(0.57-1.54)	**0.77**	(0.33-1.77)	**8.5** (-17.9-35.3)	**0.92**	(0.72-1.17)	0.490

All HRs are adjusted for age, sex, cohort, education, marital status, household amenities score, smoking, physical activity, total energy intake, vitamin supplement intake

[1] 1SD=2.3 MDS points in the Czech, 2.2 MDS points in the Polish and 2.0 points in the Russian cohort

[2] Preventable proportion of death if participants in the lowest two categories increased their adherence to Mediterranean diet one category upward

Table 4.24. **Results of Cox regression analysis between the most frequently used MDS[1] and mortality outcomes in the pooled sample**

Cause of death	dead/n	model	MDS categories: Low (0-3p) HR	Moderate (4-5p) HR	Moderate (4-5p) (95% CI)	High (6-9p) HR	High (6-9p) (95% CI)	Per 1 SD[2] increase in MDS score: HR	(95%CI)	p-value
Any-cause	1314/19,333	model1	**1.00**	**0.83**	(0.73-0.94)	**0.77**	(0.66-0.90)	**0.91**	(0.86-0.96)	0.001
		model2	**1.00**	**0.90**	(0.79-1.02)	**0.88**	(0.76-1.03)	**0.95**	(0.90-1.01)	0.108
CVD	438/19,263	model1	**1.00**	**0.79**	(0.64-0.98)	**0.69**	(0.53-0.90)	**0.86**	(0.79-0.95)	0.002
		model2	**1.00**	**0.86**	(0.69-1.07)	**0.81**	(0.62-1.06)	**0.92**	(0.83-1.01)	0.079
CHD	226/19,263	model1	**1.00**	**0.76**	(0.56-1.02)	**0.64**	(0.44-0.92)	**0.82**	(0.72-0.94)	0.004
		model2	**1.00**	**0.84**	(0.62-1.14)	**0.77**	(0.53-1.11)	**0.88**	(0.77-1.01)	0.065
Stroke	109/19,263	model1	**1.00**	**0.73**	(0.47-1.11)	**0.68**	(0.47-1.11)	**0.90**	(0.74-1.09)	0.269
		model2	**1.00**	**0.74**	(0.48-1.13)	**0.71**	(0.42-1.22)	**0.91**	(0.75-1.11)	0.369

Model 1: adjusted for sex, age, cohort

Model 2: adjusted for sex, age, cohort, smoking, education, household amenities score, marital status, energy intake, physical activity, vitamin supplement Intake

[1] Trichopoulou et al 2005

[2] 1SD=1.5 points

When the MDS components were examined separately, mortality risks decreased as component scores rose for all components except for meat and olive oil (table 4.25). However, most associations were not significant, which confirms the notion that the MDS is a better predictor of mortality than its individual components.

Table 4.25: Results of the Cox regression analysis for the association between MDS component scores and mortality outcomes

Components	Mortality outcomes							
	All-cause		CVD		CHD		Stroke	
	HR[1]	(95%CI)	**HR[1]**	(95%CI)	**HR[1]**	95%CI	**HR[1]**	95%CI
Vegetables	**0.95**	(0.87-1.04)	**0.88**	(0.76-1.03)	**0.90**	(0.73-1.11)	**0.75**	(0.55-1.03)
Fruits and nuts	**0.97**	(0.90-1.05)	**0.82**	(0.71-0.95)*	**0.87**	(0.70-1.07)	**0.68**	(0.50-0.94)*
Legumes	**0.95**	(0.89-1.01)	**0.96**	(0.86-1.08)	**1.04**	(0.89-1.22)	**0.84**	(0.66-1.05)
Cereals	**0.91**	(0.82-1.00)	**0.94**	(0.79-1.12)	**0.90**	(0.70-1.16)	**1.13**	(0.76-1.66)
Fish	**0.95**	(0.88-1.03)	**0.96**	(0.83-1.10)	**0.87**	(0.72-1.05)	**1.12**	(0.85-1.49)
Meat and meat products	**1.00**	(0.93-1.09)	**1.07**	(0.93-1.22)	**1.02**	(0.84-1.23)	**1.25**	(0.95-1.62)
Dairy products	**0.98**	(0.92-1.05)	**0.97**	(0.86-1.09)	**0.98**	(0.83-1.16)	**0.92**	(0.73-1.16)
Alcohol	**0.90**	(0.80-1.00)	**0.89**	(0.74-1.07)	**0.82**	(0.64-1.05)	**0.86**	(0.58-1.28)
Olive oil usage	**1.11**	(0.92-1.34)	**1.07**	(0.73-1.55)	**1.23**	(0.71-2.16)	**1.58**	(0.72-3.50)

All HRs are adjusted for age, sex, cohort, education, marital status, household amenities score, smoking, physical activity, total energy intake, vitamin supplement intake

1 per 1-point increase in the component score;

* $p<0.05$

- CHAPTER 5 -

DISCUSSION

This chapter discusses the interpretation and implications of the thesis` results in light of strengths and weaknesses of the available data, and in the context of the existing evidence. First, the most important findings of the thesis are summarised (section 5.1), then the strengths and limitations of the work are presented and considered in details (section 5.2). Finally, the results of the four main analyses of the thesis are interpreted and put into context in section 5.3.

5.1 Summary of main findings

The key findings of the thesis were the follows. Firstly, using dietary data collected by the same FFQ methodology across four samples, dietary intakes in the Czech, Polish and Russian cohorts of the HAPIEE study and the British Whitehall II cohort were found to be fully comparable only for a subset of foods, drinks and nutrients. The median fruit and vegetable intakes were significantly lower in the pooled Eastern European sample than in the British cohort, and there was large variation in average consumption of these foods between the Czech, Polish and Russian cohorts. Although the consumption of animal fats, including saturated fatty acids and cholesterol, was only partially comparable between cohorts, the figures suggest that intakes were significantly higher in the Eastern European cohorts compared to the British sample.

Second, some of the differences in CVD, CHD and stroke mortality rates between the Russian participants of the HAPIEE study and the British civil servants of the Whitehall II cohort were partially explained by the variation in fruit intake levels. The results indicated that approximately 10% of the excess CVD and 14% of the excess stroke mortality in the Russian sample was probably due to their inadequate fruit consumption, after several other risk factors were accounted for. Compared to the British sample, lower vegetable intake also seems to have been contributed to the higher stroke mortality rates in the Czech and Polish cohorts by approximately 8% and 7%, respectively.

Third, total, CVD, CHD and stroke mortality in the HAPIEE cohorts was inversely associated with fruit and vegetable intake, although most associations were not statistically significant. The impact of fruit and vegetable consumption was the largest for stroke mortality: the proportion of stroke deaths which could be prevented if fruit and vegetable intake was increased in the sample was approximately 16%. The inverse associations between fruit/vegetable intake and mortality outcomes were found to be stronger among smokers, reaching statistical significance for total, CVD and stroke mortality in the multivariable adjusted categorical analysis. Blood pressure lowering effect of fruit and vegetable intake appeared to be an important mediator for CVD mortality.

Fourth, in the pooled HAPIEE sample, the healthy diet indicator score, which measures the adherence to the WHO dietary recommendations published in 2003 (WHO 2003a), was found to be inversely and statistically significantly associated with mortality from CVD and CHD, but not with stroke, cancer or non-CVD-non-cancer causes of death. The association with total mortality was inverse but statistically not significant. The proportion of deaths which could be prevented if the participants` adherence to the WHO dietary guidelines increased was the highest for CVD (10%) and CHD (14%) mortality.

Finally, a recently proposed modified Mediterranean diet score (Sofi *et al.* 2014) was found to be inversely associated with deaths from all-causes, CVD, CHD and stroke in the pooled HAPIEE sample, reaching statistical significance for total and CVD mortality. The analysis also suggested that high adherence to the Mediterranean diet in this Eastern European sample was rare.

5.2 Limitations and strengths

This section describes the limitations and strengths of the work which are relevant to all performed analyses and need to be taken into account when interpreting the results of the thesis. Strengths and weaknesses which pertain to a specific analysis will be acknowledged in the following section of this chapter in which the results are interpreted separately.

5.2.1 Limitations

There are a number of limitations which need to be considered when interpreting the results of the thesis.

5.2.1.1 Generalisability of findings (selection bias)

Firstly, the selection of specific cities, restricted age range, lack of participants from rural areas, moderate response rates and the fact that most socio-demographic and lifestyle factors differed significantly between participants who were included and excluded from the analytical samples have affected the external validity of the results and their generalizability to national trends.

As the main focus of the HAPIEE study is on chronic diseases and ageing, the recruited subjects were between 45 and 69 years of age at baseline. The restricted age range means that the results can be interpreted only to adult and elderly populations. Dietary habits and mortality rates of younger individuals can be substantially different from those included in the current analyses. For example, previous study found that the probability of inadequate fruit and vegetable consumption increased with age in FSU population samples (Goryakin *et al.* 2015).

The restriction of the cohorts to selected urban centres (due to logistic reasons), and the consequent absence of rural population samples, means that the sampling frames were not representative for the respective countries as a whole. Although levels and trends in mortality in the participating study centres reflect national level data (WHO 2013), dietary habits of individuals who live in the larger towns and cities included in the HAPIEE study may be different from rural populations and do not fully represent national nutritional status. A recent study in Poland reported that hypertensive adults who live in rural areas consumed more fat and cholesterol but less carbohydrates and fibre than urban inhabitants (Suliburska *et al.* 2012). Particularly high fat intake was also reported in a rural Lithuanian sample in the CINDI survey (Petkevičiene *et al.* 2012). Insufficient fruit and vegetable consumption was also found to be more common in rural FSU population samples compared to those who live in cities (Abe *et al.* 2013; Goryakin *et al.* 2015). This suggests that the average intakes of most foods and nutrients may have been different if the HAPIEE cohorts had included rural participants. Beyond nutritional status, other lifestyle factors and socio-economic characteristics probably also differ between urban and rural inhabitants in this region (McKee *et al.* 1998).

The overall response rate of 59% in the HAPIEE study suggests that the non-response bias cannot be dismissed. However, it was found that considerable proportion of non-response occurred due to incorrect addresses, so it is likely that the actual response rates were higher than those reported here (Peasey *et al.* 2006). Furthermore, the proportion of responders was similar to other surveys conducted in Central and Eastern Europe or the Former Soviet Union (Kartashov *et al.* 1991; McKee *et al.* 1998). Previous analysis of non-responder data in the HAPIEE study showed that non-responders were more likely to be males, smokers, and had lower level of education and worse self-rated health (Peasey *et al.* 2006).

Participants of the HAPIEE study who were excluded from the analyses differed significantly from the analytical samples (table 4.3), which further reduces the generalisability of the findings. These differences between included and excluded participants were observed despite the application of multiple imputation techniques which minimised the number of subjects who had to be excluded due to missing data.

As a result of the selection bias which occurred due to the above mentioned limitations, the results of this thesis cannot be considered fully representative to the Czech, Polish and Russian populations as a whole. Furthermore, although the Czech, Polish and Russian populations are good indicators of the CEE and FSU in many aspects, the results of the analyses cannot be automatically interpolated to the whole Eastern European region. On the other hand, the lack of national and regional representativeness does not affect the internal validity of the findings, particularly of the associations between dietary variables and mortality within the cohort. Comparisons between cohorts (including comparisons with the Whitehall II study), however, may not be completely reliable.

5.2.1.2 Measurement error and bias

The second major issue, common to most nutritional epidemiological studies, relates to the measurement of diet. Although FFQ is the most commonly used method to assess habitual diet, it has well known limitations. Firstly, it tends to be semi-quantitative, rather than fully quantitative, which means that the absolute intake levels of the various foods and nutrients may be imprecise and energy intakes underestimated, however, it is likely that the ranking of subjects is adequate (Willett 2013b). Secondly, it tends to systematically overestimate the dietary intakes of foods and nutrients which are considered healthy by the individual (i.e. fruits and vegetables) and underestimate those which are considered unhealthy (i.e. meat, alcohol, energy) (Bingham 1991; Bingham *et al.* 2001; Cade *et al.* 2002; Prentice 2003; Michels 2003; Willett 2013b). Furthermore, as the questionnaire usually refers to the dietary habits over the previous months or year, the memory of the participants is required and the impact of recall bias can be considerable (Willett 2013b). The combined effect of random and systematic measurement errors usually leads to increased standard deviation of intakes, and consequent loss of statistical power and underestimation of the effects of diet on disease outcomes (Kipnis *et al.* 2002; Willett 2013a). This may be one of the explanations for the relatively weak associations of fruit and vegetable intakes and diet quality scores with mortality found in this thesis and in other published studies.

In order to assess the relative validity of the available dietary data, it is recommended to compare the FFQ measurements with data from other assessment tools such as 24-hr recalls or diet records, or with biomarker concentrations (Willett and Lenart 2013). Due to limited resources, no dietary assessment tool, apart from the FFQ, was used to collect nutritional data in the HAPIEE study. However, plasma biomarker concentrations were measured on a subsample of the study population, and these data were compared against FFQ data regarding fruit, vegetable and selected micronutrient intakes (table 4.10). The correlation coefficients were found to be somewhat lower but generally comparable to many other studies regarding most intake-concentration pairs (Al-Delaimy *et al.* 2005; Henriquez-Sanchez *et al.* 2009).

As diet was measured only at baseline, changes in dietary habits of participants could not be taken into account in the current analyses. Although major changes in diet are not expected in this age group, the occurrence of chronic diseases such as diabetes, or changes in socio-economic factors (i.e. marital or employment status) can have an impact on the individuals` diet (Bernstein *et al.* 2011; Conklin *et al.* 2014), which means that the available dietary data may not reflect the actual eating habits throughout the entire follow up period. In order to track changes of diet in participants, repeated dietary data collection is planned in future waves of the HAPIEE study.

For the calculation of nutrient intakes in the HAPIEE study the British McCance and Widdowson food composition table/database was used. This solution is not ideal as the composition of foods in the UK might be different from Eastern European countries. However, national food composition tables differ in completeness, accuracy and they often use different analytical methods to measure nutrient content of foods, which technical differences can lead to biased comparisons if nutrient intake levels are compared internationally (Ireland *et al.* 2002; Vaask *et al.* 2004). There are plans to produce standardised European food composition tables but at this point such dataset is not available (EuroFIR 2015). The application of the same food composition table in the HAPIEE and Whitehall II cohorts avoided this problem and allowed a reasonably valid comparison of nutrient intakes across cohorts.

Most socio-economic and lifestyle characteristics of participants were also assessed using self-reports. Although the validity and reliability of most questions in the study questionnaire are considered to be adequate, they are probably less accurate than objective measures. Lifestyle factors which are regarded socially desirable (i.e. physical activity) or undesirable (i.e. alcohol intake, smoking) were likely to be over- or underreported. Potential residual confounding cannot therefore be excluded.

Finally, mortality data were ascertained through linkage with mortality registers. While death registers are reliable sources of information, they are not without limitations. Potential errors can occur, for example, due to inaccurate coding of cause of death. If participants move to a different region or country, loss of follow up can be difficult to avoid. However, in the current analysis, only few individuals were lost during follow (0.7%).

5.2.1.3 Further limitations

The possibility that unmeasured socio-economic, lifestyle or dietary factors may have affected (and confounded) the examined associations between dietary intakes and mortality cannot be ruled out. For example, salt intake, which is difficult to be measured accurately with FFQ (Freedman *et al.* 2015), may be related to fruit, vegetable consumption and overall dietary patterns, as well as mortality outcomes (Aburto *et al.* 2013). However, the fact that the associations were adjusted for a large number of possible confounders, including other dietary habits in case of the association between fruit, vegetable intake and mortality, reduced the possibility of such confounding.

Although the sample size was adequate to provide sufficient statistical power for the cross-cohort comparison of dietary habits and for the analysis of the associations between dietary habits and total and CVD mortality outcomes in the pooled sample, the wide confidence intervals often limited the efforts to draw meaningful conclusions in case of CHD or stroke mortality, or when the associations were examined on cohort-specific subsamples. Associations with these less common outcomes should therefore be interpreted with caution.

5.2.2 Strengths

This PhD work also has a number of important strengths. The HAPIEE study is by far the largest study with available dietary and mortality data in any CEE and FSU population samples to date. Given the high mortality and lack of individual level evidence on dietary habits in Eastern Europe, this thesis has the potential to fill in important gaps in what is known about nutrition and health in the region.

The prospective cohort design of the HAPIEE study is one of the major advantages of this work. This setting made it possible to investigate the associations of dietary habits with mortality outcomes which are generally more reliable endpoints, in terms of the validity of data, than incidence rates, given the problems with follow up and classification of non-fatal events. It also made the temporality of the associations clear, and allowed to estimate the relative risk more accurately by taking into account the time a person spent at risk. Although measurement of diet has its limitations, and is difficult to be equally (im)precise across different settings, the multicentre design of the HAPIEE study have maximised standardisation of study protocol and study procedures across cohorts. Although it is likely that there were some differences in the execution of the study between countries, these differences were small compared to situations that different studies are harmonized and compared.

The study was sufficiently large to provide good statistical power for the cross-cohort comparison of dietary habits and to detect significant associations with most mortality outcomes in the pooled sample. In order to avoid the exclusion of participants with missing covariate data, multiple random imputation procedures were applied. As a result, sample size was larger and the impact of selection bias was smaller than it would have been with the listwise deletion approach.

Although FFQ is not a flawless instrument, the version used was very similar to those used in other major cohort studies, and, as explained above, given the central protocol across all centres for this study, the measurements were generally comparable across cohorts.

5.3 Interpretation of the results

This section presents interpretation of the results described in chapter 4. Findings are explained separately for the thesis` four main analyses. In all subsections, additional limitations which are specific for the respective analysis are described, and answers for the following key questions are sought: (1) How do the results fit into the larger context of evidence provided by previous studies? (2) What are the possible underlying reasons for the findings?

5.3.1 Objective 1: comparison of dietary intakes between the HAPIEE and Whitehall II cohorts

Ecological data and previous cross-regional studies with individual-level evidence suggested that fruit consumption is lower in CEE/FSU compared to Western Europe, however, there is probably no major difference in vegetable intake (FAO 2015). Although the findings of this thesis support these previous findings, they also suggest that important differences exist between countries within the Eastern European region regarding the consumption of fruits and vegetables. The results also support previous ecological-level data that the average consumption of animal fat foods, saturated fat and cholesterol is higher in the Czech Republic, Poland and Russia than in the UK (FAO 2015). The results further indicate that low fruit consumption partially explains the higher mortality from CVD, and particularly from stroke, of Russian urban inhabitants when compared with British civil servants.

In addition to the general limitations discussed in the previous section, there are some important issues specific to this analysis.

First, the fact that neither study populations were fully representative to their respective countries as a whole, let alone the entire Eastern and Western European regions, means that the findings can only provide a crude indication of the existing situation. The various reasons for selection bias which may have affected the dietary intake results in HAPIEE study have been described in details in section 5.2.1.1. Most of these factors, including the restricted age range (35-55 years), moderate response rate (73% in Whitehall II study) and the fact that many socio-demographic and lifestyle factors differed between participants who were included and excluded from the analytical sample (table VI-1 in appendix), are relevant in the Whitehall II study too. In addition, individuals in non-manual occupations tend to have a better quality diet than manual workers (Bolton-Smith *et al.* 1991), which suggests that participants of the Whitehall II cohort of civil servants probably have healthier diet than the general UK population.

Secondly, although the FFQ is a cost-effective instrument to provide information on habitual diet in large studies, the method has weaknesses of imprecision and information bias, as it was described in section 5.2.1.2. The relative validity of the FFQ in the Whitehall II study has been assessed previously using 7-day diet diary and plasma biomarker concentrations as reference (Brunner *et al.* 2001). The correlation coefficients between intakes and plasma concentrations of beta-carotene were somewhat higher (Spearman`s rho = 0.25 and 0.26 for males and females, respectively) than the values measured in the HAPIEE study. This suggests that the extent of measurement error regarding beta-carotene intake was likely to be higher in the HAPIEE cohorts than in Whitehall II participants. Due to the fact that the relative validity of the FFQ was not compared with other dietary assessment methods in the HAPIEE study, it is not possible to say whether there was any difference in the extent of measurement error for other food and nutrient intakes. However, as cross-cohort comparisons of dietary intakes were adjusted for energy consumption, the impact of measurement error on the comparison results was reduced (Willett 2013c). Further, cross-cohort comparability of the dietary intake data was maximised since all FFQs used the same 9-point scale answer-options for all food and drink items, and strong emphasis was put on data harmonisation in the analytical phase. Despite these efforts, many foods, drinks and nutrients were only partially comparable across cohorts. Regarding these, the interpretation of results is limited because a significant proportion of intake was unknown.

Third, when the contribution of fruit and vegetable intake to the mortality differences between cohorts was examined, only a restricted number of non-dietary risk factors were included in the analysis because of the different data collection methods in the HAPIEE and Whitehall II studies. For example, further indicators of the individual`s socio-economic position and psychosocial stress (i.e.: job insecurity, social support) would produce a more robust comparison. Similarly, due to methodological differences between studies, only leisure time physical activity was comparable across cohorts and total physical activity could not be assessed. However, as many of the participants were retired, this limitation probably did not have major impact on the results. Finally, the reported alcohol intake in the Whitehall II study and HAPIEE cohorts are likely to be different due to the different level of misreporting (see below in more details). Consequently, the contribution of alcohol intake to the mortality differences between cohorts could not be estimated adequately.

The current analysis has important strengths as well. First of all, no previous studies have compared individual-level dietary intakes between Eastern and Western populations on a large sample size which is similar to the HAPIEE and Whitehall II studies. Additionally, this is the first study that estimated the contribution of fruit and vegetable intakes to the excess total and CVD mortality rates of Eastern European population samples in relation to Western Europeans using individual-level data.

There are several specific results of this analysis which deserve further consideration. In Russia, the very low fruit and relatively high reported vegetable intake is consistent with finding from previous survey. Using data which was collected from more than 18,000 people who lived in FSU countries in 2001, including nearly 4000 Russians, Abe and colleagues found that the proportion of individuals who consumed fruit every day was lower in Russia than in any other FSU states included in the study (Abe *et al.* 2013). On the other hand, daily consumption of vegetables was found to be more common here than in most neighbouring countries. However, this study also indicated increasing trend for daily fruit intake but reduction in vegetable consumption between 2001 and 2010, which meant that the Russian figures got closer to the regional average in the more recent wave of data collection.

A number of possible explanations can be mentioned for these findings. Due to the climatic conditions, large areas of Russia, including the Novosibirsk region, are not ideal for agricultural cultivation of fruits. Probably the only exceptions are some specific types of berries, such as raspberry, strawberry, blueberry, gooseberry, redcurrant and blackcurrant, which thrive well in continental or subarctic climates. In addition, fruits which are produced in household gardens during the summer months are often made into jam or kompot in order to preserve them for year-round consumption, instead of eating them fresh. Although the import of fresh fruits to Russia from other countries has increased substantially between 1995 and 2010 (FAO 2015), their availability in this country is probably still lower than in other European states with more temperate climate. Apart from availability, societal, socio-economic and lifestyle factors also likely to play a role in the low fruit intake figures in this population. Before the era of large scale international trade, inhabitants of any geographic area consumed primarily locally produced foods. Eating fresh fruits was a rare occasion for most Russians for many centuries, consequently, these food items are usually not considered part of the traditional Russian diet (Zibart 2001). This means that if people`s diet is strongly influenced by traditional habits and cultural norms, which is suggested to be common amongst Russians (Abbott *et al.* 2006), than their fruit consumption will stay low even if fruits become widely available in the shops. Unhealthy diet also found to be related to low income and unhealthy lifestyle habits, such as smoking and heavy alcohol consumption, in Russian and other population samples (Schuit *et al.* 2002; Abe *et al.* 2013; Goryakin *et al.* 2015). Although some of these factors were taken into account as potential confounders in the current analysis, it remains a possibility that the high prevalence of socio-economic deprivation and unhealthy lifestyle in this country also contribute to the findings. Previous analysis also showed that, in contrast to British citizens, health is not amongst the main motives for food choices for Russians, but for example, sensory preferences, availability and price are more important determinants in this regard (Honkanen and Frewer 2009). On the whole, the possible reasons for the low fruit intake figures in Russian population samples would worth investigating further by qualitative and quantitative epidemiological studies in the future. If the underlying reasons were explored in details, more effective public health interventions could be designed.

Home-grown food production has a long standing tradition in Russia (Seeth *et al.* 1998). In addition, with the increasing level of economic uncertainty after 1991, the share of home production to the total food supply has grown substantially during the 1990s. It is estimated that more than 40 million Russian households owned garden plots in the mid-90s (Seeth *et al.* 1998; Southworth 2006). While many of the plots are located in rural villages, they are also often situated in the outskirts of large cities and towns (dachas), providing home-grown food products to urban populations. Although fruits are also cultivated in these gardens, their main products are potatoes and vegetables (Pallot and Nefedova 2003; Southworth 2006). In fact, according to the Russian Statistical Office, 69% of vegetables produced in the country in 2012 came from household gardens (Russian Federation`s Federal State Statistical Service 2014). This trend can explain the relatively high reported vegetable intake in the Russian cohort of the HAPIEE study.

The lower vegetable intake in the Czech and Polish cohorts compared to Russians can be explained with the smaller contribution of home-grown products to the overall diet in these countries. The higher fruit intake, on the other hand, may be due to the higher availability and relative affordability of these items, or potentially the result of the more extensive implementation of public health nutritional policies in Poland and the Czech Republic (WHO 2015a).

The observation of significantly higher intakes of animal fats (including the nutrients of saturated fat and cholesterol) in the Eastern European cohorts compared to the British cohort confirms previous data and supports the hypothesis that their consumption play an important role in the high CVD rates in these countries.

Between 1960 and 1990 livestock and meat production increased by more than 50% in the Soviet Union, in line with governmental efforts to increase the population`s meat consumption (Brainerd and Cutler 2004). The Communist leaders considered fat and protein intake necessary for the maintenance of health and aimed to establish a diet which was similar to the “Western diet of progress” (Dore *et al.* 2003). As similar economic approaches were adopted in other Eastern-bloc countries, the rise in meat intake occurred across the whole Eastern European region. Reduction in bread and potato consumption and increase in dairy product and sugar intake was also seen during this period (Jahns *et al.* 2003; Lunze *et al.* 2015). Despite the dramatic decrease in agricultural production during the economic crisis in the mid-1990s, total food consumption and energy intake hardly changed in the Russian population. This was mainly due to the increase in home-grown food production and the fact that households spent relatively larger proportion of their income on food (Jahns *et al.* 2003; Lunze *et al.* 2015). In terms of food types, people seem to have reduced their meat intake and increased the consumption of the cheaper starchy roots and vegetables during these years (Lunze *et al.* 2015). Zatonski also suggested that substitution of animal fats with vegetable oils during the 1990s was one of the main reasons for the rapid decline in ischemic heart disease mortality rates in Poland, and data from the Czech Republic showed similar pattern (Bobak *et al.* 1997; Zatonski *et al.* 1998). Although the comparability of fat intake, as well as the generalizability of our findings, is limited, the results indicate that the gap in animal fat intake between East and West still existed in the first half of the 2000s. Considering also the central role of meat and animal fat in the traditional Eastern European cuisine (Zibart 2001), this area of diet should be probably one of the most important targets of public health interventions in these countries.

There is emerging evidence that intake of foods and drinks with high added sugar content are related to increased risk of obesity, diabetes and CVD (Malik *et al.* 2013). Although sugar intake (including all mono- and disaccharides) was the highest in British subjects, this result is probably due to the large contribution of fructose consumed via fruits and vegetables in this country cohort. The intakes of sweets and confectioneries were especially high in Poles and Russians. As sweet desserts are considered to be popular ingredients of the traditional Eastern European diet (Zibart 2001), the hypothesis that added sugar consumption contributes to the high CVD rates in these countries would worth examining in further studies.

One unexpected finding of the analysis is the substantially lower reported alcohol intake of Eastern Europeans compared to the British subjects. This result is especially surprising because most previous research suggested that high alcohol consumption in Eastern European countries is one potential explanation for their poor health (Zaridze *et al.* 2009; Tomkins *et al.* 2012). However, cross-national comparison of self-reported alcohol intake has serious limitations because the extent of under-reporting may vary greatly between inhabitants of different countries depending on their cultural background. For example, people might be more willing to admit their true drinking habits in some countries compared to others, or the term “never drinking” might be interpreted differently in different cultures (Leifman 2002; Pomerleau, McKee, *et al.* 2005). In fact, previous studies showed that, compared to other countries, British tend to be more honest when reporting drinking habits (Leifman 2002). On the other hand, the extent of underreporting seems to be especially high in Russian females (Laatikainen *et al.* 2002). In addition to measurement bias, the impact of selection bias might have been also greater in the HAPIEE study compared to the Whitehall II cohort due to the lower response rates. However, no large differences were found in other lifestyle habits (smoking, physical activity, vitamin supplement intake) between the Whitehall II and HAPIEE subjects, which suggests that the unexpected findings are probably not due to selection bias.

As a conclusion of the section dealing with the descriptive dietary comparison, despite the limitations, the findings support previous ecological data suggesting that fruit intake was lower and animal fat consumption was higher in Eastern Europe compared to Western European populations. The results also indicate that there are important differences in dietary habits within CEE and FSU, and public health interventions need to be population specific.

In the second part of the comparative analysis, mortality rates in the four cohorts were assessed, and the results suggested that the inadequate fruit intake may explain approximately 10% and 14% of the excess CVD and stroke mortality of Russian subjects compared to British civil servants. Low vegetable intake seemed to play a (smaller) role in the higher stroke mortality of the Czech and Polish cohorts. No such analysis has been carried out previously in Eastern and Western European population samples but the findings are consistent with the WHO Global Burden of Disease data, which estimated that inadequate fruit and vegetable intake was responsible for larger proportion of disease burden in CEE and the FSU than in Western European populations (Lim *et al.* 2012).

The age- and sex-adjusted hazard ratios of all-cause, CVD, CHD and stroke mortality rates between the British and the Eastern European cohorts were found to be higher than the ratios between national-level age-standardized death rates (WHO Regional Office for Europe 2014). The larger East-West mortality gap in the current sample was probably due to the fact that the Whitehall II study included civil servants who have better health status and lower mortality rates than the national average, while participants in the HAPIEE study came from the general population. The contribution of traditional risk factors to the mortality differences between cohorts was found to be somewhat smaller than expected, which can be explained by the lack of adjustment for alcohol intake, as alcohol consumption measurements in the Whitehall II and HAPIEE cohorts were not comparable (as described earlier).

The results also indicate that the largest proportion of mortality difference which can be explained by low fruit and vegetable intake in the Eastern European cohorts was for stroke. Previous studies suggested that one of the most important pathway how fruit and vegetable consumption can protect health is through its blood pressure lowering effect (Appel *et al.* 1997; John *et al.* 2002). Although high blood pressure is a modifiable risk factor for all types of CVD, its strongest association is with stroke (Lewington *et al.* 2002). This biological pathway offers a plausible explanation for the finding.

High levels of alcohol consumption and high prevalence of smoking are suggested as the main lifestyle factors that contribute to the poor health status and high CVD mortality rates of Eastern European populations compared to Western Europe (Zaridze *et al.* 2009; Leon *et al.* 2009; Rechel *et al.* 2013). The current analysis supports the hypothesis that low fruit and vegetable consumption has also played a role. This finding has implications for preventive programmes focusing on CVD and other chronic diseases (as discussed in the Conclusions and implications chapter).

5.3.2 Objective 2: association between fruit, vegetable intake and mortality

The observed inverse associations of fruit and vegetable intake with mortality outcomes are consistent with most previous studies in other parts of the world (Dauchet *et al.* 2005; Dauchet *et al.* 2006; Wang *et al.* 2014). However, in our data the associations were stronger in smokers, which finding is less consistent with the existing literature (Genkinger *et al.* 2004; Dauchet *et al.* 2010). The results confirmed earlier reports that blood pressure is a potential mediator between fruit, vegetable intake and mortality (John *et al.* 2002; Bazzano *et al.* 2002), and that the consumption of these foods have a stronger association with stroke mortality rates compared to CHD (Dauchet *et al.* 2005; Dauchet *et al.* 2006).

There are several limitations specific to this part of the thesis. First, there are the issues of measurement bias and residual confounding, already described in section 5.2.1. The correlations between fruit, vegetable intake and plasma biomarker concentrations (relative validity analysis) were somewhat weaker than values reported by many other studies (Chiplonkar *et al.* 2002; Jansen *et al.* 2004; Al-Delaimy *et al.* 2005). In large scale validation studies the correlation coefficients ranged from 0.10 to 0.62 between fruit, vegetable intake and vitamin C or beta carotene plasma concentrations, with a crude mean of 0.29 (Stefler 2011), compared to the range of 0.07 to 0.29 in our pooled data. This may indicate that the extent of measurement error relating to fruit and vegetable intake in the HAPIEE sample was larger than in other studies (although not too different from the average). The validation results also suggest that some subgroups were more affected by measurement error than others. For example, vegetable intake in Czech males and Polish females seems to be less precise. Considering the fact that the FFQs were self-administered in the Czech and Polish cohort but nurse-administered in Russia, it is not surprising that measurement error was smaller in Russians. As noted earlier, measurement error tends to reduce the strength of the association between dietary habits and health outcomes (Kipnis *et al.* 2002; Willett 2013d). Therefore, the higher level of imprecision in HAPIEE study`s dietary data is probably one of the main underlying reasons for the relatively weak associations between fruit, vegetable intakes and mortality outcomes seen in these data.

The second limitation is the relatively short follow up of our cohorts, compared to other cohort studies which followed up their subjects for more than 10 or even 20 years (Bazzano *et al.* 2002; Hung *et al.* 2004; Leenders *et al.* 2013). Short follow up time can affect the results in a number of ways. It may lead to low number of observed deaths which, in turn, can result in insufficient statistical power for the analysis. Due to the relatively high death rates in these Eastern European cohorts, the 7 years of average follow up time provided adequate power for the analysis when the three samples were pooled together. However, longer follow up or larger sample size would have been required to estimate meaningful results with sufficient statistical power in cohort-specific analyses and for specific causes of death. Secondly, the health protective effects of fruit and vegetable intake may require long-term consumption. Atherosclerosis, and consequently most types of CVD, is a slowly progressing multifactorial disease which can take several decades to develop. This means that the effect of any risk or protective factor, including fruit and vegetable intake, on CVD mortality can be detected only if the time difference between the exposure and death is sufficiently long. On the other hand, long-term studies of dietary habits require repeated assessment of nutrition, to take into account changes in diet over time.

Despite the limitations, this is the first large scale study which examined the relationship between fruit and vegetable consumption and all-cause and CVD mortality in a large Eastern European population sample, and the results are consistent with the literature.

The most recent meta-analysis found that the pooled HRs (95%CIs) of all-cause and CVD mortality per one serving/day increase in fruit and vegetable intake was 0.95

(0.92-0.98) and 0.96 (0.92-0.99), respectively (Wang *et al.* 2014). This and previous studies indicated similar values for CHD and stroke, or when the associations were assessed separately for fruits and for vegetables (Dauchet *et al.* 2006; He *et al.* 2006; Wang *et al.* 2014). The results in the pooled HAPIEE sample suggest somewhat weaker link for many intake-outcome pairs which difference is most likely the result of the less precise measurement of dietary intakes.

In the HAPIEE cohorts, the inverse association between fruit and vegetable intake and mortality was significantly stronger in current smokers than non-smokers, suggesting that smokers would benefit the most if their consumption was increased. Similar effect of fruit and vegetable intake in smokers has been described in some (Hung *et al.* 2004; Dauchet *et al.* 2010; Leenders *et al.* 2013) but not all (Genkinger *et al.* 2004) previous studies. There are a number of possible explanations for this interaction. For example, as smokers are subject of increased levels of oxidative stress, the protective effect of antioxidants in fruits and vegetables might be more pronounced for them compared to non-smokers. However, the lack of association between antioxidant vitamins and health outcomes in experimental trials does not support this hypothesis (Bjelakovic *et al.* 2007). Fruits and vegetables contain large amounts of polyphenols as well, and their vasodilator, anti-inflammatory and antithrombotic effects can also counteract the harmful effects of tobacco smoke (Quiñones *et al.* 2013). On the other hand, it cannot be excluded completely that this finding was due to residual confounding, as smokers who consume lots of fruits and vegetables might be more health conscious, smoke less or quit more often than other smokers (Dauchet *et al.* 2010). However, the associations remained statistically significant when the results were further adjusted for the number of cigarettes smoked per day and the number of years has smoked. Nevertheless, because of the measurement error related to smoking (Skuladottir *et al.* 2004) and other factors not taken into account in this analysis, residual confounding may still present even after this additional adjustment.

A recently published analysis of the EPIC study, involving more than 450,000 Western European inhabitants, found that 2.68% of total and 4.24% of CVD deaths could be prevented if individuals in the lower three fruit and vegetable intake quartiles of the population shifted their consumption one quartile upwards (Leenders *et al.* 2013). Using the same formula to calculate the preventable proportions (PP%) of death, point estimates in the HAPIEE study indicate a similar proportion for total mortality (2.4%) but nearly twice higher fraction for CVD (7.7%). However, while the figures offer an attractive East-West comparison of the possible public health implications of improved fruit and vegetable consumption across the population distribution, direct comparison of the values need to be treated with caution. In addition to the problem that neither studies are representative for the Eastern and Western European regions, there are several differences between the two studies. Firstly, dietary data collection in the 23 EPIC centres were carried out using a wide range of dietary assessment tools (Riboli *et al.* 2002), and while the intake values within the EPIC study were corrected for between-centre measurement error, no such correction was possible in relation to the HAPIEE data. Secondly, although both EPIC and HAPIEE PP% figures are multivariable adjusted, the covariates included in the analyses differed between studies. Although the results seem to indicate that public health interventions which aim to improve fruit and vegetable consumption in Eastern European countries would potentially have a larger impact on population health, further research, using comparable dietary assessment methods in Eastern and Western European samples, would be needed to clarify this question.

The PP% values show the potential benefits of an overall positive shift in the fruit and vegetable intake distribution, which is a realistic model of dietary change in a population (Wahrendorf 1987). When the population attributable risk fraction (PARF%) was calculated with the traditional formula (Bhopal 2008), using 600g/day fruit and vegetable intake as the threshold between exposed and non-exposed population, the proportion of death which could be attributed to inadequate intake of fruits and vegetables was found to be 2.3%, 23.2%, 20.0% and 33.5% for all-cause, CVD, CHD and stroke mortality, respectively. These figures are very similar to previous estimations which were calculated for the EU member Eastern European states based on primarily ecological-level data (Pomerleau *et al.* 2006). The only significant discrepancy was for stroke, for which the current analysis suggested considerably higher PARF%.

Because the cut-off values between the fruit and vegetable intake quartiles were specific to the three HAPIEE cohorts, it is not possible to translate the preventable proportion figures directly into absolute numbers on a population level. Nonetheless, if the cohort-specific PP% values regarding CVD mortality were applied to the 2012 national death rates (WHO 2013), about 5672, 3848 and 85,762 CVD deaths could be prevented in the Czech Republic, Poland and Russia, respectively, if individuals in the lower three fruit and vegetable intake quartiles of the population shifted their consumption one quartile upwards. These are hypothetical figures but they do provide some indication of the potential importance of fruit and vegetable consumption for population health in CEE/FSU.

The relative impact of fruit and vegetable intake was found to be larger on stroke than on CHD mortality, reflecting the stronger association with stroke. This is consistent with previous studies. For example, meta-analyses by Dauchet showed that the risk of stroke was reduced by 11% and 5% per one portion per day increase in fruit or combined fruit and vegetable intake, respectively, however the corresponding values for CHD were 7% and 4% (Dauchet *et al.* 2005; Dauchet *et al.* 2006). Meta-analyses by He also confirmed this pattern (He *et al.* 2006; He *et al.* 2007). As blood pressure has shown to be stronger related to stroke than CHD (Lewington *et al.* 2002), the potential antihypertensive effect of fruit and vegetable intake, which is suggested by previous studies and confirmed by the current analysis, is a possible explanation for this result.

The finding that fruit and vegetable intake was related to decreased blood pressure, which, in turn, contributed to the CVD risk reduction, has been reported in a number of observational and interventional epidemiological studies (John *et al.* 2002; Bazzano *et al.* 2002; Steffen *et al.* 2005). Although there is no consensus about the mechanism by which the intake of fruits and vegetable reduces blood pressure, there are several nutrients in these foods which might be responsible for the antihypertensive effect. For example, fruits and vegetables are rich sources of potassium and magnesium. The evidence regarding these compounds` association with reduced blood pressure seems to be fairly strong (Zhao *et al.* 2011). Some authors also suggest that antioxidants might affect arterial stiffness too, thus contributing to the antihypertensive effect (Czernichow *et al.* 2004).

It is also possible that unmeasured or inadequately measured lifestyle factors may confound this relationship. Salt intake, for example, is a particularly important potential confounder for two main reasons. Firstly, the evidence for its relationship with blood pressure is strong (Aburto *et al.* 2013; He *et al.* 2013), and its association with fruit and vegetable intake is also highly likely: as fruits and vegetables contain small amount of salt, its intake is likely to be lower in high fruit and vegetable consumers. Because of this connection, the blood pressure lowering effect of fruits and vegetables in interventional trials might be also partly due to the reduced sodium intake (John *et al.* 2002). Secondly, due to the methodological difficulties to measure its intake accurately, salt consumption remains unmeasured and unadjusted for in most studies, including the current analysis. Significant association between fruit, vegetable intake and blood pressure, independently from salt intake, was found in children of pre-puberty age and adolescent females (Shi *et al.* 2014; Krupp *et al.* 2014). However, the studies in adults have not adjusted for well measured salt intake. It is important that future studies of fruit, vegetable intake and blood pressure include 24-hr urinary sodium excretion measurement, which would allow more accurate assessment of salt intake`s potential confounder role.

Investigating the relationships between fruit, vegetable consumption and socio-economic characteristics of individuals was not the main aim of this study, and more detailed analysis of this topic in the HAPIEE study has already been published (Boylan *et al.* 2009). Nevertheless, the results of this thesis are consistent with this previous analysis which showed clear socio-economic gradient for fruit intake. Other studies in Eastern Europe and elsewhere also suggest that high intakes of fruits and vegetables are more common in people with more advantageous socio-economic position and healthier lifestyle (Darmon and Drewnowski 2008; Schneider *et al.* 2009; Mayén *et al.* 2014; Goryakin *et al.* 2015). Considering the high levels of socio-economic inequality in Eastern European countries (The World Bank 2000), it is important to monitor the social gradient of fruit and vegetable intake of individuals and populations in the region.

5.3.3 Objective 3: healthy diet indicator and mortality

A priori diet quality scores are suitable tools to characterise the overall diet of individuals and populations (Slattery 2010). The HDI was designed to measure the adherence to the WHO`s dietary recommendations, and those with higher scores have been shown to have lower risk of mortality in a number of observational epidemiological studies (Huijbregts *et al.* 1997; Knoops *et al.* 2006; Jankovic *et al.* 2014). This thesis confirmed the applicability of HDI in Eastern European populations and suggested that stronger adherence to the WHO nutritional guidelines may help reducing the risk of CVD mortality in this region.

There are several limitations which are specific to this part of the thesis, including most weaknesses which are pertinent to measurements of diet and dietary patterns (described in details in the background chapter, section 1.5.3). While the relative validity of the FFQ data regarding fruit and vegetable intake has been assessed using biomarkers, no other components of the HDI have been validated against another dietary assessment method or biomarkers. This means that the extent of measurement error for most HDI components, and consequently its impact on the overall HDI score, is unknown.

Other than the fact that participants might have misreported their food intake levels, measurement error may also stem from the inaccuracy of the applied food composition tables. This issue affects the HDI score more than other dietary exposures which were investigated in this thesis (i.e. fruit, vegetable intake and MDS), because this score is based on mainly nutrient intakes. As described earlier, the impact of measurement error on the results is likely to be the reduction of risk estimates and the consequent underestimation of the association between HDI and mortality outcomes.

A further important limitation, probably common to most of nutritional epidemiology, is residual confounding. While the associations between dietary patterns and health outcomes are usually less prone to residual confounding than analyses of nutrient and food intakes, its impact on the results cannot be excluded entirely. For example, the reduction of salt intake is part of the WHO dietary guidelines but it is not included in the HDI due to the difficulties of its measurement. Similarly to fruit and vegetable consumption, further studies with appropriate assessment of (and adjustment for) salt intake would be recommended.

Finally, one may also speculate about the cultural suitability of HDI. Although it was developed to provide international guidance, it may not be fully applicable to all populations. Dietary recommendations and food based dietary guidelines are not completely similar in the three examined countries, and they also show some differences from those in Western Europe (WHO Regional Office for Europe 2003). Local guidelines take local dietary habits into account, and therefore may be more strongly associated with mortality than the global guidelines by the WHO. It is possible that adapting the score to country-specific nutritional guidelines may further improve its ability to predict mortality.

The adherence to the WHO 1990 dietary recommendations, measured with the dichotomous scoring system, and its relationship with all-cause mortality has been assessed in three previous large-scale prospective studies. A study by Huijbregts et al, carried out on Finnish, Italian and Dutch participants, found significantly reduced risk of total mortality in the highest compared to the lowest HDI tertiles (HR=0.87; 95%CI=0.77-0.98) after twenty years of follow-up (Huijbregts *et al.* 1997). The HALE (Healthy Ageing: a Longitudinal study in Europe) project, which included more than 3000 individuals from ten European countries, also showed inverse link between HDI and mortality (above vs. below median of HDI: HR=0.89; 95%CI=0.81-0.98) (Knoops *et al.* 2006). In contrast to these results, an adapted HDI score showed inverse but statistically not significant relationship with mortality in Swedish men (per one SD increase in the score: HR=0.96; 95%CI=0.77-1.19) (Sjogren *et al.* 2010). More recently, a meta-analysis of 11 European and North-American cohorts, which included nearly 400,000 participants above the age of 60 years, found that adherence to the WHO 2003 dietary guidelines, measured using an HDI with seven components and continuous scoring system, was significantly associated with reduced risk of all-cause mortality (per 10 point increase in the score: HR=0.90; 95%CI=0.87-0.93) (Jankovic *et al.* 2014). The results of the current analysis regarding total mortality are comparable with these previous studies, however direct comparison is not possible due to differences in the way the HDI score was constructed and the applied statistical methods.

In respect to cause specific mortality, some previous studies showed stronger association of HDI with CVD than with other causes of death or total mortality. For example, in the study by Huijbregts et al, the risk of CVD mortality decreased by 18% in the highest vs. lowest HDI tertiles (Huijbregts *et al.* 1997). However, more recently, a large multicentre study of elderly individuals found significant association between HDI and CVD mortality only in specific geographical regions but not in the overall study sample (Jankovic *et al.* 2015). The international literature on the association of HDI and other diet quality scores with cancer mortality is not consistent and the lack of association with this outcome has often been reported (Kant 2004). Possible reasons for such inconsistencies may be the heterogeneity of aetiology of different cancer types, the length of follow up needed for cancer to develop and low statistical power to assess site-specific cancers. In addition, this finding can also occur due to the composition of the score. For example, HDI contains only one component (fruits/vegetables) for which the relationship with cancer is considered `probable` according to WHO`s criteria, while the evidence for none of the other components is seen as `convincing` (WHO 2003a). For CVD, the strength of the evidence is considered `convincing` for fruit/vegetables and `probable` for two other components (NSP, cholesterol). This suggests that HDI is likely to be more adequate to predict CVD than cancer mortality. A score based on the World Cancer Research Fund (WCRF) recommendations would be probably more suitable to use for the investigation of the link between overall diet and cancer (WCRF 2007). Although the lack of significant relationship between HDI and deaths from non-CVD-non-cancer causes is not surprising, some previous studies indicated that death from other chronic diseases, such as gastrointestinal and respiratory conditions, might be also linked with unhealthy diet (Park *et al.* 2011; Drake *et al.* 2013).

Due to the continuous scoring method, the newly constructed HDI provided greater variation in the individual scores than the original dichotomous HDI, which may explain why it predicted mortality outcomes better. However, the low correlation between the two scores suggests that they classified participants differently. This may reflect the differences between the 1990 and 2003 WHO dietary recommendations, but it can also indicate the high sensitivity of the HDI to the applied scoring methods. The latter explanation represents a clear weakness of the score and a common limitation of the “a priori” dietary pattern approach.

The very low median HDI component score for saturated fatty-acids in all three country-cohorts is due to the higher than recommended intake of this nutrient in most participants, which finding is consistent with previous ecological data (FAO 2015). The implications of this finding has been discussed earlier when intakes were directly compared with data from the Whitehall II study (section 5.3.1).

The comparison of the overall HDI score across covariate categories suggests that healthy diet, measured here as the adherence to the WHO dietary guidelines, is more common in people who lead a generally healthier lifestyle, which result is similar to most previous research (Schröder *et al.* 2008; Moreno-Gómez *et al.* 2012). However, the association with socio-economic characteristics (education, household amenities score) does not seem to follow the direction which is showed by many earlier studies (Giskes *et al.* 2006; Giskes *et al.* 2009; Backholer *et al.* 2015). Similarly to these international studies, previous analysis of the HAPIEE data also found generally positive associations between healthy food intake habits and various indicators of socio-economic position, however discrepancies between foods and countries were also observed (Boylan *et al.* 2009). This suggests that the current results are probably due to the insufficient adjustment for potential confounders. The higher mean HDI amongst obese compared to non-obese and hypertensive compared to normotensive participants are also unexpected findings. In case of BMI and blood pressure, the cross-sectional nature of the data and the consequent possibility of reverse causation, or the reporting bias related to social desirability might contribute to these results.

The fact that this analysis included only cohorts from CEE and FSU populations should be considered when interpreting the finding that HDI did not explain any of the between-cohort mortality differences. Wider selection of populations with larger variation in diet, and other instruments to assess diet quality would help to clarify the extent of which unhealthy diet contributes to the East-West mortality divide.

One SD increase in the HDI score was approximately equal to incorporating one additional element of the WHO dietary guideline into someone`s diet. Based on the point estimates of HDI effect, adhering to one additional guideline has the potential to reduce CVD and CHD mortality in the population by 10% and 15%, respectively. The preventable proportion of death calculations indicated that similar figures could be achieved if the lowest three quartiles of the population, in terms of the adherence to the WHO recommendations, improved their diet quality one quartile upwards. These results suggest that overall diet quality is an important risk factor for CVD mortality in these Eastern European population samples and public health dietary interventions have the potential to substantially reduce CVD burden in the CEE and FSU regions.

One of the important disadvantages of the HDI is that it is primarily based on nutrients and not foods, which can make the results difficult to interpret for public health promotion purposes. Further studies focusing on individual foods, food groups or food-based diet quality scores in relation to health outcomes are necessary to identify which area of the diet needs special attention, so that more effective public health campaigns can be designed in this region.

On the whole, although HDI may not be the perfect measure of diet quality, the current results suggest that poor diet has an impact on CVD mortality in CEE and FSU countries. These findings are consistent with existing evidence that diet quality is associated with CVD, and they support the hypothesis that diet has played a role in the high mortality in Eastern Europe.

5.3.4 Objective 4: Mediterranean diet score and mortality

There is good epidemiological evidence that the Mediterranean-style diet is protective against CVD and other chronic diseases (Sofi *et al.* 2008; Estruch *et al.* 2013; Rees *et al.* 2013; Chiva-Blanch *et al.* 2014; Sofi *et al.* 2014). The results of the current analysis are similar to most previous studies which were carried out in Mediterranean and other non-Mediterranean population samples, and confirm that the inverse relationship also exist in Eastern Europeans. The literature-based MDS with absolute cut-offs for component scores, developed by Sofi and colleagues, seems to be a good indicator of healthy diet and predicts mortality outcomes well in the examined Czech, Polish and Russian population samples. The findings also indicate a relatively low adherence to the Mediterranean diet in Eastern European populations.

The most important limitations which are relevant for this analysis have already been presented in the previous sections. Selection bias affects the results in the manner as described in the general discussion section, and the impact of measurement bias and residual confounding regarding the MDS is similar to what was detailed for the HDI. One issue which needs special attention here is that MDS might have serious limitations in measuring the adherence to the exact Mediterranean-style diet in non-Mediterranean population samples. The primary reason for this problem is the difference in the composition of food groups. This means that the actual foods of which intake are measured by the various MDS components might differ between Mediterranean and Eastern European populations. For example, while apple was by far the most frequently consumed fruit in the currently analysed Eastern European population sample, consumption of summer fruits such as grapefruits, figs, pomegranates and grapes are more popular in Mediterranean countries (Hoffman and Gerber 2013). Or, as another example, spirits are popular alcoholic beverages in Eastern Europe, in contrast to Southern European countries where the primary source of alcohol is wine. In addition to the discrepancies in food group composition, food preparation techniques and meal patterns might be also different in Mediterranean and non-Mediterranean populations. For example, vegetables are often eaten raw in Greece, Italy and Spain, but usually cooked in Eastern Europe. In Mediterranean countries lunch is the main meal of the day, it is often shared with family members or colleagues, and snacking is rare (Tessier and Gerber 2005), while these characteristics are probably less typical for Eastern Europeans. Although these differences can have important health effects, they are not reflected in the MDS (Hoffman and Gerber 2013). These issues will need to be taken into account in future attempts to assess more precisely the extent to which the Mediterranean-style diet is followed by non-Mediterranean populations. However, the consistent results of the studies which used the MDS as an a priori diet quality index, including the current analysis, suggest that the dietary pattern which is characterised by this score is healthy, even if it does not follow all principles of the traditional Mediterranean-style diet in every respect.

Considering the fact that the evidence which supports the protective effect of the Mediterranean diet against chronic diseases is one of the strongest from all dietary risk factors, and that its association with mortality outcomes in large Eastern European population samples has not been investigated before, the current analysis fills an important gap in the literature. Further advantage is that this is the first study that applied the MDS with absolute component cut-off values, developed by Sofi (Sofi *et al.* 2014), for such assessment in any population. Its application makes it possible to estimate the adherence to the Mediterranean diet in individuals irrespectively of other participants` dietary habits, and allows comparison of MDS across studies.

The most recent meta-analysis of observational studies involving more than 4 million subjects from 35 prospective cohort studies found that for 2-point increase in the MDS the pooled RR (95%CI) of total and CVD mortality was 0.92 (0.91-0.93) and 0.90 (0.87-0.92), respectively (Sofi *et al.* 2014). Regarding the association with stroke in specific, pooled estimate from observational studies showed RR (95%CI) = 0.71 (0.57-0.89) for high versus low adherence to the Mediterranean diet (Psaltopoulou *et al.* 2013). In the PREDIMED (Prevencion con Dieta Mediterranea) multicentre randomized trial, healthy subjects who received dietary interventions based on the principles of the Mediterranean diet had a reduced risk for major CVD events with a HR (95%CI) of 0.71 (0.56-0.90) compared to the control group (Estruch *et al.* 2013). From the different CVD endpoints, the association with stroke was found to be the strongest (HR (95%CI) = 0.61 (0.44-0.86)). Although some of its methodological details have been criticised (Ornish 2013) and it clearly needs to be replicated in other populations, the PREDIMED trial`s significance is unquestionable. It is the first, and so far the only, primary prevention trial that has been conducted in this topic, and it suggests that the inverse association which has been showed consistently in observational studies between the Mediterranean diet and health outcomes is likely to reflect a causal effect of diet.

The direction and extent of the association between MDS and the examined mortality outcomes in the current analysis was similar to previous observational studies, which implies that the potential beneficial effect of the Mediterranean diet for Eastern Europeans is probably not different from any other populations. The non-significant results regarding CHD and stroke mortality and for the association with the traditional MDS using relative component scores can be explained with the inadequate statistical power and by measurement error.

The main associations with mortality outcomes were also largely similar to what was found for the HDI, and described in previous sections of the thesis. The only notable difference was for stroke, suggesting stronger link with MDS than HDI for this outcome, although the results were not significant in ether analyses. One possible explanation for this finding is that fruit and vegetable intake, which are strongly related to stroke mortality as discussed earlier, had a proportionally larger weight in the MDS. The differences between the components are likely to be the reason for other important discrepancies between the MDS and HDI results as well. For example, the highest mean MDS from the three country cohorts was found to be in the Polish sample while Russians scored the lowest in this respect. On the other hand, the average HDI was significantly higher for Russians compared to Czechs and Poles. The high PUFA but low fruit and legume intake in the Russian cohort, the popularity of olive oil in Poland, and the different weight of these components in the MDS and the HDI may explain this contradictory pattern. Although these two diet quality scores are focused on different aspects of the diet, their similar associations with mortality outcomes suggest that both of them are good indicators of the healthy diet in the examined population. This issue also points out the complexity and some of the limitations of the dietary pattern analysis method.

MDS which applies sex-specific medians to distinguish between high or low component scores is dependent on the characteristics of the specific study sample (Trichopoulou *et al.* 2014). Consequently, generalisation of the findings and comparison of the results across studies is not feasible. Furthermore, while this relative approach provides good statistical power for the analysis, the median intakes do not necessarily represent the cut-offs between healthy and unhealthy consumption levels (Waijers *et al.* 2007). Although there have been earlier attempts to compile MDS with absolute cut-offs (Martínez-González *et al.* 2002; Schröder *et al.* 2011), no previous scoring systems have been constructed on such a sound evidence base as the one proposed by Sofi (Sofi *et al.* 2014). Even though the method applied by the authors have limitations, as a result of the underlying systematic literature review, the cut-off values can be seen as summary estimates derived from all previous MDS studies. The overall MDS appears to be a suitable tool to assess the participants` adherence to the Mediterranean diet.

Correlations between different versions of MDS have been reported to be weak to moderate (Mila-Villarroel *et al.* 2011). In this light, the moderate agreement between the "traditional" MDS and the literature-based adherence score by Sofi et al is satisfactory. Especially if the different cut-off values of component scores and the arbitrary thresholds of the low, moderate and high scoring categories are considered. The differences might be also partly due to the component which differed between the two scores (olive oil usage vs. unsaturated/saturated fatty acid ratio). Most effect estimates were stronger with the version using absolute cut-offs. The larger variation between individual MDSs, which is the result of the 3-tier scoring system, is probably one of the primary reasons for this difference.

The fact that only one quarter of the pooled study sample scored more than 10-points (about 60% of the maximum score) suggest that the adherence to the Mediterranean diet in the currently analysed Eastern European cohorts was low, and among the three cohorts, dietary habits of Russians were the furthest from this pattern. In light of previous findings of this thesis, this is not an unexpected result. However, estimation of the MDS in population samples from other countries or regions, and comparison with our results would be necessary to test the hypothesis that the adherence to Mediterranean diet in Eastern Europe is indeed lower than other populations. If the main aim is unbiased cross-study comparison, than the same methods need to be used for data collection and analysis in all respective samples, which is not always feasible. Furthermore, as the FFQ is not an ideal tool to estimate absolute intake levels, measurements with more precise methods, such as repeated 24-hour diet recall or 7-days diet diary, would be required. Nevertheless, the MDS with absolute component cut-offs offers an attractive tool to compare the adherence to the Mediterranean diet across studies and populations.

The results are also in agreement with previous studies which showed that the MDS is strongly correlated with socio-economic and other lifestyle characteristics (Panagiotakos *et al.* 2008; Katsarou *et al.* 2010). In contrast, the links with CVD risk factors were inverse but statistically not significant, which contradicts previous evidence (Kastorini *et al.* 2010; Rees *et al.* 2013). Similar to the fruit, vegetable and HDI analysis, these associations were not adjusted fully which means that the results can only be treated as preliminary findings, and further analysis in Eastern European population samples would be recommended. Nevertheless, these interrelationships suggest that healthy diet, lifestyle and high socio-economic position often cluster together, and they emphasise the importance of multivariable adjustment when the associations of these factors with health outcomes are examined.

The fact that most components of the MDS were inversely but not significantly related to mortality outcomes suggests that the overall MDS is a better predictor of mortality than the components individually, and supports the application of diet pattern scores as opposed to single foods. The positive direction of the associations regarding meat and olive oil usage components is unexpected. In case of olive oil, the binary nature of this component and the small number of participants who reported to use it for cooking can contribute to the findings. The explanation for the meat component is probably more complex. However, a recently conducted detailed analysis in the HAPIEE sample found no significant association between meat intake and mortality (KilBridge *et al.*, unpublished manuscript, n.d.). This suggests that the positive trend regarding the meat component here is likely to be a random finding.

One great advantage of the MDS over the HDI is its food-based nature which makes the results much easier to translate into public health recommendations. In addition, the absolute cut-off approach further increases this scores` public health applicability and makes this newly constructed version appealing for policy makers, even more than the previously used MDSs. Given the strong evidence which supports the Mediterranean diet`s health protective effect and the apparently low adherence to this eating pattern in the examined population samples, dietary interventions which are designed based on the currently applied MDS would be especially advantageous in the Czech Republic, Poland, Russia, and probably other Eastern European countries as well. The score provides clear targets for the ideal intakes of food groups which are typical (fruits, vegetables, legumes, fish, cereal) or not typical (meat, dairy) to the Mediterranean-style diet, as well as for alcohol intake and olive oil usage. The results of this analysis regarding the preventable proportion of deaths (PP%) suggest that large number of deaths due to CVD, and particularly due to stroke, could be avoided if the adherence to the Mediterranean diet increased in the population.

On the whole, the current analysis further confirmed that unhealthy diet, as approximated by the MDS, is an important risk factor for total and CVD mortality in three large Eastern European population samples. The results also support the hypothesis that unhealthy diet has played a role in the high Eastern European CVD rates, and that dietary interventions have considerable potential to improve the health of populations in this region.

- CHAPTER 6 -

CONCLUSIONS AND IMPLICATIONS

The previous chapter discussed the results by separately focusing on the four main objectives of the thesis and the findings of the respective analyses. This section summarizes the overarching conclusions of the work. It also provides some recommendations for future research and considers the implications for public health policy.

6.1 Overall conclusions

By directly comparing dietary habits between Eastern and Western European population samples and examining the relationships of fruit and vegetable intake and overall dietary patterns, such as the HDI and the MDS, with total and cause-specific mortality in Eastern European individuals, this work explored the role of diet in the health of Eastern European populations.

Due to the limitations, such as measurement error, selection bias, residual confounding and geographical restriction of the data to only two CEE and one FSU countries, this PhD work, per se, cannot give definite answers to the broader scientific questions. However, it adds important individual-level information and knowledge to what is already known. To date, no previous studies have investigated the relationships between dietary habits and mortality in Eastern European individuals on such a large sample size as the current analysis, and very few studies compared dietary intakes between Eastern and Western Europeans on a similar scale.

Although there seem to be large differences in dietary habits between populations of the various Eastern European countries, the findings of this work support previous evidence which suggest that unhealthy diet in this region is likely to be common. The results also indicate that the inverse associations between fruit, vegetable intakes, overall dietary patterns and all-cause and CVD mortality outcomes are similar in Eastern Europeans than in other populations. These two core findings lead directly to the conclusion that disease burden due to unhealthy diet in this region is substantial. This is, in principle, individual-level confirmation of the WHO Global Burden of Disease Project estimations (WHO 2009; Lim *et al.* 2012). In other words, this thesis supports the hypothesis that unhealthy diet has contributed to the high CVD mortality rates and poor health of Eastern European populations, as well as to the large health gap between Eastern and Western Europe.

Growing body of evidence, including the current thesis, implies that dietary interventions, if successful, have the potential to significantly reduce CVD burden in Eastern European countries and to decrease health inequalities across Europe. These public health nutritional interventions may put special emphasis on increasing fruit intake in Russia, but they should include other components of the healthy diet in all countries. The WHO dietary recommendations for the prevention of chronic diseases or the Mediterranean-style diet pattern could be used as guidelines in the development of such dietary interventions, or when existing nutritional policies are redesigned. The results further indicate that large proportion of deaths could be prevented if fruit and vegetable intake increased or overall diet quality improved across the population distribution. However, it is also likely that there are specific population subgroups, for example smokers or individuals with high blood pressure, who would benefit the most from better diet, especially from increased fruit and vegetable consumption.

Another value of the results is that they also contribute to the general discussion on the relationship between diet and health. By confirming the inverse associations of fruit, vegetable consumption, HDI and MDS with mortality outcomes in populations which had not been involved in such analyses before and which have different covariate structure compared to the more frequently studied Western European and North American populations, the evidence which supports the health protective effects of these dietary factors became stronger. More generally, the results confirmed the value of the "a priori" diet pattern approach and the applicability of two specific diet quality scores in nutritional research.

On the whole, this thesis addressed some of the existing gaps in our knowledge on diet and health in Eastern Europe, thus providing evidence-based foundation for potential dietary interventions in the region. It also improved our insight into general nutritional epidemiological issues which helped to strengthen the evidence for specific diet-disease relationships.

6.2 Implications for further research and policy

Based on the thesis` findings, it is possible to formulate several recommendations for future research and public health policy. In this section I first summarise the suggestions regarding research. Subsequently, recommendations which are relevant for Eastern European public health policy makers will be detailed.

6.2.1 Recommendations for future research

My suggestions are divided in two parts. In the first part I list recommendations which are specific to the HAPIEE study, while more general scientific proposals are described in the second part.

6.2.1.1 Recommendations for future research in the HAPIEE study

1. Further validation of the HAPIEE FFQ data. One of the most important weaknesses of the dietary intake data in the HAPIEE study is that, other than the reported comparison with biomarker concentrations, the FFQ data was not validated against other dietary assessment methods at baseline. Data collection with multiple 24-hr diet recall or 3-day diet record, parallel with a repeated wave of FFQ, in a subsample of the study population would be recommended. The adequate sample size of the subsample in such a validation study should be between 100 and 200 and it should be as representative to the whole sample as possible (Willett and Lenart 2013). Alternatively, application of the HAPIEE FFQ in other Czech, Polish or Russian population samples together with a more accurate dietary assessment method could also provide important information on the reliability of the questionnaire. As a result of this validation study, the extent of measurement error in the overall HAPIEE sample could be estimated, and the HRs between dietary intakes and disease outcomes could be corrected using the regression calibration approach or other statistical techniques (Willett 2013a).

2. Repeated dietary data collection. In order to investigate the change in food and nutrient intakes or the shift in overall diet quality of participants, a second wave of dietary data collection, using the baseline FFQs in all three country cohorts of the

HAPIEE study would be recommended. Ecological data and some limited individual level evidence suggest that fruit intake has increased and animal fat consumption has reduced in Russia, Poland and other Eastern European countries since the early-2000s (Abe *et al.* 2013; Lunze *et al.* 2015; FAO 2015). Although distinguishing between time, age and cohort effect may prove to be challenging, longitudinal analysis of the dietary data in the HAPIEE study could add important new information to the existing evidence.

3. Explore the relationships between other dietary factors and disease/mortality outcomes. Fruit and vegetable intake is the most often hypothesised dietary factor in relation to poor health in Eastern Europe (Ginter 1998; Zatonski 2011). However, there are further foods and nutrients of which relationship with health outcomes in this study would be of special interest. For example, considering previous hypotheses or some specific results of the current study, animal fat or meat intake, or the consumption of sugars (mono- or disaccharides) would warrant investigation in similar depth as fruits and vegetables in this thesis. Subject to satisfactory validation of non-fatal outcomes, using incident CVD along with mortality would provide improved statistical power.

4. "A posteriori" diet pattern analysis of the HAPIEE data. Data driven (or "a posteriori") dietary pattern analysis is a suitable method to identify the inherent nutritional characteristics and dietary patterns of a population (Newby and Tucker 2004; Kant 2004; Tucker 2010). While this method has been adopted in a growing number of large-scale studies (Kant 2010), it has never been used in dietary data collected from Eastern European populations. Application of this approach in the HAPIEE study could be recommended for a number of reasons. Most studies which applied this method identified two distinctive dietary patterns: healthy ("Prudent") and unhealthy ("Western") (Kant 2004). If these patterns could be detected in the HAPIEE dataset as well, that would mean that the fundamental eating habits of Eastern Europeans are probably not too different from other populations. In some previous studies additional eating patterns, often labelled as "traditional", were also recognised, characterised by food items specific to the given region or ethnic group (Tucker 2010). This offers the possibility to identify a traditional Eastern European dietary pattern or separate country specific patterns typical for the Czech, Polish and Russian cohorts. Examination of these patterns` relationships with mortality outcomes, other lifestyle or socio-economic factors and "a priori" diet quality scores would be also possible. The most often applied statistical techniques to carry out such analyses are principal component analysis, cluster analysis or, more recently, reduced rank regression (Tucker 2010).

6.2.1.2 Recommendations for future research in Eastern European populations and elsewhere

1. To investigate the possible reasons for the unhealthy diet in Eastern European populations. In previous sections of the thesis a number of potential explanations were suggested for the observed low fruit and high meat and animal fat consumption in the examined populations. However, due to the lack of research, most of these hypotheses are not supported by solid evidence. For example, it is highly probable that local traditions and other societal factors play an important role in the food choices of individuals (Shepherd 2005; Abbott *et al.* 2006; Honkanen and Frewer 2009). But it is unknown whether people today (as did government leaders in the 1960s) believe that diet high in protein and fat is necessary to maintain health, or perhaps they are aware of the current principles of healthy nutrition but decide to ignore them and they just follow the “traditional” Russian-, Polish-, Czech-, etc. style diet because it is the local habit. Regarding the inadequate fruit intake, it would be important to clarify whether the main issue is the lack of knowledge, the lack of availability or other reasons. Furthermore, the contribution of socio-economic factors, for example whether people can afford fresh fruits, is also an important domain for more detailed examination. The few studies which have been conducted in this area suggest that individuals in Eastern Europe often believe that their health depends predominantly on health-care rather than on their own lifestyle, and that this attitude is one of the reasons for the high prevalence of unhealthy behaviour, including poor diet, in these countries (Palosuo 2000; Abbott *et al.* 2006; McKee 2007). Limited knowledge on healthy diet, limited availability of healthy food products and material obstacles of healthy lifestyle choices in Russian and Ukrainian individuals have also been reported (Palosuo 2000; Abbott *et al.* 2006; Honkanen and Frewer 2009). Nevertheless, further investigation of this topic is clearly needed.

It has been shown that psychosocial factors, such as job stress, social support or depression, can have a significant impact on diet (Lallukka *et al.* 2004; Kawakami *et al.* 2006; Nicklett *et al*. 2012). Unfavourable psychosocial factors appear common in Eastern Europeans (Bobak, Pikhart *et al.* 1998; Kopp *et al.* 2006; Lundberg *et al.* 2007), which might also contribute to the poor diet in this region. In order to investigate this hypothesis, further analyses should examine the association between psychosocial factors and diet quality in Eastern European populations.

In addition to large scale quantitative studies with structured and validated questionnaires, qualitative studies would be useful to explore these questions. The results of these studies would have significant implications in terms of the design of public health nutritional campaigns. If the main problem is the lack of knowledge, than the emphasis should be put on education. If, on the other hand, the problem is the lack of availability or affordability, than more upstream components of the food-supply need to be targeted. However, as the explanations are likely to be complex with multiple contributing factors, effective dietary interventions will probably need to be as comprehensive as possible.

2. Examine the relationship between fruit, vegetable consumption and health outcomes with particular attention to the potential cofounding effect of salt intake. Strong evidence from observational and interventional epidemiological studies supports the positive association of salt intake with blood pressure and CVD (Aburto *et al.* 2013; He *et al.* 2013; Aaron and Sanders 2013), while fruit and vegetable intake is likely to be inversely related with the consumption of salt. This means that the confounding role of salt for the association between fruit, vegetable intake and CVD is possible. However, due to the technical difficulties to measure salt intake accurately with traditional dietary assessment methods (Freedman *et al.* 2015), empirical test of this question is not straightforward. Although some previous observational studies adjusted for sodium intake measured by FFQ (Dauchet *et al.* 2007; Nagura *et al.* 2009), the measurement error of FFQs regarding salt consumption is large (Freedman *et al.* 2015) and the adjustment is likely to be incomplete. (No such adjustment was done in the current PhD work.) Large scale nutritional epidemiological studies with 24-hr urinary sodium measurements on adult population samples are clearly needed to clarify this question and separate the beneficial health effects of fruits and vegetables from salt.

3. Application of the Mediterranean diet score with absolute cut-offs for component scores in other population samples. MDS with absolute cut-offs for component scores, developed by Sofi and colleagues after systematically reviewing all previous MDS studies (Sofi *et al.* 2014), appears to be a useful “a priori” diet quality score that overcomes many of the limitations affecting the traditional MDS with distributional cut-offs. The respective analysis in this thesis confirmed the validity of the score in Eastern European population samples but replication of the study and further confirmation of this score`s applicability is needed.

6.2.2 Implications for public health and policy

Because diet is a modifiable risk factor for CVD and other chronic diseases, nutritional epidemiological research has important implications for public health policy. The findings of this thesis confirmed previous reports suggesting high prevalence of unhealthy diet in CEE and FSU countries, and the results indicate an important effect on the consequent disease burden. This suggests that effective nutritional interventions could have a large impact on the health status of these populations.

Systematic reviews of interventional studies have confirmed that dietary advice can be effective to increase fruit and vegetable intake and reduce CVD risk at population level (Pomerleau, Lock, *et al.* 2005; Rees *et al.* 2013). There are several examples of comprehensive community-based dietary interventions, applied on their own or in combination with measures targeting other lifestyle habits, which substantially improved the health of the general population. For example, in the early 1970s, population-wide campaigns were introduced in North Karelia, Finland, in order to improve diet quality and reduce smoking prevalence through a variety of policy change and educational programs (Puska *et al.* 1983). As a result of this project, CHD mortality rates decreased by 73% in 20 years, and North Karelia became one of the healthiest regions of Europe (Puska *et al.* 1998; Papadakis and Moroz 2008). More recently, the Beijing Fangshan community-based intervention project managed to achieve significant reduction in stroke morbidity and mortality through activities which aimed to produce population-level change in dietary habits and blood pressure (Chen *et al.* 2008). As an example for targeted fruit and vegetable interventions, the “5 a day” campaign can be mentioned. It was first introduced in the US but similar programmes were also adopted in several European states during the 1990s (Havas *et al.* 1995; WHO 2003b; WHO and FAO 2005; Ungar *et al.* 2013). As a result, awareness and consumption of fruits and vegetables increased in these populations; however, the impact was significant only in some subgroups (Stables *et al.* 2002; WHO 2003b). Although not all community-based intervention programs were successful, the overall evidence supporting the effectiveness of this approach is fairly strong (Papadakis and Moroz 2008).

There are a number of national, international and non-governmental organizations which provide guidelines and frameworks that can help to design effective public health nutritional interventions and strategies (WHO Regional Office for Europe 2004; WCRF 2015; McColl and Lobstein 2015). These guidelines emphasise the importance of both educating the individual and modifying the environment. In fact, changing the food environment in a way to help the individual make healthier choices is often more effective than education (Willett 2013e). There are several specific measures which target upstream factors of the food supply chain and can have a large impact on the population`s eating habits. These may include appropriate food labelling, improving/limiting availability of healthy or unhealthy food items, improving/limiting affordability of these items via taxes and subsidies, regulating advertisements, modifying nutrient (i.e.: salt) content of foods on production level or modifying the menus in schools or workplaces. Effective dietary programmes also require collaboration between governmental and non-governmental organisation, as well as with the private sector.

According to the WHO`s Global Database on the Implementation of Nutrition Action (GINA), most Eastern European countries, including the Czech Republic, Poland and Russia, have developed public health nutritional policies (including strategies, action plans and legislation) to tackle diet related non-communicable diseases and obesity. However, the number of specific programmes which implement these national-level policies remains low, especially in the FSU (WHO 2015a).

While this thesis confirms previous evidence suggesting that effective implementation of nutritional policies would have important beneficial effects across the Eastern European region, it also shows that the programmes and actions should be tailored to the dietary characteristics of the individual countries. For example, it is clear that dietary interventions need to put specific emphasis on fruit intake in Russia, while this dietary factor seems less of a problem in the Czech Republic. On the other hand, vegetable consumption needs to be emphasised stronger in Poland and the Czech Republic, but less so in Russia where home grown vegetables probably fulfil the population needs. This also means that it would be recommended to target fruit and vegetable intakes separately in these campaigns, which is consistent with suggestions by previous authors (Naska *et al.* 2000; WHO and FAO 2005). Increased fruit and vegetable intake should be part of the overall promotion of healthy diet, and if it is feasible, they should be also combined with programmes that aim to improve other health behaviours or risk factors in the population, such as alcohol consumption, smoking or hypertension.

One potential goal of dietary interventions in Eastern European countries could be to increase the proportion of individuals in the populations whose diet follows closely the WHO dietary guidelines or the Mediterranean eating pattern. The specific components of these dietary patterns, and their recommended (or “ideal”) intake values, could be built into policies and programmes and could be set as targets for individuals. The Mediterranean diet, for example, is especially suitable to use in educational campaigns. It is simple, due to its food-based nature it is easy to translate into everyday life, and, by applying the absolute scoring system, any person`s adherence to the Mediterranean diet can be calculated relatively easily. It is also easy to extend further, so that findings from other nutritional research, such as the recommendation to consume whole grains rather than refined grains or reduce processed meat, red meat and salt intake, can be incorporated in it directly (Ye *et al.* 2012; Aaron and Sanders 2013; Larsson and Orsini 2014; Abete *et al.* 2014). As the WHO dietary guidelines are based primarily on nutrients, they are more suited to be used in interventions which target upstream components of the food supply chain. For example, they can help to plan healthier meals in schools or workplaces, or they may help to design better food labels.

The transition of knowledge from research to policy is a complex process which is often influenced by economic or political interests. In addition, dietary change in individuals and populations is usually a slow process which can be also driven by factors other than public health policy. Nevertheless, providing reliable scientific evidence is a crucial first step on this path; if the recommendations are applied in practice, this work has the potential to improve diet and health in Eastern European populations.

BIBLIOGRAPHY

Aaron KJ and Sanders PW (2013). Role of Dietary Salt and Potassium Intake in Cardiovascular Health and Disease: A Review of the Evidence. *Mayo Clinic Proceedings* **88**:987–995.

Abbott PA, Turmov S and Wallace C (2006). Health world views of post-soviet citizens. *Social Science and Medicine* **62**:228–238.

Abe SK, Stickley A, Roberts B, Richardson E, Abbott P, Rotman D and McKee M (2013). Changing patterns of fruit and vegetable intake in countries of the former Soviet Union. *Public Health Nutrition* **16**:1924–1932.

Abete I, Romaguera D, Vieira AR, Lopez de Munain A and Norat T (2014). Association between total, processed, red and white meat consumption and all-cause, CVD and IHD mortality: a meta-analysis of cohort studies. *British Journal of Nutrition* **112**:762–775.

Aburto NJ, Ziolkovska A, Hooper L, Elliott P, Cappuccio FP and Meerpohl JJ (2013). Effect of lower sodium intake on health: systematic review and meta-analyses. *The British Medical Journal* **346**:f1326.

Agudo A, Slimani N, Ocke MC, Naska A, Miller AB, Kroke A, Bamia C, *et al.* (2002). Consumption of vegetables, fruit and other plant foods in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts from 10 European countries. *Public Health Nutrition* **5**:1179–1196.

Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett Jr DR, Tudor-Locke C, Greer JL, *et al.* (2011). 2011 Compendium of Physical Activities: a second update of codes and MET values. *Medicine and Science in Sports and Exercise* **43**:1575–1581.

Al-Delaimy WK, van Kappel AL, Ferrari P, Slimani N, Steghens JP, Bingham S, Johansson I, *et al.* (2004). Plasma levels of six carotenoids in nine European countries: report from the European Prospective Investigation into Cancer and Nutrition (EPIC). *Public Health Nutrition* **7**:713–722.

Al-Delaimy WK, Slimani N, Ferrari P, Key T, Spencer E, Johansson I, Johansson G, *et al.* (2005). Plasma carotenoids as biomarkers of intake of fruits and vegetables: ecological-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). *European Journal of Clinical Nutrition* **59**:1397–1408.

Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, Sacks FM, Bray GA, *et al.* (1997). A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. *The New England Journal of Medicine* **336**:1117–1124.

Australian Government (2013). *Eat for Health. Australian Dietary Guidelines.* Available at: https://www.eatforhealth.gov.au/sites/default/files/files/the_guidelines/n55_australian_dietary_guidelines.pdf

Bach A, Serra-Majem L, Carrasco JL, Roman B, Ngo J, Bertomeu I and Obrador B (2006). The use of indexes evaluating the adherence to the Mediterranean diet in epidemiological studies: a review. *Public Health Nutrition* **9**:132–146.

Backholer K, Spencer E, Gearon E, Magliano DJ, McNaughton SA, Shaw JE and Peeters A (2015). The association between socio-economic position and diet quality in Australian adults. *Public Health Nutrition* 20th May:1–9. [Epub ahead of print]

Bazzano LA, He J, Ogden LG, Loria CM, Vupputuri S, Myers L and Whelton PK (2002). Fruit and vegetable intake and risk of cardiovascular disease in US adults: the first National Health and Nutrition Examination Survey Epidemiologic Follow-up Study. *The American Journal of Clinical Nutrition* **76**:93–99.

Bendinelli B, Masala G, Saieva C, Salvini S, Calonico C, Sacerdote C, Agnoli C, *et al.* (2011). Fruit, vegetables, and olive oil and risk of coronary heart disease in Italian women: the EPICOR Study. *The American Journal of Clinical Nutrition* **93**:275–283.

Bernstein AM, Rosner B and Willett WC (2011). Cereal fiber and coronary heart disease: A comparison of modeling approaches for repeated dietary measurements, intermediate outcomes, and long follow-up. *European Journal of Epidemiology* **26**:877–886.

Bhaskaran K. and Smeeth L (2014). What is the difference between missing completely at random and missing at random? *International Journal of Epidemiology* **43**:1336–1339.

Bhopal R (2008). Presentation and interpretation of epidemiological data on risk. In: Bhopal R (ed.) *Concepts of Epidemiology*. 2nd edn. Oxford: Oxford University Press, pp. 235–284.

Bhupathiraju SN and Tucker KL (2011). Coronary heart disease prevention: nutrients, foods, and dietary patterns. *Clinica Chimica Acta; International Journal of Clinical Chemistry* **412**:1493–1514.

Bingham SA (1991). Limitations of the various methods for collecting dietary intake data. *Annals of Nutrition & Metabolism* **35**:117–127.

Bingham SA, Welch AA, McTaggart A, Mulligan AA, Runswick SA, Luben R, Oakes S, *et al.* (2001). Nutritional methods in the European Prospective Investigation of Cancer in Norfolk. *Public Health Nutrition* **4**:847–58.

Bjelakovic G, Nikolova D, Gluud LL, Simonetti RG and Gluud C (2007). Mortality in randomized trials of antioxidant supplements for primary and secondary prevention: systematic review and meta-analysis. *The Journal of the American Medical Association* **297**:842–857.

Bland JM and Altman DG (1995). Multiple significance tests: the Bonferroni method. *The British Medical Journal* **310**:170.

Bobak M, Brunner E, Miller NJ, Skodova Z and Marmot M (1998). Could antioxidants play a role in high rates of coronary heart disease in the Czech Republic? *European Journal of Clinical Nutrition* **52**:632–636.

Bobak M, Hense HW, Kark J, Kuch B, Vojtisek P, Sinnreich R, Gostomzyk J, *et al.* (1999). An ecological study of determinants of coronary heart disease rates: A comparison of Czech, Bavarian and Israeli men. *International Journal of Epidemiology* **28**:437–444.

Bobak M and Marmot M (1996). East-West mortality divide and its potential explanations: proposed research agenda. *The British Medical Journal* **312**:421–425.

Bobak M and Marmot M (2009). Societal transition and health. *Lancet* **373**:360–362.

Bobak M, Murphy M, Rose R and Marmot M (2007). Societal characteristics and health in the former communist countries of Central and Eastern Europe and the former Soviet Union: a multilevel analysis. *Journal of Epidemiology and Community Health* **61**:990–996.

Bobak M, Pikhart H, Hertzman C, Rose R and Marmot M (1998). Socioeconomic factors, perceived control and self-reported health in Russia. A cross-sectional survey. *Social Science & Medicine* **47**:269–279.

Bobak M, Skodova Z, Pisa Z, Poledne R and Marmot M (1997). Political changes and trends in cardiovascular risk factors in the Czech Republic, 1985-92. *Journal of Epidemiology and Community Health* **51**:272–277.

Boeing H, Bechthold A, Bub A, Ellinger S, Haller D, Kroke A, Leschik-Bonnet E, *et al.* (2012). Critical review: vegetables and fruit in the prevention of chronic diseases. *European Journal of Nutrition* **51**:637–663.

de Boer EJ, Slimani N, van ’t Veer P, Boeing H, Feinberg M, Leclercq C, Trolle E, *et al.* (2011). Rationale and methods of the European Food Consumption Validation (EFCOVAL) Project. *European Journal of Clinical Nutrition* **65 Suppl 1**:S1–4.

Bolton-Smith C, Smith WC, Woodward M and Tunstall-Pedoe H (1991). Nutrient intakes of different social-class groups: results from the Scottish Heart Health Study (SHHS). *The British Journal of Nutrition* **65**:321–335.

Boylan S, Welch A, Pikhart H, Malyutina S, Pajak A, Kubinova R, Bragina O, *et al.* (2009). Dietary habits in three Central and Eastern European countries: the HAPIEE study. *BMC Public Health* **9**:439.

Brainerd E and Cutler DM (2004). Autopsy on an Empire: Understanding Mortality in Russia and the Former Soviet Union. *National Bureau of Economic Research Working paper series* #10868. Available at: http://www.nber.org/papers/w10868

Britton A and McKee M (2000). The relation between alcohol and cardiovascular disease in Eastern Europe: explaining the paradox. *Journal of Epidemiology and Community Health* **54**:328–332.

Brunner E and Marmot M (2006). Social organization, stress and health. In: Marmot M and Wilkins R (eds.) *Social Determinants of Health*. 2nd edn. Oxford: Oxford University Press, pp. 6–31.

Brunner E, Stallone D, Juneja M, Bingham S and Marmot M (2001). Dietary assessment in Whitehall II: comparison of 7 d diet diary and food-frequency questionnaire and validity against biomarkers. *The British Journal of Nutrition* **86**:405–414.

Burisch J, Pedersen N, Cukovic-Cavka S, Turk N, Kaimakliotis I, Duricova D, Bortlik M, *et al.* (2014). Environmental factors in a population-based inception cohort of inflammatory bowel disease patients in Europe - An ECCO-EpiCom study. *Journal of Crohn's and Colitis* **8**:607–616.

van Buuren S (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research* **16**:219–242.

Cade J, Thompson R, Burley V and Warm D (2002). Development, validation and utilisation of food-frequency questionnaires - a review. *Public Health Nutrition* **5**:567–587.

Capita R and Alonso-Calleja C (2005). Differences in reported winter and summer dietary intakes in young adults in Spain. *International Journal of Food Sciences and Nutrition* **56**:431–443.

Carter P, Gray LJ, Troughton J, Khunti K and Davies MJ (2010). Fruit and vegetable intake and incidence of type 2 diabetes mellitus: systematic review and meta-analysis. *The British Medical Journal* **341**:c4229.

Central Statistical Office of Poland (2015). *Basic Data.* Available at: http://stat.gov.pl/en/basic-data/

Charrondiere UR, Vignat J, Moller A, Ireland J, Becker W, Church S, Farran A, *et al.* (2002). The European Nutrient Database (ENDB) for Nutritional Epidemiology. *Journal of Food Composition and Analysis* **15**:435–451.

Chen J, Wu X and Gu D (2008). Hypertension and cardiovascular diseases intervention in the capital steel and iron company and Beijing Fangshan community. *Obesity Reviews* **9 Suppl 1**:142–5.

Chenet L, McKee M, Fulop N, Bojan F, Brand H, Hort A and Kalbarczyk P (1996). Changing life expectancy in central Europe: is there a single reason? *Journal of Public Health Medicine* **18**:329–336.

Chiplonkar SA, Agte VV, Mengale SS and Tarwadi KV (2002). Are lifestyle factors good predictors of retinol and vitamin C deficiency in apparently healthy adults? *European Journal of Clinical Nutrition* **56**:96–104.

Chiva-Blanch G, Badimon L and Estruch R (2014). Latest evidence of the effects of the Mediterranean diet in prevention of cardiovascular disease. *Current Atherosclerosis Reports* **16**:446.

Conklin AI, Forouhi NG, Surtees P, Khaw K-T, Wareham NJ and Monsivais P (2014). Social relationships and healthful dietary behaviour: Evidence from over-50s in the EPIC cohort, UK. *Social Science and Medicine* **100**:167–175.

Crispim SP, Geelen A, Souverein OW, Hulshof PJM, Ruprich J, Dofkova M, Huybrechts I, *et al.* (2011). Biomarker-based evaluation of two 24-h recalls for comparing usual fish, fruit and vegetable intakes across European centers in the EFCOVAL Study. *European Journal of Clinical Nutrition* **65 Suppl 1**:S38–S47.

Cust AE, Smith BJ, Chau J, van der Ploeg HP, Friedenreich CM, Armstrong BK and Bauman A (2008). Validity and repeatability of the EPIC physical activity questionnaire: a validation study using accelerometers as an objective measure. *The International Journal of Behavioral Nutrition and Physical Activity* **5**:33.

Czernichow S, Blacher J and Hercberg S (2004). Antioxidant vitamins and blood pressure. *Current Hypertension Reports* **6**:27–30.

Darmon N and Drewnowski A (2008). Does social class predict diet quality? *The American Journal of Clinical Nutrition* **87**:1107–1117.

Dauchet L, Amouyel P and Dallongeville J (2005). Fruit and vegetable consumption and risk of stroke: a meta-analysis of cohort studies. *Neurology* **65**:1193–1197.

Dauchet L, Amouyel P and Dallongeville J (2009). Fruits, vegetables and coronary heart disease. *Nature Reviews. Cardiology* **6**:599–608.

Dauchet L, Amouyel P, Hercberg S and Dallongeville J (2006). Fruit and vegetable consumption and risk of coronary heart disease: a meta-analysis of cohort studies. *The Journal of Nutrition* **136**:2588–2593.

Dauchet L, Kesse-Guyot E, Czernichow S, Bertrais S, Estaquio C, Péneau S, Vergnaud AC, *et al.* (2007). Dietary patterns and blood pressure change over 5-y follow-up in the SU.VI.MAX cohort. *The American Journal of Clinical Nutrition* **85**:1650–1656.

Dauchet L, Montaye M, Ruidavets JB, Arveiler D, Kee F, Bingham A, Ferrieres J, *et al.* (2010). Association between the frequency of fruit and vegetable consumption and cardiovascular disease in male smokers and non-smokers. *European Journal of Clinical Nutrition* **64**:578–586.

Dore AR, Adair LS and Popkin BM (2003). Low income Russian families adopt effective behavioral strategies to maintain dietary stability in times of economic crisis. *The Journal of Nutrition* **133**:3469–3475.

Drake I, Gullberg B, Sonestedt E, Wallstrom P, Persson M, Hlebowicz J, Nilsson J, *et al.* (2013). Scoring models of a diet quality index and the predictive capability of mortality in a population-based cohort of Swedish men and women. *Public Health Nutrition* **16**:468–478.

Druesne-Pecollo N, Latino-Martel P, Norat T, Barrandon E, Bertrais S, Galan P and Hercberg S (2010). Beta-carotene supplementation and cancer risk: a systematic review and metaanalysis of randomized controlled trials. *International Journal of Cancer* **127**:172–184.

EFSA - European Food Safety Authority (2011a). *Comprehensive European Food Consumption Database*. 2011th ed. Available at: http://www.efsa.europa.eu/en/datexfoodcdb/datexfooddb.htm

EFSA - European Food Safety Authority (2011b). *FoodEx2*. Available at: http://www.efsa.europa.eu/en/datex/datexfoodclass

EFSA - European Food Safety Authority (2015). *The Food Classification and Description System FoodEx 2 (revision 2)*. Parma.

EFSA - European Food Safety Authority (2011c). Use of the EFSA Comprehensive European Food Consumption Database in Exposure Assessment. *EFSA Journal* **9**:2097.

EFSA - European Food Safety Authority (2010). *What`s on the Menu in Europe? - a Pan-European Food Consumption Survey (EU Menu)*. Seville.

El Ansari W, Stock C and Mikolajczyk RT (2012). Relationships between food consumption and living arrangements among university students in four European countries - a cross-sectional study. *Nutrition Journal* **11**:28.

Elmadfa I (2009). *European Nutrition and Health Report 2009.*

Estruch R, Ros E, Salas-Salvado J, Covas MI, Corella D, Aros F, Gomez-Gracia E, *et al.* (2013). Primary prevention of cardiovascular disease with a Mediterranean diet. *The New England Journal of Medicine* **368**:1279–1290.

EuroFIR - European Food Information Resource (2015). *EuroFIR AISBL.* Available at: http://www.eurofir.org/

European Commission (2013). Eurostat. Available at: http://ec.europa.eu/eurostat

FAO - Food and Agriculture Organization of the United Nations (2015). FAOSTAT. Available at: http://faostat3.fao.org/home/E

Feachem, R. (1994). Health decline in eastern Europe. *Nature* **367**:313–314.

Food Standards Agency (2002). *Food Portion Sizes*. 3rd edn. London: Her Majesty`s Stationery Office.

Food Standards Agency (2006). *FSA Nutrient and Food Based Guidelines for UK Institutions*. Available at: http://www.food.gov.uk/sites/default/files/multimedia/pdfs/nutrientinstitution.pdf

Forouzanfar MH, Moran AE, Flaxman AD, Roth G, Mensah GA, Ezzati M, Naghavi M, *et al.* (2012). Assessing the global burden of ischemic heart disease, part 2: analytic methods and estimates of the global epidemiology of ischemic heart disease in 2010. *Global Heart* **7**:331–342.

Frackiewicz J, Roszkowski W, Brzozowska A and Kaluza J (2010). [Diet quality and mortality in elderly people living in Warsaw Region]. *Przeglad Epidemiologiczny* **64**:119–125.

Fransen HP and Ocke MC (2008). Indices of diet quality. *Current Opinion in Clinical Nutrition and Metabolic Care* **11**:559–565.

Freedman LS, Commins JM, Moler JE, Willett W, Tinker LF, Subar AF, Spiegelman D, *et al.* (2015). Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for potassium and sodium intake. *American Journal of Epidemiology* **181**:473–487.

Friedenreich C, Norat T, Steindorf K, Boutron-Ruault MC, Pischon T, Mazuir M, Clavel-Chapelon F, *et al.* (2006). Physical activity and risk of colon and rectal cancers: the European prospective investigation into cancer and nutrition. *Cancer Epidemiology, Biomarkers & Prevention* **15**:2398–2407.

Gelbard A, Haub C and Kent MM (1999). World population beyond six billion. *Population Bulletin* **54**:1–44.

Genkinger JM, Platz EA, Hoffman SC, Comstock GW and Helzlsouer KJ (2004). Fruit, vegetable, and antioxidant intake and all-cause, cancer, and cardiovascular disease mortality in a community-dwelling population in Washington County, Maryland. *American Journal of Epidemiology* **160**:1223–1233.

Gilmore A, Pomerleau J, McKee M, Rose R, Haerpfer CW, Rotman D and Tumanov S (2004). Prevalence of smoking in 8 countries of the former Soviet Union: results from the living conditions, lifestyles and health study. *American Journal of Public Health* **94**:2177–2187.

Ginter E (1998). Cardiovascular disease prevention in eastern Europe. *Nutrition* **14**:452–457.

Ginter E (1995). Cardiovascular risk factors in the former communist countries. Analysis of 40 European MONICA populations. *European Journal of Epidemiology* **11**:199–205.

Giskes K, Avendaňo M, Brug J and Kunst AE (2009). A systematic review of studies on socioeconomic inequalities in dietary intakes associated with weight gain and overweight/obesity conducted among European adults. *Obesity Reviews* **11**:413–429.

Giskes K, Turrell G, van Lenthe FJ, Brug J and Mackenbach JP (2006). A multilevel study of socio-economic inequalities in food choice behaviour and dietary intake among the Dutch population: the GLOBE study. *Public Health Nutrition* **9**:75–83.

Del Gobbo LC, Khatibzadeh S, Imamura F, Micha R, Shi P, Smith M, Myers SS, *et al.* (2015). Assessing global dietary habits: a comparison of national estimates from the FAO and the Global Dietary Database. *The American Journal of Clinical Nutrition* **101**:1038–1046.

Goryakin Y, Rocco L, Suhrcke M, Roberts B and McKee M (2015). Fruit and vegetable consumption in the former Soviet Union: the role of individual- and community-level factors. *Public Health Nutrition* **18:**2825-2835.

Hall JN, Moore S, Harper SB and Lynch JW (2009). Global variability in fruit and vegetable consumption. *American Journal of Preventive Medicine* **36**:402–409 e5.

Hamer M. and Chida Y (2007). Intake of fruit, vegetables, and antioxidants and risk of type 2 diabetes: systematic review and meta-analysis. *Journal of Hypertension* **25**:2361–2369.

Hartley L, Igbinedion E, Holmes J, Flowers N, Thorogood M, Clarke A, Stranges S, *et al.* (2013). Increased consumption of fruit and vegetables for the primary prevention of cardiovascular diseases. *The Cochrane Database of Systematic Reviews* **6**:CD009874.

Havas S, Heimendinger J, Damron D, Nicklas TA, Cowan A, Beresford SA, Sorensen G, *et al.* (1995). 5 A Day for better health--nine community research projects to increase fruit and vegetable consumption. *Public Health Reports* **110**:68–79.

He FJ, Li J and Macgregor GA (2013). Effect of longer term modest salt reduction on blood pressure: Cochrane systematic review and meta-analysis of randomised trials. *The British Medical Journal* **346**:f1325.

He FJ, Nowson CA, Lucas M and MacGregor GA (2007). Increased consumption of fruit and vegetables is related to a reduced risk of coronary heart disease: meta-analysis of cohort studies. *Journal of Human Hypertension* **21**:717–728.

He FJ, Nowson CA and MacGregor GA (2006). Fruit and vegetable consumption and stroke: meta-analysis of cohort studies. *Lancet* **367**:320–326.

He Y (2010). Missing data analysis using multiple imputation: getting to the heart of the matter. *Circulation. Cardiovascular quality and outcomes* **3**:98–105.

Henriquez-Sanchez P, Sanchez-Villegas A, Doreste-Alonso J, Ortiz-Andrellucchi A, Pfrimer K and Serra-Majem L (2009). Dietary assessment methods for micronutrient intake: a systematic review on vitamins. *The British Journal of Nutrition* **102 Suppl**:S10–37.

Von Hippel PT (2007). Regression with missing Ys: an improved strategy for analyzing multiply imputed data. *Sociological Methodology* **37**:83–117.

Hoffman R and Gerber M (2013). Evaluating and adapting the Mediterranean diet for non-Mediterranean populations: a critical appraisal. *Nutrition Reviews* **71**:573–584.

Hoffmann K, Schulze MB, Schienkiewitz A, Nöthlings U and Boeing H (2004). Application of a new statistical method to derive dietary patterns in nutritional epidemiology. *American Journal of Epidemiology* **159**:935–944.

Honkanen P and Frewer L (2009). Russian consumers' motives for food choice. *Appetite* **52**:363–371.

Hu FB (2002). Dietary pattern analysis: a new direction in nutritional epidemiology. *Current Opinion in Lipidology* **13**:3–9.

Hu FB and Willett WC (2002). Optimal diets for prevention of coronary heart disease. *The Journal of the American Medical Association* **288**:2569–2578.

Huang J, Frohlich J and Ignaszewski AP (2011). The impact of dietary changes and dietary supplements on lipid profile. *The Canadian Journal of Cardiology* **27**:488–505.

Huijbregts P, Feskens E, Rasanen L, Fidanza F, Nissinen A, Menotti A and Kromhout D (1997). Dietary pattern and 20 year mortality in elderly men in Finland, Italy, and The Netherlands: longitudinal cohort study. *The British Medical Journal* **315**:13–17.

Hung HC, Joshipura KJ, Jiang R, Hu FB, Hunter D, Smith-Warner SA, Colditz GA, *et al.* (2004). Fruit and vegetable intake and risk of major chronic disease. *Journal of the National Cancer Institute* **96**:1577–1584.

Ireland J, van Erp-Baart AM, Charrondiere UR, Moller A, Smithers G and Trichopoulou A (2002). Selection of a food classification system and a food composition database for future food consumption surveys. *European Journal of Clinical Nutrition* **56 Suppl 2**:S33–45.

Jacques PF and Tucker KL (2001). Are dietary patterns useful for understanding the role of diet in chronic disease? *The American Journal of Clinical Nutrition* **73**:1–2.

Jahns L, Baturin A and Popkin BM (2003). Obesity, diet, and poverty: trends in the Russian transition to market economy. *European Journal of Clinical Nutrition* **57**:1295–1302.

Jankovic N, Geelen A, Streppel MT, de Groot LC, Kiefte-de Jong JC, Orfanos P, Bamia C, *et al.* (2015). WHO guidelines for a healthy diet and mortality from cardiovascular disease in European and American elderly: the CHANCES project. *The American Journal of Clinical Nutrition* **102**:745–756.

Jankovic N, Geelen A, Streppel MT, de Groot LC, Orfanos P, van den Hooven EH, Pikhart H, *et al.* (2014). Adherence to a healthy diet according to the World Health Organization guidelines and all-cause mortality in elderly adults from Europe and the United States. *American Journal of Epidemiology* **180**:978–988.

Jansen MCJF, Van Kappel AL, Ocké MC, Van't Veer P, Boshuizen HC, Riboli E and Bueno-de-Mesquita HB (2004). Plasma carotenoid levels in Dutch men and women, and the relation with vegetable and fruit consumption. *European Journal of Clinical Nutrition* **58**:1386–1395.

Jenab M, Salvini S, van Gils CH, Brustad M, Shakya-Shrestha S, Buijsse B, Verhagen H, *et al.* (2009). Dietary intakes of retinol, beta-carotene, vitamin D and vitamin E in the European Prospective Investigation into Cancer and Nutrition cohort. *European Journal of Clinical Nutrition* **63 Suppl 4**:S150–178.

Jenab M, Slimani N, Bictash M, Ferrari P and Bingham SA (2009). Biomarkers in nutritional epidemiology: applications, needs and new horizons. *Human Genetics* **125**:507–525.

Joffe M and Robertson A (2001). The potential contribution of increased vegetable and fruit consumption to health gain in the European Union. *Public Health Nutrition* **4**:893–901.

John JH, Ziebland S, Yudkin P, Roe LS and Neil HA (2002). Effects of fruit and vegetable consumption on plasma antioxidant concentrations and blood pressure: a randomised controlled trial. *Lancet* **359**:1969–1974.

Kant AK (2004). Dietary patterns and health outcomes. *Journal of the American Dietetic Association* **104**:615–635.

Kant AK (2010). Dietary patterns: biomarkers and chronic disease risk. *Applied Physiology, Nutrition, and Metabolism* **35**:199–206.

Karamanos B, Thanopoulou A, Angelico F, Assaad-Khalil S, Barbato A, Del Ben M, Dimitrijevic-Sreckovic V, *et al.* (2002). Nutritional habits in the Mediterranean Basin. The macronutrient composition of diet and its relation with the traditional Mediterranean diet. Multi-centre study of the Mediterranean Group for the Study of Diabetes (MGSD). *European Journal of Clinical Nutrition* **56**:983–991.

Karanikolos M, Leon DA, Smith PC and McKee M (2012). Minding the gap: changes in life expectancy in the Baltic States compared with Finland. *Journal of Epidemiology and Community Health* **66**:1043–1049.

Kardinaal A, Kok FJ, Ringstad J, Gomez-Aracena J, Mazaev VP, Kohlmeier L, Martin BC, *et al.* (1993). Antioxidants in adipose tissue and risk of myocardial infarction: The EURAMIC study. *Lancet* **342**:1379–1384.

Kartashov AI, Buzin VV, Glazunov IS and Abol'ian LV (1991). [Results of testing the basic questionnaire 'SINDI' in the evaluation of prophylaxis program development]. *Sovetskoe Zdravookhranenie / Ministerstvo Zdravookhraneniia SSSR*:45–49.

Kastorini CM, Milionis HJ, Esposito K, Giugliano D, Goudevenos JA and Panagiotakos DB (2011). The effect of Mediterranean diet on metabolic syndrome and its components: a meta-analysis of 50 studies and 534,906 individuals. *Journal of the American College of Cardiology* **57**:1299–1313.

Kastorini CM, Milionis HJ, Goudevenos JA and Panagiotakos DB (2010). Mediterranean diet and coronary heart disease: is obesity a link? - A systematic review. *Nutrition, Metabolism, and Cardiovascular Diseases* **20**:536–551.

Katsarou A, Tyrovolas S, Psaltopoulou T, Zeimbekis A, Tsakountakis N, Bountziouka V, Gotsis E, *et al.* (2010). Socio-economic status, place of residence and dietary habits among the elderly: the Mediterranean islands study. *Public Health Nutrition* **13**:1614–1621.

Kawakami N, Tsutsumi A, Haratani T, Kobayashi F, Ishizaki M, Hayashi T, Fujita O, et al (2006). Job strain, worksite support, and nutrient intake among employed Japanese men and women. *Journal of Epidemiology* **16**:79-89.

Kesteloot H, Sans S and Kromhout D (2006). Dynamics of cardiovascular and all-cause mortality in Western and Eastern Europe between 1970 and 2000. *European Heart Journal* **27**:107–113.

Keys A (1980). Coronary heart disease, serum cholesterol, and the diet. *Acta medica Scandinavica* **207**:153–160.

KilBridge AL, Stefler D, Malyutina S, Andrzej P, Kubinova R, Peasey A, Pikhart H, *et al.* Meat Consumption and All Cause, Cardiovascular Disease and Cancer mortality in Eastern Europe. Non-published manuscript (*Under review)*.

Kinsella KG (1992). Changes in life expectancy 1900-1990. *The American Journal of Clinical Nutrition* **55**:1196S–1202S.

Kipnis V, Midthune D, Freedman L, Bingham S, Day NE, Riboli E, Ferrari P*, et al.* (2002). Bias in dietary-report instruments and its implications for nutritional epidemiology. *Public Health Nutrition* **5**:915–923.

Knoops KT, Groot de LC, Fidanza F, Alberti-Fidanza A, Kromhout D and van Staveren WA (2006). Comparison of three different dietary scores in relation to 10-year mortality in elderly European subjects: the HALE project. *European Journal of Clinical Nutrition* **60**:746–755.

Koenker R, Portnoy S, Ng PT, Zeileis A, Grosjean P and Ripley BD (2013). Quantile Regression. *The Journal of Economic Perspectives* **15**:143–156.

Koloverou E, Esposito K, Giugliano D and Panagiotakos D (2014). The effect of Mediterranean diet on the development of type 2 diabetes mellitus: a meta-analysis of 10 prospective studies and 136,846 participants. *Metabolism* **63**:903–911.

Kopp M, Skrabski A, Szanto Z and Siegrist J (2006). Psychosocial determinants of premature cardiovascular mortality differences within Hungary. *Journal of Epidemiology and Community Health* **60**:782–788.

Kristenson M, Ziedén B, Kucinskienë Z, Elinder LS, Bergdahl B, Elwing B, Abaravicius A, *et al.* (1997). Antioxidant state and mortality from coronary heart disease in Lithuanian and Swedish men: concomitant cross sectional study of men aged 50. *The British Medical Journal* **314**:629–633.

Kromhout D, Keys A, Aravanis C, Buzina R, Fidanza F, Giampaoli S, Jansen A, *et al.* (1989). Food consumption patterns in the 1960s in seven countries. *The American Journal of Clinical Nutrition* **49**:889–894.

Krupp D, Shi L, Egert S, Wudy SA and Remer T (2014). Prospective relevance of fruit and vegetable consumption and salt intake during adolescence for blood pressure in young adulthood. *European Journal of Nutrition*. 20th Nov. [Epub ahead of print]

Laatikainen T, Alho H, Vartiainen E, Jousilahti P, Sillanaukee P and Puska P (2002). Self-reported alcohol consumption and association to carbohydrate-deficient transferrin and gamma-glutamyltransferase in a random sample of the general population in the Republic of Karelia, Russia and in North Karelia, Finland. *Alcohol and Alcoholism* **37**:282–288.

Lallukka T, Sarlio-Lähteenkorva S, Roos E, Laaksonen M, Rahkonen O, Lahelma E (2004). Working conditions and health behaviours among employed women and men: the Helsinki Health Study. *Preventive Medicine* **38**:48-56.

Larsson SC and Orsini N (2014). Red meat and processed meat consumption and all-cause mortality: a meta-analysis. *American Journal of Epidemiology* **179**:282–289.

Laszlo KD, Pikhart H, Kopp MS, Bobak M, Pajak A, Malyutina S, Salavecz G, *et al.* (2010). Job insecurity and health: a study of 16 European countries. *Social Science & Medicine* **70**:867–874.

Lawlor DA, Davey Smith G, Kundu D, Bruckdorfer KR and Ebrahim S (2004). Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? *Lancet* **363**:1724–1727.

Leenders M, Sluijs I, Ros MM, Boshuizen HC, Siersema PD, Ferrari P, Weikert C, *et al.* (2013). Fruit and vegetable consumption and mortality: European prospective investigation into cancer and nutrition. *American Journal of Epidemiology* **178**:590–602.

Leifman H (2002). The six-country survey of the European comparative alcohol study: comparing patterns and assessing validity. *Contemporary Drug Problems*:477–500.

Leon DA, Chenet L, Shkolnikov VM, Zakharov S, Shapiro J, Rakhmanova G, Vassin S, *et al.* (1997). Huge variation in Russian mortality rates 1984-94: artefact, alcohol, or what? *Lancet* **350**:383–388.

Leon DA, Saburova L, Tomkins S, Andreev E, Kiryanov N, McKee M and Shkolnikov VM (2007). Hazardous alcohol drinking and premature mortality in Russia: a population based case-control study. *Lancet* **369**:2001–2009.

Leon DA, Shkolnikov VM and McKee M (2009). Alcohol and Russian mortality: a continuing crisis. *Addiction* **104**:1630–1636.

Lesser S, Pauly L, Volkert D and Stehle P (2008). Nutritional situation of the elderly in Eastern/Baltic and Central/Western Europe - the AgeingNutrition project. *Annals of Nutrition & Metabolism* **52 Suppl 1**:62–71.

Lewington S, Clarke R, Qizilbash N, Peto R and Collins R (2002). Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* **360**:1903–13.

Lewis RA (2002). *CRC Dictionary of Agricultural Sciences*. Boca Raton, Florida, US: CRC Press.

Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, Amann M, *et al.* (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**:2224–2260.

Lixandru D, Mohora M, Coman A, Stoian I, Van Gils C, Aerts P and Manuel-Y-Keenoy B. (2010). Diet and paraoxonase 1 Enzymatic activity in diabetic foot patients from romania and belgium: Favorable association of high flavonoid dietary intake with arylesterase activity. *Annals of Nutrition and Metabolism* **56**:294–301.

Lock K, Pomerleau J, Causer L, Altmann DR and McKee M (2005). The global burden of disease attributable to low consumption of fruit and vegetables: implications for the global strategy on diet. *Bulletin of the World Health Organization* **83**:100–108.

de Lorgeril M, Salen P, Martin JL, Mamelle N, Monjaud I, Touboul P and Delaye J (1996). Effect of a mediterranean type of diet on the rate of cardiovascular complications in patients with coronary artery disease. Insights into the cardioprotective effect of certain nutriments. *Journal of the American College of Cardiology* **28**:1103–1108.

Lundberg J, Bobak M, Malyutina S, Kristenson M and Pikhart H (2007). Adverse health effects of low levels of perceived control in Swedish and Russian community samples. *BMC Public Health* **7**:314.

Lunze K, Yurasova E, Idrisov B, Gnatienko N and Migliorini L (2015). Food security and nutrition in the Russian Federation a health policy analysis. *Global Health Action* **8:**27537.

Mackenbach JP, Karanikolos M and McKee M (2013). The unequal health of Europeans: successes and failures of policies. *Lancet* **381**:1125–1134.

Malik VS, Pan A, Willett WC and Hu FB (2013). Sugar-sweetened beverages and weight gain in children and adults: a systematic review and meta-analysis. *The American Journal of Clinical Nutrition* **98**:1084–1102.

Malyutina S, Bobak M, Kurilovitch S, Ryizova E, Nikitin Y and Marmot M (2001). Alcohol consumption and binge drinking in Novosibirsk, Russia, 1985-95. *Addiction* **96**:987–995.

Malyutina S, Simonova G and Nikitin Y (2002). Coronary heart disease and cardiovascular mortality in the urban Siberian population: gender specific findings from a 10-year cohort study. In: Weidner G, Kopp M and Kristenson M (eds) *Heart Disease: Environment, Stress and Gender* Amsterdam: IOS Press, pp. 69–79.

Marmot M and Brunner E (2005). Cohort Profile: the Whitehall II study. *International journal of epidemiology* **34**:251–256.

Marmot M, Smith GD, Stansfeld S, Patel C, North F, Head J, White I, *et al.* (1991). Health inequalities among British civil servants: the Whitehall II study. *Lancet* **337**:1387–1393.

Marrie RA, Dawson NV and Garland A (2009). Quantile regression and restricted cubic splines are useful for exploring relationships between continuous variables. *Journal of Clinical Epidemiology* **62**:511–517.

Martínez ME, Marshall JR and Sechrest L (1998). Invited commentary: Factor analysis and the search for objectivity. *American Journal of Epidemiology* **148**:17–19.

Martínez-González MA, Fernández-Jarne E, Serrano-Martínez M, Marti A, Martinez JA and Martín-Moreno JM (2002). Mediterranean diet and reduction in the risk of a first acute myocardial infarction: an operational healthy dietary score. *European Journal of Nutrition* **41**:153–160.

Mayén AL, Marques-Vidal P, Paccaud F, Bovet P and Stringhini S (2014). Socioeconomic determinants of dietary patterns in low- and middle-income countries: a systematic review. *The American Journal of Clinical Nutrition* **100**:1520–1531.

McCance R and Widdowson E (2002). *McCance & Widdowson`s The Composition of Foods: Summary Edition*. 6th edn. Cambridge: Royal Society of Chemistry.

McColl K and Lobstein T (2015). *Nutrient Profiling: Changing the Food of Britain*. London.

McKee M (2007). Cochrane on Communism: the influence of ideology on the search for evidence. *International Journal of Epidemiology* **36**:269–273.

McKee M, Bobak M, Rose R, Shkolnikov V, Chenet L and Leon D (1998). Patterns of smoking in Russia. *Tobacco Control* **7**:22–26.

Mente A, de Koning L, Shannon HS and Anand SS (2009). A systematic review of the evidence supporting a causal link between dietary factors and coronary heart disease. *Archives of Internal Medicine* **169**:659–669.

Micha R, Khatibzadeh S, Shi P, Fahimi S, Lim S, Andrews KG, Engell RE, *et al.* (2014). Global, regional, and national consumption levels of dietary fats and oils in 1990 and 2010: a systematic analysis including 266 country-specific nutrition surveys. *The British Medical Journal* **348**:g2272.

Michels KB (2003). Nutritional epidemiology--past, present, future. *International Journal of Epidemiology* **32**:486–488.

Miere D, Filip L, Indrei LL, Soriano JM, Molto JC and Manes J (2007). [Nutritional assessment of the students from two European university centers]. *Revista Medico-chirurgicala a Societatii de Medici si Naturalisti din Iasi* **111**:270–275.

Mila-Villarroel R, Bach-Faig A, Puig J, Puchal A, Farran A, Serra-Majem L and Carrasco JL (2011). Comparison and evaluation of the reliability of indexes of adherence to the Mediterranean diet. *Public Health Nutrition* **14**:2338–2345.

Moreno-Gómez C, Romaguera-Bosch D, Tauler-Riera P, Bennasar-Veny M, Pericas-Beltran J, Martinez-Andreu S and Aguilo-Pons A (2012). Clustering of lifestyle factors in Spanish university students: the relationship between smoking, alcohol consumption, physical activity and diet quality. *Public Health Nutrition* **15**:2131–2139.

Muraki I, Imamura F, Manson JE, Hu FB, Willett WC, van Dam RM and Sun Q (2013). Fruit consumption and risk of type 2 diabetes: results from three prospective longitudinal cohort studies. *The British Medical Journal* **347**:f5001.

Nagura J, Iso H, Watanabe Y, Maruyama K, Date C, Toyoshima H, Yamamoto A, *et al.* (2009). Fruit, vegetable and bean intake and mortality from cardiovascular disease among Japanese men and women: the JACC Study. *The British Journal of Nutrition* **102**:285–292.

Naska A, Vasdekis VG, Trichopoulou A, Friel S, Leonhäuser IU, Moreiras O, Nelson M, *et al.* (2000). Fruit and vegetable availability among ten European countries: how does it compare with the 'five-a-day' recommendation? DAFNE I and II projects of the European Commission. *The British Journal of Nutrition* **84**:549–556.

National and Kapodistrian University of Athens (2005). Data Food Networking (DAFNE) databank. Available at: http://www.nut.uoa.gr/dafnesoftweb/

National Research University Higher School of Economics, ZAO Demoscope and Carolina Population Center, University of North Carolina (2013). *The Russia Longitudinal Health Survey-Higher School of Economics*. Available at: http://www.cpc.unc.edu/projects/rlms-hse

Newby PK and Tucker KL (2004). Empirically derived eating patterns using factor or cluster analysis: a review. *Nutrition Reviews* **62**:177–203.

Nicholson A, Bobak M, Murphy M, Rose R and Marmot M (2005). Alcohol consumption and increased mortality in Russian men and women: a cohort study based on the mortality of relatives. *Bulletin of the World Health Organization* **83**:812–819.

Nicklett EJ, Semba RD, Simonsick EM, Szanton S, Bandeen-Roche K, Ferrucci L, Guralnik JM, et al (2012). Diet quality and social support: factors associated with serum carotenoid concentrations among older disabled women (the Women's Health and Aging Study). *The Journal of Nutrition, Health & Aging* **16**:511-518.

Novakovic R, Cavelaars AE, Bekkering GE, Roman-Vinas B, Ngo J, Gurinovic M, Glibetic M, *et al.* (2013). Micronutrient intake and status in Central and Eastern Europe compared with other European countries, results from the EURRECA network. *Public Health Nutrition* **16**:824–840.

Ornish D. (2013). Mediterranean diet for primary prevention of cardiovascular disease. *The New England Journal of Medicine* **369**:675–676.

Paalanen L, Prattala R, Palosuo H and Laatikainen T (2011). Socio-economic differences in the consumption of vegetables, fruit and berries in Russian and Finnish Karelia: 1992-2007. *European Journal of Public Health* **21**:35–42.

Pajak A, Szafraniec K, Kubinova R, Malyutina S, Peasey A, Pikhart H, Nikitin Y, *et al.* (2013). Binge drinking and blood pressure: cross-sectional results of the HAPIEE study. *PloS One* **8**:e65856.

Pallot J and Nefedova T (2003). Geographical Differentiation in Household Plot Production in Rural Russia. *Eurasian Geography and Economics* **44**:40–64.

Palosuo H (2000). Health-related lifestyles and alienation in Moscow and Helsinki. *Social Science & Medicine (1982)* **51**:1325–1341.

Panagiotakos DB, Pitsavos C, Chrysohoou C, Vlismas K, Skoumas Y, Palliou K and Stefanadis C (2008). Dietary habits mediate the relationship between socio-economic status and CVD factors among healthy adults: the ATTICA study. *Public Health Nutrition* **11**:1342–1349.

Papadakis S and Moroz I (2008). Population-level interventions for coronary heart disease prevention: what have we learned since the North Karelia project? *Current Opinion in Cardiology* **23**:452–461.

Park Y, Subar AF, Hollenbeck A and Schatzkin A (2011). Dietary fiber intake and mortality in the NIH-AARP diet and health study. *Archives of Internal Medicine* **171**:1061–1068.

Passmore R and Eastwood M (1986). Foods from the vegetable kingdom. In: Passmore R and Eastwood M (ed.) *Davidson and Passmore Human Nutrition and Dietetics*. 8th edn. Edinburgh: Churchill Livingstone, pp. 193–205.

Peasey A, Bobak M, Kubinova R, Malyutina S, Pajak A, Tamosiunas A, Pikhart H, *et al.* (2006). Determinants of cardiovascular disease and other non-communicable diseases in Central and Eastern Europe: rationale and design of the HAPIEE study. *BMC Public Health* **6**:255.

Petkevičiene J, Klumbiene J, Ramažauskiene V, Kriaučioniene V, Šakyte E and Grabauskas V (2012). Diet and dyslipidemias in a lithuanian rural population aged 25-64: The CINDI survey. *Medicina (Lithuania)* **48**:211–217.

Petkeviciene J, Simila M, Becker W, Kriaucioniene V and Valsta LM (2009). Validity and reproducibility of the NORBAGREEN food frequency questionnaire. *European Journal of Clinical Nutrition* **63**:141–149.

Poledne R and Skodova Z (2000). Changes in nutrition, cholesterol concentration, and cardiovascular disease mortality in the Czech population in the past decade. *Nutrition* **16**:785–786.

Pomerleau J, Lock K, Knai C and McKee M (2005). Interventions designed to increase adult fruit and vegetable intake can be effective: a systematic review of the literature. *The Journal of Nutrition* **135**:2486–2495.

Pomerleau J, Lock K and McKee M (2003). Discrepancies between ecological and individual data on fruit and vegetable consumption in fifteen countries. *The British Journal of Nutrition* **89**:827–834.

Pomerleau J, Lock K and McKee M (2006). The burden of cardiovascular disease and cancer attributable to low fruit and vegetable intake in the European Union: differences between old and new Member States. *Public Health Nutrition* **9**:575–983.

Pomerleau J, McKee M, Lobstein T and Knai C (2003). The burden of disease attributable to nutrition in Europe. *Public Health Nutrition* **6**:453–461.

Pomerleau J, McKee M, Rose R, Haerpfer CW, Rotman D and Tumanov S (2005). Drinking in the Commonwealth of Independent States - Evidence from eight countries. *Addiction* **100**:1647–1668.

Powles JW, Day NE, Sanz MA and Bingham SA (1996). Protective foods in winter and spring: a key to lower vascular mortality? *Lancet* **348**:898–899.

Prattala R, Hakala S, Roskam AJ, Roos E, Helmert U, Klumbiene J, Van Oyen H, *et al.* (2009). Association between educational level and vegetable use in nine European countries. *Public Health Nutrition* **12**:2174–2182.

Prättälä R, Paalanen L, Grinberga D, Helasoja V, Kasmel A and Petkeviciene J (2007). Gender differences in the consumption of meat, fruit and vegetables are similar in Finland and the Baltic countries. *European Journal of Public Health* **17**:520–525.

Prentice RL (2003). Dietary assessment and the reliability of nutritional epidemiology reports. *Lancet* **362**:182–183.

Psaltopoulou T, Sergentanis TN, Panagiotakos DB, Sergentanis IN, Kosti R and Scarmeas N (2013). Mediterranean diet, stroke, cognitive impairment, and depression: A meta-analysis. *Annals of Neurology* **74**:580–591.

Pudule I, Grinberga D, Kadziauskiene K, Abaravicius A, Vaask S, Robertson A and McKee M (1999). Patterns of smoking in the Baltic Republics. *Journal of Epidemiology and Community Health* **53**:277–282.

Puska P, Helasoja V, Prattala R, Kasmel A and Klumbiene J (2003). Health behaviour in Estonia, Finland and Lithuania 1994-1998. Standardized comparison. *European Journal of Public Health* **13**:11–17.

Puska P, Salonen JT, Nissinen A, Tuomilehto J, Vartiainen E, Korhonen H, Tanskanen A, *et al.* (1983). Change in risk factors for coronary heart disease during 10 years of a community intervention programme (North Karelia project). *The British Medical Journal* **287**:1840–1844.

Puska P, Vartiainen E, Tuomilehto J, Salomaa V and Nissinen A (1998). Changes in premature deaths in Finland: successful long-term prevention of cardiovascular diseases. *Bulletin of the World Health Organization* **76**:419–425.

Quiñones M, Miguel M. and Aleixandre A (2013). Beneficial effects of polyphenols on cardiovascular disease. *Pharmacological Research* **68**:125–131.

Ramazauskiene V, Petkeviciene J, Klumbiene J, Kriaucioniene V and Sakyte E (2011). Diet and serum lipids: changes over socio-economic transition period in Lithuanian rural population. *BMC Public Health* **11**:447.

Rechel B, Roberts B, Richardson E, Shishkin S, Shkolnikov VM, Leon DA, Bobak, M, *et al.* (2013). Health and health systems in the Commonwealth of Independent States. *Lancet* **381**:1145–1155.

Rees K, Hartley L, Flowers N, Clarke A, Hooper L, Thorogood M and Stranges S (2013). ’ Mediterranean ’ dietary pattern for the primary prevention of cardiovascular disease (Review). *The Cochrane Database of Systematic Reviews* **8**:CD009825.

Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondière UR, *et al.* (2002). European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutrition* **5**:1113–1124.

Roark RA and Niederhauser VP (2013). Fruit and vegetable intake: issues with definition and measurement. *Public Health Nutrition* **16**:2–7.

Rolls BJ, Ello-Martin JA and Tohill BC (2004). What can intervention studies tell us about the relationship between fruit and vegetable consumption and weight management? *Nutrition Reviews* **62**:1–17.

Russian Federation`s Federal State Statistical Service (2014). *Production of Basic Agricultural Products by Types of Enterprises* [Online]. Available at: http://www.gks.ru/bgd/regl/b13_12/IssWWW.exe/stg/d01/15-04.htm

Salavecz G, Chandola T, Pikhart H, Dragano N, Siegrist J, Jockel KH, Erbel R, *et al.* (2010). Work stress and health in Western European and post-communist countries: an East-West comparison study. *Journal of Epidemiology and Community Health* **64**:57–62.

Schneider S, Huy C, Schuessler M, Diehl K and Schwarz S (2009). Optimising lifestyle interventions: identification of health behaviour patterns by cluster analysis in a German 50+ survey. *European Journal of Public Health* **19**:271–277.

Schoenfeld D. (1980). Chi-square goodness-of-fit test for the proportional hazards regression model. *Biometrika*:145–153.

Schofield WN (1985). Predicting basal metabolic rate, new standards and review of previous work. *Human Nutrition.* **39 Suppl 1**:5–41.

Schröder H, Covas M, Elosua R, Mora J and Marrugat J (2008). Diet quality and lifestyle associated with free selected low-energy density diets in a representative Spanish population. *European Journal of Clinical Nutrition* **62**:1194–1200.

Schröder H, Fitó M, Estruch R, Martínez-González MA, Corella D, Salas-Salvadó J, Lamuela-Raventós R, *et al.* (2011). A short screener is valid for assessing Mediterranean diet adherence among older Spanish men and women. *The Journal of Nutrition* **141**:1140–1145.

Schroll K, Carbajal A, Decarli B, Martins I, Grunenberger F, Blauw YH and de Groot CP (1996). Food patterns of elderly Europeans. SENECA Investigators. *European Journal of Clinical Nutrition* **50 Suppl 2**:S86–100.

Schuit J, van Loon JM, Tijhuis M and Ocké M (2002). Clustering of lifestyle risk factors in a general adult population. *Preventive Medicine* **35**:219–224.

Seeth HT, Chachnov S, Surinov A and Von Braun J (1998). Russian poverty: muddling through economic transition with garden plots. *World Development* **26**:1611–1623.

Serra-Majem L, MacLean D, Ribas L, Brule D, Sekula W, Prattala R, Garcia-Closas R, *et al.* (2003). Comparative analysis of nutrition data from national, household, and individual levels: results from a WHO-CINDI collaborative project in Canada, Finland, Poland, and Spain. *Journal of Epidemiology and Community Health* **57**:74–80.

Sesso HD, Stampfer MJ, Rosner B, Hennekens CH, Gaziano JM, Manson JE and Glynn RJ (2000). Systolic and diastolic blood pressure, pulse pressure, and mean arterial pressure as predictors of cardiovascular disease risk in Men. *Hypertension* **36**:801–807.

Shepherd R (2005). Influences on food choice and dietary behavior. *Forum of Nutrition*:36–43.

Shi L, Krupp D and Remer T (2014). Salt, fruit and vegetable consumption and blood pressure development: a longitudinal investigation in healthy children. *The British Journal of Nutrition* **111**:662–671.

Shkolnikov V, McKee M and Leon DA (2001). Changes in life expectancy in Russia in the mid-1990s. *Lancet* **357**:917–921.

Sjogren P, Becker W, Warensjo E, Olsson E, Byberg L, Gustafsson IB, Karlstrom B, *et al.* (2010). Mediterranean and carbohydrate-restricted diets and mortality among elderly men: a cohort study in Sweden. *The American Journal of Clinical Nutrition* **92**:967–974.

Skuladottir H, Tjoenneland A, Overvad K, Stripp C, Christensen J, Raaschou-Nielsen O and Olsen JH (2004). Does insufficient adjustment for smoking explain the preventive effects of fruit and vegetables on lung cancer? *Lung Cancer* **45**:1–10.

Slattery ML (2010). Analysis of dietary patterns in epidemiological research. *Applied Physiology, Nutrition, and Metabolism* **35**:207–210.

Slinker BK and Glantz SA (1985). Multiple regression for physiological data analysis: the problem of multicollinearity. *The American Journal of Physiology* **249**:R1–12.

Sofi F, Cesari F, Abbate R, Gensini GF and Casini A (2008). Adherence to Mediterranean diet and health status: meta-analysis. *The British Medical Journal* **337**:a1344.

Sofi F, Macchi C, Abbate R, Gensini GF and Casini A (2014). Mediterranean diet and health status: an updated meta-analysis and a proposal for a literature-based adherence score. *Public Health Nutrition* **17**:2769–2782.

Southworth C (2006). The Dacha Debate: Household Agriculture and Labor Markets in Post-Socialist Russia. *Rural Sociology* **71**:451–478.

Stables GJ, Subar AF, Patterson BH, Dodd K, Heimendinger J, Van Duyn MAS and Nebeling L (2002). Changes in vegetable and fruit consumption and awareness among US adults: results of the 1991 and 1997 5 A Day for Better Health Program surveys. *Journal of the American Dietetic Association* **102**:809–817.

Steffen LM, Kroenke CH, Yu X, Pereira MA, Slattery ML, Van Horn L, Gross MD, *et al.* (2005). Associations of plant food, dairy product, and meat intakes with 15-y incidence of elevated blood pressure in young black and white adults: the Coronary Artery Risk Development in Young Adults (CARDIA) Study. *The American Journal of Clinical Nutrition* **82**:1169–1177.

Stefler D (2011). Relationships between fruit, vegetable consumptions, nutrient intakes and plasma antioxidant levels in three Central and Eastern European countries. MSc thesis. University College London.

Steptoe A, Tsuda A, Tanaka Y and Wardle J (2007). Depressive symptoms, socio-economic background, sense of control, and cultural factors in university students from 23 countries. *International Journal of Behavioral Medicine* **14**:97–107.

Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, *et al.* (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *The British Medical Journal* **338**:b2393.

Su LC, Bui M, Kardinaal A, Gomez-Aracena J, Martin-Moreno J, Martin B, Thamm M, *et al.* (1998). Differences between plasma and adipose tissue biomarkers of carotenoids and tocopherols. *Cancer Epidemiology, Biomarkers & Prevention* **7**:1043–1048.

Suliburska J, Bogdański P, Duda G, Pupek-Musialik D, Piątek J and Żukiewicz-Sobczak W (2012). An assessment of dietary intake and state of nutritional in hypertensive patients from rural and urban areas of Greater Poland. *Annals of Agricultural and Environmental Medicine* **19**:339–343.

Tessier S and Gerber M (2005). Comparison between Sardinia and Malta: the Mediterranean diet revisited. *Appetite* **45**:121–126.

The World Bank (2000). Making Transition Work for Everyone. Poverty and Inequality in Europe and Central Asia. Washington DC.

Tomkins S, Collier T, Oralov A, Saburova L, McKee M, Shkolnikov V, Kiryanov N, *et al.* (2012). Hazardous alcohol consumption is a major factor in male premature mortality in a typical Russian city: prospective cohort study 2003-2009. *PloS One* **7**:e30274.

Trichopoulou A (1992). Monitoring food intake in Europe: a food data bank based on household budget surveys. *European Journal of Clinical Nutrition* **46 Suppl 5**:S3–8.

Trichopoulou A, Kouris-Blazos A, Wahlqvist ML, Gnardellis C, Lagiou P, Polychronopoulos E, Vassilakou T, *et al.* (1995). Diet and overall survival in elderly people. *The British Medical Journal* **311**:1457–1460.

Trichopoulou A, Martínez-González M, Tong TY, Forouhi NG, Khandelwal S, Prabhakaran D, Mozaffarian D, *et al.* (2014). Definitions and potential health benefits of the Mediterranean diet: views from experts around the world. *BMC Medicine* **12**:112.

Trichopoulou A, Naska A and Costacou T (2002). Disparities in food habits across Europe. *The Proceedings of the Nutrition Society* **61**:553–558.

Trichopoulou A, Orfanos P, Norat T, Bueno-de-Mesquita B, Ocke MC, Peeters PH, van der Schouw YT, *et al.* (2005). Modified Mediterranean diet and survival: EPIC-elderly prospective cohort study. *The British Medical Journal* **330**:991.

Tucker KL (2010). Dietary patterns, approaches, and multicultural perspective. *Applied Physiology, Nutrition, and Metabolism* **35**:211–218.

Uemura K and Pisa Z (1988). Trends in cardiovascular disease mortality in industrialized countries since 1950. *World Health Statistics Quarterly* **41**:155–178.

Ungar N, Sieverding M and Stadnitski T (2013). Increasing fruit and vegetable intake. ‘Five a day’ versus ‘just one more’. *Appetite* **65**:200–204.

USDA - U.S. Department of Agriculture (2014). A series of systematic reviews on the relationship between dietary patterns and health outcomes. pp. 1–501. Available at: http://www.nel.gov/vault/2440/web/files/DietaryPatterns/DPRptFullFinal.pdf

USDA - U.S. Department of Agriculture and U.S. Department of Health and Human Services (2010). *Dietary Guidelines for Americans 2010*. 7th Edition, Washington, DC: Government Printing Office.

USDA - U.S. Department of Agriculture (2015). *Scientific Report of the 2015 Dietary Guidelines Advisory Committee*. Available at: http://health.gov/dietaryguidelines/2015-scientific-report/pdfs/scientific-report-of-the-2015-dietary-guidelines-advisory-committee.pdf

Vaask S, Pomerleau J, Pudule I, Grinberga D, Abaravicius A, Robertson A and McKee M (2004). Comparison of the Micro-Nutrica Nutritional Analysis program and the Russian Food Composition Database using data from the Baltic Nutrition Surveys. *European Journal of Clinical Nutrition* **58**:573–579.

Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, Poole C, *et al.* (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology* **18**:805–835.

Vivekananthan DP, Penn MS, Sapp SK, Hsu A and Topol EJ (2003). Use of antioxidant vitamins for the prevention of cardiovascular disease: meta-analysis of randomised trials. *Lancet* **361**:2017–2023.

Voutilainen S, Nurmi T, Mursu J and Rissanen TH (2006). Carotenoids and cardiovascular health. *The American Journal of Clinical Nutrition* **83**:1265–1271.

Wahrendorf J (1987). An estimate of the proportion of colo-rectal and stomach cancers which might be prevented by certain changes in dietary habits. *International Journal of Cancer* **40**:625–628.

Waijers PM, Feskens EJ and Ocke MC (2007). A critical review of predefined diet quality scores. *The British Journal of Nutrition* **97**:219–231.

Wang X, Ouyang Y, Liu J, Zhu M, Zhao G, Bao W and Hu FB (2014). Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies. *The British Medical Journal* **349**:g4490.

Wardle J, Steptoe A, Bellisle F, Davou B, Reschke K, Lappalainen R and Fredrikson M (1997). Healthy dietary practices among European students. *Health Psychology* **16**:443–450.

Waskiewicz A, Piotrowski W, Sygnowska E, Rywik S and Jasiñski B (2006). Did favourable trends in food consumption observed in the 1984-2001 period contribute to the decrease in cardiovascular mortality? Pol-MONICA Warsaw Project. *Kardiologia Polska* **64**:16–23.

WCRF - World Cancer Research Fund (2015). *NOURISHING Framework.* Available at: http://www.wcrf.org/int/policy/nourishing-framework

WCRF - World Cancer Research Fund and American Institute for Cancer Research (2007). Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective. Washington DC.

White IR, Royston P and Wood AM (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine* **30**:377–399.

WHO (2003a). *Diet, Nutrition and the Prevention of Chronic Diseases*. Geneva: WHO Press.

WHO (1990). *Diet, Nutrition and the Prevention of Chronic Diseases*. Geneva: WHO Press.

WHO (2003b). *Fruit and Vegetable Promotion Initiative: A Meeting Report (25-27/08/2003).*

WHO (2015a). *Global Database on the Implementation of Nutrition Action (GINA).* Available at: https://extranet.who.int/nutrition/gina/en.

WHO (2013). Global Health Observatory. Available at: http://www.who.int/gho/en/

WHO (2009). *Global Health Risks: Mortality and Burden of Disease Attributable to Selected Major Risks*. Geneva.

WHO (2015b). *International Classification of Diseases (ICD)*. Available at: http://www.who.int/classifications/icd/en/

WHO (2015c). *Obesity and Overweight (Key Facts)*. Available at: http://www.who.int/mediacentre/factsheets/fs311/en/

WHO and FAO (2005). *Fruit and Vegetables for Health: Report of a Joint FAO/WHO Workshop 1-3 September 2004, Kobe, Japan*.

WHO Regional Office for Europe (2014). *European Health for All Database*. Available at: http://data.euro.who.int/hfadb/

WHO Regional Office for Europe (2004). *Food and Health in Europe: A New Basis for Action*. Copenhagen.

WHO Regional Office for Europe (2003). *Food Based Dietary Guidelines in the WHO European Region*. Copenhagen.

Willett W (2013a). Correction for the Effects of Measurement Error. In: Willett W (ed.) *Nutritional Epidemiology*. 3rd edn. Oxford University Press, pp. 287–304.

Willett W (2013b). Food frequency methods. In: Willett W (ed.) *Nutritional Epidemiology*. 3rd edn. Oxford University Press, pp. 70–95.

Willett W (2013c). Implications of Total Energy Intake for Epidemiological Analyses. In: Willett W (ed.) *Nutritional Epidemiology*. 3rd edn. Oxford University Press, pp. 260–286.

Willett W (2013d). Issues in analysis and presentation of dietary data. In: Willett W (ed.) *Nutritional Epidemiology*. Third edit. Oxford: Oxford University Press, pp. 305–334.

Willett W (2013e). Policy Applications. In: Willett W (ed.) *Nutritional Epidemiology*. 3rd edn. Oxford University Press, pp. 357–379.

Willett W and Lenart E (2013). Reproducibility and Validity of Food Frequency Questionnaires. In: Willett W (ed.) *Nutritional Epidemiology*. 3rd edn. Oxford University Press, pp. 96–141.

Willett W, Howe GR and Kushi LH (1997). Adjustment for total energy intake in epidemiologic studies. *The American Journal of Clinical Nutrition* **65**:1220S–1228S

Willett W, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, Hennekens CH, *et al.* (1985). Reproducibility and validity of a semiquantitative food frequency questionnaire. *American Journal of Epidemiology* **122**:51–65.

Winkler G, Brasche S, Doring A and Heinrich J (1998). Dietary intake of middle-aged men from an East and a West German city after the German reunification: do differences still exist? *European Journal of Clinical Nutrition* **52**:98–103.

Winkler G, Holtz H and Döring A (1992). Comparison of food intakes of selected populations in former East and West Germany: results from the MONICA Projects Erfurt and Augsburg. *Annals of Nutrition & Metabolism* **36**:219–234.

Woodside JV, Young IS, Gilchrist SECM, Vioque J, Chakravarthy U, De Jong PTVM, Rahu M, *et al.* (2013). Factors associated with serum/plasma concentrations of vitamins A, C, e and carotenoids in older people throughout Europe: The EUREYE study. *European Journal of Nutrition* **52**:1493–1501.

Ye EQ, Chacko SA, Chou EL, Kugizaki M and Liu S (2012). Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease, and weight gain. *The Journal of Nutrition* **142**:1304–1313.

Zaridze D, Brennan P, Boreham J, Boroda A, Karpov R, Lazarev A, Konobeevskaya I, *et al.* (2009). Alcohol and cause-specific mortality in Russia: a retrospective case-control study of 48,557 adult deaths. *Lancet* **373**:2201–2214.

Zatonski W, Campos H and Willett W. (2008). Rapid declines in coronary heart disease mortality in Eastern Europe are associated with increased consumption of oils rich in alpha-linolenic acid. *European Journal of Epidemiology* **23**:3–10.

Zatonski WA (2011). Epidemiological analysis of health situation development in Europe and its causes until 1990. *Annals of Agricultural and Environmental Medicine* **18**:194–202.

Zatonski WA, McMichael AJ and Powles JW (1998). Ecological study of reasons for sharp decline in mortality from ischaemic heart disease in Poland since 1991. *The British Medical Journal* **316**:1047–1051.

Zhang X, Shu XO, Xiang YB, Yang G, Li H, Gao J, Cai H, *et al.* (2011). Cruciferous vegetable consumption is associated with a reduced risk of total and cardiovascular disease mortality. *The American Journal of Clinical Nutrition* **94**:240–246.

Zhao D, Qi Y, Zheng Z, Wang Y, Zhang XY, Li HJ, Liu HH, *et al.* (2011). Dietary factors associated with hypertension. *Nature Reviews. Cardiology* **8**:456–465.

Zibart, E. (2001). Russia, Poland and Eastern Europe. In: *The Ethnic Food Lover`s Companion*. Birmingham, U.S.: Menasha Ridge Press, pp. 105–121.

APPENDICES

APPENDIX I.

Systematic literature review of cross-national studies: additional information

Table I-1: Study quality assessment using the STROBE checklist

STROBE criteria 1 – criterion met 0 – criterion not met	**Reviewed studies**										
	Kromhout 1989	Winkler 1992	Kardinaal 1993	Schroll 1996	Wardle 1997	Kristenson 1997	Bobak 1998	Bobak 1999	Karamanos 2002	Serra-Majem 2003	Prattala 2007
1. Title and abstract	0	0	1	1	0	1	1	1	1	1	1
Introduction											
2. Background/rationale	1	1	1	1	1	1	1	1	1	1	1
3. Objectives	0	0	1	1	1	1	0	1	1	1	1
Methods											
4. Study design	1	1	1	1	1	1	1	1	1	1	1
5. Setting	1	1	1	1	1	1	1	1	1	1	1
6. Participants	1	1	1	0	1	1	1	1	1	1	1
8. Data sources/measurement	1	1	1	1	1	1	1	1	1	1	1
9. Bias	0	1	1	0	0	0	1	1	1	1	1
10. Study size	0	0	0	0	0	0	0	1	0	0	0
12. Statistical methods	0	1	1	1	0	1	0	1	1	0	1
Results											
13. Participants	1	1	1	1	1	1	1	1	1	1	1
14. Descriptive data	1	1	1	0	1	1	1	1	1	1	1
17. Other analyses	0	0	1	1	1	0	1	0	0	0	1
Discussion											
18. Key results	0	0	1	1	1	1	1	1	0	1	1
19. Limitations	0	0	0	1	1	0	0	1	0	0	0
20. Interpretation	1	1	1	1	1	1	1	1	1	1	1
21. Generalisability	0	0	1	1	1	1	1	1	1	1	1
Other information											
22. Funding	1	1	1	0	0	1	1	1	1	1	1
TOTAL SCORE	**9**	**11**	**16**	**13**	**13**	**14**	**14**	**17**	**14**	**14**	**16**

STROBE criteria 1 – criterion met 0 – criterion not met	**Reviewed studies**									
	Miere 2007	Hall 2009	Petkeviciene 2009	Prattala 2009	Lixandru 2010	Palaanen 2011	Crispim 2011	El Ansari 2012	Woodside 2013	Burisch 2014
1. Title and abstract	1	1	1	1	1	1	1	1	1	1
Introduction										
2. Background/rationale	1	1	1	1	1	1	1	1	1	1
3. Objectives	1	0	1	1	1	1	1	1	1	1
Methods										
4. Study design	0	1	0	1	0	1	0	1	1	1
5. Setting	0	1	1	1	0	1	1	1	0	1
6. Participants	0	1	1	1	1	1	1	1	1	1
8. Data sources/measurement	1	1	1	1	1	1	1	1	1	1
9. Bias	0	0	1	0	0	1	1	0	0	0
10. Study size	0	0	0	0	0	0	0	0	0	0
12. Statistical methods	1	1	1	1	1	1	1	1	1	1
Results										
13. Participants	0	1	1	1	0	1	1	1	1	1
14. Descriptive data	1	1	0	0	1	1	1	1	1	1
17. Other analyses	0	1	1	1	1	1	1	1	1	1
Discussion										
18. Key results	1	1	1	1	1	1	1	0	1	1
19. Limitations	0	1	1	1	0	0	1	1	1	1
20. Interpretation	1	1	1	1	1	1	1	1	1	1
21. Generalisability	0	1	1	1	1	1	1	1	1	1
Other information										
22. Funding	0	1	1	1	1	1	1	0	1	1
TOTAL SCORE	**8**	**15**	**15**	**15**	**12**	**16**	**16**	**14**	**15**	**16**

<u>Table I-2:</u> Grouping of Central and Eastern European (CEE)/ Former Soviet Union (FSU) and Western European (WE) countries

Region	Sub-region	Countries
CEE/FSU	North	Armenia, Azerbaijan, Belarus, Czech Republic, Estonia, Georgia, Hungary, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Poland, Republic of Moldova, Romania, Russian Federation, Slovakia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan
	South	Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Montenegro, Serbia, Slovenia, TFYR Macedonia
WE	North	Austria, Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Liechtenstein, Luxembourg, Netherlands, Norway, Sweden, Switzerland, United Kingdom
	South	Andorra, Greece, Italy, Portugal, San Marino, Spain

APPENDIX II.

Food frequency questionnaires used in the HAPIEE and Whitehall II studies

1. HAPIEE FFQ:

1. We would like to ask you to estimate your average food use. Please cross the appropriate square in each row of the tables below a number indicating how often, on average, you have eaten the specified amount during the last 3 months.

	Amount	6+ per day	4-5 per day	2-3 per day	1 per day	5-6 per week	2-4 per week	1 per week	1-3 per month	Never or less than 1 per month
Bread and cereals										
F1 White bread, rolls	Medium slice, 1 roll	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F2 Dark bread, rolls	Medium slice, 1 roll	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F3 Cereals	Medium bowl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Potatoes, rice, pasta, dumplings										
F4 Potatoes boiled or mashed	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F5 Potatoes fried (chips) or roasted	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F6 Rice	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F7 Pasta (spaghetti, noodles)	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F8 Pizza	Medium slice	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F9 Roll-dumplings	4 slices	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F10 Potato-dumplings	4 slices	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F11 Groats	Medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F15 Pirog with meat (Ru)										
F16 Pirog with vegetables										
F17 Sweet pirog										
Dairy products and fats										
F20 Cream, sour cream	50 ml	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F21 White yoghurt	1 carton (100-150 ml)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F22 Fruit yoghurt	1 carton (100-150ml)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F23 Milk desserts	1 carton (100-150 ml)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F24 Soft cottage cheese	Medium serving (about 30 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F25 Hard cottage cheese	Medium serving (about 30 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F26 Low fat soft cheese	Medium serving (about 30 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F27 High fat soft cheese	Medium serving (about 30g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F28 Hard cheese, processed cheese – low fat	Medium serving (about 30 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F29 Hard cheese, processed	Medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9

	Amount	*6+ per day*	*4-5 per day*	*2-3 per day*	*1 per day*	*5-6 per week*	*2-4 per week*	*1 per week*	*1-3 per month*	*Never or less than 1 per month*
cheese – high fat	(about 30 g)									
F39 cottage cheese (ru)										
F40 cottage cheese										
F41 brindza (Ru)										
F30 Eggs	1 egg	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F31Margarine (on bread)	1 teaspoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F32Margarine (in food)	1 teaspoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F33 Butter (on bread)	1 teaspoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F34Butter (in food)	1 teaspoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Mixture of margarine and butter (on bread)	1 teaspoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Mixture of margarine and butter (in food)	1 teaspoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F35 Vegetable oil	1 tablespoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F36 Lard (on bread)	1 teaspoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F37 Lard (in food)	1 teaspoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F38 Mayonnaise	1 tablespoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Soups, sauces and spreads										
F50 Borsch, shiee, vegetable soup	Medium serving (about 250 ml)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F51 Bouillon	Medium serving (about 250 ml)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Beetroot soup, white borsch	Medium serving (about 250 ml)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Cabbage soup	Medium serving (about 250 ml)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F52 Other soups	Medium serving (about 250 ml)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F53 Ketchup	1 tablespoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F54 Sauces with meat, pasta, groats (such as gravy or white sauces)	Medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F55 Marmalade, jam, honey	1 teaspoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Sweets and snacks										
F60 Biscuits	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F61 Cakes, pies (sweet)	medium slice	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F62 Buns, pastries, doughnuts, muffins	1 piece	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F63 Sweets	1 bonbon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F64 Chocolate	1 bar	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F65 Ice cream	one scoop	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9

	Amount	6+ per day	4-5 per day	2-3 per day	1 per day	5-6 per week	2-4 per week	1 per week	1-3 per month	Never or less than 1 per month
F66 Milk pudding	medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F67 Sweet rice	medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F68 Pancakes	1 pancake	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F69 Sweet (fruit) dumplings	4 pieces	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F70 Crisps, crackers and other packet-snacks	1 small packet (25 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F71 Peanuts and other nuts	1 small packet (50 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F72 Sugar into coffee, tea	1 teaspoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F73 Sweetener into coffee, tea	1 capsule, 1 tablet	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Drinks										
F80 Milk	2 dl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F 94Cocoa (not Cz)	2 dl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F83 Fruit juice	2 dl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F81 Fizzy drinks (lemonade, coke, fanta)	2 dl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F82 Diet/low calorie fizzy drinks	2 dl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F84 Squash	one tablespoon	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F87 Coffee	2 dl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F88 Tea	2 dl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F89 Wine	1 dl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F90 Beer	0.25 l	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F91 Port, sherry, vermouth	1 dl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F92 Liqueurs	0.5 dl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F93 Spirits	0.25 dl	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F100 total amount of drinks										
Meat and fish										
F101 Beef : roast, steak, mince, stew or casserole	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F102 Lamb: roast, chops or stew	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F103 Pork: roast, chops or stew	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F104 Poultry	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F105 Rabbit	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F106 Offals (heart, kidney, liver)	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F107 Soft sausages	Medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9

	Amount	*6+ per day*	*4-5 per day*	*2-3 per day*	*1 per day*	*5-6 per week*	*2-4 per week*	*1 per week*	*1-3 per month*	*Never or less than 1 per month*
	(about 100 g)									
F108Hard sausages	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F109 Soft salami	50 g	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F110 Hard salami	50 g	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F111 Ham	about 50 g	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F112 Bacon	2 slices	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F113 Pate	50 g	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Meat pie	Medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F114Luncheon meat	50 g	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F115Canned meat	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F122 ? Meat ravioli (ru)	Serving (10 pieces)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F121 fat - lard from bacon (Ru)										
F125 Polish – meat PIROGI										
Fish – fresh, frozen or canned (not in oil)										
F117 Fresh water fish (e.g. carp, pike)	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F118 Salt water white fish (e.g. cod of haddock)	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F119 Oily fish (e.g. mackerel, tuna, salmon, sardines, herring, kippers)	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Other fish										
F116Fish canned in oil	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F123 (ru) Fish fingers, fish Afilé	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F124 Salted fish (Ru)	25 g	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F120 Crab, prawns, mussels (sea food)	Medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Fresh fruit										
F130Apples	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F131Pears	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F132Oranges	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F135 Grapefruit	½ medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F133 Mandarins	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F134Lemons	½ medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F136 Peaches	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9

	Amount	*6+ per day*	*4-5 per day*	*2-3 per day*	*1 per day*	*5-6 per week*	*2-4 per week*	*1 per week*	*1-3 per month*	*Never or less than 1 per month*
F137 Apricots	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F138 Plums	about 100 g	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F139 Cherries	about 100 g	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F140 Strawberries	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F141 Raspberries	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F142 Red currant	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F143 Black currant	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F144 Blueberries	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F145 Gooseberry	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F146 Kiwi	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F147 Melon	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F148 Pineapple	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F149 Bananas	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F150 Grapes	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F151 **Tinned or bottled fruit**	medium serving (about 100g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F152 **Dried fruit** (e.g. raisins, apricots, apples)	medium serving (about 50g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
Vegetables										
F160 Green salad (lettuce)	Medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F161 Spinach	Medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F189 Brussels sprouts	5 sprouts	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F162 Cabbage	Medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F163 Beans	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F164 Lentils	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F165 Dried peas	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F166 Green beans	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F167 Green peas	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F170 Turnips, swedes, parsnips	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9

	Amount	*6+ per day*	*4-5 per day*	*2-3 per day*	*1 per day*	*5-6 per week*	*2-4 per week*	*1 per week*	*1-3 per month*	*Never or less than 1 per month*
F171 Radish	4 radishes	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F173 Celeriac	50 g	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F174 Parsley ? F190 in Ru	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F168 Cauliflower	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F169 Broccoli	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F172 Carrots	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F175 Onion	½ medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F176 Leeks	½ medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F177Garlic	1 clove	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F181 Peppers	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F182 Tomatoes	1 medium	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F178 Cucumbers	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F179Aubergine	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F180 Courgette/marrow	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F183 Corn	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F191 Beet-root cooked Russian salad (RU)	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F184 Sauerkraut	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F185 Pickled vegetables, gherkins	Medium serving (about 50 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F186 Mushrooms	Medium serving	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F188 Soya meat	Medium serving (about 100g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9
F 187 Mixed frozen vegetables	Medium serving (about 100 g)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐8	☐9

2. Whitehall II FFQ:

1 Please estimate your average food use as best you can, and please answer every question. **DO NOT LEAVE ANY LINES BLANK.**

FOODS AND AMOUNTS ***AVERAGE USE IN LAST 12 MONTHS***

	MEAT AND FISH (include meat, fish & poultry eaten in sandwiches)	**Amount**	Never or less than once/mth	1 - 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 - 3 per day	4 – 5 per day	6 + per day
MBEEF	Beef: roast, steak, mince, stew or casserole	Medium serving									
	Beefburgers **MBEEFBUR**	One medium burger									
MPORK	Pork: roast, chops or stew	Medium serving									
MLAMB	Lamb: roast, chops or stew	Medium serving									
MCHICK	Chicken or other poultry	Medium serving									
	Bacon **MBACON**	Two rashers									
	Ham **MHAM**	One medium thick slice									
	Corned beef, Spam or luncheon meats **MCORNBF**	One medium thick slice									
	Sausages **MSAUSAG**	Two medium									
	Savoury pies, eg meat pie, pork pie, pasties, steak & kidney pie **MSAVPIES**	One individual pie									
	Liver, liver pate, liver sausage **MLIVER**	Medium serving									
	Fried fish in batter, as in fish and chips **MBATFISH**	One medium fillet									
	Fish fingers or fish cakes **MFISHFIN**	Two pieces									
	Other white fish, fresh or frozen, eg cod, haddock plaice, sole, halibut **MWHIFISH**	One medium fillet or serving									
	Oily fish, fresh or canned, eg mackerel, kippers, tuna, salmon, sardines, herring **MOILFISH**	One medium fillet									
	Shellfish, eg crab, pawns, mussels **MSHEFISH**	Medium serving									
		Amount	Never or less than once/mth	1 - 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day

Please answer every question. **DO NOT LEAVE ANY LINES BLANK.**

FOODS AND AMOUNTS		***AVERAGE USE IN LAST 12 MONTHS***								
BREAD & SAVOURY BISCUITS (include bread eaten in sandwiches) MWHIBRD	**Amount**	Never or less than once/mth	1 - 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day
White bread and rolls	One slice or roll	☐	☐	☐	☐	☐	☐	☐	☐	☐
MBROBRD Brown bread and rolls	One slice or roll	☐	☐	☐	☐	☐	☐	☐	☐	☐
MWHOLBRD Wholemeal bread and rolls	One slice or roll	☐	☐	☐	☐	☐	☐	☐	☐	
MCRACKER Cream crackers, cheese bisc.	One biscuit	☐	☐	☐	☐	☐	☐	☐	☐	
Crispbread, eg Ryvita MCRISBRD	One slice	☐	☐	☐	☐	☐	☐	☐	☐	☐

CEREALS

Do you eat cereals? **MCEREALS** Yes ☐ No ☐

If no, please go to 'POTATOES, RICE AND PASTA'

If yes, please indicate which brand(s) (eg Kellogg's) and type(s) (eg. Corn Flakes) and the amount used in the last 12 months.

	Amount (one medium cereal bowl)								
	Never or less than once/mth	1-3 per mth	Once a week	2-4 per week	5-6 per week	Once a day	2-3 per day	4-5 per day	6+ per day
1. Brand: MCEREAL1 / Type:	☐	☐	☐	☐	☐	☐	☐	☐	☐
2. Brand: MCEREAL2 / Type:	☐	☐							

POTATOES, RICE AND PASTA

Boiled, mashed, instant or jacket potatoes **MBOILPOT**	One medium potato/serving	☐	☐	☐	☐	☐	☐	☐	☐	☐
Chips or french fries **MCHIPS**	Medium serving	☐	☐	☐	☐	☐	☐	☐	☐	☐
Roast potatoes **MROASPOT**	One medium potato	☐	☐	☐	☐	☐	☐	☐	☐	☐
Potato salad **MPOTSALD**	Half cup	☐	☐	☐	☐	☐	☐	☐	☐	☐
White rice **MWRICE**	Half cup cooked	☐	☐	☐	☐	☐	☐	☐	☐	☐
Brown rice **MBRICE**	Half cup cooked	☐	☐	☐	☐	☐	☐	☐	☐	☐
	Amount	Never or less than once/mth	1 - 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day

Please answer every question. **DO NOT LEAVE ANY LINES BLANK.**

FOODS AND AMOUNTS ***AVERAGE USE IN LAST 12 MONTHS***

POTATOES, RICE & PASTA (Continued)	**Amount**	Never or less than once/mth	1 – 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day
Wholemeal pasta **MPASTA**	One cup cooked	☐	☐	☐	☐	☐	☐	☐	☐	☐
White or green pasta eg spaghetti, macaroni, noodles	One cup cooked **MWHPASTA**				☐	☐	☐	☐	☐	☐
Lasagne **MLASAGNE**	3" x 3" square	☐	☐	☐	☐	☐	☐	☐	☐	☐
Pizza **MPIZZA**	One medium slice	☐	☐	☐	☐	☐	☐	☐	☐	☐
DAIRY PRODUCTS & FATS										
Single cream **MSCREAM**	Tablespoon	☐	☐	☐	☐	☐	☐	☐	☐	☐
Double or clotted cream	Tablespoon **MDCREAM**				☐	☐	☐	☐	☐	☐
Yoghurt **MYOGHURT**	5 oz. carton	☐	☐	☐	☐	☐	☐	☐	☐	☐
Cheese, eg Cheddar, Brie, Edam **MCHEESE**	1 oz/30 g piece (matchbox size)	☐	☐	☐	☐	☐	☐	☐	☐	☐
Cottage Cheese, low fat soft cheese **MCOMCHE**	One tablespoon	☐	☐	☐	☐	☐	☐	☐	☐	☐
Eggs as boiled, fried, scrambled, etc **MEGGS**	One	☐	☐	☐	☐	☐	☐	☐	☐	☐
Quiche **MQUICHE**	Medium slice	☐	☐	☐	☐	☐	☐	☐	☐	☐
Salad cream, mayonnaise **MMAYO**	Tablespoon	☐	☐	☐	☐	☐	☐	☐	☐	☐
French dressing/vinaigrette **MVINAIGR**	Tablespoon	☐	☐	☐	☐	☐	☐	☐	☐	☐
The following on bread, vegetables, sandwiches etc:										
Butter **MBUTTER**	Teaspoon	☐	☐	☐	☐	☐	☐	☐	☐	☐
Hard margarine in wrapper eg Stork, Krona **MHARDMAR**	Teaspoon	☐	☐	☐	☐	☐	☐	☐	☐	☐
Polyunsaturated margarine, eg Flora, sunflower **MPOLYUNS**	Teaspoon	☐	☐	☐	☐	☐	☐	☐	☐	☐
Other soft margarine in tub eg Blue Band, Stork S.B.	Teaspoon **MSOFTMAR**				☐	☐	☐	☐	☐	☐
Low fat spread, et Outline, Gold	Teaspoon **MLFSPREA**				☐	☐	☐	☐	☐	☐
	Amount	Never or less than once/mth	1 – 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day

Please answer every question. **DO NOT LEAVE ANY LINES BLANK.**

FOODS AND AMOUNTS ***AVERAGE USE IN LAST 12 MONTHS***

SWEETS & SNACKS	Amount	Never or less than once/mth	1 – 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day
Sweet biscuits, eg Nice, digestive chocolate **MBISCUIT**	One									
Cakes **MCAKES**	Medium slice									
Buns & pastries **MBUNS**	One									
Fruit pies, tarts, crumbles	Medium slice/serving **MTARTS**									
Milk puddings, eg rice, semolina, tapioca	Medium serving **MMILKPUD**									
Sponge puddings **MSPONGE**	Medium serving									
Ice cream, choc ices	One scoop **MICECREA**									
Chocolate, chocolate bars, eg Mars, Crunchy	One bar/ four chocolates **MCHOC**									
Sweets, toffees, mints **MSWEETS**	One									
Sugar added to tea, coffee, cereal **MSUGAR**	Teaspoon									
Crisps or other packet snacks, eg Wotsits **MCRISPS**	1small (25g) packet									
Peanuts or other nuts **MNUTS**	10 whole									
SOUPS, SAUCES AND SPREADS										
Vegetable soups **MVEGSOUP**	Medium soup bowl									
Meat soups **MMEATSOU**	Medium soup bowl									
Sauces, eg white sauce, cheese sauce, gravy **MSAUCE**	Tablespoon									
Tomato ketchup **MKETCHUP**	Tablespoon									
Pickles, chutney **MPICKLES**	Tablespoon									
Marmite, Bovril **MMARMITE**	Teaspoon									
Jam, marmalade, honey **MJAM**	Teaspoon									
Peanut butter **MPEANUTB**	Teaspoon									
	Amount	Never or less than once/mth	1 – 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day

Please answer every question. **DO NOT LEAVE ANY LINES BLANK.**

FOODS AND AMOUNTS		***AVERAGE USE IN LAST 12 MONTHS***								
DRINKS 'Average glass' means 200 ml/7oz 'Wine glass' means 125 ml/ 4.5 oz	**Amount**	Never or less than once/mth	1 – 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day
Tea **MTEA**	Cup	☐	☐	☐	☐	☐	☐	☐	☐	☐
Coffee, regular **MCOFFEE**	Cup	☐	☐	☐	☐	☐	☐	☐	☐	☐
Coffee, decaffeinated **MDECAFF**	Cup	☐	☐	☐	☐	☐	☐	☐		☐
Coffee whitener eg. Coffee-mate **MCOFFWH**	Teaspoon	☐	☐	☐	☐	☐	☐	☐	☐	☐
Cocoa, hot chocolate **MCOCOA**	Cup	☐	☐	☐	☐	☐	☐	☐	☐	☐
Horlicks, Ovaltine **MHORLI**	Cup	☐	☐	☐	☐	☐	☐	☐	☐	☐
Wine **MWINE**	Wine glass	☐	☐	☐	☐	☐	☐	☐	☐	☐
Beer, lager or cider **MBEER**	Half pint	☐	☐	☐	☐	☐	☐	☐	☐	☐
Port, sherry or vermouth **MPORT**	Measure (50 ml)	☐	☐	☐	☐	☐	☐	☐	☐	☐
Liqueurs eg Baileys **MLIQU**	Measure (50 ml)	☐	☐	☐	☐	☐	☐	☐	☐	☐
Spirits, eg gin, brandy, whisky, vodka **MSPIRITS**	Single (25ml)	☐	☐	☐	☐	☐	☐	☐	☐	☐
Fizzy soft drinks, eg Coca Cola, lemonade **MFIZZY**	Average glass	☐	☐	☐	☐	☐	☐	☐	☐	☐
Low calorie or diet fizzy soft drinks **MLOWCAL**	Average glass	☐	☐	☐	☐	☐	☐	☐	☐	☐
Real fruit juice (100%) eg orange, apple juice **MFJUICE**	Average glass	☐	☐	☐	☐	☐	☐	☐	☐	☐
Fruit squash or cordial **MSQUASH**	Average glass	☐	☐	☐	☐	☐	☐	☐	☐	☐
FRUIT For very seasonal fruits such as strawberries, please estimate your average use when the fruit is in season										
Apples **MAPPLES**	One Medium	☐	☐	☐	☐	☐	☐	☐	☐	☐
Pears **MPEARS**	One medium	☐	☐	☐	☐	☐	☐	☐	☐	☐
Oranges, satsumas, mandarins **MORANGES**	One medium	☐	☐	☐	☐	☐	☐	☐	☐	☐
Grapefruit **MGRAPEFR**	Half medium	☐	☐	☐	☐	☐	☐	☐	☐	☐
Bananas **MBANANAS**	One medium	☐	☐	☐	☐	☐	☐	☐	☐	☐
Grapes **MGRAPES**	Small bunch	☐	☐	☐	☐	☐	☐	☐	☐	☐
	Amount	Never or less than once/mth	1 – 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day

Please answer every question. **DO NOT LEAVE ANY LINES BLANK.**

FOODS AND AMOUNTS ***AVERAGE USE IN LAST 12 MONTHS***

FRUIT (Continued)	**Amount**	Never or less than once/mth	1 – 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day
Melon **MMELON**	Half medium									
Peaches, plums, apricots	One **MPEACHES**									
Strawberries, raspberries	Medium serving **MSTRAWB**									
Tinned fruit **MTNFRUIT**	Medium serving									
Dried fruit, eg raisins, prunes **MDRIEDFR**	Medium serving									
VEGETABLES – FRESH, FROZEN OR TINNED										
Carrots **MCARROTS**	One medium									
Spinach **MSPINACH**	Medium serving									
Broccoli **MBROCCOL**	Medium serving									
Spring greens, kale **MGREENS**	Medium serving									
Brussels sprouts **MSPROUTS**	Five sprouts									
Cabbage **MCABBAGE**	Quarter small									
Peas **MPEAS**	One tablespoon									
Green beans, broad beans runner beans **MBEANS**	Medium serving									
Marrow, courgettes **MMARROW**	Medium serving									
Cauliflower **MCAULIFL**	Medium serving									
Parsnips, turnips, swedes **MPARSNIP**	One medium									
Leeks **MLEEKS**	One medium									
Onion **MONIONS**	One medium									
Garlic **MGARLIC**	One Clove									
Mushrooms **MMUSHROO**	Medium serving									
Sweet peppers **MPEPPERS**	One medium									
	Amount	Never or less than once/mth	1 – 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day

Please answer every question. **DO NOT LEAVE ANY LINES BLANK.**

FOODS AND AMOUNTS ***AVERAGE USE IN LAST 12 MONTHS***

VEGETABLES (Continued)	**Amount**	Never or less than once/mth	1 – 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day
Green salad **MSALAD**	Medium serving									
Tomatoes **MTOMATO**	One medium									
Coleslaw **MCOLESL**	One tablespoon									
Baked beans **MBAKEDB**	One tablespoon									
Dried lentils, beans, peas **MLENTILS**	One tablespoon cooked									
Tofu or soya bean curd **MTOFU**	?" x 2" x 1" piece									
Soya meat, TVP, vegeburger **MTVP**	One burger									
	Amount	Never or less than once/mth	1 – 3 per mth	Once a week	2 – 4 per week	5 – 6 per week	Once a day	2 – 3 per day	4 – 5 per day	6 + per day

APPENDIX III.

Association between fruit, vegetable intake and mortality: additional analyses

<u>Table III-1:</u> Results of the Cox regression analysis on the association between fruit, vegetable intake and mortality on the pooled sample, including participants with no missing data only (n=17,858)

			Cohort-specific fruit and vegetable intake quartiles										**Per 100g/day increase[1]**	
			Q1		**Q2**		**Q3**		**Q4**					
Cause of death	**Deaths/n**	**Model**	**HR**		**HR**	**(95%CI)**	**HR**	**(95%CI)**	**HR**	**(95%CI)**	**p-value (trend)**	**PP% (95%CI)[2]**	**HR**	**(95%CI)**
All-cause	1201/17,858	model1	**1.00**	ref.	**0.81**	(0.70-0.94)	**0.77**	(0.66-0.90)	**0.66**	(0.56-0.78)	<0.001	**10.5** (6.1-15.1)	**0.90**	(0.86-0.93)
		model2	**1.00**	ref.	**0.97**	(0.83-1.13)	**0.97**	(0.82-1.14)	**0.89**	(0.74-1.07)	0.268	**2.9** (-1.6-7.7)	**0.97**	(0.93-1.02)
CVD	404/17,789	model1	**1.00**	ref.	**0.68**	(0.53-0.88)	**0.64**	(0.49-0.84)	**0.52**	(0.39-0.70)	<0.001	**16.9** (8.8-25.3)	**0.87**	(0.81-0.93)
		model2	**1.00**	ref.	**0.84**	(0.65-1.09)	**0.82**	(0.62-1.09)	**0.71**	(0.51-0.98)	0.035	**8.6** (0.5-17.6)	**0.95**	(0.88-1.02)
CHD	213/17,789	model1	**1.00**	ref.	**0.64**	(0.45-0.91)	**0.60**	(0.42-0.87)	**0.57**	(0.38-0.84)	0.002	**15.3** (4.4-27.6)	**0.87**	(0.79-0.95)
		model2	**1.00**	ref.	**0.83**	(0.58-1.19)	**0.85**	(0.58-1.26)	**0.87**	(0.57-1.35)	0.481	**3.7** (-7.3-15.8)	**0.98**	(0.89-1.09)
Stroke	106/17,789	model1	**1.00**	ref.	**0.63**	(0.38-1.05)	**0.66**	(0.40-1.10)	**0.51**	(0.29-0.91)	0.021	**17.5** (2.2-34.3)	**0.88**	(0.77-1.00)
		model2	**1.00**	ref.	**0.69**	(0.41-1.16)	**0.70**	(0.41-1.20)	**0.55**	(0.29-1.03)	0.065	**15.3** (-0.7-33.6)	**0.91**	(0.78-1.05)

Model 1: adjusted for sex, age, cohort

Model 2: adjusted for sex, age, cohort, alcohol intake, smoking, education, household amenities score, marital status, energy intake, physical activity, vitamin supplement Intake, HDI (without F&V component)

[1] per one unit increase across six intake categories (<100g/d, 1-200g/d, 2-300g/d, 3-400g/d, 4-500g/d, >500g/d)

[2] Preventable proportion of death if participants in the lowest three quartiles increased their intake one quartile upward

Table III-2: Relationship between fruit, vegetable intake and mortality in the pooled sample if subjects who died in the first two years of follow up were excluded from the analysis (n=19,047)

			Cohort-specific fruit and vegetable intake quartiles										Per 100g/day increase[1]	
			Q1		Q2		Q3		Q4					
Cause of death	**Deaths/n**	**Model**	**HR**		**HR**	**(95%CI)**	**HR**	**(95%CI)**	**HR**	**(95%CI)**	**p-value (trend)**	**PP% (95%CI)[2]**	**HR**	**(95%CI)**
All-cause	1028/19,047	model1	**1.00**	ref.	**0.81**	(0.68-0.95)	**0.79**	(0.67-0.93)	**0.74**	(0.62-0.88)	0.001	**7.8** (3.2-12.8)	**0.92**	(0.89-0.96)
		model2	**1.00**	ref.	**0.95**	(0.81-1.13)	**0.96**	(0.81-1.15)	**0.95**	(0.78-1.15)	0.600	**1.3** (-3.4-6.5)	**0.99**	(0.94-1.04)
CVD	354/18,998	model1	**1.00**	ref.	**0.69**	(0.53-0.92)	**0.72**	(0.55-0.95)	**0.58**	(0.43-0.79)	0.001	**14.0** (5.7-22.7)	**0.90**	(0.83-0.96)
		model2	**1.00**	ref.	**0.85**	(0.64-1.13)	**0.90**	(0.67-1.21)	**0.78**	(0.55-1.10)	0.205	**6.2** (-2.3-15.7)	**0.98**	(0.90-1.06)
CHD	179/18,998	model1	**1.00**	ref.	**0.63**	(0.42-0.94)	**0.74**	(0.50-1.09)	**0.64**	(0.42-0.98)	0.051	**12.0** (0.5-24.8)	**0.91**	(0.82-1.01)
		model2	**1.00**	ref.	**0.79**	(0.53-1.20)	**1.01**	(0.66-1.52)	**0.96**	(0.60-1.54)	0.983	**1.1** (-10.3-14.3)	**1.02**	(0.91-1.14)
Stroke	88/18,998	model1	**1.00**	ref.	**0.59**	(0.33-1.04)	**0.75**	(0.44-1.28)	**0.53**	(0.28-1.00)	0.072	**16.4** (0.0-35.1)	**0.92**	(0.80-1.07)
		model2	**1.00**	ref.	**0.65**	(0.36-1.16)	**0.79**	(0.44-1.39)	**0.56**	(0.30-1.12)	0.145	**14.7** (-2.6-33.3)	**0.96**	(0.81-1.13)

Model 1: adjusted for sex, age, cohort

Model 2: adjusted for sex, age, cohort, alcohol intake, smoking, education, household amenities score, marital status, energy intake, physical activity, vitamin supplement Intake, HDI (without the fruit and vegetable component)

[1] Per one unit increase across six intake categories (<100g/d, 1-200g/d, 2-300g/d, 3-400g/d, 4-500g/d, >500g/d)

[2] Preventable proportion of death if participants in the lowest three quartiles increased their intake one quartile upward

Table III-3: **Relationship between fruit, vegetable intake and cause-specific mortality in the pooled sample using competing risk regression models**[1]

			Cohort-specific fruit and vegetable intake quartiles										Per 100g/day increase[2]	
			Q1		Q2		Q3		Q4					
Cause of death	**Deaths/n**	**Model**	**SHR**		**SHR**	**(95%CI)**	**SHR**	**(95%CI)**	**SHR**	**(95%CI)**	**p-value (trend)**	**PP% (95%CI)**[3]	**SHR**	**(95%CI)**
CVD	438/19,263	model1	**1.00**	ref.	**0.66**	(0.52-0.85)	**0.66**	(0.51-0.84)	**0.55**	(0.42-0.73)	<0.001	**15.7** (7.9-23.7)	**0.87**	(0.82-0.93)
		model2	**1.00**	ref.	**0.80**	(0.62-1.03)	**0.83**	(0.63-1.09)	**0.75**	(0.55-1.03)	0.083	**7.4** (-0.7-16.1)	**0.95**	(0.90-1.02)
CHD	226/19,263	model1	**1.00**	ref.	**0.62**	(0.44-0.88)	**0.61**	(0.43-0.87)	**0.61**	(0.42-0.89)	0.007	**13.7** (3.0-25.3)	**0.88**	(0.80-0.97)
		model2	**1.00**	ref.	**0.79**	(0.55-1.12)	**0.86**	(0.58-1.27)	**0.94**	(0.62-1.43)	0.697	**1.7** (-8.9-13.8)	**1.00**	(0.89-1.10)
Stroke	109/19,263	model1	**1.00**	ref.	**0.62**	(0.38-1.04)	**0.71**	(0.43-1.16)	**0.51**	(0.29-0.90)	0.028	**17.3** (2.4-33.8)	**0.89**	(0.78-1.00)
		model2	**1.00**	ref.	**0.67**	(0.40-1.13)	**0.74**	(0.45-1.24)	**0.53**	(0.28-1.01)	0.066	**16.0** (-0.2-33.8)	**0.91**	(0.78-1.05)

Model 1: adjusted for sex, age, cohort

Model 2: adjusted for sex, age, cohort, alcohol intake, smoking, education, household amenities score, marital status, energy intake, physical activity, vitamin supplement Intake, HDI (without the fruit and vegetable component)

[1] Competing risk events were cancer and non-CVD-non-cancer deaths for CVD mortality; stroke, cancer and non-CVD-non-cancer deaths for CHD mortality; and CHD, cancer and non-CVD-non-cancer deaths for stroke mortality

[2] Per one unit increase across six intake categories (<100g/d, 1-200g/d, 2-300g/d, 3-400g/d, 4-500g/d, >500g/d)

[3] Preventable proportion of death if participants in the lowest three quartiles increased their intake one quartile upward

Table III-4: Relationship between fruit, vegetable intake and mortality among smokers: adjustment for the number of cigarettes smoked and the number of years has been smoked

Cause of death	Death/n	Model	Q1 HR		Q2 HR	Q2 (95%CI)	Q3 HR	Q3 (95%CI)	Q4 HR	Q4 (95%CI)	p-value (trend)	PP% (95%CI)[2]	Per 100g/day increase[1] HR	Per 100g/day increase[1] (95%CI)
All-cause	638/5905	model 1	**1.00**	ref.	**0.90**	(0.73-1.11)	**0.87**	(0.70-1.09)	**0.70**	(0.53-0.91)	0.011	**8.8** (2.2-15.9)	**0.93**	(0.87-0.98)
		model 2	**1.00**	ref.	**0.91**	(0.74-1.11)	**0.90**	(0.72-1.12)	**0.72**	(0.55-0.94)	0.023	**7.9** (1.4-15.0)	**0.94**	(0.88-0.99)
CVD	226/5871	model 1	**1.00**	ref.	**0.75**	(0.53-1.06)	**0.78**	(0.54-1.14)	**0.62**	(0.40-0.97)	0.037	**11.9** (0.7-24.3)	**0.94**	(0.85-1.04)
		model 2	**1.00**	ref.	**0.75**	(0.53-1.06)	**0.80**	(0.55-1.16)	**0.64**	(0.41-0.99)	0.048	**11.3** (0.2-23.7)	**0.95**	(0.86-1.04)
CHD	125/5871	model 1	**1.00**	ref.	**0.72**	(0.44-1.16)	**0.82**	(0.50-1.37)	**0.76**	(0.43-1.35)	0.340	**7.3** (-7.2-24.1)	**0.98**	(0.86-1.12)
		model 2	**1.00**	ref.	**0.71**	(0.44-1.16)	**0.83**	(0.50-1.38)	**0.77**	(0.43-1.37)	0.372	**6.9** (-7.5-24.1)	**0.98**	(0.86-1.12)
Stroke	50/5871	model 1	**1.00**	ref.	**0.76**	(0.37-1.56)	**0.66**	(0.30-1.46)	**0.30**	(0.10-0.94)	0.038	**25.6** (1.2-50.8)	**0.85**	(0.68-1.06)
		model 2	**1.00**	ref.	**0.75**	(0.36-1.52)	**0.67**	(0.31-1.49)	**0.31**	(0.10-0.97)	0.044	**25.3** (0.6-50.8)	**0.86**	(0.69-1.07)

Columns Q1–Q4: Cohort-specific fruit and vegetable intake quartiles

Model 1: adjusted for sex, age, cohort, alcohol intake, education, household amenities score, marital status, energy intake, physical activity, vitamin supplement intake, HDI (without the fruit and vegetable component)

Model 2: In addition to all variables in model 1, HRs were adjusted for the number of cigarettes smoked per day and number of years the participant had smoked

[1] Per one unit increase across six intake categories (<100g/d, 1-200g/d, 2-300g/d, 3-400g/d, 4-500g/d, >500g/d)

[2] Preventable proportion of death if participants in the lowest three quartiles increased their intake one quartile upward

Table III-5: Results of Cox regression analysis by intake of fruit and vegetable subgroups

		Cohort-specific tertiles of intake						Per 30g/day increase[1]	
		T1		T2		T3			
Fruit and vegetable subgroups	Cause of death	HR		HR	(95%CI)	HR	(95%CI)	HR	(95%CI)
Citrus fruits	All-cause	**1.00**	ref.	**0.93**	(0.81-1.06)	**1.07**	(0.93-1.23)	**1.00**	(0.95-1.05)
	CVD	**1.00**	ref.	**0.99**	(0.79-1.25)	**1.10**	(0.86-1.41)	**1.00**	(0.91-1.09)
	CHD	**1.00**	ref.	**1.04**	(0.75-1.43)	**1.28**	(0.90-1.81)	**1.03**	(0.91-1.17)
	Stroke	**1.00**	ref.	**0.91**	(0.58-1.41)	**0.72**	(0.42-1.22)	**0.86**	(0.71-1.04)
Berries	All-cause	**1.00**	ref.	**0.87**	(0.71-1.08)	**1.00**	(0.87-1.14)	**1.01**	(0.92-1.10)
	CVD	**1.00**	ref.	**0.88**	(0.61-1.27)	**0.94**	(0.74-1.19)	**1.03**	(0.88-1.21)
	CHD	**1.00**	ref.	**0.78**	(0.45-1.37)	**0.90**	(0.65-1.26)	**1.13**	(0.92-1.38)
	Stroke	**1.00**	ref.	**1.14**	(0.57-2.29)	**0.96**	(0.60-1.54)	**1.10**	(0.81-1.49)
Green/leafy vegetables	All-cause	**1.00**	ref.	**0.96**	(0.85-1.10)	**0.91**	(0.78-1.05)	**0.98**	(0.91-1.05)
	CVD	**1.00**	ref.	**0.90**	(0.72-1.11)	**0.82**	(0.63-1.06)	**0.89**	(0.79-1.02)
	CHD	**1.00**	ref.	**1.01**	(0.75-1.37)	**0.98**	(0.68-1.40)	**0.95**	(0.80-1.13)
	Stroke	**1.00**	ref.	**0.68**	(0.44-1.06)	**0.56**	(0.33-0.97)	**0.70**	(0.53-0.93)
Processed fruits and vegetables	All-cause	**1.00**	ref.	**0.83**	(0.73-0.95)	**0.96**	(0.84-1.10)	**1.08**	(0.90-1.30)
	CVD	**1.00**	ref.	**0.81**	(0.64-1.03)	**0.99**	(0.78-1.25)	**1.11**	(0.81-1.53)
	CHD	**1.00**	ref.	**0.64**	(0.46-0.89)	**0.75**	(0.54-1.04)	**1.10**	(0.70-1.72)
	Stroke	**1.00**	ref.	**1.11**	(0.68-1.81)	**1.55**	(0.97-2.49)	**1.40**	(0.82-2.40)

All HRs are adjusted for sex, age, cohort, alcohol intake, education, household amenities score, marital status, energy intake, physical activity, vitamin supplement intake, HDI (without the fruit and vegetable component). Further, all fruit and vegetable subgroups were mutually adjusted for each-other.

[1] Per one unit increase across four intake categories (<30g/d, 30-60g/d, 60-90g/d, >90g/d)

APPENDIX IV.

Comparison of dietary intakes between the HAPIEE and the Whitehall II cohorts: subgroup analyses

<u>Table IV-1:</u> Average intake of foods and drinks in the British, Czech, Polish, Russian cohorts and the pooled Eastern European sample, including only those who were still employed at the time of the questionnaire in Whitehall II, and those in sedentary occupation in HAPIEE study

Food groups and subgroups (FoodEx2)	UK (n=2662)	CZE (n=1622)		POL (n=1824)		RUS (n=1332)		POOLED Czech, Polish and Russian sample (n=4778)	
	Median[1] (IQR)	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]
Fully comparable foods and drinks[3]									
Animal fresh meat / animal offals	**74.2** (49.0-102.0)	**83.8** (50.4-120.0)	<0.0001	**76.8** (60.0-111.6)	<0.0001	**120.0** (83.8-161.8)	<0.0001	**85.2** (60.0-125.6)	<0.0001
Eggs	**7.0** (3.5-21.5)	**7.0** (7.0-21.5)	1.0	**21.5** (7.0-21.5)	<0.0001	**21.5** (7.0-21.5)	<0.0001	**21.5** (7.0-21.5)	<0.0001
Fruits and fruit products	**250.4** (150.2-380.9)	**264.3** (144.8-454.1)	0.0018	**225.8** (133.7-371.0)	<0.0001	**161.2** (89.0-266.6)	<0.0001	**212.6** (120.1-366.0)	<0.0001
Fresh fruits	**229.0** (132.5-353.1)	**248.6** (130.8-425.6)	0.0063	**209.6** (121.5-345.1)	<0.0001	**128.0** (61.9-218.3)	<0.0001	**189.3** (99.7-337.3)	<0.0001
Processed fruit products	**14.2** (4.8-28.5)	**14.0** (7.0-22.3)	<0.0001	**10.5** (2.5-21.2)	<0.0001	**21.5** (7.7-48.5)	<0.0001	**14.2** (7.0-25.0)	0.0004
Vegetables *(all non-products)*[4]	**238.6** (163.8-329.1)	**178.2** (107.3-281.9)	<0.0001	**203.0** (134.5-309.9)	<0.0001	**294.7** (227.1-381.3)	<0.0001	**225.0** (139.5-328.4)	<0.0001
Starchy roots or tubers	**98.3** (75.3-149.8)	**86.8** (53.3-101.1)	<0.0001	**86.8** (75.3-141.1)	0.0262	**86.8** (47.6-145.6)	<0.0001	**86.8** (75.3-115.6)	<0.0001
Sugars, confectionery and water-based sweet dessert	**8.1** (3.5-26.0)	**9.5** (4.1-22.6)	0.5742	**19.1** (7.0-36.0)	<0.0001	**36.0** (19.5-49.3)	<0.0001	**19.6** (7.0-37.1)	<0.0001
Alcoholic beverages (portion/day)	**1.0** (0.4-2.5)	**0.4** (0.1-1.1)	<0.0001	**0.1** (0.0-0.4)	<0.0001	**0.1** (0.1-0.6)	<0.0001	**0.2** (0.1-0.7)	<0.0001
Coffee, cocoa, tea and infusions	**883.0** (513.3-1055.0)	**675.0** (390.0-975.0)	<0.0001	**690.0** (581.7-975.0)	<0.0001	**675.0** (489.0-883.0)	<0.0001	**675.0** (489.0-975.0)	<0.0001

Food groups and subgroups (FoodEx2)	UK (n=2662)	CZE (n=1622)		POL (n=1824)		RUS (n=1332)		POOLED Czech, Polish and Russian sample (n=4778)	
	Median[1] (IQR)	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]
Partially comparable foods and drinks[5]									
All meat and meat products	**90.6** (60.0-123.5)	**100.3** (66.9-137.9)	<0.0001	**108.0** (80.1-141.8)	<0.0001	**146.1** (104.1-193.3)	<0.0001	**114.6** (80.0-155.1)	<0.0001
Grains and grain based products	**188.0** (125.7-266.8)	**169.3** (117.4-237.1)	0.9410	**175.9** (127.7-255.3)	0.8830	**213.6** (135.4-294.8)	0.0341	**181.8** (125.1-262.7)	0.9594
Legumes, nuts, oilseeds, spices	**30.1** (16.1-48.5)	**12.3** (6.3-19.6)	<0.0001	**11.2** (6.3-19.6)	<0.0001	**11.2** (4.9-18.2)	<0.0001	**11.2** (6.3-18.9)	<0.0001
Animal fats and oils	**0.7** (0.0-4.3)	**1.4** (0.7-10.0)	<0.0001	**7.9** (0.0-25.0)	<0.0001	**4.3** (1.4-10.0)	<0.0001	**4.3** (0.7-10.0)	<0.0001
Seasoning, sauces, condiments	**10.8** (4.3-26.7)	**13.6** (8.6-29.5)	<0.0001	**10.8** (4.9-21.6)	0.2562	**21.8** (8.5-39.6)	<0.0001	**12.9** (7.0-30.2)	<0.0001
Fruit and vegetable juices and nectars	**86.0** (14.0-200.0)	**14.0** (0.0-28.0)	<0.0001	**86.0** (14.0-158.0)	<0.0001	**28.0** (14.0-86.0)	<0.0001	**28.0** (14.0-86.0)	<0.0001

[1] Values are g/day intakes except for alcoholic beverages where portion/day intake is shown

[2] All p-values were calculated with quantile regression using the intake values in the UK cohort as reference category, adjusted for sex, age, energy intake, smoking, education, marital status, leisure time physical activity, CVD/diabetes in medical history

[3] On average, more than 80% of their intake was provided by the common items (n=81) in all four cohorts

[4] Including: brassica vegetables; bulb, stalk and stem vegetables; fruiting vegetables; leafy vegetables; legume greens, sprouts; non-starchy root and tuber vegetables; fungi; marine algae, aromatic herbs or flowers

[5] On average, 60-80% of their intake was provided by the common items (n=81) in at least one of the cohorts, and more than 80% in the other cohorts

Table IV-2: **Average intake of nutrients in the British, Czech, Polish, Russian cohorts and the pooled Eastern European sample, including only those who were still employed at the time of the questionnaire in Whitehall II, and those in sedentary occupation in HAPIEE study**

Nutrients	UK (n=2662)	CZE (n=1622)		POL (n=1824)		RUS (n=1332)		POOLED Czech, Polish and Russian sample (n=4778)	
	Median[1] (IQR)	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]
Fully comparable nutrients[3]									
Alcohol (g/day)	**11.4** (4.9-28.4)	**4.3** (1.3-10.6)	<0.0001	**1.1** (0.0-4.3)	<0.0001	**1.7** (0.6-5.5)	<0.0001	**1.9** (0.6-6.6)	<0.0001
Beta-carotene (mg/day)	**6.0** (3.5-8.2)	**4.6** (3.4-7.0)	<0.0001	**6.8** (4.0-9.6)	<0.0001	**11.1** (7.2-13.7)	<0.0001	**6.8** (4.1-10.8)	<0.0001
Partially comparable nutrients[4]									
Total carbohydrate (g/day)	**232.5** (202.1-259.8)	**217.4** (194.0-244.4)	<0.0001	**221.7** (197.5-244.9)	<0.0001	**219.2** (194.8-241.9)	<0.0001	**219.5** (195.5-244.1)	<0.0001
Sugar (g/day)	**133.7** (92.0-136.7)	**107.0** (82.3-135.5)	<0.0001	**105.8** (84.6-129.9)	<0.0001	**111.6** (90.4-132.1)	<0.0001	**107.6** (85.4-132.2)	<0.0001
Protein (g/day)	**72.6** (63.6-81.6)	**80.0** (70.2-89.5)	<0.0001	**82.1** (73.9-91.1)	<0.0001	**82.8** (72.8-93.6)	<0.0001	**81.5** (72.5-91.3)	<0.0001
Total fat (g/day)	**67.2** (58.4-76.3)	**75.9** (67.5-83.8)	<0.0001	**79.0** (68.7-88.4)	<0.0001	**78.6** (70.0-88.2)	<0.0001	**77.7** (68.7-86.6)	<0.0001
Saturated fat (g/day)	**25.2** (21.2-30.2)	**30.9** (26.7-35.5)	<0.0001	**33.3** (27.9-39.2)	<0.0001	**30.1** (26.1-34.6)	<0.0001	**31.3** (26.8-36.5)	<0.0001
Polyunsaturated fat (g/day)	**11.6** (9.6-14.3)	**11.3** (9.7-13.3)	<0.0001	**10.7** (9.2-12.8)	<0.0001	**14.4** (11.5-18.3)	<0.0001	**11.6** (9.8-14.4)	0.0125
Cholesterol (mg/day)	**214.5** (167.6-269.1)	**303.8** (254.7-360.8)	<0.0001	**341.9** (289.9-397.9)	<0.0001	**322.1** (274.0-385.7)	<0.0001	**324.7** (271.5-383.7)	<0.0001
Non-starch polysaccharides (g/day)	**16.4** (13.8-19.6)	**15.0** (11.9-18.7)	<0.0001	**14.6** (12.1-17.8)	<0.0001	**14.2** (12.0-16.6)	<0.0001	**14.5** (12.0-17.8)	<0.0001

Nutrients	UK (n=2662)	CZE (n=1622)		POL (n=1824)		RUS (n=1332)		POOLED Czech, Polish and Russian sample (n=4778)	
	Median[1] (IQR)	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**
Vitamin C (mg/day)	**144.5** (102.4-199.5)	**133.1** (88.2-215.0)	0.0093	**126.0** (88.2-191.2)	<0.0001	**95.8** (64.0-154.3)	<0.0001	**120.9** (79.4-189.1)	<0.0001
Total energy (MJ/day)	**7.3** (6.1-8.8)	**6.6** (5.3-8.3)	<0.0001	**7.0** (5.7-8.6)	0.1683	**8.1** (6.5-10.1)	<0.0001	**7.1** (5.7-8.8)	0.8930

[1] All values are energy standardized around 8MJ/day, except for alcohol and total energy intake for which absolute intakes are shown

[2] All p-values were calculated with quantile regression using the intake values in the UK cohort as reference category, adjusted for sex, age, energy intake, smoking, education, marital status, leisure time physical activity, CVD/diabetes in medical history

[3] On average, more than 80% of their intake was provided by the common items (n=81) in all four cohorts

[4] On average, 60-80% of their intake was provided by the common items (n=81) in at least one of the cohorts, and more than 80% in the other cohorts

Table IV-3: Average intake of foods and drinks in the British, Czech, Polish, Russian cohorts and the pooled Eastern European sample, including only male participants

Food groups and subgroups (FoodEx2)	UK (n=3921)	CZE (n=3665)		POL (n=4847)		RUS (n=4149)		POOLED Czech, Polish and Russian sample (n=12,661)	
	Median[1] (IQR)	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**
Fully comparable foods and drinks[3]									
Animal fresh meat / animal offals	**75.4** (49.0-102.0)	**85.2** (50.4-120.0)	<0.0001	**83.8** (65.6-111.6)	<0.0001	**125.6** (85.2-161.8)	<0.0001	**92.2** (67.0-127.0)	<0.0001
Eggs	**7.0** (3.5-21.5)	**7.0** (7.0-21.5)	1.0	**21.5** (7.0-21.5)	<0.0001	**21.5** (7.0-39.5)	<0.0001	**21.5** (7.0-21.5)	<0.0001
Fruits and fruit products	**246.1** (148.4-367.7)	**212.4** (117.6-369.2)	0.0023	**182.1** (106.3-314.2)	<0.0001	**112.0** (60.2-192.8)	<0.0001	**162.0** (86.7-295.8)	<0.0001
Fresh fruits	**220.3** (127.4-336.8)	**194.2** (101.7-342.6)	0.0095	**163.1** (94.0-294.8)	0.0009	**74.9** (33.8-151.3)	<0.0001	**139.6** (64.3-275.1)	<0.0001
Processed fruit products	**17.7** (7.0-35.5)	**14.0** (7.0-23.5)	<0.0001	**7.7** (1.3-18.0)	<0.0001	**21.5** (7.7-48.5)	<0.0001	**14.2** (4.8-28.5)	0.0195
Vegetables *(all non-products)*[4]	**240.6** (165.3-320.9)	**160.0** (100.9-252.1)	<0.0001	**189.2** (122.2-290.0)	<0.0001	**282.7** (215.0-367.4)	<0.0001	**215.8** (132.0-314.8)	<0.0001
Starchy roots or tubers	**101.2** (78.1-152.6)	**86.8** (75.3-101.2)	<0.0001	**89.6** (75.3-151.8)	<0.0001	**98.3** (75.3-146.2)	<0.0001	**89.6** (75.3-146.2)	<0.0001
Sugars, confectionery and water-based sweet dessert	**9.5** (3.5-27.0)	**10.1** (4.5-22.6)	0.0009	**22.6** (9.5-36.6)	<0.0001	**36.0** (18.4-42.9)	<0.0001	**22.0** (8.1-37.4)	<0.0001
Alcoholic beverages (portion/day)	**1.2** (0.6-2.6)	**0.8** (0.2-1.9)	<0.0001	**0.1** (0.0-0.5)	<0.0001	**0.5** (0.1-1.2)	<0.0001	**0.4** (0.1-1.0)	<0.0001
Coffee, cocoa, tea and infusions	**883.0** (526.6-1055.0)	**526.6** (350.1-690.0)	<0.0001	**675.0** (503.0-975.0)	<0.0001	**633.0** (475.0-883.0)	<0.0001	**675.0** (475.0-941.0)	<0.0001

Food groups and subgroups (FoodEx2)	UK (n=3921)	CZE (n=3665)		POL (n=4847)		RUS (n=4149)		POOLED Czech, Polish and Russian sample (n=12,661)	
	Median[1] (IQR)	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]
Partially comparable foods and drinks[5]									
All meat and meat products	**91.8** (62.5-124.8)	**104.5** (69.7-142.1)	<0.0001	**114.3** (86.1-148.1)	<0.001	**152.2** (110.7-194.1)	<0.0001	**122.5** (85.6-163.9)	<0.0001
Grains and grain based products	**196.2** (135.0-276.0)	**169.2** (118.3-238.3)	0.5209	**207.7** (141.3-278.9)	0.0001	**251.7** (174.9-329.8)	<0.0001	**212.3** (140.2-286.1)	<0.0001
Legumes, nuts, oilseeds, spices	**33.6** (16.8-50.6)	**11.2** (6.3-18.2)	<0.0001	**11.2** (6.3-18.2)	<0.0001	**8.4** (3.5-14.7)	<0.0001	**11.2** (4.9-17.5)	<0.0001
Animal fats and oils	**0.0** (0.0-4.3)	**4.3** (0.7-10.0)	<0.0001	**4.3** (0.0-25.0)	<0.0001	**7.9** (1.4-10.0)	<0.0001	**4.3** (0.7-10.0)	<0.0001
Seasoning, sauces, condiments	**10.8** (4.3-28.1)	**13.6** (8.7-30.2)	<0.0001	**10.1** (4.3-24.3)	0.1266	**17.9** (5.7-37.4)	<0.0001	**12.9** (6.4-30.0)	<0.0001
Fruit and vegetable juices and nectars	**86.0** (14.0-200.0)	**14.0** (0.0-28.0)	<0.0001	**28.0** (0.0-86.0)	<0.0001	**14.0** (0.0-28.0)	<0.0001	**14.0** (0.0-86.0)	<0.0001

[1] Values are g/day intakes except for alcoholic beverages where portion/day intake is shown

[2] All p-values were calculated with quantile regression using the intake values in the UK cohort as reference category, adjusted for age, energy intake, smoking, education, employment status, marital status, leisure time physical activity, CVD/diabetes in medical history

[3] On average, more than 80% of their intake was provided by the common items (n=81) in all four cohorts

[4] Including: brassica vegetables; bulb, stalk and stem vegetables; fruiting vegetables; leafy vegetables; legume greens, sprouts; non-starchy root and tuber vegetables; fungi; marine algae, aromatic herbs or flowers

[5] On average, 60-80% of their intake was provided by the common items (n=81) in at least one of the cohorts, and more than 80% in the other cohorts

Table IV-4: Average intake of nutrients in the British, Czech, Polish, Russian cohorts and the pooled Eastern European sample, including only male participants

Nutrients	UK (n=3921)	CZE (n=3665)		POL (n=4847)		RUS (n=4149)		POOLED Czech, Polish and Russian sample (n=12,661)	
	Median[1] (IQR)	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]
Fully comparable nutrients[3]									
Alcohol (g/day)	**12.8** (5.7-29.6)	**7.7** (2.0-17.9)	<0.0001	**1.3** (0.0-4.9)	<0.0001	**4.3** (0.6-11.2)	<0.0001	**3.8** (0.6-10.0)	<0.0001
Beta-carotene (mg/day)	**6.2** (3.5-8.4)	**4.5** (3.3-6.5)	<0.0001	**6.8** (4.2-9.5)	<0.0001	**10.8** (6.9-13.2)	<0.0001	**6.9** (4.2-11.0)	<0.0001
Partially comparable nutrients[4]									
Total carbohydrate (g/day)	**233.1** (203.7-259.5)	**210.2** (182.7-237.0)	<0.0001	**220.4** (195.8-243.1)	<0.0001	**223.4** (197.7-246.2)	<0.0001	**218.5** (192.7-242.7)	<0.0001
Sugar (g/day)	**113.7** (92.0-135.8)	**96.9** (74.5-121.3)	<0.0001	**96.0** (77.4-117.3)	<0.0001	**97.9** (79.5-117.0)	<0.0001	**96.9** (77.4-118.2)	<0.0001
Protein (g/day)	**71.5** (63.1-80.1)	**79.4** (69.3-89.4)	<0.0001	**81.8** (73.4-90.8)	<0.0001	**82.4** (72.6-92.7)	<0.0001	**81.2** (71.8-91.1)	<0.0001
Total fat (g/day)	**66.5** (58.2-75.2)	**75.6** (66.6-84.3)	<0.0001	**78.9** (69.1-88.3)	<0.0001	**74.0** (65.9-82.5)	<0.0001	**76.3** (67.2-85.3)	<0.0001
Saturated fat (g/day)	**25.3** (21.3-30.0)	**30.9** (26.6-35.8)	<0.0001	**32.4** (27.1-38.9)	<0.0001	**28.0** (24.1-32.4)	<0.0001	**30.4** (25.7-35.8)	<0.0001
Polyunsaturated fat (g/day)	**11.3** (9.4-14.0)	**11.0** (9.4-12.9)	<0.0001	**10.7** (9.1-12.7)	<0.0001	**13.0** (10.3-16.3)	<0.0001	**11.4** (9.5-13.8)	0.8598
Cholesterol (mg/day)	**216.5** (172.7-271.8)	**309.4** (258.6-369.1)	<0.0001	**356.1** (305.1-412.5)	<0.0001	**329.4** (276.1-402.3)	<0.0001	**334.4** (280.9-398.3)	<0.0001
Non-starch polysaccharides (g/day)	**16.3** (13.8-19.4)	**14.2** (11.6-17.6)	<0.0001	**14.2** (11.9-16.9)	<0.0001	**13.8** (12.0-15.9)	<0.0001	**14.1** (11.8-16.7)	<0.0001

Nutrients	UK (n=3921)	CZE (n=3665)		POL (n=4847)		RUS (n=4149)		POOLED Czech, Polish and Russian sample (n=12,661)	
	Median[1] (IQR)	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**
Vitamin C (mg/day)	**138.9** (98.7-189.5)	**113.7** (77.8-179.0)	<0.0001	**99.1** (66.2-146.9)	<0.0001	**72.0** (51.1-111.0)	<0.0001	**93.0** (62.2-143.8)	<0.0001
Total energy (MJ/day)	**7.6** (6.3-9.2)	**6.7** (5.4-8.3)	<0.0001	**7.2** (5.9-8.8)	0.0003	**8.4** (6.8-10.1)	<0.0001	**7.4** (5.9-9.1)	0.0074

[1] All values are energy standardized around 8MJ/day, except for alcohol and total energy intake for which absolute intakes are shown

[2] All p-values were calculated with quantile regression using the intake values in the UK cohort as reference category, adjusted for age, energy intake, smoking, education, employment status, marital status, leisure time physical activity, CVD/diabetes in medical history

[3] On average, more than 80% of their intake was provided by the common items (n=81) in all four cohorts

[4] On average, 60-80% of their intake was provided by the common items (n=81) in at least one of the cohorts, and more than 80% in the other cohorts

Table IV-5: Average intake of foods and drinks in the British, Czech, Polish, Russian cohorts and the pooled Eastern European sample, including only female participants

Food groups and subgroups (FoodEx2)	UK (n=1512)	CZE (n=4199)		POL (n=5053)		RUS (n=4993)		POOLED Czech, Polish and Russian sample (n=14,245)	
	Median[1] (IQR)	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]
Fully comparable foods and drinks[3]									
Animal fresh meat / animal offals	**68.4** (42.0-102.0)	**68.4** (40.6-103.2)	0.0289	**75.4** (51.6-92.2)	0.2812	**103.2** (64.4-143.8)	<0.0001	**76.8** (51.6-111.6)	<0.0001
Eggs	**7.0** (3.5-21.5)	**7.0** (7.0-21.5)	1.0	**21.5** (7.0-21.5)	<0.0001	**21.5** (3.5-21.5)	0.0007	**21.5** (7.0-21.5)	<0.0001
Fruits and fruit products	**292.2** (179.3-422.2)	**332.9** (193.9-553.2)	<0.0001	**246.0** (141.2-387.2)	<0.0001	**146.7** (79.9-244.4)	<0.0001	**220.3** (123.1-385.0)	<0.0001
Fresh fruits	**266.3** (163.2-393.6)	**308.2** (173.0-530.4)	<0.0001	**222.9** (128.1-362.4)	0.0001	**110.1** (52.5-204.8)	<0.0001	**193.1** (100.8-355.2)	<0.0001
Processed fruit products	**14.7** (6.0-29.8)	**15.3** (8.3-28.5)	<0.0001	**11.2** (3.5-21.5)	0.1181	**21.5** (7.7-48.5)	<0.0001	**15.3** (7.0-32.0)	<0.0001
Vegetables *(all non-products)*[4]	**264.8** (182.8-380.6)	**211.9** (132.2-330.8)	<0.0001	**204.6** (132.6-314.3)	<0.0001	**299.7** (233.6-393.6)	<0.0001	**246.6** (156.1-350.0)	0.0001
Starchy roots or tubers	**89.6** (53.2-143.9)	**80.8** (75.3-98.3)	<0.0001	**78.1** (75.3-138.3)	<0.0001	**86.8** (41.6-138.3)	<0.0001	**80.8** (75.3-103.9)	<0.0001
Sugars, confectionery and water-based sweet dessert	**7.0** (3.4-21.5)	**7.6** (3.5-18.5)	0.0198	**15.8** (6.0-31.1)	<0.0001	**28.1** (15.0-42.9)	<0.0001	**16.4** (6.0-35.0)	<0.0001
Alcoholic beverages (portion/day)	**0.4** (0.1-1.1)	**0.1** (0.0-0.4)	<0.0001	**0.0** (0.0-0.1)	<0.0001	**0.1** (0.0-0.1)	<0.0001	**0.1** (0.0-0.1)	<0.0001
Coffee, cocoa, tea and infusions	**690.0** (489.0-975.0)	**675.0** (390.0-900.0)	<0.0001	**675.0** (503.0-975.0)	0.0657	**561.0** (475.0-690.0)	<0.0001	**675.0** (475.0-855.0)	<0.0001

Food groups and subgroups (FoodEx2)	UK (n=1512)	CZE (n=4199)		POL (n=5053)		RUS (n=4993)		POOLED Czech, Polish and Russian sample (n=14,245)	
	Median[1] (IQR)	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]
Partially comparable foods and drinks[5]									
All meat and meat products	**83.7** (51.6-117.0)	**82.4** (52.8-117.2)	0.0519	**97.9** (74.1-123.5)	<0.0001	**122.1** (77.9-165.7)	<0.0001	**99.9** (68.4-136.1)	<0.0001
Grains and grain based products	**158.7** (112.2-234.8)	**156.8** (104.0-221.1)	0.6375	**175.4** (129.2-240.7)	<0.0001	**185.8** (118.7-268.0)	0.5983	**170.8** (116.3-242.6)	0.0021
Legumes, nuts, oilseeds, spices	**25.9** (13.3-45.3)	**11.2** (6.3-18.2)	<0.0001	**11.2** (4.9-17.5)	<0.0001	**9.8** (4.9-16.8)	<0.0001	**11.2** (4.9-17.5)	<0.0001
Animal fats and oils	**0.7** (0.0-7.9)	**1.4** (0.7-10.0)	<0.0001	**7.9** (0.0-25.0)	<0.0001	**4.3** (1.4-10.0)	<0.0001	**4.3** (0.7-10.0)	<0.0001
Seasoning, sauces, condiments	**8.7** (4.2-24.4)	**11.5** (7.8-22.5)	<0.0001	**8.6** (3.5-17.2)	0.0637	**12.9** (4.2-30.9)	<0.0001	**11.3** (4.3-26.4)	<0.0001
Fruit and vegetable juices and nectars	**86.0** (14.0-200.0)	**14.0** (0.0-28.0)	<0.0001	**28.0** (0.0-86.0)	<0.0001	**14.0** (0.0-86.0)	<0.0001	**14.0** (0.0-86.0)	<0.0001

[1] Values are g/day intakes except for alcoholic beverages where portion/day intake is shown

[2] All p-values were calculated with quantile regression using the intake values in the UK cohort as reference category, adjusted for age, energy intake, smoking, alcohol consumption, education, vitamin supplement intake, employment status, marital status, leisure time physical activity, CVD/diabetes in medical history

[3] On average, more than 80% of their intake was provided by the common items (n=81) in all four cohorts

[4] Including: brassica vegetables; bulb, stalk and stem vegetables; fruiting vegetables; leafy vegetables; legume greens, sprouts; non-starchy root and tuber vegetables; fungi; marine algae, aromatic herbs or flowers

[5] On average, 60-80% of their intake was provided by the common items (n=81) in at least one of the cohorts, and more than 80% in the other cohorts

<u>Table IV-6:</u> Average intake of nutrients in the British, Czech, Polish, Russian cohorts and the pooled Eastern European sample, including only female participants

Nutrients	UK (n=1512)	CZE (n=4199)		POL (n=5053)		RUS (n=4993)		POOLED Czech, Polish and Russian sample (n=14,245)	
	Median[1] (IQR)	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**	**Median[1]** (IQR)	**p-value[2]**
Fully comparable nutrients[3]									
Alcohol (g/day)	**4.9** (0.8-11.5)	**1.2** (0.0-3.9)	<0.0001	**0.0** (0.0-0.7)	<0.0001	**0.6** (0.0-1.3)	<0.0001	**0.6** (0.0-1.4)	<0.0001
Beta-carotene (mg/day)	**6.9** (4.0-10.1)	**5.6** (4.0-9.4)	<0.0001	**7.8** (5.0-11.3)	<0.0001	**12.3** (8.8-15.2)	<0.0001	**8.7** (5.2-12.9)	<0.0001
Partially comparable nutrients[4]									
Total carbohydrate (g/day)	**240.0** (210.6-267.8)	**230.4** (205.2-255.5)	<0.0001	**230.3** (206.3-253.8)	<0.0001	**227.9** (202.4-252.3)	<0.0001	**229.5** (204.5-253.7)	<0.0001
Sugar (g/day)	**125.3** (102.6-150.1)	**119.9** (94.5-148.6)	<0.0001	**111.1** (90.2-135.1)	<0.0001	**115.7** (95.4-138.0)	<0.0001	**115.2** (93.2-140.1)	<0.0001
Protein (g/day)	**75.8** (65.8-86.5)	**77.2** (67.4-87.0)	0.4917	**81.1** (73.0-90.5)	<0.0001	**81.5** (70.0-93.3)	<0.0001	**80.1** (70.1-90.5)	<0.0001
Total fat (g/day)	**67.7** (58.8-77.8)	**76.4** (67.8-85.6)	<0.0001	**77.3** (67.8-86.6)	<0.0001	**78.2** (69.9-87.5)	<0.0001	**77.4** (68.6-86.6)	<0.0001
Saturated fat (g/day)	**25.4** (21.0-30.9)	**31.8** (27.2-36.6)	<0.0001	**32.6** (27.2-38.7)	<0.0001	**30.2** (26.0-34.7)	<0.0001	**31.3** (26.7-36.5)	<0.0001
Polyunsaturated fat (g/day)	**11.9** (9.8-14.8)	**11.4** (9.8-13.3)	<0.0001	**10.6** (9.0-12.7)	<0.0001	**14.6** (11.6-18.5)	<0.0001	**11.9** (9.9-14.8)	0.3215
Cholesterol (mg/day)	**223.9** (168.2-276.3)	**307.6** (251.8-371.5)	<0.0001	**339.4** (286.4-394.7)	<0.0001	**311.9** (255.1-375.0)	<0.0001	**320.8** (263.7-381.6)	<0.0001
Non-starch polysaccharides (g/day)	**18.2** (14.9-21.6)	**17.5** (14.1-21.7)	0.8066	**15.8** (13.0-18.9)	<0.0001	**15.0** (12.8-17.5)	<0.0001	**15.8** (13.2-19.2)	<0.0001

Nutrients	UK (n=1512)	CZE (n=4199)		POL (n=5053)		RUS (n=4993)		POOLED Czech, Polish and Russian sample (n=14,245)	
	Median[1] (IQR)	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]	Median[1] (IQR)	p-value[2]
Vitamin C (mg/day)	**161.9** (115.6-232.9)	**164.5** (106.2-256.8)	0.0405	**118.8** (80.9-180.5)	<0.0001	**91.7** (63.7-148.1)	<0.0001	**120.0** (77.9-192.1)	<0.0001
Total energy (MJ/day)	**6.7** (5.4-8.1)	**6.1** (4.9-7.8)	<0.0001	**6.6** (5.4-7.9)	0.5719	**7.1** (5.6-8.7)	<0.0001	**6.6** (5.3-8.2)	0.7872

[1] All values are energy standardized around 8MJ/day, except for alcohol and total energy intake for which absolute intakes are shown

[2] All p-values were calculated with quantile regression using the intake values in the UK cohort as reference category, adjusted for age, energy intake, smoking, education, employment status, marital status, leisure time physical activity, CVD/diabetes in medical history

[3] On average, more than 80% of their intake was provided by the common items (n=81) in all four cohorts

[4] On average, 60-80% of their intake was provided by the common items (n=81) in at least one of the cohorts, and more than 80% in the other cohorts

APPENDIX V.

Healthy diet indicator and mortality: additional analyses

Table V-1: Results of Cox regression analysis between HDI and mortality outcomes, after applying different exclusion criteria for energy misreporting

Cause of death	Sample	Exclusion criteria for energy misreporting	Dead/n	Model 1		Model 2	
				HR/SD (95%CI)[1]	p-value	HR/SD (95%CI)[1]	p-value
All-cause	Pooled	**Excl. 1**	1209/18,559	**0.94** (0.89-1.00)	0.055	**0.95** (0.89-1.00)	0.068
		Excl. 2	1087/17,100	**0.94** (0.89-1.01)	0.076	**0.95** (0.89-1.01)	0.079
		Excl. 3	1216/18,637	**0.95** (0.89-1.00)	0.070	**0.95** (0.89-1.00)	0.067
		Excl. 4	1230/18,718	**0.95** (0.90-1.01)	0.102	**0.95** (0.90-1.01)	0.119
CVD	Pooled	**Excl. 1**	423/18,494	**0.89** (0.81-0.99)	0.030	**0.90** (0.81-0.99)	0.030
		Excl. 2	381/17,048	**0.89** (0.80-0.98)	0.023	**0.89** (0.81-0.99)	0.030
		Excl. 3	423/18,573	**0.89** (0.81-0.99)	0.028	**0.89** (0.81-0.98)	0.022
		Excl. 4	431/18,653	**0.90** (0.82-1.00)	0.040	**0.90** (0.82-1.00)	0.040

Model 1: adjusted for age, sex, cohort

Model 2: adjusted for age, sex, cohort, education, household amenities score, marital status, smoking, alcohol intake, energy intake, vitamin supplement intake, physical activity, medical history

Excl. 1: Participants in the top and bottom 1% of the energy intake (EI) vs. basal metabolic rate (BMR) ratio were excluded from the analysis

Excl. 2: Participants in the top and bottom 5% of the EI vs. BMR ratio were excluded from the analysis

Excl. 3: Males and females with more than 5000/4500 kcal/day or less than 800/500 kcal/day reported energy intake, respectively, were excluded

Excl. 4: Participants who reported to consume more than 65 items or less than 5 items a day were excluded

[1] effect of one standard deviation (SD) increase in the score;

Table V-2: Results of Cox regression analysis with categorical HDI groups (categories are based on distance from sample mean)

			Distance of HDI score from sample mean								
			≤ -1 SD (n=2934)		> -1 SD and ≤ mean (n=6754)		> mean and ≤ 1 SD (n=5956)		>1 SD (n=2915)		
Cause of death	Sample	Model	**HR**		**HR**	(95% CI)	**HR**	(95% CI)	**HR**	(95% CI)	**p-value for trend**
All cause	Pooled	model 1	**1.00**	(ref.)	**0.77**	(0.66-0.91)	**0.84**	(0.71-1.00)	**0.77**	(0.62-0.94)	0.067
		model 2	**1.00**	(ref.)	**0.84**	(0.71-0.98)	**0.90**	(0.76-1.06)	**0.78**	(0.63-0.96)	0.085
CVD	Pooled	model 1	**1.00**	(ref.)	**0.89**	(0.67-1.19)	**0.85**	(0.63-1.14)	**0.66**	(0.46-0.94)	0.025
		model 2	**1.00**	(ref.)	**0.98**	(0.74-1.30)	**0.91**	(0.68-1.22)	**0.66**	(0.46-0.95)	0.025
CHD	Pooled	model 1	**1.00**	(ref.)	**0.86**	(0.59-1.25)	**0.69**	(0.46-1.04)	**0.56**	(0.34-0.93)	0.011
		model 2	**1.00**	(ref.)	**0.96**	(0.66-1.40)	**0.75**	(0.50-1.13)	**0.56**	(0.34-0.92)	0.009
Stroke	Pooled	model 1	**1.00**	(ref.)	**0.83**	(0.46-1.52)	**0.88**	(0.48-1.62)	**0.88**	(0.44-1.74)	0.848
		model 2	**1.00**	(ref.)	**0.91**	(0.50-1.65)	**0.94**	(0.51-1.73)	**0.93**	(0.46-1.86)	0.903
Cancer	Pooled	model 1	**1.00**	(ref.)	**0.74**	(0.56-0.97)	**0.87**	(0.66-1.15)	**0.89**	(0.64-1.24)	0.915
		model 2	**1.00**	(ref.)	**0.78**	(0.60-1.03)	**0.91**	(0.69-1.21)	**0.91**	(0.65-1.27)	0.952
Non-CVD-non-cancer	Pooled	model 1	**1.00**	(ref.)	**0.78**	(0.56-1.09)	**0.86**	(0.60-1.22)	**0.84**	(0.55-1.29)	0.629
		model 2	**1.00**	(ref.)	**0.85**	(0.61-1.19)	**0.93**	(0.65-1.32)	**0.84**	(0.55-1.29)	0.609

Model 1: adjusted for age, sex, cohort

Model 2: adjusted for age, sex, cohort, education, household amenities score, marital status, smoking, alcohol intake, energy intake, vitamin supplement intake, physical activity

<u>Table V-3:</u> Results of Cox regression analysis between HDI and mortality outcomes on participants including those with prevalent CVD, cancer and diabetes (n=25,858)

			Model 1		Model 2	
Cause of death	**Sample**	**Dead/n**	**HR/SD (95%CI)[1]**	**p-value**	**HR/SD (95%CI)[1]**	**p-value**
All-cause	Pooled	2332/25,858	**0.96** (0.92-1.01)	0.104	**0.96** (0.92-1.00)	0.045
CVD	Pooled	954/25,740	**0.98** (0.92-1.05)	0.547	**0.96** (0.90-1.03)	0.225
CHD	Pooled	529/25,740	**0.97** (0.89-1.06)	0.546	**0.95** (0.87-1.03)	0.220
Stroke	Pooled	220/25,740	**1.01** (0.89-1.16)	0.830	**1.00** (0.87-1.14)	0.943
Cancer	Pooled	801/25,740	**0.96** (0.89-1.04)	0.323	**0.97** (0.90-1.04)	0.406
Non-CVD-non-cancer	Pooled	459/25,740	**0.91** (0.83-1.01)	0.075	**0.90** (0.82-1.00)	0.040

Model 1: adjusted for age, sex, cohort

Model 2: adjusted for age, sex, cohort, education, household amenities score, marital status, smoking, alcohol intake, energy intake, vitamin supplement intake, physical activity, medical history

[1] effect of one standard deviation (SD) increase in the score;

Table V-4: Results of Cox regression analysis using the "original" HDI score[1] (n=18,559)

			Model 1		Model 2	
Cause of death	**Sample**	**Dead/n**	**HR/SD (95%CI)[2]**	**p-value**	**HR/SD (95%CI)[2]**	**p-value**
All-cause	Pooled	1209/18,559	**0.94** (0.88-1.00)	0.043	**0.97** (0.92-1.04)	0.403
CVD	Pooled	423/18,494	**0.94** (0.85-1.04)	0.247	**0.97** (0.88-1.08)	0.616
CHD	Pooled	220/18,494	**0.97** (0.84-1.12)	0.701	**1.01** (0.87-1.16)	0.928
Stroke	Pooled	105/18,494	**1.00** (0.81-1.23)	0.998	**1.03** (0.84-1.27)	0.781
Cancer	Pooled	437/18,494	**0.99** (0.90-1.10)	0.910	**1.03** (0.93-1.14)	0.561
Non-CVD-non-cancer	Pooled	284/18,494	**0.86** (0.76-0.98)	0.027	**0.91** (0.80-1.04)	0.156

Model 1: adjusted for age, sex, cohort

Model 2: adjusted for age, sex, cohort, education, household amenities score, marital status, smoking, alcohol intake, energy intake, vitamin supplement intake, physical activity

[1] Huibregts et al 1997

[2] effect of one standard deviation (SD) increase in the score;

APPENDIX VI.

Comparison of participants who were included and excluded from the analysis in the Whitehall II study

Table VI-1: Comparison of Whitehall II study participants who were included in and excluded from the analyses

Covariate	Category	Included (n=5433)	Excluded (n=1534)	p value[1]
Mean age, years (SD)		61.2 (6.0)	61.4 (6.2)	0.346
		%	%	
Sex	Males	72.2	63.4	
	Females	27.8	36.6	<0.001
Marital status	Single/divorced/wid.	23.3	28.6	
	Married/cohabiting	76.3	69.9	<0.001
	Missing	0.5	1.6	
Education	Primary or less	8.7	8.0	
	O-level/vocational	22.7	20.4	
	A-level/secondary	25.2	19.5	
	BA/BSc or higher	32.1	27.3	0.347
	Missing	11.3	24.9	
Employment status	Employed	49.0	49.0	
	Retired	45.5	42.3	
	Non-employed-non-retired	5.2	7.9	0.151
	Missing	0.3	0.9	
Smoking habits	Never smoker	49.3	46.4	
	Ex-smoker	42.9	39.6	
	Current smoker	7.3	11.6	<0.001
	Missing	0.5	2.4	
Leisure time physical activity	Inactive	15.4	19.7	
	Moderately active	43.6	40.8	
	Active	39.9	29.2	<0.001
	Missing	1.2	10.3	
Medical history (CVD, diabetes)	Negative	89.9	78.2	
	Positive	8.6	12.9	<0.001
	Missing	1.5	8.9	

[1] All p values were calculated with logistic regression using inclusion/exclusion as outcome variable and the covariates as explanatory variables

APPENDIX VII.

Published papers of the thesis

The European Journal of Public Health Advance Access published December 4, 2015
European Journal of Public Health, 1–7

doi:10.1093/eurpub/ckv216

Comparison of food and nutrient intakes between cohorts of the HAPIEE and Whitehall II studies

Denes Stefler[1], Andrzej Pajak[2], Sofia Malyutina[3,4], Ruzena Kubinova[5], Martin Bobak[1], Eric J. Brunner[1]

1 Department of Epidemiology and Public Health, University College London, London, UK
2 Department of Epidemiology and Population Sciences, Jagiellonian University, Krakow, Poland
3 Institute of Internal and Preventive Medicine, Siberian Branch of the Russian Academy of Medical Sciences, Novosibirsk, Russia
4 Novosibirsk State Medical University, Novosibirsk, Russia
5 Centre for Health Monitoring, National Institute of Public Health, Prague, Czech Republic

*Correspondence: Denes Stefler, Department of Epidemiology and Public Health, University College London, 1-19 Torrington Place, London WC1E 6BT, UK, Tel: +44 (0) 20 7679 1680, Fax: +44 (0)20 7813 0242, e-mail: denes.stefler.10@ucl.ac.uk

Background: Differences in dietary habits have been suggested as an important reason for the large health gap between Eastern and Western European populations. Few studies have compared individual-level nutritional data directly between the two regions. This study addresses this hypothesis by comparing food, drink and nutrient intakes in four large population samples. **Methods:** Czech, Polish and Russian participants of the Health, Alcohol and Psychosocial Factors in Eastern Europe (HAPIEE) study, and British participants in the Whitehall II study, altogether 29 972 individuals aged 45–73 years, were surveyed in 2002–2005. Dietary data were collected by customised food frequency questionnaires. Reported food, drink and nutrient intake data were harmonised and compared between cohorts using multivariable adjusted quantile regression models. **Results:** Median fruit and vegetable intakes were lower in the pooled Eastern European sample, but not in all country cohorts, compared with British subjects. Median daily consumption of fruits were 275, 213, 130 and 256 g in the Czech, Polish, Russian and Whitehall II cohort, respectively. The respective median daily intakes of vegetables were 185, 197, 292 and 246 g. Median intakes of animal fat foods and saturated fat, total fat and cholesterol nutrients were significantly higher in the Czech, Polish and Russian cohorts compared with the British; for example, median daily intakes of saturated fatty acids were 31.3, 32.5, 29.2 and 25.4 g, respectively. **Conclusion:** Our findings suggest that there are important differences in dietary habits between and within Eastern and Western European populations which may have contributed to the health gap between the two regions.

Introduction

High prevalence of unhealthy diets in Central and Easter Europe (CEE) and the former Soviet Union (FSU) has been suggested to play an important role in the high cardiovascular disease (CVD) mortality rates in these regions.[1–3] Ecological data indicate that people in CEE and FSU consume less fruits and vegetable oils but more animal fats than individuals in Western Europe.[4] However, comparison of individual-level dietary data between Eastern and Western European countries is rare.[5–7]

Nationally representative, individual-level nutritional surveys are conducted regularly in many European countries in order to monitor the population's dietary habits. Although they provide good evidence for public health recommendations in the specific countries, their applicability for international comparison is limited because the dietary assessment methods differ between countries.[8] Methods differ, to varying degrees, in terms of data collection tools, food classification, portion sizes and nutrient composition tables.[9–13]

The Health, Alcohol and Psychosocial Factors in Eastern Europe (HAPIEE) study is one of the largest and most recent studies with data on dietary habits of general population samples from the Czech Republic, Poland and the Russian Federation.[14]

In the current analysis, we compared individual-level food, drink and nutrient intakes between participants of the three HAPIEE cohorts and the UK-based Whitehall II cohort using identical methods for data analysis in both studies. Country-customised food frequency questionnaires (FFQ) with closely analogous design and layout were used for dietary data collection in the four cohorts.[15,16]

Methods

Study participants and dietary data collection

The design, recruitment process and dietary assessment of the HAPIEE and Whitehall II studies has been described previously.[14,17]

In brief, the HAPIEE study is a prospective cohort study, which is designed to investigate the relationship between traditional, non-conventional and psychosocial risk factors and chronic non-communicable diseases, particularly CVD, in CEE and FSU.[14] The baseline survey in 2002–2005 recruited randomly selected population samples in Novosibirsk (Russia), Krakow (Poland) and six cities in the Czech Republic. Overall, 28 945 men and women (8857 Czechs, 10 728 Poles, 9360 Russians) aged 45–69 years at baseline were included in the study (overall response rate of 59%).

The Whitehall II study is a prospective cohort study of civil servants set up in 1985–1988 with the central aim to examine the impact of social inequalities on physical and mental health.[17] Participants were recruited from 20 civil service departments in London; they undergo medical examination every 5 years and complete postal questionnaires between the screening phases. In the current analysis, we used dietary data from the seventh wave of the study which took place between 2002 and 2004, the same time as the baseline data collection of the HAPIEE study. In this phase, 6967 participants aged 50–73 years took part (68% of phase 1 responders).

In both studies, dietary data collection was carried out using a semi-quantitative FFQ. The FFQ used in the HAPIEE study was constructed on the basis of the Whitehall II study questionnaire.

Participants could indicate how frequently they consumed a particular food or drink item using a 9-point scale ranging from 'never, or less than once a month' to 'more than 6-times a day'.[15,16]

The FFQs completed by the Czech, Polish, Russian and UK cohorts consisted of 136, 147, 142 and 116 food and drink items, respectively. There were two reasons for the discrepancies: (1) Some food products were combined into one FFQ item in one country, but asked separately in others. For example, apricots, peaches and plums were combined in one question in the UK but in three separate questions in the HAPIEE cohorts. (2) Certain items were not included in all FFQs, because some of them were country-specific foods (e.g. pirogi, borscht). However, the majority of these FFQ-specific items (77, 66, 67 and 59% in the Czech, Polish, Russian and British questionnaires, respectively) were consumed in all four countries (e.g. pineapple, aubergine, cucumber, lasagne).

In all cohorts, participants who answered <90% of the FFQ questions and those who stated that the FFQ was not representative of their diet were excluded from the analysis. Participants with implausible food intake values, i.e. the bottom and top 1% of the cohort-specific energy intake/BMR ratio, were omitted.

Participants with missing data in any of the confounder variables were also excluded. Overall, 4473 British, 7298 Czech, 9098 Polish and 9103 Russian participants were included in the current analysis.

Dietary data harmonisation

Measured intake of a given food group is likely to be proportional to the number of relevant items in the FFQ. Unless the differences between the FFQs represent country-specific differences in dietary habits (i.e. country-specific food items), which is not the case in the current comparison as described above, these discrepancies in the number of FFQ items may introduce reporting bias and need to be taken into account.

Firstly, we excluded those items from the analysis which were not common in all four FFQs. Secondly, regarding food and drink items which were asked separately in one but in combination in other FFQs, the portion/day intake levels were summarised and the data on the combined intakes were used in all cohorts. Overall, dietary intake data from 81 single or combined food and drink items were used in the current analysis.

Participants had to estimate their intake habits regarding an *average* portion or *medium*-sized food or drink item in all four FFQs. In order to calculate g day^{-1} intake of a specific item, standard portion sizes, provided by local dieticians, were used in previous analyses.[15,16] These country-specific portion sizes were identical or similar for most items, however, for 29 (36%) of 81 items the difference was >50%. Although some of the small differences might reflect real regional differences, large discrepancies are likely due to arbitrary choices made by local dieticians during the construction of the FFQs. To avoid information bias due to different portion sizes, the g day^{-1} intake of each food and drink items was recalculated by substituting identical portion sizes in all cohort-specific datasets, using the portion sizes published by the UK's Food Standard Agency.[18] Alcoholic drink sizes were an exception, because the size of a standard drink clearly differs between countries and the questions on the FFQs were asked in line with the local habits. (i.e. 1 beer is 1/2 pint = 287 ml in the UK but 1 glass = 250 ml in CEE/FSU.)

In the HAPIEE cohorts, participants were asked to estimate their eating habits over the past 3 months. In contrast, the questions referred to the previous year in Whitehall II study, and regarding seasonal foods (i.e. fruits, vegetables), participants were asked to estimate their intakes in the time period when that particular item is in season. In order to eliminate the differences due to the different reference periods of the FFQs, we compared weighted intake data for fresh fruits and vegetables: for those participants of the HAPIEE cohorts who completed the FFQ during winter or spring, the intake of fresh fruits and vegetables were multiplied by the within-cohort summer–autumn vs. winter–spring ratio of median fresh fruit and vegetable intake.

National Food Composition tables and databases (FCDs) differ in completeness, accuracy and may use different analytical methods to measure nutrient content of foods. Because these technical differences in FCDs can lead to biased international comparisons of nutrient intake levels,[10,19] we used the McCance and Widdowson's FCD to estimate nutrient intake levels in both Whitehall II and HAPIEE cohorts.

Further data preparation and statistical analysis

The food and drink items listed in the FFQs were categorised into food/drink groups and subgroups according to the European Food Safety Authority's Foodex2 food classification system.[20] The comparisons were carried out on absolute intake values for food/drink groups and subgroups, and on energy standardised intake values (calculated by the residual method) for nutrients.[21]

To take account of possible information bias, food/drink groups and nutrients were categorised as fully, partially or not comparable between cohorts, according to the contribution of the 81 identical items to their total intake. Food/drink groups and nutrients were considered fully comparable if >80% of intake was provided by common items in all cohorts. If the contribution was 60–80% in one or more of the cohorts, they were considered partially comparable. If the contribution was <60% of intake in one or more of the cohorts then the food, drink or nutrient was not considered comparable and results were not shown.

In the multivariable adjusted models, quantile regression method was used because of the non-normal distribution of food, drink and nutrient intake data. All comparisons were adjusted for age (continuous), sex, energy intake (kJ day^{-1}, continuous), marital status (married/cohabiting; single/widowed/divorced), highest level of education (primary or less; O-level/vocational; A-level/secondary; BA/BSc or higher), employment status (employed; retired; not employed/not retired), alcohol intake (abstainers; moderate drinkers: <15 g day^{-1} for women, <30 g day^{-1} for men; heavy drinkers: ≥15 g day^{-1} for women, ≥30 g day^{-1} for men), smoking (non-; ex-; current smokers), vitamin supplement usage (regular users; irregular or not users), leisure time physical activity (high: >15 MET-hours day^{-1}; moderate: 5–15 MET-hours day^{-1}; low: <5 MET-hours day^{-1}) and medical history (CVD or DM in medical history; no CVD or DM in medical history).

All statistical analyses were carried out using STATA 13.1 statistical software (StataCorp., College Station, TX, USA).

Results

On average, ~75% of total food/drink and energy intakes were captured by the 81 identical items in each cohort (tables 1 and 2). However, this proportion varied across food/drink groups, nutrients and cohorts. For example, on average, 2.2% of vegetable oil intake was provided by the common item in the Russian sample, while nearly all (96.1–100%) of the fresh meat intake came from identical items in all four cohorts (table 1).

Table 3 shows the medians (IQR) g day^{-1} intakes of foods and drinks which were considered fully or partially comparable across cohorts. Multivariable adjusted cross-cohort comparisons, using the UK values as reference, are also shown. Average total and fresh fruit intake was significantly lower in Russian and Polish participants but higher in Czechs compared with the UK cohort. Russians had the lowest fresh fruit intakes, with average consumption less than half of any other cohort. In contrast, vegetable intake was significantly higher in Russians but lower in Poles and Czechs compared with the British sample. British participants reported higher consumption of starchy roots, alcohol, coffee, tea, legumes and fruit juices, but less meat products, sweets and animal fats than any of the Eastern European cohorts.

Table 1 Comparison of the FFQs used in the British, Czech, Polish and Russian cohorts

Overall food and drink categories	Food and drink groups and subgroups (FoodEx2)	No. items in FFQ UK	CZE	POL	RUS	No. items identical across the four FFQs	Mean percentage of food and drink intakes from the identical items[a] UK	CZE	POL	RUS
Foods of animal origin	Meat and meat products	9	15	14	15	8	98.2	76.2	81.5	86.2
	Animal fresh meat/animal offals	5	6	6	7	5	100.0	96.2	98.9	98.9
	Processed meat products/sausages and comminuted meat	4	9	8	8	3	92.1	40.5	56.2	53.7
	Milk and dairy products	9	13	15	12	6	25.4	49.4	50.2	59.8
	Eggs and egg products	1	1	1	1	1	100.0	100.0	100.0	100.0
	Fish, seafood, amphibians, reptiles and invertebrates	5	5	7	7	3	75.6	37.0	54.2	36.3
Foods of plant origin	Grains and grain-based products	15	10	10	10	7	72.6	74.1	72.1	66.1
	Fruits and fruit products	11	23	22	23	11	100.0	86.7	85.4	86.8
	Fresh fruits	8	20	19	20	8	100.0	85.5	84.1	81.6
	Processed fruit products	3	3	3	3	3	100.0	100.0	100.0	100.0
	Vegetables and vegetable products	18	25	28	26	16	94.9	79.9	72.5	87.2
	Vegetables (all non-products)[b]	18	22	24	23	16	94.9	89.0	86.2	94.2
	Vegetable products	0	3	4	3	0	na.	0.0	0.0	0.0
	Legumes, nuts, oilseeds and spices	6	6	4	6	4	87.9	60.4	100.0	78.5
	Starchy roots or tubers and products	4	3	3	3	3	84.2	100.0	100.0	100.0
	Sugar, confectionery and water-based sweet desserts	3	4	5	4	3	100.0	94.5	96.3	98.1
Foods of mixed origin	Animal and vegetable fats and oils	5	7	9	7	3	38.7	60.4	58.3	32.7
	Animal fats and oils	1	4	4	4	1	100.0	78.9	86.5	95.2
	Vegetable fats and oils	2	2	2	2	1	8.3	31.9	23.8	2.2
	Fats and oils of mixed origin	2	1	3	1	1	11.8	100.0	48.7	100.0
	Seasoning, sauces and condiments	6	3	4	3	3	64.2	100.0	95.4	100.0
	Composite dishes	10	8	13	13	3	58.5	64.7	47.9	41.0
Drinks	Alcoholic beverages	5	5	5	5	5	100.0	100.0	100.0	100.0
	Water and water-based beverages	2	4	2	2	2	100.0	25.0	100.0	100.0
	Coffee, cocoa, tea and infusions	5	2	3	3	2	89.3	100.0	98.4	99.2
	Fruit and vegetable juices and nectars	2	2	2	2	1	80.1	65.8	66.2	88.7
Total		116	136	147	142	81[c]	80.4	68.3	79.1	78.6

a: Values were calculated for each participant (in g day^{-1}) as follows: Intake from the 81 identical FFQ items∗100/Intake from all items in the original FFQs, for each food/drink group and overall.
b: Including: brassica vegetables; bulb, stalk and stem vegetables; fruiting vegetables; leafy vegetables; legume greens, sprouts; non-starchy root and tuber vegetables; fungi; marine algae, aromatic herbs or flowers.
c: Including nine which included more than one items each (combined items).
na.—not applicable.

Table 2 Mean percentage of nutrient and energy intake from the identical items compared with the original FFQs in the four cohorts[a]

Nutrients/energy	UK	CZE	POL	RUS
Total carbohydrate (g day^{-1})	76.4	76.7	75.8	74.7
Sugar (g day^{-1})	81.0	78.2	76.5	83.9
Protein (g day^{-1})	75.1	75.3	74.2	72.1
Total fat (g day^{-1})	73.4	70.9	69.5	63.3
Saturated fat (g day^{-1})	74.8	76.9	75.3	71.0
Polyunsaturated fat (g day^{-1})	65.5	65.2	64.9	60.7
Trans fat (g day^{-1})	57.2	76.9	78.0	79.3
Cholesterol (mg day^{-1})	83.7	84.2	81.6	77.1
Alcohol (g day^{-1})	100.0	100.0	100.0	100.0
Non-starch polysaccharides (g day^{-1})	78.6	79.0	73.5	76.8
Vitamin C (mg day^{-1})	86.8	80.1	72.3	66.8
Beta-carotene (μg day^{-1})	91.7	89.7	89.8	94.9
Total energy (kJ day^{-1})	76.7	75.0	73.4	70.4

a: Values were calculated for each participant as follows.
Intake from the 81 identical FFQ items∗100/Intake from all items in the original FFQs, for each nutrient and energy.

Table 4 shows the medians (IQR) of energy-standardised nutrient intakes in the four cohorts, as well as the results of the quantile regression analysis. Only alcohol and beta-carotene intakes were fully comparable across cohorts. There was higher intake of beta-carotenes but lower intake of vitamin C in Russians compared with the other cohorts, in line with the high vegetable and low fruit intake in this sample. Total fat, saturated fat and cholesterol intake were significantly higher in all three Eastern European cohorts than in the British sample, consistent with the food intake data. Alcohol consumption of British participants was the highest of any cohort.

An important difference between the Whitehall II and HAPIEE study participants was that the British cohort was based on civil

Table 3 Average intake of foods and drinks in the British, Czech, Polish, Russian cohorts and the pooled Eastern European sample

Food groups and subgroups (FoodEx2)	UK (n = 4 473)	CZE (n = 7298)		POL (n = 9098)		RUS (n = 9103)		POOLED Czech, Polish and Russian sample (n = 25 499)	
	Median[a] (IQR)	Median[a] (IQR)	P-value[b]	Median[a] (IQR)	P-value[b]	Median[a] (IQR)	P-value[b]	Median[a] (IQR)	P-value[b]
Fully comparable foods and drinks[c]									
Animal fresh meat/animal offals	74.2 (47.6–102.0)	76.8 (47.6–111.6)	<0.001	76.8 (60.0–103.2)	<0.001	117.2 (68.4–154.8)	<0.001	85.2 (58.6–120.0)	<0.001
Eggs	7.0 (3.5–21.5)	7.0 (7.0–21.5)	1.0	21.5 (7.0–21.5)	<0.001	21.5 (7.0–21.5)	<0.001	21.5 (7.0–21.5)	<0.001
Fruits and fruit products	256.1 (158.8–382.2)	275.0 (152.4–477.3)	<0.001	212.6 (124.4–346.6)	<0.001	130.0 (70.1–219.7)	<0.001	188.0 (102.7–335.9)	<0.001
Fresh fruits	231.8 (137.7–350.0)	256.0 (138.2–451.4)	<0.001	190.2 (114.6–325.1)	<0.001	91.4 (43.1–180.0)	<0.001	162.8 (78.0–308.1)	<0.001
Processed fruit products	16.5 (7.0–32.0)	14.7 (7.7–25.2)	<0.001	9.5 (2.5–20.0)	<0.001	21.5 (7.7–48.5)	<0.001	14.7 (7.0–31.7)	<0.001
Vegetables (all non-products)[d]	246.1 (170.6–337.5)	185.0 (113.7–293.8)	<0.001	197.3 (128.1–303.6)	<0.001	291.6 (225.6–381.0)	<0.001	235.9 (145.6–334.1)	<0.001
Starchy roots or tubers	98.3 (75.3–152.6)	86.8 (75.3–101.2)	<0.001	86.8 (75.3–141.1)	<0.001	86.8 (73.8–146.2)	<0.001	86.8 (75.3–138.3)	<0.001
Sugars, confectionery and water-based sweet dessert	8.1 (3.5–24.9)	8.8 (3.5–21.5)	<0.001	19.6 (7.0–35.1)	<0.001	31.1 (15.6–42.9)	<0.001	19.1 (7.0–36.6)	<0.001
Alcoholic beverages (portion day^{-1})	1.0 (0.4–2.5)	0.3 (0.1–1.0)	<0.001	0.1 (0.0–0.2)	<0.001	0.1 (0.0–0.5)	<0.001	0.1 (0.0–0.5)	<0.001
Coffee, cocoa, tea and infusions	869.0 (503.0–1055.0)	581.7 (390.0–690.0)	<0.001	675.0 (503.0–975.0)	<0.001	561.0 (475.0–855.0)	<0.001	675.0 (475.0–883.0)	<0.001
Partially comparable foods and drinks[e]									
All meat and meat products	90.1 (59.8–122.7)	92.2 (59.8–130.9)	<0.001	105.2 (80.0–136.4)	<0.001	135.5 (92.2–179.3)	<0.001	110.2 (76.3–151.5)	<0.001
Grains and grain based products	188.1 (127.8–267.0)	162.0 (107.7–228.4)	0.981	190.7 (134.8–263.3)	<0.001	218.5 (137.2–296.3)	0.002	190.5 (127.2–268.6)	<0.001
Legumes, nuts, oilseeds, spices	30.1 (16.1–49.7)	11.2 (6.3–18.2)	<0.001	11.2 (6.3–18.2)	<0.001	8.4 (4.9–14.7)	<0.001	11.2 (4.9–17.5)	<0.001
Animal fats and oils	0.0 (0.0–4.3)	1.4 (0.7–10.0)	<0.001	7.9 (0.0–25.0)	<0.001	4.3 (1.4–10.0)	<0.001	4.3 (0.7–10.0)	<0.001
Seasoning, sauces, condiments	10.8 (4.3–26.7)	12.2 (7.8–28.1)	<0.001	8.7 (4.3–19.4)	0.114	15.7 (4.3–33.7)	<0.001	12.2 (5.7–28.8)	<0.001
Fruit and vegetable juices and nectars	86.0 (14.0–200.0)	14.0 (0.0–28.0)	<0.001	28.0 (0.0–86.0)	<0.001	14.0 (0.0–86.0)	<0.001	14.0 (0.0–86.0)	<0.001

a: Values are g day^{-1} intakes except for alcoholic beverages where portion/day intake is shown
b: All P-values were calculated with quantile regression using the intake values in the UK cohort as reference category, adjusted for sex, age, energy intake, smoking, alcohol consumption, education, vitamin supplement intake, employment status, marital status, leisure time physical activity, CVD/diabetes in medical history.
c: On average, more than 80% of their intake was provided by the common items (n = 81) in all four cohorts.
d: Including: brassica vegetables; bulb, stalk and stem vegetables; fruiting vegetables; leafy vegetables; legume greens, sprouts; non-starchy root and tuber vegetables; fungi; marine algae, aromatic herbs or flowers.
e: On average, 60–80% of their intake was provided by the common items (n = 81) in at least one of the cohorts, and more than 80% in the other cohorts.

service office workers, while large proportions of the Eastern European cohorts were engaged in physical occupations. In a sensitivity analysis restricting the comparisons to office workers the results were substantially similar (Supplementary tables S1 and S2). Further, the results of comparisons were similar to the main findings when the analysis was carried out separately in males or females (Supplementary tables S3, S4, S5 and S6).

Discussion

Main findings

In this study, using data collection based on the same FFQ methodology across four samples, dietary intakes in the HAPIEE and Whitehall II cohorts were fully comparable only for a subset of foods, drinks and nutrients. Median fruit and vegetable intakes were significantly lower in the pooled Eastern European sample than in the British cohort. Notably, we found large variation in average consumption of these foods between the Czech, Polish and Russian cohorts, such that vegetable rather than fruit consumption was important in the Russian diet while fruit was important in the Czech diet. Although the consumption of animal fats, including saturated fatty acids and cholesterol, was only partially comparable between cohorts, the figures suggest that intakes were significantly higher in Eastern European participants compared with the British.

Strengths and limitations

Our study has a number of limitations which needs to be taken into account when interpreting the results. First, none of the included cohorts are fully representative of their respective national populations as a whole. The sampling frame included only urban inhabitants in the HAPIEE cohorts and London-based civil servants in the Whitehall II study. Second, there was a relatively low response rate in the Eastern European cohorts and some loss of baseline participants by Phase 7 of Whitehall II study which reduces the generalisability of our findings. A study in Poland recently found that hypertensive adults who live in rural areas consumed more fat and cholesterol

Table 4 Average intake of nutrients in the British, Czech, Polish, Russian cohorts and the pooled Eastern European sample

Nutrients	UK (n = 4 473)	CZE (n = 7298)		POL (n = 9098)		RUS (n = 9103)		POOLED Czech, Polish and Russian sample (n = 25 499)	
	Median[a] (IQR)	Median[a] (IQR)	P-value[b]	Median[a] (IQR)	P-value[b]	Median[a] (IQR)	P-value[b]	Median[a] (IQR)	P-value[b]
Fully comparable nutrients[c]									
Alcohol (g day^{-1})	10.9 (3.4–28.4)	2.6 (0.6–9.8)	<0.001	0.6 (0.0–2.4)	<0.001	1.1 (0.0–4.8)	<0.001	1.2 (0.0–4.9)	<0.001
Beta-carotene (mg day^{-1})	6.3 (3.7–8.7)	5.1 (3.6–8.0)	<0.001	7.3 (4.6–10.4)	<0.001	11.5 (7.8–14.3)	<0.001	7.8 (4.7–12.1)	<0.001
Partially comparable nutrients[d]									
Total carbohydrate (g day^{-1})	234.8 (205.1–261.3)	220.4 (193.8–247.8)	<0.001	225.6 (201.1–249.3)	<0.001	225.5 (200.1–249.7)	<0.001	224.4 (198.6–249.0)	<0.001
Sugar (g day^{-1})	116.1 (94.4–139.1)	108.3 (83.3–136.9)	<0.001	103.6 (83.5–127.2)	<0.001	107.4 (86.9–129.0)	<0.001	106.2 (84.8–130.4)	<0.001
Protein (g day^{-1})	72.3 (63.9–81.7)	78.4 (68.3–88.1)	<0.001	81.7 (73.3–90.7)	<0.001	82.0 (71.4–93.0)	<0.001	80.8 (71.1–90.8)	<0.001
Total fat (g day^{-1})	66.8 (58.4–76.0)	76.1 (67.2–85.1)	<0.001	78.0 (68.4–87.4)	<0.001	76.4 (67.8–85.2)	<0.001	76.8 (67.9–85.9)	<0.001
Saturated fat (g day^{-1})	25.4 (21.3–30.1)	31.3 (26.9–36.2)	<0.001	32.5 (27.1–38.7)	<0.001	29.2 (25.0–33.7)	<0.001	30.8 (26.2–36.1)	<0.001
Polyunsaturated fat (g day^{-1})	11.4 (9.5–14.2)	11.2 (9.6–13.2)	<0.001	10.7 (9.0–12.7)	<0.001	13.8 (10.9–17.5)	<0.001	11.7 (9.7–14.4)	0.715
Cholesterol (mg day^{-1})	218.3 (172.2–272.3)	308.9 (255.7–371.0)	<0.001	348.1 (295.2–403.8)	<0.001	320.0 (263.5–387.2)	<0.001	327.6 (272.0–389.3)	<0.001
Non-starch polysaccharides (g day^{-1})	16.6 (14.0–19.8)	15.8 (12.6–19.9)	<0.001	14.9 (12.4–18.0)	<0.001	14.4 (12.4–16.7)	<0.001	14.9 (12.4–18.0)	<0.001
Vitamin C (mg day^{-1})	143.6 (102.1–197.6)	136.5 (90.1–219.6)	0.003	109.3 (73.6–163.7)	<0.001	81.8 (56.7–131.0)	<0.001	105.5 (69.4–167.4)	<0.001
Total energy (MJ day^{-1})	7.4 (6.1–8.9)	6.4 (5.1–8.1)	<0.001	6.9 (5.6–8.3)	0.315	7.7 (6.2–9.5)	<0.001	7.0 (5.6–8.7)	0.892

a: All values are energy standardised around 8 MJ day^{-1}, except for alcohol and total energy intake for which absolute intakes are shown.
b: All P-values were calculated with quantile regression using the intake values in the UK cohort as reference category, adjusted for sex, age, energy intake, smoking, alcohol consumption, education, vitamin supplement intake, employment status, marital status, leisure time physical activity, CVD/diabetes in medical history.
c: On average, more than 80% of their intake was provided by the common items (n = 81) in all four cohorts.
d: On average, 60–80% of their intake was provided by the common items (n = 81) in at least one of the cohorts, and >80% in the other cohorts.

but less carbohydrates and fibre than urban inhabitants.[22] Particularly high-fat intake was also reported in a rural Lithuanian sample in the CINDI survey.[23] This suggests that in the Polish sample, and probably in the other two Eastern European cohorts as well, the average intake of fats and other nutrients may have been higher if the HAPIEE cohorts had included rural participants. Individuals in non-manual occupations tend to have a better-quality diet than manual workers,[24] indicating that participants of the Whitehall II cohort probably have healthier dietary patterns than the general UK population.

The FFQ is a cost-effective instrument to provide information on habitual diet in large studies. While the method has weaknesses of imprecision and information bias,[25,26] the extent of random and systematic error stemming from these weaknesses is likely to be similar in all the cohorts we studied. Thus, the major impact on between-country comparisons was probably to reduce power to detect small differences in intake. Further, cross-cohort comparability of the dietary intake data was maximised since all FFQs used the same 9-point scale answer–options for all food and drink items, and strong emphasis was put on data harmonisation in the analytical phase. On the other hand, despite these efforts, many foods, drinks and nutrients were only partially comparable across cohorts. Regarding these, the interpretation of results is limited because a significant proportion of intake was unknown.

Further strengths of our study were the large sample sizes and contemporaneous data collections, between 2002 and 2005, in all four cohorts.

Interpretation

Ecological data suggested that, on the aggregate level, fruit consumption is lower in CEE/FSU countries compared with Western Europe; however, there is probably no large difference in vegetable intake.[4] Although this study confirms these previous findings, it also shows that important differences exist between countries *within* the Eastern European region. In Russia, the very low reported fruit intake is consistent with FAO data[4] and it adds to the evidence that public health campaigns focusing on fruit consumption may be useful. On the other hand, high vegetable intake in this cohort is a favourable finding. To some extent it is probably due to widespread consumption of low-cost home-grown products. According to the Russian Statistical Office, 69% of vegetables produced in the country in 2012 came from household gardens, including dachas.[27]

The observation of significantly higher intakes of animal fat in the Eastern European cohorts compared with the British cohort confirms previous data and supports the hypothesis that its consumption plays an important role in the high CVD rates in these countries. Zatonski *et al.*[28] suggested that substitution of animal fats with vegetable oils during the 1990s was one of the main reasons for the rapid decline in ischemic heart disease mortality rates in Poland. Although the comparability of fat intake, as well as the generalisability of our findings, is limited, the results indicate that the gap in animal fat intake between East and West still existed in the first half of the 2000s. This area of diet should probably be one of the central targets of the public health interventions in the Czech Republic, Poland and Russia.

Research suggests that intake of foods and drinks with high added sugar content are related to increased risk of obesity, diabetes and CVD.[29–31] Although sugar intake (including all mono- and disaccharides) was the highest in British subjects, this result is probably due to the large contribution of fructose consumed via fruits and vegetables in this country cohort. The intakes of sweets and confectioneries were especially high in Poles and Russians. Added sugar consumption in Eastern European countries and its contribution to the high CVD rates would be worth examining in further studies.

Conclusion

Despite the limited direct international comparability of many food groups and nutrients, our study supports hypotheses proposing that inadequate fruit and high animal fat consumption contributed to poor vascular and metabolic health status in several Eastern European countries in the early 2000s. The results indicate that there are important differences in dietary habits within CEE and FSU, such that dietary and nutritional recommendations are relevant across the whole region, but public health interventions need to be tailored to specific countries.

Acknowledgements

We would like to thank all researchers, interviewers and participants of the HAPIEE and Whitehall II studies.

Supplementary data

Supplementary data are available at EURPUB online.

Key points

- Important differences in dietary habits exist between Eastern and Western European populations.
- This study found significantly lower fruit and vegetable intake but higher animal fat consumption in a pooled sample of Czech, Polish and Russian individuals compared with British civil servants. However, large variation in intake levels between the three Eastern European subsamples was also seen.
- The results support the hypothesis that unhealthy diet contributes to the high CVD rates of Eastern European populations.

Funding

The HAPIEE study was supported by the Wellcome Trust [grant numbers WT064947, WT081081]; the US National Institute of Aging [grant number 1RO1AG23522] and the MacArthur Foundation Initiative on Social Upheaval and Health. The Whitehall II was supported by the British Heart Foundation [grant number RG/13/2/30098]; British Medical Research Council [grant number K013351]; the British Health and Safety Executive; the British Department of Health, the British Stroke Association [grant number TSA 2008/05]; the US National Heart, Lung, and Blood Institute [grant number R01HL036310] and the US National Institute on Aging [grant numbers R01AG013196, R01AG034454]. Denes Stefler was supported by the British Heart Foundation.

Conflicts of interest

None declared.

References

1 Bobak M, Marmot M. East-west mortality divide and its potential explanations: proposed research agenda. *BMJ* 1996;312:421–5.

2 Kesteloot H, Sans S, Kromhout D. Dynamics of cardiovascular and all-cause mortality in Western and Eastern Europe between 1970 and 2000. *Eur Heart J* 2006;27:107–13.

3 Zatonski WA, HEM project team Epidemiological analysis of health situation development in Europe and its causes until 1990. *Ann Agric Environ Med* 2011;18:194–202.

4 Food and Agriculture Organisation of the United Nation. FAOSTAT, 2014. http://faostat3.fao.org/faostat-gateway/go/to/home/E (06 11 2014, date last accessed).

5 Serra-Majem L, MacLean D, Ribas L, et al. Comparative analysis of nutrition data from national, household, and individual levels: results from a WHO-CINDI collaborative project in Canada, Finland, Poland, and Spain. *J Epidemiol Commun Health* 2003;57:74–80.

6 Paalanen L, Prattala R, Palosuo H, Laatikainen T. Socio-economic differences in the consumption of vegetables, fruit and berries in Russian and Finnish Karelia: 1992–2007. *Eur J Public Health* 2011;21:35–42.

7 Crispim SP, Geelen A, Souverein OW, et al. Biomarker-based evaluation of two 24-h recalls for comparing usual fish, fruit and vegetable intakes across European centers in the EFCOVAL Study. *Eur J Clin Nutr* 2011;65:S38–47.

8 European Food Safety Authority. Use of the EFSA comprehensive European food consumption database in exposure assessment. *EFSA J* 2011;9:2097.

9 Charrondiere UR, Vignat J, Moller A, et al. The European Nutrient Database (ENDB) for Nutritional Epidemiology. *J Food Compos Anal* 2002;15:435–51.

10 Ireland J, van Erp-Baart AM, Charrondiere UR, et al. Selection of a food classification system and a food composition database for future food consumption surveys. *Eur J Clin Nutr* 2002;56:S33–45.

11 Lesser S, Pauly L, Volkert D, Stehle P. Nutritional situation of the elderly in Eastern/Baltic and Central/Western Europe—the AgeingNutrition project. *Ann Nutr Metab* 2008;52:62–71.

12 de Boer EJ, Slimani N, van't Veer P, et al. Rationale and methods of the European Food Consumption Validation (EFCOVAL) Project. *Eur J Clin Nutr* 2011;65:S1–4.

13 Novakovic R, Cavelaars AE, Bekkering GE, et al. Micronutrient intake and status in Central and Eastern Europe compared with other European countries, results from the EURRECA network. *Public Health Nutr* 2013;16:824–40.

14 Peasey A, Bobak M, Kubinova R, et al. Determinants of cardiovascular disease and other non-communicable diseases in Central and Eastern Europe: rationale and design of the HAPIEE study. *BMC Public Health* 2006;6:255.

15 Brunner E, Stallone D, Juneja M, et al. Dietary assessment in Whitehall II: comparison of 7 d diet diary and food-frequency questionnaire and validity against biomarkers. *Br J Nutr* 2001;86:405–14.

16 Boylan S, Welch A, Pikhart H, et al. Dietary habits in three Central and Eastern European countries: the HAPIEE study. *BMC Public Health* 2009;9:439.

17 Marmot M, Brunner E. Cohort profile: the Whitehall II study. *Int J Epidemiol* 2005;34:251–6.

18 Food Standards Agency. *Food Portion Sizes*, 3rd edn. London: Her Majesty's Stationery Office, 2002.

19 Vaask S, Pomerleau J, Pudule I, et al. Comparison of the micro-nutrica nutritional analysis program and the russian food composition database using data from the Baltic Nutrition Surveys. *Eur J Clin Nutr* 2004;58:573–9.

20 European Food Safety Authority. *The Food Classification and Description System FoodEx 2 (Draft-Revision 1)*. Parma: Supporting Publications, 2011.

21 Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* 1997;65:1220S–8S.

22 Suliburska J, Bogdanski P, Duda G, et al. An assessment of dietary intake and state of nutritional in hypertensive patients from rural and urban areas of Greater Poland. *Ann Agric Environ Med* 2012;19:339–43.

23 Petkeviciene J, Klumbiene J, Ramazauskiene V, et al. Diet and Dyslipidemias in a Lithuanian rural population aged 25–64: the CINDI Survey. *Medicina (Kaunas)* 2012;48:211–7.

24 Bolton-smith C, Smith WCS, Woodward M, Tunstall-pedoe H. Nutrient intakes of different social-class groups: results from the Scottish Heart Health Study (SHHS). *Br J Nutr* 1991;65:321–35.

25 Bingham SA. Limitations of the various methods for collecting dietary intake data. *Ann Nutr Metab* 1991;35:117–27.

26 Prentice RL. Dietary assessment and the reliability of nutritional epidemiology reports. *Lancet* 2003;362:182–3.

27 Russian Federation's Federal State Statistical Service. Production of basic agricultural products by types of enterprises. 2014. http://www.gks.ru/bgd/regl/b13_12/IssWWW.exe/stg/d01/15-04.htm (23 October 2014, date last accessed).

28 Zatonski WA, McMichael AJ, Powles JW. Ecological study of reasons for sharp decline in mortality from ischaemic heart disease in Poland since 1991. *BMJ* 1998;316:1047–51.

29 Yang Q, Zhang Z, Gregg EW, et al. Added sugar intake and cardiovascular diseases mortality among US adults. *JAMA Intern Med* 2004;174:516–24.

30 Malik VS, Pan A, Willett WC, Hu FB. Sugar-sweetened beverages and weight gain in children and adults: a systematic review and meta-analysis. *Am J Clin Nutr* 2013;98:1084–102.

31 Xi B, Li S, Liu Z, et al. Intake of fruit juice and incidence of type 2 diabetes: a systematic review and meta-analysis. *PLoS One* 2014;9:e93471.

Stefler and Bobak *Archives of Public Health* (2015) **73**:29
DOI 10.1186/s13690-015-0078-8

ARCHIVES OF PUBLIC HEALTH

SYSTEMATIC REVIEW Open Access

Does the consumption of fruits and vegetables differ between Eastern and Western European populations? Systematic review of cross-national studies

Denes Stefler* and Martin Bobak

Abstract

Background: Difference in fruit and vegetable consumption has been suggested as a possible reason for the large gap in cardiovascular disease (CVD) mortality rates between Eastern and Western European populations. However, individual-level dietary data which allow direct comparison across the two regions are rare. In this systematic review we aimed to answer the question whether cross-national studies with comparable individual-level dietary data reveal any systematic differences in fruit and vegetable consumption between populations in Central and Eastern Europe (CEE) and the Former Soviet Union (FSU) compared to Western Europe (WE).

Methods: Studies were identified by electronic search of MEDLINE, EMBASE and Web of Science databases from inception to September 2014, and hand search. Studies which reported data on fruit, vegetable consumption or carotene and vitamin C intake or tissue concentrations of adult participants from both CEE/FSU and WE countries were considered for inclusion. Quality of the included studies was assessed by a modified STROBE statement. Power calculation was performed to determine the statistical significance of the comparison results.

Results: Twenty-two studies fulfilled the inclusion criteria. Fruit consumption was found to be consistently lower in CEE/FSU participants compared to Western Europeans. Results on vegetable intake were less unambiguous. Antioxidant studies indicated lower concentration of beta-carotene in CEE/FSU subjects, but the results for vitamin C were not consistent.

Conclusion: This systematic review suggests that populations in CEE and FSU consume less fruit than Western Europeans. The difference in the consumption of fruit may contribute to the CVD gap between the two regions.

Keywords: Fruit and vegetable consumption, Central and Eastern Europe, Former Soviet Union, Cross-national studies

Background

Cardiovascular disease (CVD) mortality rates are considerably higher in countries of Central and Eastern Europe (CEE) and Former Soviet Union (FSU) compared Western Europe (WE) [1]. Differences in diet quality between the two regions, fruit and vegetable consumption in particular, has been one of the proposed explanations for this health gap [2–5].

The lack of internationally comparable, individual-level dietary data in Europe is a well-known problem in public health nutrition [6–9]. In 2011, the European Food Safety Authority (EFSA) published the Comprehensive European Food Consumption Database of food consumption data for most EU member states collected by national dietary surveys of individual-level intakes. However, the authors emphasised that due to the differences in data collection methods, the database was not suitable for international comparisons [10]. Other than the differences in dietary assessment methods, the lack of uniform food-grouping and coding system, and differences

* Correspondence: denes.stefler.10@ucl.ac.uk
Department of Epidemiology and Public Health, University College London, 1-19 Torrington Place, London WC1E 6BT, UK

in estimated portion sizes and food composition tables also make the nationally collected and analysed dietary data inadequate for direct country-to-country comparison [7, 8, 11].

Previous systematic reviews of fruit, vegetable and micronutrient intakes in CEE, FSU and WE countries used data from studies which had been conducted separately in the two regions [12, 13]. These reviews found that the methodological differences between studies seriously limited the interpretation of the results, and emphasised that the lack of comparable data was especially important in CEE and FSU countries. In this respect, cross-national studies which include participants from both CEE/FSU and WE countries, and collect and analyse dietary data in a standardized way, may be therefore more suitable for direct comparisons of food intakes between the two regions.

The aim of this work was to systematically review cross-national studies which reported individual-level data on consumption of fruits, vegetables, or their indicators, such as vitamin C and carotenoids, of participants from CEE/FSU and WE populations using identical methods for data collection and analysis in the two regions.

Methods

Search strategy

MEDLINE, EMBASE and Web of Science databases were searched from inception to September 2014, using search terms described in Appendix 1. References and citation lists of selected papers were studied for additional papers, and hand search of key journals (*Public Health Nutrition, European Journal of Clinical Nutrition, European Journal of Public Health*) was also performed. No restriction on language was applied.

Inclusion and exclusion criteria

Original, quantitative, observational epidemiological studies which described fruit, vegetable, antioxidant intakes or antioxidant status of adult participants who live in CEE or FSU countries and provided comparison populations from Western Europe were included in the review. Based on the data collection methods and reported dietary data, the following studies were considered for inclusion: (1) Dietary surveys: studies which reported data on fruit and vegetable intake levels using established nutritional assessment methods such as food frequency questionnaire (FFQ), diet history, dietary record and 24-h diet recall. (2) Health behavioural surveys: reporting data on fruit and vegetable intakes using lifestyle questionnaires with questions regarding fruit or vegetable consumption habits. (3) Antioxidant studies: reporting data on average vitamin C or carotenoid intakes or status (including plasma, serum and adipose tissue concentrations).

Studies were excluded if data collection methods or the inclusion criteria of participants differed substantially between the two regions. Studies which compared dietary habits between the former East and West Germany were used only if their data collection took place before 1991, because food consumption patterns of East Germans seem to have changed rapidly after the reunification [14].

To avoid bias towards studies which reported more than one exposure of interest from the same participants, we included only one set of data from these studies in the review: data on carotenoid and vitamin C intake or status were included only if no data on fruit or vegetable consumption were available. If both antioxidant intake and status were reported, only intake data was used, and if data on more than one type of carotenoid concentration were available, only beta-carotene was extracted.

Quality assessment

Quality of the included studies was assessed by a shortened version of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement [15]. Modification of the checklist was necessary because several studies described only the nutritional characteristics of the subjects and the analysis of the relationship with disease outcomes was not reported. Therefore four items of the statement, which refer to the variables and outcome results of an analytic study (item nos. 7, 11, 15 and 16), were omitted and the assessment was carried out using the remaining 18 items.

Data analysis

Most studies described dietary data of participants from more than one country within a certain region. For these studies, the average values for CEE/FSU and WE were calculated and reported in the review.

To take into account the well-documented difference in fruit and vegetable consumption between Northern and Southern European countries [16, 17], both CEE/FSU and WE regions were divided into "south" and "north" subregions (Table 1). If a study reported g/day intake levels of fruits or vegetables of participants from opposite subregions, north/south weighting was applied: the intake figure of the "south" country was multiplied with a weighting factor calculated from FAO data [18] by dividing the average fruit or vegetable supply of all northern countries of that region between 1970 and 2009 by the specific country's average supply over the same time period. For studies reporting data on the percentages of participants eating daily fruits or vegetables, or antioxidant data, no such weighting was carried out because appropriate weighting factors were not available.

If data were collected in winter or spring months in one region and during summer or autumn in the other, seasonal weighting of the CEE/FSU data was applied: the

Table 1 Grouping of Central and Eastern European (CEE)/former Soviet Union (FSU) and Western European (WE) countries

Region	Sub-region	Countries
CEE/FSU	North	Armenia, Azerbaijan, Belarus, Czech Republic, Estonia, Georgia, Hungary, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Poland, Republic of Moldova, Romania, Russian Federation, Slovakia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan
	South	Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Montenegro, Serbia, Slovenia, TFYR Macedonia
WE	North	Austria, Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Liechtenstein, Luxembourg, Netherlands, Norway, Sweden, Switzerland, United Kingdom
	South	Andorra, Greece, Italy, Portugal, San Marino, Spain

intake figures were multiplied with a weighting factor which was calculated from the Health Alcohol and Psychosocial Factors in Eastern Europe (HAPIEE) study, which is the largest study in CEE/FSU with dietary data [19]. The weighting factor was determined as the ratio of the energy standardized mean intake level between participants who completed the questionnaire in the summer/autumn months and those who completed it during the winter or spring months. Weighting for seasonal variation was applied only in CEE/FSU because seasonal differences in this region are more substantial than in Western Europe [5, 20, 21].

Most reviewed studies did not report statistical significance of the differences between CEE/FSU and WE. In order to assess whether the reported differences were statistically significant, power calculation was applied. If a study had more than 80 % power to show the described difference as statistically significant on the 0.05 significance level, we considered the reported difference statistically significant. If the power was between 20 % and 80 %, we considered that the observed difference was non-significant but the trend was worth noting, and if the power was lower than 20 %, the difference was considered negligible. Power calculations were carried out using STATA 12.1 statistical software (StataCorp Texas, USA).

If standard deviation (SD) value was required for power calculation but it was not available from the specific study [22–27], the average SD of fruit, vegetable, vitamin C and beta-carotene intake and concentration levels reported in the European Prospective Investigation into Cancer and Nutrition (EPIC) study cohorts was assumed [16, 28]. We considered this assumption appropriate because EPIC is the largest international study with such data available and its results suggest that SD values vary in a narrow range irrespectively of study size and mean intake level. In the study which measured adipose tissue beta-carotene concentration [29] the SD reported on a subsample of the same study participants were used [30]. In studies where south/north or seasonal weighting was applied, SD values were multiplied with the same figures as the mean values.

Results

Characteristics of the reviewed studies

Twenty-two studies met the inclusion criteria: ten dietary surveys [22–26, 31–35], six health behavioural surveys [36–41] and six antioxidant studies [27, 29, 42–45]. Fig. 1 shows the study selection process and Table 2 (see Additional file 1) describes the main features of the included studies. Most studies were cross-sectional in design or reported cross-sectional data from cohort studies. In two studies [29, 32], data were extracted from case–control setting. Participants from 18 CEE/FSU countries and 18 WE states were included in the comparisons and most countries were covered by more than one study. The earliest study [22] reported data from the early 1960s, while the latest data collection took place in 2010 [41]. Sample sizes ranged from 30 to 85 921 per region. Five studies [22, 29, 31, 42, 43] recruited only males but the majority gave dietary data for both genders. More than half of the studies applied random sampling method at recruitment and eight [26, 33, 37–40, 43, 45] used the general population as the sampling frame.

Overall, the quality of the reviewed studies was good. 15 studies scored 14 or more points on the 18 point scale and only two [22, 44] scored less than ten points. Quality of one study [40] was not assessed because it was published as an online database, with no peer-reviewed research paper available.

Findings of the reviewed studies

Table 3 (see Additional file 2) shows the average intake, percentage and concentration values of CEE/FSU and WE participants regarding fruit, vegetable and antioxidants reported by the reviewed studies. The directions of the observed differences and the extent of their significance, determined by power calculation, are also summarised.

Most studies reported their results separately for fruits and vegetables and for males and females. Majority of dietary surveys gave average fruit or vegetable consumption values as mean gram per day intakes, and most of the health behavioural surveys as the percentage of the sample who eat these foods at least once a day.

Regarding fruit intake, both dietary and health behavioural surveys showed consistently lower intakes in CEE/FSU compared to WE. Although six out of nine dietary survey comparisons with adequate power found higher vegetable intake in CEE/FSU countries, the estimates were consistently lower in health behavioural

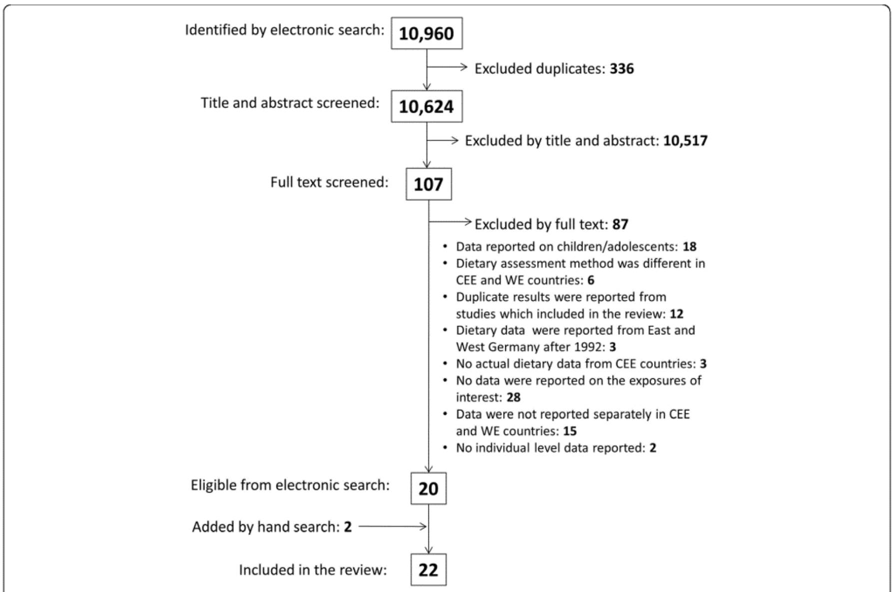


Fig. 1 Flow diagram of the study selection process

surveys. All antioxidant studies indicated lower concentration of beta-carotene in CEE/FSU subjects, but the results for vitamin C were not consistent. No consistent difference was found between males and females.

Discussion

This systematic review of cross-national studies on fruit and vegetable intake found consistently lower fruit intake figures in CEE/FSU populations compared to WE, but no consistent difference for vegetable intake between the two regions.

Our results are congruent with ecological dietary data of food availability based on food balance sheets (FBS) and household budgetary surveys (HBS). Comparison of average fruit and vegetable supply in CEE/FSU and WE countries between 1970 and 2009 suggests clear difference only for fruits but not for vegetables [18]. Similarly, comparison of HBS data from DAFNE database indicates that, on average, the availability of fruits is lower but vegetables is higher in CEE/FSU countries [46].

The inconsistency of our findings regarding vegetable intake can be due to the lack of north/south weighting of health behavioural survey results. For example, in the European Health Interview Survey (EHIS), the largest health behavioural survey included in the review, most participants came from southern countries of Western Europe and northern part of CEE/FSU. If, as a sensitivity analysis, we applied the weighting factors calculated from FAO database for the EHIS results, the comparison showed that the proportion of individuals who consumed vegetables at least once a day was higher in CEE/FSU countries, which is similar to most dietary surveys.

On the other hand, most health behaviour surveys had larger sample size than the dietary surveys, and they are also less prone to measurement error. Furthermore, since the main food sources of beta-carotene are vegetables [47], the findings of the antioxidant studies are also in support of the health behavioural survey results and the lower vegetable intake in Eastern Europe.

On the whole, we cannot exclude the possibility that the reason for the inconsistent results regarding vegetable consumption is that there is no actual difference in intake between CEE/FSU and WE populations.

Our review has several limitations. Firstly, it is possible that further published or non-published studies exist

Table 2 Characteristics of included studies

1st author, year of publication	Name of study	Examined food or antioxidant	Dietary assessment	Participants' country of origin	Year of data collection	Month of data collection	Sample size	Response rate (%)	Females (%)	Age range or mean (years)	Sampling method	Basis of sample	Quality score* (max:18)
1. DIETARY SURVEYS													
Kromhout 1989 [22]	Seven Countries Study	Fruits, vegetables	7d record	CEE: Yugoslavia	1960-64	Jan-May, Sep	150	nd	0	40-59	random	farm/factory workers, academics	9
				WE: Finland, Italy, Greece Netherlands	1959-65	Feb-Sep	286	nd	0	40-59	random	village inhabitants, railroad workers	
Winkler 1992[31]		Fruits, vegetables	3d record	CEE: GDR	1987	Oct-Dec	132	73	0	45-64	random	urban inhabitants	11
				WE: FDR	1984-85	Oct-May	424	70	0	45-64	cluster	urban inhabitants	
Schroll 1996 [23]	SENECA	Fruits, vegetables	Diet history	CEE: Poland	1993	Jan-Jun	120	51†	61	74-79	random	urban inhabitants	13
				WE: Belgium, Denmark, France, Italy, Netherlands, Portugal, Spain, UK, Switzerland	1993	Jan-Jun	1237	51†	51	74-79	random	urban inhabitants	
Karamanos 2002 [24]		Fruits, vegetables	Diet history	CEE: Bulgaria	nd	nd	288	nd	50	35-60	random	urban inhabitants	14
				WE: Italy, Greece	nd	nd	1058	nd	54	35-60	random	urban and rural inhabitants	
Serra-Majem 2003 [25]	WHO-CINDI	Fruits, vegetables	24hr recall	CEE: Poland	1991-94	nd	4440	nd	50	20-65	random	factory workers	14
				WE: Spain	1992	nd	2757	69	nd	6-75	random	general population	
Petkeviciene 2009 [26]	NORBAGREEN	Fruits, vegetables	FFQ	CEE: Lithuania	2002	Apr	99	68	57	19-75	random	general population	15
				WE: Finland	2002	Jan-May	125	91	nd	25-64	random	general population	
Lixandru 2010 [32]		Fruits, vegetables	FFQ	CEE: Romania	2005	Apr-Nov	40	nd	30	63	convenience	diabetic patients	12
				WE: Belgium	2005	Apr-Nov	30	nd	20	62	convenience	diabetic patients	
Paalanen 2011 [33]		Fruits, vegetables	FFQ	CEE: Russia	1992-07	Mar-May	2672	45-92	57	25-64	random	general population	16
				WE: Finland	1992-02	Mar-May	4365	67-81	53	25-64	random	general population	
Crispim 2011 [34]	EFCOVAL	Fruits, vegetables	24hr recall	CEE: Czech Republic	2007-08	Oct-Apr	118	nd	51	45-65	convenience	healthy individuals	16
				WE: Belgium, France, Norway Netherlands,	2007-08	Apr-Jul, Oct-Apr	482	nd	50	45-65	convenience	healthy individuals	
El Ansari 2012 [35]	CNSHS	Fruits, vegetables	FFQ	CEE: Bulgaria, Poland	2005	nd	1143	95	70	21	convenience	university students	14
				WE: Denmark, Germany	2005	nd	1236	85-92	53	21	convenience	university students	
2. HEALTH BEHAVIOUR SURVEYS													
Wardle 1997 [36]	EHBS	Fruits	na	CEE: Poland, Hungary, GDR	1989-92	nd	2293	90-100	51	22	convenience	university students	13

Table 2 Characteristics of included studies *(Continued)*

				WE:	Austria, Belgium, FDR, UK Denmark, Finland, Spain, France, Greece, Iceland, Ireland, Italy, Sweden, Netherlands, Norway, Portugal, Switzerland	1989-92	nd	14192	90-100	56	21	convenience	university students	
Prattala 2007 [37]	Finbalt Health Monitor project	Fruits	na	CEE:	Estonia, Latvia, Lithuania	1998-02	Apr-May	15740	62-80	57	20-64	random	general population	16
				WE:	Finland	1998-02	Apr-May	9354	65-70	53	20-64	random	general population	
Prattala 2009 [38]	EUROTHINE	Vegetables	na	CEE:	Estonia, Latvia, Lithuania	2000-04	nd	14219	60-73	58	20-64	random	general population	15
				WE:	Finland, Denmark, Spain, Germany, France, Italy	1998-04	nd	86924	61-87	51	20-64	random	general population	
Hall 2009 [39]	WHS	Fruits, vegetables	na	CEE:	Bosnia and Herzegovina, Croatia, Czech Republic, Estonia, Georgia, Hungary, Kazakhstan, Latvia, Russia, Slovakia, Slovenia, Ukraine	2002-03	nd	22475	69-100	53	18-99	random	general population	15
				WE:	Spain	2002-03	nd	5448	86	60	18-99	random	general population	
European Commission 2013 [40]	EHIS	Fruits, vegetables	na	CEE:	Bulgaria, Czech Republic, Estonia, Latvia, Hungary, Poland, Romania, Slovakia, Slovenia	2006-09	nd	85921	56-89	53	15-99	random	general population	na
				WE:	Belgium, Greece, Spain, France	2006-09	nd	62700	60-96	55	15-99	random	general population	
Burisch 2014 [41]	ECCO-EpiCom	Fruits, Vegetables	na	CEE:	Croatia, Czech Rep, Estonia, Hungary, Lithuania, Moldova, Romania, Russia	2010	Jan-Dec	249	76†	42	15+	convenience	IBD patients (at diagnosis)	16
				WE:	Cyprus, Denmark, Finland, Greece, Iceland, Ireland, Israel, Italy, Portugal, Spain, Sweden, UK	2010	Jan-Dec	933	76†	46	15+	convenience	IBD patients (at diagnosis)	
3. ANTIOXIDANT STUDIES														
Kardinaal 1993 [29]	EURAMIC	Beta-carotene in adipose tissue	na	CEE:	Russia	1991-92	nd	200	79-97	0	51	convenience	hospital patients, healthy controls	16
				WE:	Finland, Germany, Netherlands, Norway, UK, Spain, Switzerland	1991-92	nd	1180	50-98	0	54	convenience	hospital patients, healthy controls	

Stefler and Bobak *Archives of Public Health* (2015) **73**:29

Table 2 Characteristics of included studies *(Continued)*

Kristenson 1997 [42]	LiVicordia	Beta-carotene in plasma	na	CEE:	Lithuania	1993-94	Oct-Jun	100	83	0	50	random	urban inhabitants	14
				WE:	Sweden	1993-94	Oct-Jun	95	83	0	50	random	urban inhabitants	
Bobak 1998 [27]		Beta-carotene in plasma	na	CEE:	Czech Republic	1992	Sep-Nov	136	70	49	40-59	random	urban inhabitants	14
				WE:	UK	1991-93	nd	358	73	31	40-59	random	civil servants	
Bobak 1999 [43]		Beta-carotene in plasma	na	CEE:	Czech Republic	1995	Apr-Jun	188	70	0	45-64	random	general population	17
				WE:	Germany	1995	Apr-Jun	153	70	0	45-64	random	general population	
Miere 2007 [44]		Vitamin C intake	24h recall	CEE:	Romania	nd	nd	312	nd	87	21	convenience	university students	8
				WE:	Spain	nd	nd	918	nd	58	22	convenience	university students	
Woodside 2013 [45]	EUREYE	Vitamin C and Beta-carotene in plasma	na	CEE:	Estonia	2000-03	nd	833	58.6	66	65+	random	general population	15
				WE:	Norway, UK, France, Italy, Greece, Spain	2000-03	nd	3300	36-56	52	65+	random	general population	

WHO-CINDI, World Health Organization Countrywide Integrated Non-communicable Disease Intervention; NORBAGREE, Consumption of vegetables and fruits and other dietary health indicator foods in the Nordic and Baltic countries; EFCOVAL, European Food Consumption Validation; CNSHS, Cross National Student Health Survey; EHBS, European Health and Behaviour Survey; WHS, World Health Survey; EHIS, European Health Interview Survey; EURAMIC, European Community Multicentre Study on Antioxidants, Myocardial Infarction and Breast Cancer; LiVicordia, Linkoping-Vilnius Coronary Disease Risk Assessment Study; ECCO-EpiCom, European Crohn's and Colitis Organization's Epidemiological Committee study; FDR, Federal Republic of Germany; GDR, German Democratic Republic; CEE: Central and Eastern Europe (or Former Soviet Union); WE, Western Europe; FFQ, Food frequency questionnaire; na, not applicable; nd, no data available; IBD, Inflammatory bowel disease
*Based on evaluation using a modified STROBE checklist; †Overall response rate

Table 3 Summary results of the included studies

1st author, year of publication	Unit of measurement	Sex	CEE countries			WE countries			Power	Summary: CEE compared to WE‡
			Average intake, cc. or %	Range*	SD	Average intake, cc. or %	Range*	SD		
1. DIETARY SURVEYS										
FRUITS										
Kromhout 1989 [22]§‖	g/day intake	M	58.6	1.0-153.6	207.3†	132.1	21.3-310.9	178.3†	0.96	**LOWER**
Winkler 1992 [31]	g/day intake	M	98.0		145.3	101.0		164.3	0.05	no difference
Schroll 1996 [23]§	g/day intake	M	186.0		239.1†	234.0	120.0-532.5	230.2†	0.26	lower-ns
		F	162.0		210.2†	208.0	135.0-399.6	202.4†	0.43	lower-ns
Karamanos 2002 [24]	g/day intake	M	293.0		239.1†	315.0	236.0-355.0	239.1†	0.16	no difference
		F	303.0		210.2†	325.7	234.0-377.0	210.2†	0.21	lower-ns
Serra-Majem 2003 [25]§	g/day intake	M+F	137.0		224.7†	290.0		218.0†	1.00	**LOWER**
Petkeviciene 2009 [26]	p/month intake	M+F	20.8		84.3†	29.4		84.3†	0.12	no difference
Lixandru 2010 [3]	% eat daily	M	100.0		na	89.5		na	0.34	higher-ns
		F	100.0		na	100.0		na	na	no difference
Paalanen 2011 [33]	% eat daily	M	14.0	2.0-31.0	na	52.3	43.0-61.0	na	1.00	**LOWER**
		F	26.0	4.0-50.0	na	73.3	66.0-82.0	na	1.00	**LOWER**
Crispim 2011 [34]	g/day intake	M	207.0		176.7	197.0	163.0-228.0	175.1	0.07	no difference
		F	226.0		155.7	230.5	194.0-265.0	151.1	0.05	no difference
El Ansari 2012 [35]	% eat daily	M	31.6	23.8-39.4	na	30.4	28.6-32.1	na	0.05	no difference
		F	46.8	39.5-54.1	na	51.6	47.8-55.4	na	0.42	lower-ns
VEGETABLES										
Kromhout 1989 [22]§‖	g/day intake	M	240.0	159.0-276.0	198.2†	102.6	57.3-227	88.1†	1.00	**HIGHER**
Winkler 1992 [31]	g/day intake	M	126.0		154.8	124.0		154.8	0.05	no difference
Schroll 1996 [23]§	g/day intake	M	341.0		154.8†	288.0	82.4-461.0	128.1†	0.63	higher-ns
		F	297.0		143.9†	238.0	77.0-383.0	121.0†	0.92	**HIGHER**
Karamanos 2002 [24]	g/day intake	M	243.0		154.8†	189.0	168.0-214.0	154.8†	0.96	**HIGHER**
		F	291.0		143.9†	197.3	178.0-222.0	143.9†	1.00	**HIGHER**
Serra-Majem 2003 [25]§	g/day intake	M+F	288.0		149.4†	97.1		68.7†	1.00	**HIGHER**
Petkeviciene 2009 [26]	p/month intake	M+F	29.9		56.0†	29.1		56.0†	0.05	no difference
Lixandru 2010 [32]	g/day intake	M	287.0		189.4	269.9		108.1	0.07	no difference
		F	258.3		157.9	283.3		125.2	0.06	no difference
Paalanen 2011 [33]	% eat daily	M	15.0	10.0-24.0	na	48.7	44.0-54.0	na	1.00	**LOWER**
		F	22.3	11.0-35.0	na	70.7	69.0-72.0	na	1.00	**LOWER**

Stefler and Bobak *Archives of Public Health* (2015) **73**:29

Table 3 Summary results of the included studies *(Continued)*

Crispim 2011 [34]	g/day intake	M	162.0		121.1	201.0	168.0-222.0	112.8	0.60	lower-ns
		F	157.0		99.1	202.3	166.0-254.0	108.5	0.87	**LOWER**
El Ansari 2012 [35]	% eat daily	M	37.8	23.9-51.6	na	24.4	23.3-25.4	na	0.99	**HIGHER**
		F	44.9	28.0-61.8	na	42.0	37.5-46.4	na	0.18	no difference
2. HEALTH BEHAVIOUR SURVEYS										
FRUITS										
Wardle 1997 [36]	% eat daily	M	40.0	36.0-45.0	na	42.9	23.0-78.0	na	0.43	lower-ns
		F	65.0	59.0-74.0	na	61.1	36.2-86.0	na	0.72	higher-ns
Prattala 2007 [37]	% eat daily	M	11.0	10.0-12.0	na	18.0		na	1.00	**LOWER**
		F	20.3	17.0-25.0	na	36.0		na	1.00	**LOWER**
EHIS 2013 [40]	% eat daily	M	52.8	39.4-66.8	na	60.6	57.9-66.0	na	1.00	**LOWER**
		F	67.0	49.2-82.3	na	69.1	62.3-74.5	na	1.00	**LOWER**
Burisch 2014[41]	% eat daily	M+F	43.4		na	54.3		na	0.87	**LOWER**
VEGETABLES										
Prattala 2009 [38]	% eat daily	M	22.5	16.1-27.5	na	32.1	24.7-39.1	na	1.00	**LOWER**
		F	30.4	25.0-33.4	na	45.9	36.9-59.1	na	1.00	**LOWER**
EHIS 2013 [40]	% eat daily	M	54.8	44.2-71.3	na	68.6	56.0-82.7	na	1.00	**LOWER**
		F	62.5	55.0-78.6	na	74.2	65.3-87.4	na	1.00	**LOWER**
Burisch 2014 [41]	% eat daily	M+F	49.0		na	60.1		na	0.88	**LOWER**
FRUITS AND VEGETABLES										
Hall 2009 [3]	% eat >=5 p/day	M	18.1	8.0-44.5	na	22.0		na	0.98	**LOWER**
		F	23.5	9.4-49.7	na	24.9		na	0.38	lower-ns
3. ANTIOXIDANT STUDIES										
BETA CAROTENE										
Kardinaal 1993 [29]	ug/g fatty acid	M	0.51	0.45-0.56	0.80	0.42	0.18-0.59	0.80	0.31	higher-ns
Kristenson 1997 [42]	umol/l cc.	M	0.38		0.20	0.51		0.32	0.92	**LOWER**
Bobak 1998 [27]	umol/l cc.	M	0.39		0.26†	0.77		0.26†	1.00	**LOWER**
		F	0.52		0.40†	0.97		0.40†	1.00	**LOWER**
Bobak 1999 [43]	umol/l cc**	M	0.11		0.08	0.20		0.21	1.00	**LOWER**
Woodside 2013 [45]	umol/l cc	M	0.25		0.26	0.34	0.19-0.48	0.31	1.00	**LOWER**
		F	0.36		0.34	0.44	0.30-0.67	0.37	1.00	**LOWER**
VITAMIN C										
Miere 2007 [44]	mg/day intake	M	80.3		54.8	106.2		83.4	0.77	lower-ns

Stefler and Bobak *Archives of Public Health* (2015) **73**:29

Table 3 Summary results of the included studies *(Continued)*

		F	88.8	67.9	124.4		94.8	1.00	**LOWER**
Woodside 2013 [45]	umol/l cc.	M	42.0	23.8	38.0	32.7-44.4	23.1	0.74	higher-ns
		F	54.5	27.7	48.5	43.5-52.4	23.4	1.00	**HIGHER**

M, Males; F, Females; p, portion; EHIS, European Health Interview Survey; na, not applicable; cc., concentration

*Range of intake levels, percentages or concentrations if data was reported from more than one country or site

†SD assumed from EPIC study

‡**LOWER**: Intake level, percentage or concentration significantly **lower** in CEE/FSU countries compared to data from WE, (power > 0.80); **HIGHER**: Intake level, percentage or concentration significantly **higher** in CEE/FSU countries compared to data from WE, (power > 0.80); **lower-ns**: Intake level, percentage or concentration **lower** in CEE/FSU but difference not significant (power < 0.80 and >0.20); **higher-ns**: Intake level, percentage or concentration **higher** in CEE/FSU but difference not significant (power < 0.80 and >0.20); **no difference**: power < 0.20

§North–south weighting was applied

I:Seasonal weighting was applied

**Calculated from reported data using molar mass = 537 g which we did not identify during the search. However, cross-national studies tend to require substantial funding, logistics and international cooperation between institutions, which often go hand in hand with the endeavour to publish the work in internationally reputable journals which can be found in the electronic databases we searched. In addition, as we applied no language restriction in the electronic search, the possibility of finding studies from non-English speaking countries was increased.

Secondly, our data analysis involved several assumptions. The weighting factors from FAO database and HAPIEE study were the best options currently available for these purposes, and the SD values brought over from EPIC study did not influence the direction of the results, it only helped to decide whether the studies were sufficiently large to draw meaningful conclusions of their findings.

Although the reviewed studies included participants from a large number of CEE/FSU and WE countries, some of them providing nationally representative food consumption data, specific comparisons were representative only for a small proportion of the whole CEE/FSU and WE populations. Because large differences exist in fruit and vegetable intakes within the regions, the reported comparisons can only be seen as pixels of a much larger picture. The complete picture will emerge only when nationally representative, comparable dietary data is available for most European countries; in fact, this is the main aim of EFSA's on-going "EU Menu" project [48].

Conclusion

This systematic review supports previous data that people in CEE/FSU countries consume less fruit than Western Europeans, and that the difference in vegetable intake is probably less clear-cut. Since inadequate consumption of fruit is suggested as a modifiable risk factor for CVD [49, 50], the difference in fruit intake may contribute to the gap in CVD mortality rates between the two regions.

Additional files

Additional file 1: Characteristics of included studies.
Additional file 2: Summary of results of included studies.
Additional file 3: Search terms used for MEDLINE search.

Abbreviations

CV: Cardiovascular disease; CEE: Central and Eastern Europe; DAFNE: Data Food Networking; EFSA: European Food Safety Authority; EHIS: European Health Interview Survey; EPIC: European Prospective Investigation into Cancer and Nutrition study; FAO: Food and Agriculture Organization; FBS: Food balance sheet; FFQ: Food frequency questionnaire; FSU: Former Soviet Union; HAPIEE: Health Alcohol and Psychosocial Factors in Eastern Europe study; HBS: Household budgetary Survey; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; WE: Western Europe.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

DS carried out the literature search, quality assessment of reviewed studies and the data analysis. DS and MB wrote the manuscript. Both authors read and approved the final manuscript

Acknowledgements

The work was partly supported by the Russian Scientific Foundation (# 14-45-00030). D Stefler was supported by the British Heart Foundation.

Received: 24 November 2014 Accepted: 5 May 2015
Published: 15 June 2015

References

1. WHO Regional Office for Europe (2014). Health for All Database. http://data.euro.who.int/hfadb/
2. Ginter E. Cardiovascular risk factors in the former communist countries. Analysis of 40 European MONICA populations. Eur J Epidemiol. 1995;11:199–205.
3. Pomerleau J, McKee M, Lobstein T, Knai C. The burden of disease attributable to nutrition in Europe. Public Health Nutr. 2003;6:453–61.
4. Powles JW, Zatonski W, Vander Hoorn S, Ezzati M. The contribution of leading diseases and risk factors to excess losses of healthy life in Eastern Europe: burden of disease study. BMC Public Health. 2005;5:116.
5. Zatonski WA. Epidemiological analysis of health situation development in Europe and its causes until 1990. Ann Agric Environ Med. 2011;18:194–202.
6. West CE. Eurofoods: towards compatibility of nutrient data banks in Europe. Ann Nutr Metab. 1985;29 Suppl 1:1–72.
7. de Boer EJ, Slimani N, Boeing H, Feinberg M, Leclercq C, Trolle E, et al. Rationale and methods of the European Food Consumption Validation (EFCOVAL) Project. Eur J Clin Nutr. 2011;65 Suppl 1:S1–4.
8. Charrondiere UR, Vignat J, Moller A, Ireland J, Becker W, Church S, et al. The European Nutrient Database (ENDB) for Nutritional Epidemiology. J Food Compost Anal. 2002;15:435–51.
9. WHO Regional Office for Europe (2010): Report of the Workshop on integration of data on household food availability and individual dietary intakes. Copenhagen. 1–39
10. European Food Safety Authority. Use of the EFSA Comprehensive European Food Consumption Database in Exposure Assessment. EFSA J. 2011;9:2097.
11. Ireland J, van Erp-Baart AM, Charrondiere UR, Møller A, Smithers G, Trichopoulou A, et al. Selection of a food classification system and a food composition database for future food consumption surveys. Eur J Clin Nutr. 2002;56 Suppl 2:S33–45.
12. Lesser S, Pauly L, Volkert D, Stehle P, Ageing Nutrition Group. Nutritional situation of the elderly in Eastern/Baltic and Central/Western Europe - the AgeingNutrition project. Ann Nutr Metab. 2008;52 Suppl 1:62–71.
13. Novakovic R, Cavelaars AE, Bekkering GE, Roman-Viñas B, Ngo J, Gurinović M, et al. Micronutrient intake and status in Central and Eastern Europe compared with other European countries, results from the EURRECA network. Public Health Nutr. 2013;16:824–40.
14. Winkler G, Brasche S, Doring A, Heinrich J. Dietary intake of middle-aged men from an East and a West German city after the German reunification: do differences still exist? Eur J Clin Nutr. 1998;52:98–103.
15. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Epidemiology. 2007;18:805–35.
16. Agudo A, Slimani N, Ocke MC, Naska A, Miller AB, Kroke A, et al. Consumption of vegetables, fruit and other plant foods in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts from 10 European countries. Public Health Nutr. 2002;5:1179–96.
17. Trichopoulou A, Naska A, Costacou T. Disparities in food habits across Europe. Proc Nutr Soc. 2002;61:553–8.
18. Food and Agriculture Organization of the United Nations (2014): FAOSTAT. http://faostat3.fao.org/faostat-gateway/go/to/home/E
19. Peasey A, Bobak M, Kubinova R, Malyutina S, Pajak A, Tamosiunas A, et al. Determinants of cardiovascular disease and other non-communicable diseases in Central and Eastern Europe: rationale and design of the HAPIEE study. BMC Public Health. 2006;6:255.

20. Powles JW, Day NE, Sanz MA, Bingham SA. Protective foods in winter and spring: a key to lower vascular mortality? Lancet. 1996;348:898–9.
21. Capita R, Alonso-Calleja C. Differences in reported winter and summer dietary intakes in young adults in Spain. Int J Food Sci Nutr. 2005;56:431–43.
22. Kromhout D, Keys A, Aravanis C, Buzina R, Fidanza F, Giampaoli S, et al. Food consumption patterns in the 1960s in seven countries. Am J Clin Nutr. 1989;49:889–94.
23. Schroll K, Carbajal A, Decarli B, Martins I, Grunenberger F, Blauw YH, et al. Food patterns of elderly Europeans SENECA Investigators. Eur J Clin Nutr. 1996;50 Suppl 2:S86–100.
24. Karamanos B, Thanopoulou A, Angelico F, Assaad-Khalil S, Barbato A, Del Ben M, et al. Nutritional habits in the Mediterranean Basin. The macronutrient composition of diet and its relation with the traditional Mediterranean diet. Multi-centre study of the Mediterranean Group for the study of diabetes (MGSD). Eur J Clin Nutr. 2002;56:983–91.
25. Serra-Majem L, MacLean D, Ribas L, Brulé D, Sekula W, Prattala R, et al. Comparative analysis of nutrition data from national, household, and individual levels: Results from a WHO-CINDI collaborative project in Canada, Finland, Poland, and Spain. J Epidemiol Community Health. 2003;57:74–80.
26. Petkeviciene J, Simila M, Becker W, Kriaucioniene V, Valsta LM. Validity and reproducibility of the NORBAGREEN food frequency questionnaire. Eur J Clin Nutr. 2009;63:141–9.
27. Bobak M, Brunner E, Miller NJ, Skodová Z, Marmot M. Could antioxidants play a role in high rates of coronary heart disease in the Czech Republic? Eur J Clin Nutr. 1998;52:632–6.
28. Al-Delaimy WK, van Kappel AL, Ferrari P, Slimani N, Steghens JP, Bingham S, et al. Plasma levels of six carotenoids in nine European countries: report from the European Prospective Investigation into Cancer and Nutrition (EPIC). Public Health Nutr. 2004;7:713–22.
29. Kardinaal AFM, Kok FJ, Ringstad J, Gomez-Aracena J, Mazaev VP, Kohlmeier L, et al. Antioxidants in Adipose-Tissue and Risk of Myocardial-Infarction - the Euramic Study. Lancet. 1993;342:1379–84.
30. Su LC, Bui M, Kardinaal A, Gomez-Aracena J, Martin-Moreno J, Martin B, et al. Differences between plasma and adipose tissue biomarkers of carotenoids and tocopherols. Cancer Epidemiol Biomarkers Prev. 1998;7:1043–8.
31. Winkler G, Holtz H, Doring A. Comparison of food intakes of selected populations in former East and West Germany: Results from the MONICA projects Erfurt and Augsburg. Ann Nutr Metab. 1992;36:219–34.
32. Lixandru D, Mohora M, Coman A, Stoian I, van Gils C, Aerts P, et al. Diet and paraoxonase 1 Enzymatic activity in diabetic foot patients from Romania and Belgium: Favorable association of high flavonoid dietary intake with arylesterase activity. Ann Nutr Metab. 2010;56:294–301.
33. Paalanen L, Prattala R, Palosuo H, Laatikainen T. Socio-economic differences in the consumption of vegetables, fruit and berries in Russian and Finnish Karelia: 1992–2007. Eur J Public Health. 2011;21:35–42.
34. Crispim SP, Geelen A, Souverein OW, Hulshof PJ, Ruprich J, Dofkova M, et al. Biomarker-based evaluation of two 24-h recalls for comparing usual fish, fruit and vegetable intakes across European centers in the EFCOVAL Study. Eur J Clin Nutr. 2011;65 Suppl 1:S38–47.
35. El Ansari W, Stock C, Mikolajczyk RT. Relationships between food consumption and living arrangements among university students in four European countries - a cross-sectional study. Nutr J. 2012;11:28.
36. Wardle J, Steptoe A, Bellisle F, Davou B, Reschke K, Lappalainen R, et al. Healthy dietary practices among European students. Health Psychol. 1997;16:443–50.
37. Prattala R, Paalanen L, Grinberga D, Helasoja V, Kasmel A, Petkeviciene J. Gender differences in the consumption of meat, fruit and vegetables are similar in Finland and the Baltic countries. Eur J Public Health. 2007;17:520–5.
38. Prattala R, Hakala S, Roskam AJR, Roos E, Helmert U, Klumbiene J, et al. Association between educational level and vegetable use in nine European countries. Public Health Nutr. 2009;12:2174–82.
39. Hall JN, Moore S, Harper SB, Lynch JW. Global variability in fruit and vegetable consumption. Am J Prev Med. 2009;36:402–9.
40. European Commission (2013): European Health Interview Survey. *Eurostat*. http://ec.europa.eu/eurostat/data/database
41. Burisch J, Pedersen N, Cukovic-Cavka S, Turk N, Kaimakliotis I, Duricova D, et al. Environmental factors in a population-based inception cohort of inflammatory bowel disease patients in Europe–an ECCO-EpiCom study. J Crohns Colitis. 2014;8:607–16.
42. Kristenson M, Zieden B, Kucinskiene Z, Elinder LS, Bergdahl B, Elwing B, et al. Antioxidant state and mortality from coronary heart disease in Lithuanian and Swedish men: concomitant cross sectional study of men aged 50. BMJ. 1997;314:629–33.
43. Bobak M, Hense HW, Kark J, Kuch B, Vojtisek P, Sinnreich R, et al. An ecological study of determinants of coronary heart disease rates: a comparison of Czech, Bavarian and Israeli men. Int J Epidemiol. 1999;28:437–44.
44. Miere D, Filip L, Indrei LL, Soriano JM, Molto JC, Manes J. Nutritional assessment of the students from two European university centers [Romanian] Evaluarea nutritionala a studentilor din doua centre universitare Europene. Rev Med Chir Soc Med Nat Iasi. 2007;111:270–5.
45. Woodside JV, Young IS, Gilchrist SE, Vioque J, Chakravarthy U, de Jong PT, et al. Factors associated with serum/plasma concentrations of vitamins A, C, E and carotenoids in older people throughout Europe: the EUREYE study. Eur J Nutr. 2013;52:1493–501.
46. National and Kapodistrian University of Athens (2005): Data Food Networking (DAFNE) databank. http://www.nut.uoa.gr/dafnesoftweb/
47. Jenab M, Salvini S, van Gils CH, Brustad M, Shakya-Shrestha S, Buijsse B, et al. Dietary intakes of retinol, beta-carotene, vitamin D and vitamin E in the European Prospective Investigation into Cancer and Nutrition cohort. Eur J Clin Nutr. 2009;63 Suppl 4:S150–78.
48. European Food Safety Authority's Advisory Forum on the Pan-European Food Consumption Survey (2010): What's on the menu in Europe? - a Pan-European Food Consumption Survey (EU Menu). Seville. http://www.efsa.europa.eu/en/eumenu/docs/eumenudeclarationen.pdf
49. Dauchet L, Amouyel P, Dallongeville J. Fruits, vegetables and coronary heart disease. Nat Rev Cardiol. 2009;6:599–608.
50. Boeing H, Bechthold A, Bub A, Ellinger S, Haller D, Kroke A, et al. Critical review: vegetables and fruit in the prevention of chronic diseases. Eur J Nutr. 2012;51:637–63.