FRUIT, VEGETABLE, AND LEGUME CONSUMPTION AND CARDIOVASCULAR DISEASE AND MORTALITY IN AN INTERNATIONAL POPULATION

By

Victoria Miller, B.SC (Hons)

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2018)
Hamilton, Ontario (Health Research Methodology)

TITLE: Fruit, vegetable, and legume consumption and cardiovascular disease and mortality in an international population
AUTHOR: Victoria Emily Miller, B.Sc (Hons)
SUPERVISOR: Professor Andrew Mente
NUMBER OF PAGES: xxii, 192

ABSTRACT

Background: Diet is an important modifiable risk factor for cardiovascular disease. Numerous studies have examined the association between dietary intake and cardiovascular disease in North America and Europe, but little information is available on potential associations in many geographic regions including South Asia, South America, Africa, China, and the Middle East. Presently, it is unclear whether the findings from Western countries are applicable to these regions where population characteristics, background diet, and disease risk differ. This thesis aims to investigate the relationship between diet and cardiovascular disease and mortality in a heterogeneous, international population.

Methods: Baseline data from the Prospective Urban Rural Epidemiology (PURE) study was used to investigate the availability, affordability, and consumption of fruits and vegetables. Additionally, PURE baseline and follow-up was used to examine the association between foods (fruits, vegetables and legumes) and macronutrients (total and fat subtypes, carbohydrate and protein) and cardiovascular outcomes and mortality. PURE is a prospective cohort study in individuals aged 35 to 70 years in 18 high-income middle-income and low-income countries on five continents. Availability and affordability of fruits and vegetables was collected from centrally located grocery stores and market places in each PURE community. Diet was measured using country and region-specific food frequency questionnaires at baseline. Case-report forms, death certificates, medical records and verbal autopsies were used to capture data about major cardiovascular events, and death during follow-up. The cost and diversity of fruits and vegetables was documented and mean fruit and vegetable intake by their relative cost was assessed. Associations between fruit, vegetable and legume consumption with risk of cardiovascular outcomes and mortality were examined. We investigated the association between macronutrients and risk of mortality and modeled nutrient replacement using energy-adjustment and joint effect models.

Results: Results from the PURE study indicate that consumption of fruits and vegetables is low worldwide, particularly in low-income countries, and this is associated with low affordability. Higher fruit, vegetable and legume consumption was associated with a lower risk of non-cardiovascular, and total mortality and benefits appear to be maximal at three to four servings per day. This finding indicates that health benefits can be achieved at intake lower than most dietary recommendations, an approach that is likely to be more affordable in poor countries. Higher carbohydrate intake was positively associated with an increased risk of mortality, while total and fat subtypes, and protein was inversely associated with death. For the nutrient replacement analysis, the joint effect method demonstrated higher agreement with the single nutrient results compared to the conventional energy-adjustment method. This result suggests that traditional nutrient replacement modeling is not appropriate for international populations with diverse nutrient intake.

Conclusions: Dietary intake varies across geographic regions and interventions to improve diet and nutrition recommendations should be tailored to the geographic setting.

ACKNOWLEDGEMENTS

Successful completion of my doctoral studies has only been possible because of a large group of supportive individuals who have believed in me over the last four and half years. Sincere appreciation is especially extended to my doctoral thesis committee at McMaster University. I am very fortunate to have been supervised and mentored by Dr. Andrew Mente. He has provided me with guidance through each step of the research process, supported the development of my research skills, and provided encouragement all along the way. In addition, I would like to thank Dr. Koon Teo, Dr. Sonia Anand, and Dr. Noori Akhtar-Danesh for their commitment, support, and insightful comments. I am forever indebted to Dr. Salim Yusuf for his extraordinary mentorship and the countless opportunities he provided me during my graduate training. A special thank you to Dr. Mahshid Dehghan for her willingness to explain the finer details of nutrition and unwavering confidence in my abilities.

Importantly, thank you to an exceptional group of collaborators, colleagues and friends: Sumathy Rangarajan, Maha Mushtaha, Gayle Lewis, Roxanna Solano, Xiaohe Zhang, Chinthanie Ramasundarahettige, Patrick Sheridan, Dr. Shofiqul Islam, Weihong He, Leanne Dyal, Amparo Casanova, Dr. Shrikant Bangdiwala, Dr. Clara Chow, Dr. Daniel Corsi, Dr. Karen Lock, Dr. Rasha Khatib, Dr. Philip Joseph, Dr. Darryl Leong, Dr. MyLinh Duong, Dr. Russell de Souza, Dr. David Meyre, Dr. David Collister, Dr. Stephanie Ross, the PURE Project Office Staff, National Coordinators, Investigators, and Key Staff, the PURE participants, and the Health Research Methodology Program at McMaster University.

I would like to thank my extended family and friends for their never-ending supply of love, encouragement, and patience. I am forever grateful to my husband Jason for supporting my desire to continue my education and always believing in my ability to accomplish any goal.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1. Availability, affordability, and consumption of fruits and vegetables in 18 countries across income levels: findings from the Prospective Urban Rural Epidemiology (PURE) study

Chapter 2. Fruit, vegetable, and legume intake, and cardiovascular disease and deaths in 18 countries (PURE): a prospective cohort study

Chapter 3. Comparison of two methods for modeling nutrient replacement in an international population

LIST OF TABLES

Chapter 1. Availability, affordability, and consumption of fruits and vegetables in 18 countries across income levels: findings from the Prospective Urban Rural Epidemiology (PURE) study

Chapter 2. Fruit, vegetable, and legume intake, and cardiovascular disease and deaths in 18 countries (PURE): a prospective cohort study

Chapter 3. Comparison of two methods for modeling nutrient replacement in an international population

LIST OF APPENDICES

Chapter 1. Availability, affordability, and consumption of fruits and vegetables in 18 countries across income levels: findings from the Prospective Urban Rural Epidemiology (PURE) study

Supplementary methods

Tables

Figures

Chapter 2. Fruit, vegetable, and legume intake, and cardiovascular disease and deaths in 18 countries (PURE): a prospective cohort study

Supplementary methods

Tables

Figures

Chapter 3. Comparison of two methods for modeling nutrient replacement in an international population

Supplementary methods

LIST of ABBREVIATIONS (alphabetical order)

%E- Percent Energy

CI- Confidence Interval

CM- Centimeter

E.g.- Exempli Grati

FFQ- Food Frequency Questionnaire

g- Grams

HIC- High-Income Countries

HR- Hazard Ratio

I.e.- Id Est

IQR- Interquartile Range

Kcal- Kilocalorie

LDL- Low-Density Lipoprotein

LIC- Low-Income Countries

LMIC- Lower-Middle-Income Countries

MET- Metabolic Equivalent of Task

MIC- Middle-Income Countries

PURE- Prospective Urban Rural Epidemiology

Q- Quartile or Quintile

r- Pearson Correlation Coefficient

r_s- Spearman Correlation Coefficient

SD- Standard Deviation

UMIC- Upper-Middle-Income Countries

USA- United States of America

USDA- United Stated Department of Agriculture

DECLARATION OF ACADEMIC ACHIEVEMENT

This doctoral dissertation is a ‘sandwich’ thesis that consists of three chapters, two of which have been published in peer-reviewed medical journals (chapters 1 and 2), and one of which has been prepared for future submission to a peer-reviewed medical journal (chapter 3). I was the first author and main contributor for each chapter. My independent contributions involved leading each chapter’s conception and design; performing analysis and interpretation of data; drafting and revising the manuscripts; providing final approval of the versions published; and agreeing to be accountable for all aspects of each work. My co-authors also made important contributions to each chapter related to conception, interpretation of data; reviewing and revising the manuscripts; and agreeing to be accountable for all aspects of each work. I completed all aspects of this thesis between September 2013 and March 2018.

INTRODUCTION

1.1 Background

Globally, cardiovascular disease is a leading cause of death[1]. The burden of cardiovascular disease disproportionately affects low-income and middle-income countries[2], with 80% of deaths from cardiovascular disease estimated to occur in these regions[3]. Diet is an important modifiable risk factor for cardiovascular disease[4]. Numerous prospective studies have examined the association between diet and disease, but the evidence is largely limited to studies from North America and Western Europe[5-7], with very little data available from other regions of the world including South Asia, South America, Africa, China, and the Middle East. The most recent meta-analyses of observational studies, largely done in North America and Europe, observed reductions in the risk of death for higher fruit and vegetable intake[6], while no association was found for total fat and saturated fat and risk of death from coronary heart disease[8] or all-cause mortality[5]. Presently, it is not known whether the observations from North America and Europe which constitute less than 10 to 15% of the world's population can be extrapolated to other regions where cultural factors, food availability, affordability, cooking methods and patterns[9] differ markedly. Further, disease prevalence (e.g. rates of cardiovascular disease, cancers and infectious diseases) vary between regions[2,10] and other lifestyle factors[2] are markedly dissimilar.

Given the absence of evidence on the optimal levels of food and nutrient intake to minimize disease risk in many geographic regions, dietary guidelines are based largely on data from North America and Europe and may not be appropriate for non-Western populations. The Prospective Urban Rural Epidemiology (PURE) study provides the unique opportunity to examine the relationship between diet and disease risk in a large, international population from seven geographic regions: North America and Europe, South America, the Middle East, South Asia, China, Southeast Asia, and Africa. The diverse population is optimal for investigating relationships across a wide range of intake including under- and over-nutrition, which has not been possible by most previous studies. Further, the PURE cohort is an ideal population for characterizing non-linear associations (i.e., U-shape or threshold) between dietary exposures and health outcomes, which are likely to occur with a broad range of nutrient intake. Lastly, the PURE study collected data on environmental and economic factors that may affect dietary consumption including the availability and affordability of food items, which may be vital to understanding dietary behaviors in low-income and middle-income countries.

1.2 Availability, affordability and consumption of fruits and vegetables

Food cost is hypothesized to be an important predictor of consumption[11,12]. The availability and affordability of essential foods may differ among countries with different levels of economic development. Currently, most nutritional guidelines recommend a minimum of two daily servings of fruits and three daily servings of vegetables[13-15]. However, many individuals do not meet this recommendation, particularly those in low-income and middle-income countries[9]. The availability, affordability and relationship between cost and consumption of fruits and vegetables has not previously been described in many regions of the world. Higher fruit and vegetable cost may substantially constrain intake[16-18] and characterizing affordability is critical for the development of realistic and achievable dietary policies and recommendations.

1.3 Association between fruits and vegetables and disease risk

The association between consumption of fruits, vegetables, and legumes with cardiovascular disease and mortality has been examined extensively in Europe, the USA, Japan, and China, but little or no data are available from the Middle East, South America, Africa or South Asia. The most recent meta-analysis reported reductions in the risk of cardiovascular disease and all-cause mortality up to 800 g/day of fruits and vegetables, with the steepest reduction up to 400 g/day[6]. However, previous studies have shown that the consumption of fruits, vegetables, and legumes is low among many populations, predominantly those in countries outside Europe or the USA[9]. It is unclear whether these food items are beneficially associated with cardiovascular disease and mortality risk in non-Western countries. Further, vegetables might be consumed raw or cooked and the cooking process might alter the bioavailability of nutrients (such as phytochemicals, vitamins, minerals, and fibre), and digestibility[19-21]. In Europe and the USA, vegetables are eaten in both raw and cooked form, while in Asia and in other parts of the world cooked vegetables are frequently consumed. Presently, little information is available on the effect of raw versus cooked vegetable intake on cardiovascular disease and total mortality in low-income and middle-income countries.

1.4 Nutrient replacement and disease risk

Until recently, little information was available on the relationship between macronutrient intake with risk of death outside of North American and Europe populations. In a recent publication from the PURE study, higher carbohydrate intake was associated with an increased risk of total mortality, while total fat and each type of fat (saturated, monounsaturated, polyunsaturated), and protein were inversely associated with a decreased risk[22]. However, nutrients are not consumed in isolation and potential interactions between different nutrients, the concomitant consumption of some (collinearity), or replacement with other nutrients may affect disease risk. In observational studies, the conventional approach for studying the replacement of one nutrient for another nutrient on disease risk is multivariable energy-adjustment models whereby all of the nutrients expressed as a percentage of energy, except for one, are included in a regression model[23]. The regression coefficients are interpreted as the effect of substituting a certain amount of each nutrient for the same amount of nutrient excluded from the model while holding constant the intakes of energy and the other macronutrients[23]. For this method, the relationship between the nutrients and outcome are assumed to be linear, which may not be in case in a global population with a broader range of nutrient intake than observed in previous studies conducted mostly in Western populations. Currently, the appropriateness of using conventional dietary replacement modeling in a global population is unclear.

1.5 Thesis rationale and objectives

This thesis aims to investigate the relationship between diet and cardiovascular disease and mortality in a heterogeneous, international population. It consists of three chapters, the first of which examined the availability, affordability and consumption of fruits and vegetables in 18 high, middle and low-income countries. To our knowledge, this chapter is the first to describe the availability and affordability of fruits and vegetables across economic regions globally and to relate affordability to consumption. Chapter 2 assesses the relationship of fruit, vegetable and legume intake and cardiovascular disease and mortality among 135,335 participants from 18 countries. Presently, this is the only global study relating fruit, vegetable, and legume intake to cardiovascular disease events and mortality and includes a broad range of intake including very low consumption of fruits and vegetables, and high consumption of legumes. Lastly, Chapter 3 investigates the replacement of energy from saturated fat with other energy-contributing nutrients using two methods: 1) the energy-adjustment method and 2) the joint effect method, and contrast the agreement of these approaches with single nutrient models. Although nutrient replacement modeling has been used in North American and European populations, our study is the first to examine its appropriateness for international populations with a diverse range of nutrient intake.

Chapter 1

Availability, affordability, and consumption of fruits and vegetables in 18 countries across income levels: findings from the Prospective Urban Rural Epidemiology (PURE) study

Reprinted from *The Lancet Global Health;* 4; Victoria Miller, Salim Yusuf, Clara K Chow, Mahshid Dehghan, Daniel J Corsi, Karen Lock, Barry Popkin, Sumathy Rangarajan, Rasha Khatib, Scott A Lear, Prem A Mony, Manmeet Kaur, Viswanathan Mohan, Krishnapillai Vijayakumar, Rajeev Gupta, Annamarie Kruger, Lungiswa Tsolekile, Noushin Mahammadifard, Omar Rahman, Annika Rosengren, Alvaro Avezum, Andrés Orlandini, Noorhassim Ismail, Patricio Lopez-Jaramillo, Afzalhussein Yusufali, Kubilay Karsidag, Romaina Iqbal, Jephat Chifamba, Solange Martinez Oakley, Farnaza Ariffin, Katarzyna Zatonska, Paul Poirier, Li Wei, Bo Jian, Chen Hui, Liu Xu, Bai Xiulin, Koon Teo, and Andrew Mente; Availability, affordability, and consumption of fruits and vegetables in 18 countries across income levels: findings from the Prospective Urban Rural Epidemiology (PURE) study; e695-703; Copyright 2016; with permission from Elsevier. doi:10.1016/ S2214-109X(16)30186-3

Availability, affordability, and consumption of fruits and vegetables in 18 countries across income levels: findings from the Prospective Urban Rural Epidemiology (PURE) study

Victoria Miller, Salim Yusuf, Clara K Chow, Mahshid Dehghan, Daniel J Corsi, Karen Lock, Barry Popkin, Sumathy Rangarajan, Rasha Khatib, Scott A Lear, Prem Mony, Manmeet Kaur, Viswanathan Mohan, Krishnapillai Vijayakumar, Rajeev Gupta, Annamarie Kruger, Lungiswa Tsolekile, Noushin Mohammadifard, Omar Rahman, Annika Rosengren, Alvaro Avezum, Andrés Orlandini, Noorhassim Ismail, Patricio Lopez-Jaramillo, Afzalhussein Yusufali, Kubilay Karsidag, Romaina Iqbal, Jephat Chifamba, Solange Martinez Oakley, Farnaza Ariffin, Katarzyna Zatonska, Paul Poirier, Li Wei, Bo Jian, Chen Hui, Liu Xu, Bai Xiulin, Koon Teo, Andrew Mente

Lancet Glob Health 2016; 4: e695–703

See **Comment** page e664

Population Health Research Institute, Hamilton Health Sciences and McMaster University, Hamilton, ON, Canada (V Miller BSc, Prof S Yusuf DPhil, M Dehghan PhD, S Rangarajan MSc, Prof K Teo PhD, Prof A Mente PhD); **Westmead Hospital and the George Institute for Global Health, Sydney University, Sydney, NSW, Australia** (C K Chow PhD); **Ottawa Hospital Research Institute, Ottawa, ON, Canada** (D J Corsi PhD); **Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, London, UK** (Prof K Lock PhD); **Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA** (Prof B Popkin PhD); **Stritch School of Medicine, Loyola University Chicago, Maywood, IL, USA** (R Khatib PhD); **Institute of Community and Public Health, Birzeit University, Ramallah, Occupied Palestinian Territory** (R Khatib); **Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada** (Prof Scott A Lear PhD); **St John's Medical College & Research Institute, Bangalore, India** (Prof P Mony MD); **School of Public Health, Postgraduate Institute of Medical Education and Research, Chandigarh, India** (M Kaur PhD); **Madras Diabetes Research Foundation, Chennai, India** (V Mohan DSc); **Community Medicine, Health Action By People,**

Summary

Background Several international guidelines recommend the consumption of two servings of fruits and three servings of vegetables per day, but their intake is thought to be low worldwide. We aimed to determine the extent to which such low intake is related to availability and affordability.

Methods We assessed fruit and vegetable consumption using data from country-specific, validated semi-quantitative food frequency questionnaires in the Prospective Urban Rural Epidemiology (PURE) study, which enrolled participants from communities in 18 countries between Jan 1, 2003, and Dec 31, 2013. We documented household income data from participants in these communities; we also recorded the diversity and non-sale prices of fruits and vegetables from grocery stores and market places between Jan 1, 2009, and Dec 31, 2013. We determined the cost of fruits and vegetables relative to income per household member. Linear random effects models, adjusting for the clustering of households within communities, were used to assess mean fruit and vegetable intake by their relative cost.

Findings Of 143 305 participants who reported plausible energy intake in the food frequency questionnaire, mean fruit and vegetable intake was 3·76 servings (95% CI 3·66–3·86) per day. Mean daily consumption was 2·14 servings (1·93–2·36) in low-income countries (LICs), 3·17 servings (2·99–3·35) in lower-middle-income countries (LMICs), 4·31 servings (4·09–4·53) in upper-middle-income countries (UMICs), and 5·42 servings (5·13–5·71) in high-income countries (HICs). In 130 402 participants who had household income data available, the cost of two servings of fruits and three servings of vegetables per day per individual accounted for 51·97% (95% CI 46·06–57·88) of household income in LICs, 18·10% (14·53–21·68) in LMICs, 15·87% (11·51–20·23) in UMICs, and 1·85% (–3·90 to 7·59) in HICs (p_{trend}=0·0001). In all regions, a higher percentage of income to meet the guidelines was required in rural areas than in urban areas (p<0·0001 for each pairwise comparison). Fruit and vegetable consumption among individuals decreased as the relative cost increased (p_{trend}=0·00040).

Interpretation The consumption of fruit and vegetables is low worldwide, particularly in LICs, and this is associated with low affordability. Policies worldwide should enhance the availability and affordability of fruits and vegetables.

Funding Population Health Research Institute, the Canadian Institutes of Health Research, Heart and Stroke Foundation of Ontario, AstraZeneca (Canada), Sanofi-Aventis (France and Canada), Boehringer Ingelheim (Germany and Canada), Servier, GlaxoSmithKline, Novartis, King Pharma, and national or local organisations in participating countries.

Introduction

Most nutritional guidelines recommend the consumption of at least two servings of fruits and three servings of vegetables per day.[1,2] However, a large proportion of individuals do not meet these targets.[3–5] An improved understanding of the factors that affect fruit and vegetable consumption is essential to improving the diet quality of populations.

Food cost has been shown to affect dietary intake in developed countries,[6,7] but similar data for low-income countries (LICs) and middle-income countries (MICs) are sparse. High food cost might particularly affect affordability among households spending a considerable proportion of their income on food.[8,9] Increases in the cost of food have been shown to result in food-based coping strategies such as reductions in the quantity, quality, and diversity of food selections, and consumption of increased quantities of cheap, energy-dense foods.[10–12]

Determining the affordability of essential foods such as fruits and vegetables in countries with different levels of economic development is important. In this study, we aimed to document the availability cost of fruits and vegetables in community grocery stores and market places, and the affordability of meeting dietary guidelines

Thiruvananthapuram, Kerala, India (Prof K Vijayakumar MD); **Department of Medicine, Fortis Escorts Hospital, Jaipur, India** (R Gupta MD); **Faculty of Health Science North-West University, Potchefstroom Campus, Potchefstroom, South Africa** (Prof A Kruger PhD); **University of the Western Cape, Bellville, South Africa** (L Tsolekile MPH); **Isfahan Cardiovascular Research Center, Cardiovascular Research Institute, Isfahan University of Medical Sciences, Isfahan, Iran** (N Mohammadifard PhD); **Independent University, Bangladesh Bashundhara, Dhaka, Bangladesh** (Prof O Rahman MD); **Sahlgrenska Academy and Sahlgrenska University Hospital, University of Gothenburg, Gothenburg, Sweden** (Prof A Rosengren MD); **Research Division, Dante Pazzanese Institute of Cardiology, São Paulo, Brazil** (Prof A Avezum MD); **Estudios Clinicos Latino America, Rosario, Argentina** (A Orlandini MD); **Department of Community Health, University Kebangsaan Malaysia Medical Centre, Kuala Lumpur Malaysia** (N Ismail MD); **Grupo Investigaciones FOSCAL, Fundacion Oftalmologica de Santander and Medical School, Universidad de Santander, Bucaramanga, Colombia** (Prof P Lopez-Jaramillo MD); **Hatta Hospital, Dubai Health Authority, Dubai, United Arab Emirates** (Prof A Yusufali MD); **Department of Internal Medicine, Istanbul University, Istanbul, Turkey** (K Karsidag MD); **Department of Community Health Sciences and Department of Medicine, Aga Khan University, Karachi, Pakistan** (R Iqbal PhD); **Physiology Department, University of Zimbabwe College of Health Sciences, Harare, Zimbabwe** (J Chifamba MPhil); **Facultad de Medicina, Universidad de La Frontera, Temuco, Chile** (S Martinez Oakley MSc); **Faculty of Medicine, UiTM Sungai Buloh Campus, Selangor, Malaysia** (F Ariffin MBBS); **Department of Social Medicine, Medical University in Wroclaw, Wroclaw, Poland** (K Zatonska MD); **Laval University Heart and Lungs Institute, Quebec City, QC, Canada** (P Poirier MD); **and**

Research in context

Evidence before this study

We searched PubMed for articles published between Jan 1, 1960, and Jan 15, 2016, using the search term “fruit” OR “vegetable” OR “produce” OR “food” AND “cost” OR “afford*” OR “price” OR “purchasing” OR “availability” OR “diversity”. We used search terms in English but did not apply any language restrictions. We screened papers by title and abstract to identify full-text reports that were relevant to the study aims. We also screened citation lists from these full-text reports to identify other relevant articles. Papers were considered relevant if they report assessment of the relation between fruit and vegetable intake and availability or affordability. The papers cited here were selected to be representative of the existing evidence base and are not an exhaustive list of relevant research. Existing evidence was limited to the affordability of healthy food items in high-income countries. The absolute cost of food items was reported in several papers. However, information on the relative cost and proportion of individuals unable to afford the food items was not described.

Added value of this study

To our knowledge, this study is the first to describe the availability and affordability of fruits and vegetables across economic regions globally and to relate affordability to consumption. Our results show that the consumption of fruits and vegetables is low worldwide, particularly in low-income countries because of low affordability.

Implications of all the available evidence

Most dietary guidelines recommend the consumption of two servings of fruits and three servings of vegetables per day. However, purchasing this recommended amount would require a substantial proportion of household income, making fruits and vegetables unaffordable in many low-income and middle-income countries. Policies that enhance the affordability of fruits and vegetables are crucially needed to meet these recommendations.

for fruit and vegetable consumption in 18 countries with different income levels. We also aimed to relate the affordability of fruits and vegetables to their consumption.

Methods

Study design and sample selection

Between Jan 1, 2003, and Dec 31, 2013, the Prospective Urban Rural Epidemiology (PURE) study enrolled 157 254 adults aged 35–70 years in 667 communities from 18 countries on five continents. Countries were selected from four income strata according to the World Bank classification in 2006 on the basis of gross national income per person. There were four LICs (Bangladesh, India, Pakistan, and Zimbabwe), four lower-middle-income countries (LMICs; China, Colombia, Iran, Occupied Palestinian Territory), seven upper-middle-income countries (UMICs; Argentina, Brazil, Chile, Malaysia, Poland, Turkey, South Africa), and three high-income countries (HICs; Canada, Sweden, United Arab Emirates). A detailed description of participant, community, and country selection has been published elsewhere (appendix pp 4–5).[13,14] In the PURE study, 147 938 participants completed country-specific, validated semi-quantitative food frequency questionnaires (appendix p 6).[15–22] Of these individuals, we included those who had plausible energy intake (500–5000 kcal per day) in our analyses of fruit and vegetable consumption.

For analyses of food availability and affordability, we collected information on the cost of at least one fruit and one vegetable in each PURE community between Jan 1, 2009, and Dec 31, 2013. A 1 km observation walk was done by research staff in a centrally located area within each community. Within each area, non-sale prices (ie, retail prices before any discounts) were collected from the grocery store or market place located in closest proximity to the observation walk zone for the following fruits and vegetables: apples, oranges, bananas, pears, carrots, tomatoes, and cabbage. A checklist of 48 types of fruits and 59 types of vegetable was used to assess the variety of fruits and vegetables available. Additional grocery stores or market places in the 1 km area were visited if research staff were unable to collect the cost of the fruits and vegetables. The total number of types of fruit and vegetable available for sale in each community was calculated to assess the diversity (see appendix p 7 for methods used to estimate fruit and vegetable availability and affordability). Additionally, we collected household income data from participants in these communities (appendix p 8). The methods used to calculate daily income, and fruit and vegetable costs and consumption are shown in appendix p 12. The study variables and their unit of analysis are summarised in appendix pp 13–14.

Statistical analysis

The affordability of two servings of fruits and three servings of vegetables per day was assessed using the least expensive fruit and vegetable available for sale within each community. Additionally, the affordability of purchasing five servings of the cheapest fruit or vegetable was assessed to estimate the most optimistic scenario of affordability that is reflective of substituting either type of produce to reach five daily servings. To define affordability, we used a threshold of less than 20% of household income per household member required to purchase two servings of fruits and three servings of vegetables per day for every household member. We used this demarcation point for affordability because we found that few households in HICs used more than 20% of their income in the purchase of the recommended number of servings. Furthermore, when other various thresholds were explored, we found the same pattern of unaffordability across economic regions (appendix p 23). We also calculated the proportional increase in food expenditure necessary to meet the recommended intake of fruits and vegetables among individuals who did not meet this target.

We used Spearman correlation coefficients to test the strength of the association between country gross national income and mean percentage of household income spent on food. At the community level, we did an analysis of variance, with tests for linear trend, to compare the mean number of different types of fruit and vegetables (ie, diversity) and the mean cost, adjusted by purchasing price parity, of one serving of fruit and vegetables in each economic region. At the individual level, we used linear random effects models with fixed intercepts and random slope, accounting for clustering of households within communities, to examine the mean cost of one serving of fruit and one serving of vegetables in each economic region, with tests for linear trend. Additionally, linear random effects models were used to assess the mean proportion of income per household member required to purchase two servings of fruits and three servings of vegetables in each economic region. We tested for interactions between the association of availability, affordability, and income level, by urban or rural location. We did not account for clustering of individuals within households, since the mean number of participant per household was 1·4 (SD 0·6), so the degree of clustering of individuals within households would be minimal. Finally, linear random effect models with tests for linear trend were used to examine the mean intake of fruit and vegetables by their relative cost (in quartiles), adjusting for energy intake and, in a separate model, further adjusting for age as a continuous variable, and sex and economic region as categorical variables. The association between intake and relative cost was further assessed in subgroup analyses by economic region, with testing for heterogeneity in the overall sample. We used SPSS software (Armonk, NY, USA), version 22.0, for all statistical analyses.

Role of the funding source

The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Of 147 938 PURE study participants who completed the food frequency questionnaires, 143 305 (97%) had plausible energy intake and were included in our analyses of fruit and vegetable intake (table 1). These participants and the participants who were included in community assessments generally had similar characteristics (see appendix pp 16–19 for a summary of total household size

National Center for Cardiovascular Diseases, Fuwai Hospital, Beijing, China
(Prof L Wei PhD, B Jian BSc, C Hui, L Xu MSc, B Xiulin)

Correspondence to:
Prof Andrew Mente, Population Health Research Institute, Hamilton Health Sciences and McMaster University, Hamilton, ON L8L 2X2, Canada
andrew.mente@phri.ca

See Online for appendix

	Entire cohort (n=143 305)	High-income countries (n=15 193)	Upper-middle-income countries (n=36 875)	Lower-middle-income countries (n=59 392)	Low-income countries (n=31 845)
Median age, years (IQR)	50·0 (34·0–66·0)	53·0 (38·0–68·0)	51·0 (35·0–67·0)	51·0 (35·0–67·0)	47·0 (31·0–63·0)
Female sex	83 007 (58%)	8313 (55%)	22 251 (60%)	34 445 (58%)	17 998 (57%)
Education level					
Less than high school graduate	61 122 (43%)	1710 (11%)	19 417 (53%)	23 741 (40%)	16 254 (51%)
High school graduate	54 453 (38%)	4467 (29%)	11 816 (32%)	26 978 (45%)	11 192 (35%)
At least some college education	27 363 (19%)	9004 (59%)	5556 (15%)	8540 (14%)	4263 (13%)
Current smoker	29 852 (21%)	2083 (14%)	8297 (23%)	12 252 (21%)	7220 (23%)
High physical activity*	58 988/129 258 (46%)	7403/13 913 (53%)	14 938/31 491 (47%)	24 280/56 222 (43%)	12 367/27 632 (45%)
Mean body-mass index, kg/m^2 (SD)	25·8 (5·2)	27·6 (5·3)	28·4 (5·9)	25·2 (4·1)	23·3 (4·9)
Median energy intake per day, kcal (IQR)	1991 (964–3020)	2144 (1036–3252)	2057 (936–3178)	1936 (1012–2860)	1969 (891–3047)
Mean vegetable intake, servings per day (95% CI)†	2·19 (2·13–2·25)	3·58 (3·44–3·71)	1·77 (1·65–1·89)	1·96 (1·87–2·05)	1·48 (1·37–1·59)
Mean fruit intake, servings per day (95% CI)†	1·62 (1·53–1·72)	1·99 (1·79–2·19)	2·50 (2·32–2·67)	1·21 (1·06–1·35)	0·80 (0·60–1·01)
Consume ≥1 serving of vegetables per day	114 657 (80%)	14 304 (94%)	28 322 (77%)	50 761 (85%)	21 270 (67%)
Consume ≥1 serving of fruits per day	69 207 (48%)	12 364 (81%)	25 532 (69%)	22 690 (38%)	8621 (27%)

Data are n (%) or n/N (%), unless indicated otherwise. The sample comprised individuals who completed a food frequency questionnaire in the Prospective Urban Rural Epidemiology study and had an energy intake of 500–5000 kcal per day. *Defined as ≥3000 metabolic equivalent of task minutes per week; participants with missing data were excluded from analysis. †Accounting for clustering of households within communities.

***Table 1:* Participant characteristics**

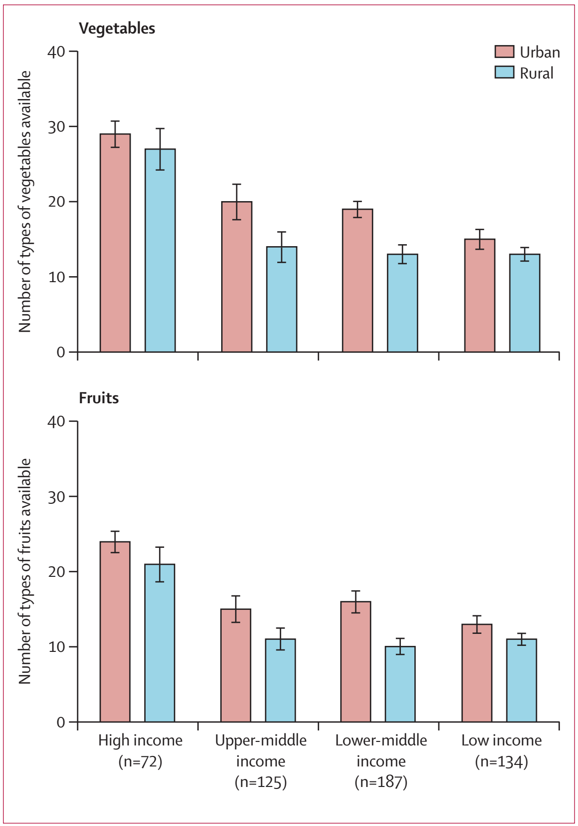


Figure 1: Mean number of types of vegetables and fruits available in urban and rural communities, by economic region
Error bars represent 95% CI.

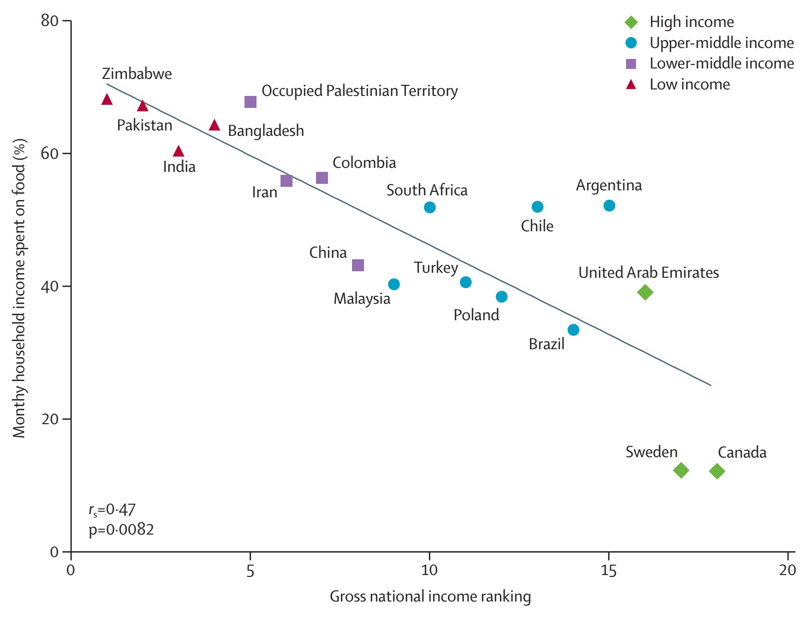


Figure 2: Mean percentage of monthly household income spent on food, by gross national income ranking

and composition, including household members not participating in the PURE study, by country and economic region). The median age of these 143 305 participants was 50·0 years (IQR 34·0–66·0), and men and women were equally represented. The mean body-mass index was 25·8 kg/m² (SD 5·2), 29 852 (21%) of participants were current smokers, and more than half (55%) had low or moderate physical activity levels (<600 or 600–3000 metabolic equivalent of task minutes per week, respectively). Median energy intake was 1991 kcal per day (IQR 964–3020).

Across participants in all countries studied, mean fruit and vegetable intake was 3·76 servings (95% CI 3·66–3·86) per day. Mean daily consumption of fruits and vegetables was 2·14 servings (1·93–2·36) in LICs, 3·17 servings (2·99–3·35) in LMICs, 4·31 servings (4·09–4·53) in UMICs, and 5·42 servings (5·13–5·71) in HICs. Per-person gross national income was positively associated with fruit and vegetable intake ($p_{\text{trend}}=0\cdot0020$; $r_s=0\cdot37$).

Data for the availability and cost of at least one fruit and one vegetable were obtained from 518 PURE communities (134 in LICs, 187 in LMICs, 125 in UMICs, and 72 in HICs). The number of different types of vegetables and fruits available for sale was greatest in HICs, intermediate in UMICs, lower in LMICs, and lowest in LICs ($p_{\text{trend}}=0\cdot00021$ for vegetables, $p_{\text{trend}}=0\cdot00064$ for fruits; figure 1).

We obtained household income data from 90 247 households in these communities, comprising 130 402 participants—29 421 in LICs, 52 090 in LMICs, 35 069 in UMICs, and 13 822 in HICs. A strong, inverse association exists between gross national income ranking and mean proportion of total household income spent on food (figure 2). Worldwide, the mean proportion of household income spent on food was 42·40% (95% CI 41·24–43·56). Households in HICs spend the smallest proportion (13·30%, 10·27–16·24) of their income purchasing food, compared with 42·15% (39·91–44·39) in UMICs, 52·30% (50·48–54·11) in LMICs, and 61·84% (59·69–64·00) in LICs.

At the community level, the absolute cost (adjusted by purchasing price parity) of one serving of vegetables was cheapest in LICs and most expensive in HICs ($p_{\text{trend}}=0\cdot0023$; table 2). Conversely, the adjusted cost of one serving of fruit was highest in LICs ($p_{\text{trend}}=0\cdot0061$; table 2). The cost of one serving of vegetables relative to income per household member was more than 19 times higher in LICs than in HICs ($p_{\text{trend}}=0\cdot00029$), and the relative cost of one serving of fruit was 50 times higher in LICs than in HICs ($p_{\text{trend}}=0\cdot00011$; table 2). The relative cost of fruit was more expensive than that of vegetables in each region (table 2). Mean daily income per household member was greatest in HICs and lowest in LICS, and greater in urban communities than rural communities across all income regions (table 2).

	High-income countries	Upper-middle-income countries	Lower-middle-income countries	Low-income countries	p_{trend}
Mean (95% CI) absolute cost of one portion (international dollars)					
Vegetables	$0·24 (0·22 to 0·25)	$0·19 (0·18 to 0·20)	$0·13 (0·12 to 0·14)	$0·11 (0·10 to 0·11)	0·0023
Fruits	$0·25 (0·24 to 0·27)	$0·26 (0·25 to 0·28)	$0·22 (0·21 to 0·23)	$0·33 (0·32 to 0·35)	0·0061
Mean (95% CI) proportion of household income spent*					
Vegetables	0·54% (–1·02 to 2·10)	3·97% (2·49 to 5·45)	3·90% (2·94 to 4·86)	10·54% (8·95 to 12·13)	0·00029
Fruits	0·59% (–2·11 to 3·29)	5·19% (3·14 to 7·24)	6·20% (4·53 to 7·87)	29·37% (26·61 to 32·13)	0·00011
Mean (95% CI) daily income per household member (international dollars)					
Urban	$68·36 (67·74 to 68·67)	$26·74 (18·81 to 19·50)	$9·60 (9·33 to 9·88)	$7·18 (6·62 to 7·73)	..
Rural	$56·27 (55·83 to 56·72)	$9·15 (8·92 to 9·38)	$5·36 (5·18 to 5·55)	$1·92 (1·57 to 2·27)	..

*Cost relative to income per household member.

Table 2: **Absolute cost, adjusted by purchasing price parity, and proportion of household income spent on one serving of vegetables and fruits, and daily income per household member, by economic region**

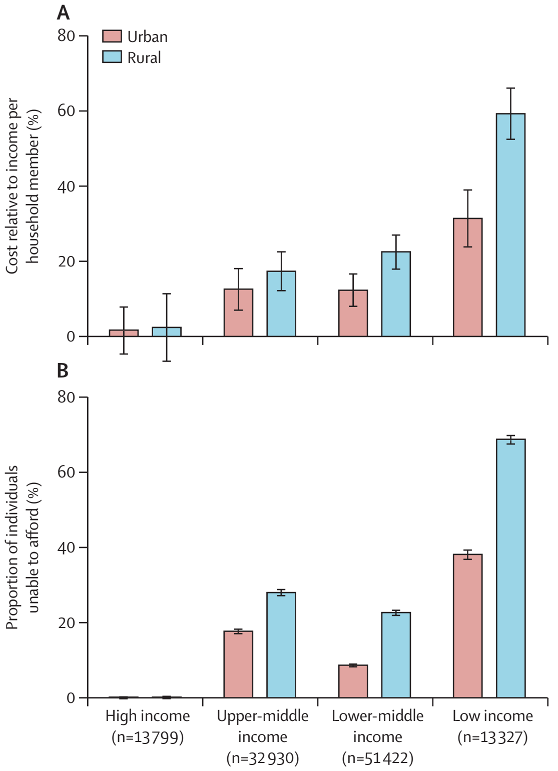


Figure 3: **(A) Mean proportion of income per household member required to purchase three servings of vegetables and two servings of fruits per day and (B) proportion of individuals who were unable to afford three servings of vegetables and two servings of fruits per day**
Error bars represent 95% CI.

Overall, 21·95% (95% CI 19–45–24·45) of income per household member was needed to purchase two servings of fruits and three servings of vegetables. Participants in LICs spend the largest proportion of their income to meet the recommendation (51·97%, 46·06–57·88), compared with 18·10% (14·53–21·68) in LMICs, 15·87% (11·51–20·23) in UMICs, and 1·85% (–3·90 to 7·59) in HICs (p_{trend}=0·0001; figure 3A). In all regions, a higher proportion of income to meet the recommended intake was required in rural areas than in urban areas ($p<0·0001$ for all pairwise comparisons), particularly in UMICs, LMICs, and LICs ($p_{heterogeneity}$=0·0048).

The proportion of individuals who could not afford the recommended daily intake was highest in LICs (57·42%, 95% CI 56·58–58·26), compared with 25·42% (24·95–25·89) in UMICs, 17·68% (17·35–18·01) in LMICs, and 0·25% (0·17–0·33) in HICs (p_{trend}=0·0082; figure 3B). In all regions, unaffordability was higher in rural areas than in urban areas (p=0·027 for all urban *vs* rural pairwise comparisons).

86 506 (60%) participants did not meet the recommended fruit and vegetable intake, and a shift in diet to meet this recommendation would increase food expenditure by 0·45% (95% CI –2·68 to 3·58) of household income in HICs, 7·71% (5·31–10·1) in UMICs, 10·3% (8·14–12·4) in LMICs, and 25·4% (22·0–28·7) in LICs. The increase would be significantly steeper in rural areas than in urban areas ($p_{heterogeneity}$=0·00024; appendix p 25).

Both vegetable and fruit consumption decreased as the relative cost per serving increased, after adjusting for energy intake, age, sex, and economic region (p_{trend}=0·00071 for vegetables and p_{trend}=0·00033 for fruit for vegetables and for fruits; figure 4). Combined fruit and vegetable intake decreased as the relative cost of two servings of fruits and three servings of vegetables per day increased, both overall (p_{trend}=0·00040) and by economic region, except in HICs (figure 5).

When we recalculated income per household member using a weighted approach (reflecting the lower energy needs of children), the association between the relative cost of one serving of vegetables and fruit with economic region persisted (appendix p 22). When examining the association between the affordability of current vegetable and fruit recommendations and economic region, the results were again similar (appendix p 22).

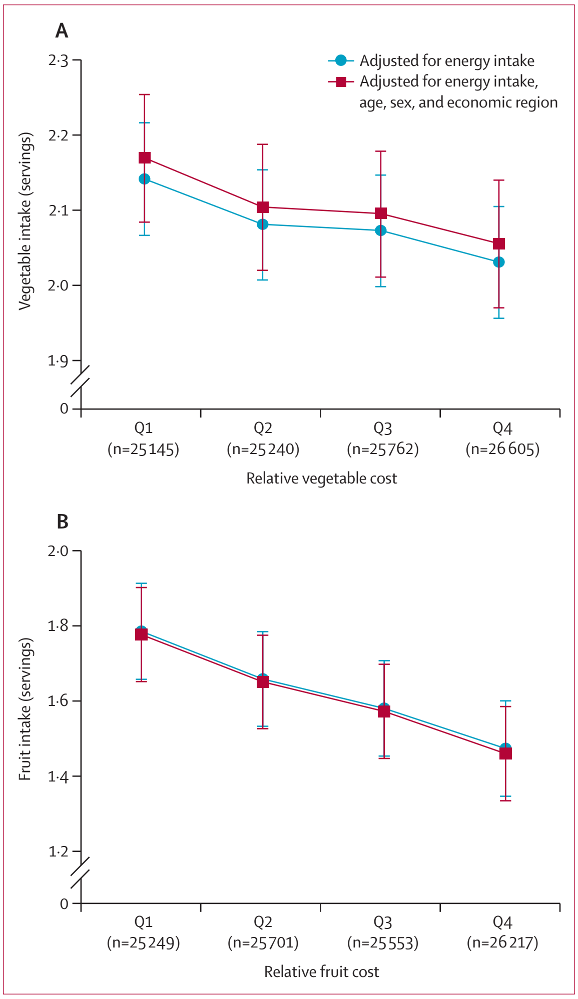


Figure 4: **Mean intake of (A) vegetables and (B) fruits per person adjusted for covariates, by relative cost**

Error bars represent 95% CI. Q=quartile.

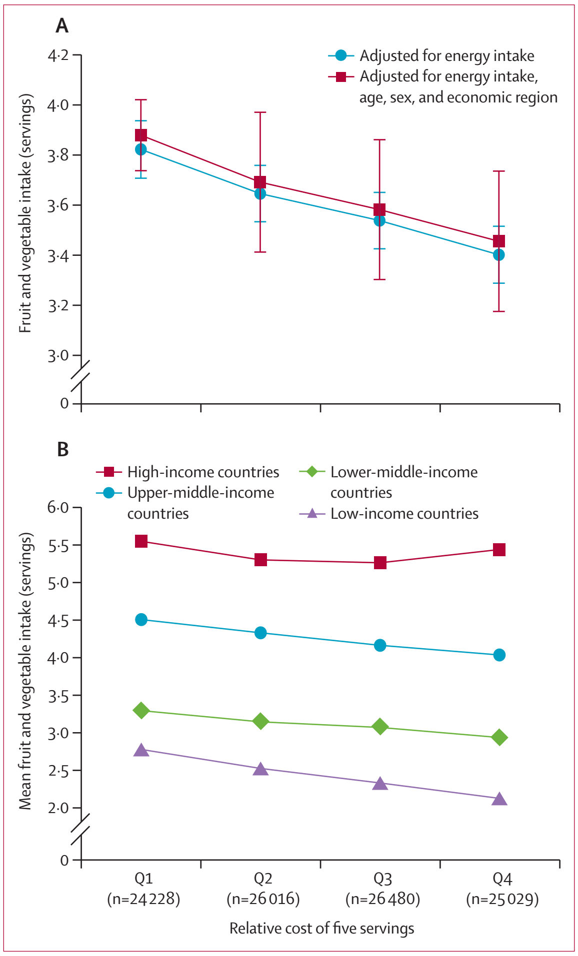


Figure 5: **Mean vegetable and fruit intake per person by relative cost of three servings of vegetables and two servings of fruits (A) in the overall sample, adjusting for covariates, and (B) stratified by economic region, adjusting for energy intake, age, and sex**

Error bars represent 95% CI. Q=quartile.

Discussion

In this study of 18 countries with a range of income levels, we found that individuals in countries with low gross national income consume fewer fruits and vegetables and spend a greater proportion of their income purchasing food than those in high-income countries. Absolute fruit cost was highest in communities of LICs, whereas vegetable cost was lowest in these communities adjusted by purchasing price parity). However, the costs of both fruits and vegetables (relative to household income) were substantially higher for individuals in countries with low gross national income than in other economic regions. Furthermore, in LICs, households spend 29% and 11% of their income to purchase one serving of fruits and vegetables, respectively, and the dietary recommendation of two servings of fruits and three servings of vegetables per day was unaffordable for 57% of individuals. Unsurprisingly, increased costs of fruits and vegetables relative to household income were associated with reduced consumption.

Households in LICs and LMICs spend a substantial proportion (roughly half) of their income on food (compared with 13% in HICs), with households in some countries (eg, Occupied Palestinian Territory, Bangladesh, Pakistan, and Zimbabwe) spending about two-thirds of their income on food (figure 2). These findings are consistent with previous work showing that food spending ranges from 35% to 65% in MICs[23] and from 55% to 77% in LICs.[23,24] However, our findings of the relative costs of fruits and vegetables could not be compared with previous work in LICs or MICs because few such countries systematically monitor the cost of food and disclose national statistics.[25] Moreover, the national estimates of the cost of major food commodities available from the World Bank[26] and the UN Food and Agriculture Organization[27] do not include fruits and vegetables.

The consumption of a variety of fruits and vegetables is important to a high-quality diet.[28,29] In the PURE study, most participants consumed fewer than the recommended five daily servings of fruits and vegetables, and mean vegetable intake was lower than the recommended three daily servings in all economics regions except HICs (table 1). In 2015, global fruit and vegetable intake was estimated to be lower than the average observed in our study.[30] Of note, previous estimates were mainly based on qualitative questionnaires, shorter dietary tools, or household surveys. These dietary tools are brief questionnaires in which a structured list of food items is absent and as few as one question might be used to estimate the consumption of a particular food type. This method does not include portion sizes to quantify level of intake and provides a less precise estimate of absolute intake than 24 h dietary recall or semi-quantitative food frequency questionnaires.[31] The household surveys are useful for monitoring food commodity use, but they might not be appropriate for measuring absolute dietary intake or energy intake because they reflect both intake and food lost through waste at the retail, food service, and household level.[32] Among studies of HICs using semi-quantitative food frequency questionnaires and with similar age and sex characteristics as the PURE study, our estimates of mean fruit and vegetable intake correspond closely with those in other similar populations (appendix p 11).[33,34]

Our study has a few limitations. First, fruit and vegetable costs were not recorded in 80 communities (11 953 participants), most of which were in LMICs. Since fruit and vegetable costs might vary across communities, imputing costs was unlikely to reflect the heterogeneity in prices. Our sample included a small representation of participants in South Africa and Zimbabwe because of missing data for fruit and vegetable costs and household income. However, the participants of the PURE study and non-participants included in our analysis were similar in baseline characteristics, so potential biases resulting from exclusion of participants were likely to be minimal. Second, a true probability sampling approach was not used to select our study population. Such a method was not feasible because of the many practical constraints of studying food cost and availability in a wide range of countries and settings. The fact that sampling was not random should be considered when interpreting the generalisability of our findings but should not compromise the internal validity. Third, the costs of the different fruits and vegetables were collected at the community level and assumed to reflect the average cost that households would pay. The costs were collected from grocery stores located centrally in each community to ensure that the costs were representative of most households. Fourth, we did not account for seasonal differences in prices, since we did not collect the cost of fruits and vegetables in each community at different times of the year. However, these data were collected over several seasons for most countries (appendix p 15). Because many of the countries have fairly uniform climate (particularly in LICs and MICs), the results are likely to provide a reasonable approximation of the average seasonal price for fruit and vegetable items in these communities.

Fifth, costs were collected for fruits and vegetables that were thought to be the most widely available in most countries, but not necessarily the cheapest or most regularly consumed items within all countries. The fruits and vegetables chosen were widely available across economic regions, with the exception of pears and cabbage in LICs (appendix p 20). Furthermore, the least expensive fruit and vegetable items in each economic region were available for sale in most communities (appendix p 21). The interpretation of the affordability of fruits and vegetables might be limited to these commonly available produce, and cheaper alternatives might have been accessible. Nevertheless, fruit and vegetable intake was assessed using country-specific food frequency questionnaires that reflected the individual food items most commonly consumed in each country, and we still found a strong graded association with fruit and vegetable cost. Additionally, the cost of fruits and vegetables were collected as non-sale prices, since sale prices might change on a daily or weekly basis, thus increasing the variability of estimates, whereas the non-sale prices would be expected to provide a more consistent estimate of costs within and across communities. Finally, the data presented are cross-sectional, and inferences cannot be made about the causal relation between affordability and consumption of fruits and vegetables.

This study provides an international comparison of fruit and vegetable costs and affordability using a standardised and validated instrument. Another important strength of this study is the large sample size and heterogeneity of the study population. Additionally, a large proportion of study participants are from MICs and LICs, for which limited information on food affordability is available.

Hunger and under-nutrition remain highly prevalent in many LICs and MICs,[35] and nutrition strategies in these countries often prioritise meeting the minimum energy intake over diet quality. The unaffordability of fruits and vegetables might be a large barrier to achieving these nutritional targets. Worldwide, 1·7 million annual deaths are estimated to be associated with low fruit and vegetable intake,[36] and many populations are unable to meet the dietary recommendations. Our results show that increased cost of fruits and vegetables relative to household income was associated with reduced consumption, highlighting the need for policies that expand affordability and availability of these foods, which might improve the diet quality of many populations, especially in LICs and LMICs.

Contributors

VM and AM designed the study, were involved in data management and statistical analysis, and wrote the first and subsequent drafts of the report. SY designed the study, conceived and initiated the Prospective Urban Rural Epidemiology (PURE) study, supervised its conduct and data analysis, and provided crucial comments on all drafts of the report. CKC, DJC, and KL conceived and initiated the Environmental Profile of a Community's Health (EPOCH) study, supervised its conduct, and provided crucial comments on all drafts of the report. MD developed and validated the country-specific food frequency questionnaires, supervised the collection of dietary information, and commented on drafts of the report. SR coordinated the worldwide PURE study and reviewed and commented on drafts of the report. KT was the coprincipal investigator of the PURE study and reviewed and commented on drafts of the report. All other authors coordinated the study, collected data in their respective countries, and provided comments on drafts of the report.

Declaration of interests

We declare no competing interests.

Acknowledgments

The main PURE study and its components are funded by the Population Health Research Institute; the Canadian Institutes of Health Research; the Heart and Stroke Foundation of Ontario (Canada); unrestricted grants from AstraZeneca (Sweden, Canada, Turkey), Sanofi-Aventis (France, Canada, Turkey), Boehringer Ingelheim (Germany and Canada), Servier, GlaxoSmithKline, Novartis, King Pharma; the Bangladesh Independent University; Mitra and Associates (Bangladesh); Unilever Health Institute (Brazil); Public Health Agency of Canada; Champlain Cardiovascular Disease Prevention Network (Canada); Universidad de la Frontera (Chile); National Center for Cardiovascular Diseases (China); Colciencias (Colombia; grant number 6566-04-18062); Indian Council of Medical Research; Malaysian Ministry of Science, Technology and Innovation (grant number 07-05-IFN-MEB010); Malaysian Ministry of Higher Education (grant number 600-RMI/ LRGS/5/3); Universiti Kebangsaan Malaysia (UKM-Hejim-Komuniti-15-2010); Polish Ministry of Science and Higher Education (grant number 290/W-PURE/2008/0); Wroclaw Medical University (Poland); North-West University (South Africa); South Africa Netherlands Research Programme on Alternatives in Development (SANPAD); National Research Foundation (South Africa); Medical Research Council of South Africa; the South Africa Sugar Association (SASA); Faculty of Community and Health Sciences (University of the Western Cape, South Africa); Council for Working Life and Social Research (Sweden); Swedish Research Council for Environment; Agricultural Sciences and Spatial Planning (Sweden); Swedish Heart and Lung Foundation; Swedish Research Council; grant from the Swedish State under LUA (LäkarUtbildningsAvtalet) agreement; grant from the Västra Götaland Region (FOUU; Sweden); Metabolic Syndrome Society; and the Sheikh Hamdan Bin Rashid Al Maktoum Award For Medical Sciences (Dubai Health Authority). AM is a recipient of a Research Early Career Award from Hamilton Health Sciences Foundation. SY is funded by the Marion Burke Chair of the Heart and Stroke Foundation of Canada and is President of the World Heart Federation. This paper does not necessarily reflect the position of the World Heart Federation or any other organisation.

References

1 Joint WHO/FAO expert consultation on diet, nutrition and the prevention of chronic diseases. Geneva: Joint FAO/WHO Expert Consultation, 2003.

2 The world health report 2002—reducing risks, promoting healthy life. Geneva: World Health Organization, 2003.

3 Hall JN, Moore S, Harper SB, Lynch JW. Global variability in fruit and vegetable consumption. *Am J Prev Med* 2009; **36:** 402–09.

4 Murphy MM, Barraj LM, Spungen JH, Herman DR, Randolph RK. Global assessment of select phytonutrient intakes by level of fruit and vegetable consumption. *Br J Nutr* 2014; **112:** 1004–08.

5 Del Gobbo LC, Khatibzadeh S, Imamura F, et al. Assessing global dietary habits: a comparison of national estimates from the FAO and the Global Dietary Database. *Am J Clin Nutr* 2015; **101:** 1038–46.

6 Power EM. Determinants of healthy eating among low-income Canadians. *Can J Public Health* 2005; **96** (suppl 3)**:** S37–42, S42–48.

7 Steenhuis IH, Waterlander WE, de Mul A. Consumer food choices: the role of price and pricing strategies. *Public Health Nutr* 2011; **14:** 2220–26.

8 Brinkman HJ, de Pee S, Sanogo I, Subran L, Bloem MW. High food prices and the global financial crisis have reduced access to nutritious food and worsened nutritional status and health. *J Nutr* 2010; **140:** 153S–61S.

9 Gustafson DJ. Rising food costs & global food security: key issues & relevance for India. *Indian J Med Res* 2013; **138:** 398–410.

10 Ruel MT, Garrett JL, Hawkes C, Cohen MJ. The food, fuel, and financial crises affect the urban and rural poor disproportionately: a review of the evidence. *J Nutr* 2010; **140:** 170S–6S.

11 Klotz C, de Pee S, Thorne-Lyman A, Kraemer K, Bloem M. Nutrition in the perfect storm: why micronutrient malnutrition will be a widespread health consequence of high food prices. *Sight and Life Magazine* 2008; **2:** 6–11.

12 D'Souza A, Joliffe D. Conflict, food price shocks, and food security: the experience of Afghan households. IZA Discussion Paper number 6621. Institute for the Study of Labor, 2012. http://ftp.iza.org/dp6621.pdf (accessed June 30, 2016).

13 Teo K, Chow CK, Vaz M, Rangarajan S, Yusuf S, PURE Investigators-Writing Group. The Prospective Urban Rural Epidemiology (PURE) study: examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. *Am Heart J* 2009; **158:** 1–7.

14 Corsi DJ, Subramanian SV, Chow CK, et al. Prospective Urban Rural Epidemiology (PURE) study: baseline characteristics of the household sample and comparative analyses with national data in 17 countries. *Am Heart J* 2013; **166:** 636–646.e4.

15 Kelemen LE, Anand SS, Vuksan V, et al. Development and evaluation of cultural food frequency questionnaires for South Asians, Chinese, and Europeans in North America. *J Am Diet Assoc* 2003; **103:** 1178–84.

16 Bharathi AV, Kurpad AV, Thomas T, Yusuf S, Saraswathi G, Vaz M. Development of food frequency questionnaires and a nutrient database for the Prospective Urban and Rural Epidemiological (PURE) pilot study in South India: methodological issues. *Asia Pac J Clin Nutr* 2008; **17:** 178–85.

17 Dehghan M, Martinez S, Zhang X, et al. Relative validity of an FFQ to estimate daily food and nutrient intakes for Chilean adults. *Public Health Nutr* 2013; **16:** 1782–88.

18 Dehghan M, del Cerro S, Zhang X, et al. Validation of a semi-quantitative Food Frequency Questionnaire for Argentinean adults. *PLoS One* 2012; **7:** e37958.

19 Dehghan M, Ilow R, Zatonska K, et al. Development, reproducibility and validity of the food frequency questionnaire in the Poland arm of the Prospective Urban and Rural Epidemiological (PURE) study. *J Hum Nutr Diet* 2012; **25:** 225–32.

20 Dehghan M, Lopez Jaramillo P, Duenas R, et al. Development and validation of a quantitative food frequency questionnaire among rural- and urban-dwelling adults in Colombia. *J Nutr Educ Behav* 2012; **44:** 609–13.

21 Dehghan M, Al Hamad N, Yusufali A, Nusrath F, Yusuf S, Merchant AT. Development of a semi-quantitative food frequency questionnaire for use in United Arab Emirates and Kuwait based on local foods. *Nutr J* 2005; **4:** 18.

22 Merchant AT, Dehghan M, Chifamba J, Terera G, Yusuf S. Nutrient estimation from an FFQ developed for a Black Zimbabwean population. *Nutr J* 2005; **4:** 37.

23 Banerjee AV, Duflo E. What is middle class about the middle classes around the world? *J Econ Perspect* 2008; **22:** 3–28.

24 FAO, IFAD, WFP. How does international price volatility affect domestic economies and food security. Rome: Food and Agriculture Organization, 2011.

25 Lee A, Mhurchu CN, Sacks G, et al. Monitoring the price and affordability of foods and diets globally. *Obes Rev* 2013; **14** (suppl 1)**:** 82–95.

26 World Bank. Food price watch. Washington, DC: World Bank, 2014.

27 FAO. World food situation: food price index. Rome: Food and Agriculture Organization, 2015.

28 Liu RH. Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals. *Am J Clin Nutr* 2003; **78:** 517S–20S.

29 Cooper AJ, Sharp SJ, Lentjues MA, et al. A prospective study of the association between quantity and variety of fruit and vegetable intake and incident type 2 diabetes. *Diabetes Care* 2012; **35:** 1293–300.

30 Micha R, Khatibzadeh S, Shi P, Andrews KG, Engell RE, Mozaffarian D. Global, regional and national consumption of major food groups in 1990 and 2010: a systematic analysis including 266 country-specific nutrition surveys worldwide. *BMJ Open* 2015; **5:** e008705.

31 Thompson FE, Subar AF. Nutrition in the prevention and treatment of disease. In: Coulston AM, Rock CL, Monsen ER, eds. Dietary assessment methodology. San Diego: Academic Press, 2001.

32 Lock K, Pomerleau J, Causer L, McKee M. Low fruit and vegetable consumption. In: Ezzati M, Lopez AD, Rodgers A, Murray CJL, eds. Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors. Geneva: World Health Organization, 2004: 597–727.

33 Rohrmann S, Giovannucci E, Willett WC, Platz EA. Fruit and vegetable consumption, intake of micronutrients, and benign prostatic hyperplasia in US men. *Am J Clin Nutr* 2007; **85:** 523–29.

34 Corley J, Kyle JA, Starr JM, McNeill G, Deary IJ. Dietary factors and biomarkers of systemic inflammation in older people: the Lothian Birth Cohort 1936. *Br J Nutr* 2015; **114:** 1088–98.

35 FAO, IFAD, WFP. The state of food insecurity in the world 2013. Rome: Food and Agriculture Organization, 2013.

36 Lim SS, Vos T, Flaxman AD, et al. 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* 2012; **380:** 2224–60.

THE LANCET Global Health

Supplementary appendix

Supplement to: Miller V, Yusuf S, Chow CK, et al. Availability, affordability, and consumption of fruits and vegetables in 18 countries across income levels: findings from the Prospective Urban Rural Epidemiology (PURE) study. *Lancet Glob Health* 2016; published online Aug 23. http://dx.doi.org/10.1016/S2214-109X(16)30186-3.

Supplementary methods

Tables

Figures

Supplementary Methods: PURE Study Participant Selection Methodology as Excerpted from Teo et al[1].

Selection of Countries

PURE country selection reflected a balance between the feasibility of long-term follow-up, the inclusion of countries at different economic levels and heterogeneity in social and economic conditions, and the feasibility of centers to successfully achieve long-term follow-up. Thus, PURE included sites in which investigators are committed to collecting good-quality data for a low-budget study over the planned 10-year follow-up period and did not aim for a strict proportionate sampling of the entire world.

Selection of Communities

Within each country, urban and rural communities were selected based on broad guidelines (see Table below). A common definition for “community” that is applicable globally is difficult to establish. In PURE, a community was defined as a group of people who have common characteristics and reside in a defined geographic area. A city or large town was not usually considered to be a single community, rather communities from low-, middle-, and high-income areas were selected from sections of the city and the community area defined according to a geographical measure (eg, a set of contiguous postal code areas or a group of streets or a village). The primary sampling unit for rural areas in many countries was the village. The reason for inclusion of both urban and rural communities is that for many countries, urban and rural environments exhibit distinct characteristics in social and physical environment, and hence, by sampling both, we ensured considerable variation in societal factors across PURE communities. The number of communities selected in each country varied, with the aim to recruit communities with substantial heterogeneity in social and economic circumstances balanced against the capacity of local investigators to maintain follow-up. In some countries (eg, India, China, Canada, and Colombia), communities from several states/provinces were included to capture regional diversity, in policy, socioeconomic status, culture, and physical environment. In other countries (eg, Iran, Poland, Sweden, and Zimbabwe), fewer communities were selected.

Selections of Households and Individuals

Within each community, sampling was designed to achieve a broadly representative sample of that community of adults aged between 35 and 70 years. The choice of sampling frame within each center was based on both “representativeness” and feasibility of long-term follow-up, following broad study guidelines. Once a community was identified, where possible, common and standardized approaches were applied to the enumeration of households, identification of individuals, recruitment procedures, and data collection. The method of approaching households differed between regions. For example, in rural areas of India and China, a community announcement was made to the village through contact of a community leader, followed by in-person door-to-door visits of all households. In contrast in Canada, initial contact was by mail followed by telephone inviting members of the households to a central clinic. For each approach, at least 3 attempts at contact were made. Households were eligible if at least 1 member of the household was between the ages of 35 and 70 years and the household members intended to continue living in their current home for a further 4 years. All individuals within these households between 35 and 70 years providing written informed consent were enrolled. When a household refused to participate, demographics and simple self-report risk factor data were recorded in a nonresponder form.

Guidelines for Selection of Countries, Communities, Households, and Individuals Recruited to PURE
Countries
1. High-income countries, middle-income countries, and low-income countries, with the bulk of the recruitment from low- and middle-income regions.
2. Committed local investigators with experience in recruiting for population studies.
Communities
1. Select both urban and rural communities. Use the national definition of the country to determine urban and rural communities.
2. Select rural communities that are isolated (distance of >50 km or lack easy access to commuter transportation) from urban centers. However, consider ability to process bloods samples, eg, villages in rural developing countries should be within 45-min drive of an appropriate facility.
3. Define community within a geographical area, eg, using postal codes, catchment area of health service/clinics, census tracts, areas bordered by specific streets or natural borders such as a river bank.
4. Consider feasibility for long-term follow-up, eg, for urban communities, choose sites that have a stable population such as residential colonies related to specific work sites in developing countries. In rural areas, choose villages that have a stable population. Villages at greater distance from urban centers are less susceptible to large migration to urban centers.
5. Enlist a community organization to facilitate contact with the community, eg, in urban areas, large employers (government and private), insurance companies, clubs, religious organizations, clinic or hospital service regions. In rural areas, local authorities such as priests or community elders, hospital or clinic, village leader, or local politician.
Individual
1. Broadly representative sampling of adults 35 to 70 years within each community unit.
2. Consider feasibility for long-term follow-up when formulating community sampling framework, eg, small percentage random samples of large communities may be more difficult to follow-up because they are dispersed by distance. In rural areas of developing countries that are not connected by telephone, it may be better to sample entire community (ie, door-to-door systematic sampling).
3. The method of approach of households/individuals may differ between sites. In MIC and HIC, mail, followed up by phone contact may be the first practical means of contact. In LIC, direct household contact through household visits may be the most appropriate means of first contact.
4. Once recruited, all individuals are invited to a study clinic to complete standardized questionnaires and have a standardized set of measurements.

Socioeconomic status of PURE participants

Households from the countries included in PURE were found to have good agreement with the United Nations (UN) World Population Prospects on age, sex, urban/rural location, education and mortality[20]. The socioeconomic status of PURE participants was found to be representative of national statistics for the countries included, with less than 7% differences for each education category (less than primary, primary complete, secondary, and higher)[2].

Supplementary Methods: Methodology Used for Estimating Dietary Intake

The PURE study completed dietary assessments of 147,938 participants using country specific, validated or newly validated semi-quantitative food frequency questionnaires (FFQs)[3-9]. An international nutrient database was developed using the US Department of Agriculture's food composition database and local food composition tables to calculate nutrients and servings. Estimates of energy intake, and daily fruit and vegetable servings were computed from the FFQs. Pearson correlation coefficients, using energy-adjusted and de-attenuated correlations, and weighted kappa were used to validate the FFQs measures against 24-hour dietary recalls[3-9]. Our validation studies demonstrated reasonable agreement between the FFQs and 24-hour dietary recalls for fruits (r_s=0.23-0.66) and vegetables (r_s=0.3-0.81) which is in keeping with previous FFQ validation studies[10-16]. Of the 147,938 individuals with FFQ data, 143,305 participants had plausible energy intakes and were included in the analyses on fruit and vegetable intake by economic region. Of these individuals, 130,402 participants had information collected on fruit and vegetable cost in their community and on household income, and were included in analyses relating food availability and affordability to consumption.

Supplementary Methods: Methodology Used for Estimating Fruit and Vegetable Availability and Affordability

The Environmental Profile of a Community's Health (EPOCH) component of PURE is a cross-sectional, standardized and validated instrument that was conducted within PURE. As part of EPOCH, a 1 km observation walk was conducted in a centrally located area within each community. Within this area, the grocery or market place that was most frequented by participants was used to collect the non-sale prices for the following fruits and vegetables: apples, oranges, bananas, pears, carrots, tomatoes and cabbage. The price per kilogram of each fruit and vegetable was collected in the local currency. A checklist of 48 types of fruits and 59 types of vegetables was used to assess the availability of a variety of fruits and vegetables. The checklists were collated in consultation with study investigators from each country in the development of the EPOCH instruments. The total number of types of fruits and vegetables available for sale within each community was calculated to assess the diversity. If one or more fruit or vegetable items were not available in the first grocery store/market, additional market places within the designated walk area were visited to collect these data. A full explanation of the development and validation of the EPOCH tool was described previously[17,18].

Defining food affordability is challenging, since affordability is largely subjective and dependent on individuals' willingness to pay. Total household income and disposable income vary markedly across countries making it difficult to establish thresholds globally. Previous studies on food affordability have been conducted in single countries or geographic regions and examined food costs at the household level in relation to household income[19-23]. However, several other groups are currently working to develop standardized methods to assess food affordability globally[24,25].

Supplementary Methods: Methodology Used for Estimating Household Income

Monthly household income and the amount of income spent on food were measured using standardized questionnaires and participant self-report in PURE. The median income collected for each country was compared to the World Bank per capita gross national income (GNI) estimates[26] ($r=0{\cdot}91$, $p<0{\cdot}001$) (Figure S1).

Supplementary Methods: Energy Intake Provided by Fruits and Vegetables

On the basis of daily energy needs, the energy intake provided by 2 servings of fruits and 3 servings of vegetables ranged from 9-19% of the energy requirement for a 2000 kcal diet, for the fruit and vegetable types included in the study.

Supplementary Material: Statistical Methodology for Sensitivity Analysis

To address the impact of lower energy needs for children in the household (and accordingly the lower income required to meet nutritional requirements), we conducted a sensitivity analysis where we recomputed the household member income using a weighted approach. Specifically, instead of dividing household income by the number of household members, we reduced the weight contributed by young household members (aged 12 years or younger) by half, and assessed mean relative vegetable and fruit costs and the affordability of meeting produce recommendations.

Supplementary Material: Statistical Methodology for Fruit and Vegetable Intake Across Studies

Recently, the Global Burden of Disease (GBD) study reported a mean global intake of 0·7 servings/day for fruit and 1·7 servings/day for vegetables[27]. These estimates were largely based on qualitative questionnaires, short dietary tools, and household surveys, which may not accurately reflect absolute levels of intake but are suited for ranking individuals according to level of intake[28-32]. Additionally, participant characteristics may influence estimates of mean intake. Women generally consume more fruit and vegetables compared to men[33-38], and older adults consume more than younger individuals[27,39-41]. In the PURE study, women and men were equally representative, and were older than the majority of other studies, apart from the Health Professionals Follow-Up Study[42] and the Lothian Birth Cohort[43]. In comparing these two studies to PURE, both studies (which also used a validated semi-quantitative FFQ as in PURE) reported similar fruit and vegetable intake (median 5·5 servings/day and mean 5·6 servings/day, respectively) as the PURE high-income region (median 5·0 servings/day and mean 5·4 servings/day). Therefore, among studies of similar age and sex characteristics as PURE, our estimates of mean fruit and vegetable intake correspond closely with other similar populations.

Table S1. Computational methods for variables used in the analysis.

Variable	**Method**
Monthly household income	Monthly household income was collected by self-report from each participant. To allow for cross-country comparisons, monthly household income was converted to International dollars using country-specific World Bank Purchasing Price Parity (PPP). In some communities in which household income was collected at a different time to food costs the Consumer Price Index (CPI) was used to correct for inflation. Further, CPI was used to adjust for variation in the year the costs were collected for each country.
Daily income per household member	Monthly household income was multiplied by 12 and divided by 365 to calculate daily household income. This value was divided by the number of individuals residing in the household to compute daily income per household member.
Percent of income spent purchasing food	The amount of income households indicated spending on food was divided by the household's monthly income and multiplied by 100. To minimize the likelihood of extreme values skewing the results, the percentage was truncated at the 5th and 95th percentiles.
PPP adjusted fruit and vegetable cost	Fruit and vegetable costs were computed per 125g serving by dividing the cost collected (1000g) by 8 servings. The cost of a 125g serving was converted to International dollars using country-specific World Bank PPP rates. The cost of a serving of fruits and vegetables in International dollars was used for all subsequent computations.
Mean PPP adjusted fruit cost	The mean cost of a 125g serving of fruit was computed by averaging the cost of a serving of apples, bananas, oranges and pears.
Mean PPP adjusted vegetable cost	The mean cost of a 125g serving of vegetables was computed by averaging the cost of a serving of cabbage, carrots and tomatoes.
Mean relative fruit cost	The mean PPP adjusted fruit cost was divided by the daily income per household member.
Mean relative vegetable cost	The mean PPP adjusted vegetable cost was divided by the daily income per household member.
Mean relative fruit and vegetable cost	The mean PPP adjusted fruit and vegetable cost was divided by the daily income per household member.
Relative cost of purchasing 2 servings of fruits and 3 servings of vegetables per day	For each community, the cheapest fruit and vegetable was used to calculate the relative cost of meeting fruit and vegetable recommendations. The sum of 2 servings of fruits and 3 servings of vegetables was divided by the daily income per household member.
Relative cost of purchasing 5 servings of fruits or vegetables per day	For each community, the cheapest fruit or vegetable was used to calculate the relative cost of consuming 5 servings of produce. The sum of 5 servings of fruits or vegetables was divided by the daily income per household member.
Percent of individuals unable to afford 2 servings of fruits and 3 servings of vegetables per day	The percentage of individuals that would have to spend greater than 20% of their income to purchase the relative cost of 2 servings of fruits and 3 servings or vegetables per household member.
Daily fruit and vegetable intake	The fruit and vegetable serving size in each country was converted to a standard serving (125 g), and daily intake was truncated at 1500g for fruits and vegetables separately. The sum of fruit and vegetable intake was computed.

Table S2. Main study variables and the source of the variable.

Variable	**Data source**
Participant characteristics	PURE study
Household income and the number of individuals residing in the household	PURE study
Fruit, vegetable and energy intake (FFQ)	PURE study
Number of types of fruits and vegetables available for sale	EPOCH
Fruit and vegetable cost	EPOCH

EPOCH: Environmental Component of a Community's Health sub-study

Table 3. Main study variables and their unit of analysis.

Variable	Unit of analysis
Number of types of vegetables and fruit available for sale (diversity)	Community
Cost of 1 serving of vegetables and fruits in communities (PPP adjusted cost)	Community
Percentage of household income spent on food	Household
Cost of 1 serving of vegetables and fruits as a percent of household member income (relative cost) †	Individual
Cost of 2 servings of fruits and 3 servings of vegetables as a percent of household member income (relative cost) †	Individual
Percentage of individuals unable to afford to meet vegetable and fruit recommendations	Individual
Fruit and vegetable consumption (from the FFQs)	Individual

† Information on the cost of vegetables and fruits were extrapolated to the individual level by assuming that the cost collected at the community level was representative of the cost each individual would pay within the community. Similarly, household income was divided by the number of individuals within the household to compute an equal estimate of income per each household member.

Table S4. Years of data collection and the World Bank conversion rates for household income and fruit and vegetable costs.

Economic Region	Country	Year of data collection		Consumer Price Index (CPI)*+	Purchasing Price Parity (PPP)*+
		EPOCH (Fruit and vegetable costs)	**Income**		
High-income countries	Canada	Fall 2009- Winter 2009, Spring 2013	2006-2009	107·0	1·3
	Sweden	Winter 2010- Summer 2011	2005-2009	107·9	9·3
	United Arab Emirates	Fall 2009- Winter 2010	2005-2009	114·0	2·7
Upper middle-income countries	Argentina	Fall 2009- Winter 2010	2006-2009	139·3	1·5
	Brazil	Spring 2010	2005-2009	125·7	1·8
	Chile	Spring 2010	2006-2009	101·4	370·3
	Malaysia	Spring 2010- Summer 2010	2007-2010	112·1	2·2
	Poland	Winter 2009	2007-2009	112·2	2·0
	South Africa	Winter 2009-Spring 2010	2005-2010	133·9	5·7
	Turkey	Fall 2009- Spring 2010	2008-2009	151·9	1·1
Lower middle-income countries	China	Winter 2010-Fall 2010	2005-2009	115·4	4·2
	Colombia	Summer 2010	2006-2009	125·5	1339·8
	Iran, Islamic Rep.	Winter 2010-Spring 2010	2005-2009	205·9	5006·8
	Palestine	Summer 2011- Winter 2013	2012-2013	119·4	2·3
Low-income countries	India	Fall 2009- Summer 2010	2003-2007	135·6	21·2
	Bangladesh	Summer 2009	2007-2008	133·7	33·0
	Pakistan	2010±	2008-2009	158·7	33·5
	Zimbabwe	Winter 2009- Winter 2010	2006-2007	24·9	0·5

* CPI and PPP rates were assumed to be equal to the previous year if values were unavailable for the year of data collection.

\+ The mean year of data collection was used for countries with data collection occurring over more than one year.

∞ Season of EPOCH data collection was defined using the following time periods: Spring: March 20-June 20, Summer: June 21-September 22, Fall: September 23-December 20, Winter: December 21-March 19.

± Season of year was not available.

Table S5. Characteristics of the EPOCH participants included in the analysis.

Characteristic	**Overall (n=130,402)**	**High-income countries (n=13,822)**	**Upper middle-income countries (n=35,069)**	**Lower middle-income countries (n=52,090)**	**Low-income countries (n=29,421)**
Median age-yrs (IQR)	50·0 (34·0-66·0)	53·0 (38·0-68·0)	51·0 (35·0-67·0)	51·0 (36·0-66·0)	47·0 (31·0-63·0)
Sex- F(%)	74,731 (57·3)	7,492 (54·2)	20,692 (59·0)	30,169 (57·9)	13,043 (55·7)
Education level- no. (%)					
Less than high-school graduate	47,131 (41·3)	1,537 (11·8)	14,631 (49·5)	18,170 (38·2)	12,793 (53·4)
High-school graduate	43,502 (38·1)	3,910 (30·0)	9,914 (33·6)	21,892 (46·1)	7,786 (32·5)
Some college or more	23,428 (20·5)	87,579 (58·2)	4,999 (16·9)	7,455 (14·3)	3,395 (14·2)
Current smoking status- no. (%)	23,293 (20·5)	1,805 (13·9)	4,912 (22·1)	9,380 (20·0)	5,616 (23·4)
High physical activity level- no. (%)	46,351 (43·8)	6,327 (53·1)	11,242 (44·6)	19,480 (42·1)	9,302 (41·3)
BMI	25·7 ± 5·1	27·6 ± 0·3	28·3 ± 5·7	25·1 ± 4·0	23·2 ± 4·8
BMI Categories- no. (%)					
Underweight	5,708 (5·3)	88 (0·7)	449 (1·9)	1,077 (2·3)	3,994 (18·1)
Healthy weight	45,718 (42·6)	4,238 (33·3)	7,042 (27·4)	23,790 (50·9)	10648 (48·3)
Overweight	37,508 (35·0)	5,142 (40·4)	9,591 (37·3)	17,259 (36·9)	5,516 (25·0)
Obese	18,365 (17·1)	3,274 (25·7)	8,533 (33·2)	4,658 (10·0)	1,900 (8·6)
Median energy intake, kcal/day (IQR) α	2,020 (966-3,074)	2,154 (1,016-3,292)	2,121 (914-3,328)	1,960 (1,035-2,885)	1,984 (882-3,086)
Vegetable intake, servings/day*	2·2 (2·1- 2·2)	3·4 (3·2-3·6)	1·8 (1·7-1·9)	2·1 (2·0-2·2)	1·4 (1·3-1·5)
Fruit intake, servings/day*	1·6 (1·5-1·7)	2·0 (1·8-2·2)	2·5 (2·3-2·7)	1·2 (1·1-1·3)	0·7 (0·6-0·9)
Consume ≥1 serving of vegetables/day- no. (%)	99,334 (82·9)	12,456 (94·2)	24,944 (81·0)	43,825 (90·5)	18,109 (66·2)
Consume ≥1 serving of fruits/day- no. (%)	59,211 (45·4)	10,782 (81·5)	22,183 (72·0)	19,204 (39·7)	7,042 (25·8)

*Mean (95% CI) accounting for clustering of households within communities.

α Plausible energy intake ≥500 kcal/day and ≤5000 kcal/day.

Table S6. Measured characteristics of participants included in the analysis versus excluded participants.

	HIC		**UMIC**		**LMIC**		**LIC**	
	Excluded	**Included**	**Excluded**	**Included**	**Excluded**	**Included**	**Excluded**	**Included**
N	2,292	13,822	8,093	35,069	10,718	52,090	5,749	29,421
Median age-yrs (IQR)	53·0 (37·0-69·0)	53·0 (38·0-68·0)	50·0 (34·0-66·0)	51·0 (35·0-67·0)	49·0 (34·0-64·0)	51·0 (36·0-66·0)	48·0 (30·0-66·0)	47·0 (31·0-63·0)
% Female	56·6	54·2	61·5	59·0	58·8	57·9	59·4	55·7
Mean vegetable intake, servings/day*	3·7 (3·5-3·9)	3·4 (3·2-3·6)	1·8 (1·6-2·0)	1·8 (1·7-1·9)	1·9 (1·8-2·0)	2·1 (2·0-2·2)	1·3 (1·1-1·5)	1·4 (1·3-1·5)
Mean fruit intake, servings/day*	2·0 (1·8-2·2)	2·0 (1·8-2·2)	1·6 (1·4-1·8)	2·5 (2·3-2·7)	1·2 (1·0-1·3)	1·2 (1·1-1·3)	0·9 (0·7-1·1)	0·7 (0·6-0·9)
Median energy intake-kcal/day (IQR) α	2,188 (1,54-3,322)	2,154 (1,016-3,292)	2,009 (890-3,128)	2,121 (914-3,328)	1,836 (871-2,801)	1,960 (1,035-2,885)	2,021 (806-3,238)	1,984 (882-3,086)
% Current smokers	13·1	13·9	25·1	22·1	24·5	20·0	25·3	23·4
% Low education	8·1	11·8	66·0	49·5	49·3	38·2	59·6	53·4

*Mean (95% CI) accounting for clustering of households within communities.

α Plausible energy intake ≥500 kcal/day and ≤5000 kcal/day.

Table S7: Countries included in the analysis with number of communities, household and participants.

		Number of Communities	Number of Households	Number of Participants
Total		518	90,247	130,402
High income countries (HIC)		72	10,245	13,822
	Canada	46	6,579	8,831
	Sweden	23	2,709	3,622
	UAE	3	957	1,369
Upper middle income countries (UMIC)		125	25,066	35,069
	Argentina	20	2,447	3,400
	Brazil	14	3,717	4,896
	Chile	5	2,228	3,316
	Malaysia	35	10,567	14,696
	Poland	4	1,499	2,031
	South Africa	9	2,090	2,685
	Turkey	38	2,518	4,045
Lower middle income countries (LMIC)		187	35,712	52,090
	China	81	25,922	39,151
	Colombia	55	4,863	6,358
	Iran	20	3,552	5,204
	Palestine	31	1,375	1,377
Low income countries (LIC)		134	19,224	29,421
	Bangladesh	42	1,554	2,185
	India	87	16,562	25,448
	Pakistan	4	1,043	1,713
	Zimbabwe	1	65	75

Table S8. Mean ± SD number of children and adults in participating households, by country.

Region	Country	# of individuals in household	Range of individuals in household	# of adults in household	Range of adults in household	# of children in household*±	Range of children in household
HIC (n=10,245)	Canada	2·75 ± 1·38	1-11	2·39 ± 1·14	1-10	0·34 ± 0·76	0-6
	Sweden	2·53 ± 1·55	1-8	2·19 ± 0·96	1-7	0·34 ± 0·73	0-4
	UAE	7·63 ± 3·87	1-29	5·60 ± 2·93	1-18	2·03 ± 1·96	0-14
UMIC (n=25,066)	Argentina	3·64 ± 1·77	1-15	3·02 ± 1·34	1-10	0·62 ± 0·96	0-6
	Brazil	3·06 ± 1·45	1-14	2·74 ± 1·28	1-13	0·29 ± 0·63	0-5
	Chile	3·71 ± 2·64	1-12	3·23 ± 1·36	1-10	0·49 ± 0·77	0-5
	Malaysia	3·99 ± 2·33	1-43	3·15 ± 1·90	1-16	0·75 ± 1·18	0-8
	Poland	2·90 ± 1·52	1-10	2·68 ± 1·31	1-7	0·21 ± 0·53	0-4
	South Africa	4·86 ± 2·92	1-24	3·71 ± 2·10	1-14	1·08 ± 1·42	0-11
	Turkey	4·17 ± 1·83	1-14	3·57 ± 1·45	1-11	0·60 ± 0·90	0-7
LMIC (n=35,712)	China	3·04 ± 1.28	1-14	2·80 ± 1·11	1-11	0·23 ± 0·50	0-6
	Colombia	4·00 ± 2.00	1-22	3·31 ± 1·62	1-16	0·65 ± 0·99	0-8
	Iran	4·15 ± 1·34	1-12	3·64 ± 1·36	1-10	0·51 ± 0·73	0-5
	Palestine	6·24 ± 2·45	1-18	4·74 ± 2·10	1-12	1·47 ± 1·56	0-7
LIC (n=19,224)	Bangladesh	4·63 ± 1·93	1-10	3·45 ± 1·52	1-16	1·17 ± 1·08	0-6
	India	4·92 ± 2·50	1-35	3·94 ± 1·90	1-35	0·97 ± 1·24	0-16
	Pakistan	5·79 ± 2·60	1-17	4·22 ± 2·13	1-13	1·56 ± 1·56	0-8
	Palestine	5·34 ± 2·29	1-12	4·03 ± 1·77	1-9	1·25 ± 1·13	0-4

*Children defined as less than or equal to 12 years of age.

±Missing age of household member assumed to be less than or equal to 12 years of age.

Table S9. Frequency of the availability (%) of fruits and vegetables in communities, by economic region.

	Overall (n=518)	**HIC (n=72)**	**UMIC (n=125)**	**LMIC (n=187)**	**LIC (n=134)**
Apple	502 (96·9)	72 (100)	121 (96·8)	180 (96·3)	129 (96·3)
Banana	506 (97·7)	72 (100)	125 (100)	176 (94·1)	133 (99·3)
Orange	449 (86·7)	70 (97·2)	112 (89·6)	148 (79·1)	119 (88·8)
Pear	337 (65·1)	70 (97·2)	99 (79·2)	139 (74·3)	29 (21·6)
Cabbage	452 (87·3)	72 (100)	109 (87·2)	176 (94·1)	95 (70·9)
Carrot	483 (93·2)	72 (100)	118 (94·4)	180 (96·3)	113 (84·3)
Tomato	505 (97·5)	70 (97·2)	125 (100)	184 (98·4)	126 (94·0)

No. (%)

Table S10. Mean (SD) PPP adjusted cost of fruits and vegetables in communities, by economic region.

	Overall (n=518)	**HIC (n=72)**	**UMIC (n=125)**	**LMIC (n=187)**	**LIC (n=134)**
Apple	0·35 ± 0·19	0·27 ± 0·07	0·29 ± 0·16	0·27 ± 0·12	0·55 ± 0·20
Banana	0·19 ± 0·10	0·21 ± 0·13	0·25 ± 0·12	0·17 ± 0·08	0·14 ± 0·03
Orange	0·23 ± 0·12	0·25 ± 0·09	0·21 ± 0·09	0·18 ± 0·11	0·32 ± 0·11
Pear	0·29 ± 0·14	0·27 ± 0·07	0·35 ± 0·12	0·26 ± 0·17	0·40 ± 0·19
Cabbage	0·14 ± 0·08	0·20 ± 0·08	0·18 ± 0·09	0·11 ± 0·06	0·08 ± 0·05
Carrot	0·14 ± 0·06	0·15 ± 0·05	0·17 ± 0·07	0·12 ± 0·06	0·13 ± 0·06
Tomato	0·19 ± 0·13	0·37 ± 0·15	0·23 ± 0·11	0·15 ± 0·05	0·10 ± 0·06

± Standard deviation

Table S11. Mean cost of 1 serving of vegetables and fruits as a percent of household member income (relative cost) using equal and unequally weighted income for adults and children, by economic region.

Region	**Mean relative vegetable cost**		**Mean relative fruit cost**	
	Equal weight*	**Unequal weight±**	**Equal weight***	**Unequal weight±**
HIC (n=13,799)	0·56 (0·49 to 0·63)	0·52 (0·46 to 0·59)	0·62 (0·40 to 0·84)	0·58 (0·45 to 0·71)
UMIC (n=32,930)	4·50 (3·47 to 5·53)	4·10 (3·17 to 5·03)	5·94 (5·80 to 6·08)	5·40 (4·29 to 6·51)
LMIC (n=51,422)	3·18 (2·74 to 3·62)	3·00 (2·59 to 3·41)	4·98 (4·87 to 5·09)	4·71 (4·08 to 5·34)
LIC (n=13,327)	8·40 (6·15 to 10·66)	7·77 (5·69 to 9·84)	21·73 (21·51 to 21·96)	20·23 (15·37 to 25·09)

*Equal weight= relative cost computed using equal income for all household members (children and adults).

±Unequal weight= relative cost computed using 50% less income for household members 12 years of age and younger compared to household members greater than 12 years of age.

Table S12. Mean cost of 2 servings of vegetables and 3 servings of fruits as a percent of household member income (relative cost), by economic region and urban/rural location.

Region	**Urban**		**Rural**	
	Equal weight*	**Unequal weight±**	**Equal weight***	**Unequal weight±**
HIC (n=13,799)	1·62 (1·36 to 1·88)	1·52 (1·29 to 1·75)	2·43 (1·41 to 3·44)	2·23 (1·18 to 3·28)
UMIC (n=32,930)	12·92 (8·80 to 17·04)	11·99 (8·13 to 15·86)	21·90 (21·38 to 22·42)	19·71 (14·30 to 25·12)
LMIC (n=51,422)	9·31 (8·40 to 10·22)	8·84 (7·99 to 9·70)	19·47 (19·05 to 19·89)	18·44 (15·09 to 21·80)
LIC (n=13,327)	25·44 (15·09 to 35·80)	23·52 (13·85 to 33·18)	49·28 (48·48 to 50·07)	45·65 (37·97 to 53·33)

*Equal weight= relative cost computed using equal income for all household members (children and adults).

±Unequal weight= relative cost computed using 50% less income for household members 12 years of age and younger compared to household members greater than 12 years of age.

Table S13. Mean number of individuals unable to afford 2 servings of vegetables and 3 servings of fruits, by economic region and urban/rural location.

Region	**Urban**		**Rural**	
	Equal weight*	**Unequal weight±**	**Equal weight***	**Unequal weight±**
HIC (n=13,799)	0·17 (0·09 to 0·25)	0·10 (0·041 to 0·16)	0·47 (0·25 to 0·69)	0·34 (0·15 to 0·53)
UMIC (n=32,930)	17·69 (17·14 to 18·24)	14·79 (14·27 to 15·31)	35·14 (34·47 to 36·01)	27·33 (26·60 to 28·06)
LMIC (n=51,422)	8·64 (8·32 to 8·96)	5·95 (5·68 to 6·22)	29·44 (28·84 to 30·04)	24·48 (23·92 to 25·04)
LIC (n=13,327)	38·08 (36·96 to 39·20)	31·44 (30·27 to 32·51)	79·93 (78·93 to 80·93)	74·89 (73·81 to 75·97)

*Equal weight= relative cost computed using equal income for all household members (children and adults).

±Unequal weight= relative cost computed using 50% less income for household members 12 years of age and younger compared to household members greater than 12 years of age.

Table S14. Percentage of individuals unable to afford fruit and vegetable recommendations at thresholds in 10% increments.

Region	N	>10%	>20%	>30%	>40%	>50%	>60%	>70%	>80%	>90%	>100%
HIC	13,799	1.46	0.23	0.08	0.02	0.01	0.01	0.00	0.00	0.00	0.00
UMIC	32,930	43.49	22.55	13.81	9.28	6.48	4.94	3.77	2.84	2.31	1.81
LMIC	51,422	40.61	15.31	8.21	4.82	3.38	2.58	1.86	1.42	1.10	0.88
LIC	13,327	73.27	54.47	39.82	28.92	21.57	16.27	12.07	9.00	7.67	6.00

Figure S1

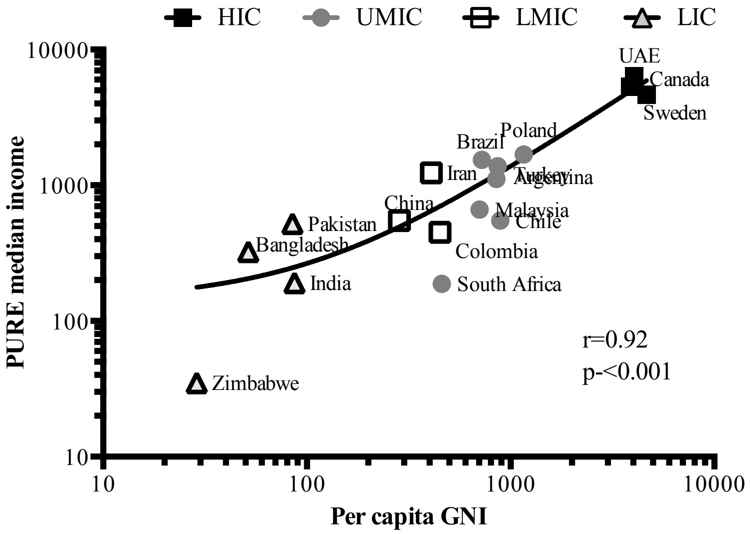


Figure S1. Log median income per country in PURE compared to the World Bank Gross National Income (GNI) estimates.

Figure S2

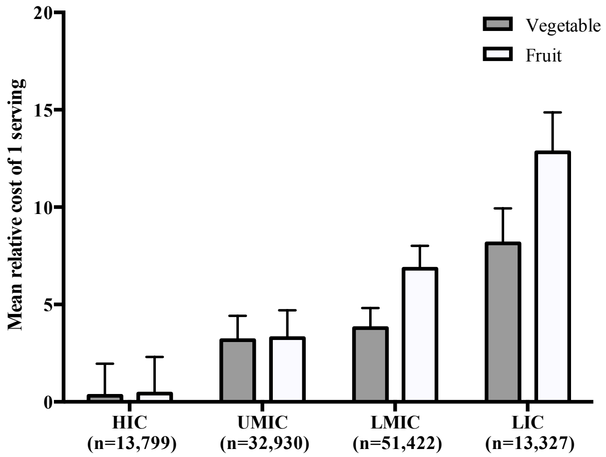


Figure S2.

Mean (95% CI) percent of income required to purchase 1 serving of vegetables and fruits per day (relative cost) using the lowest vegetable and fruit cost per community to compute the relative cost.

Figure S3A

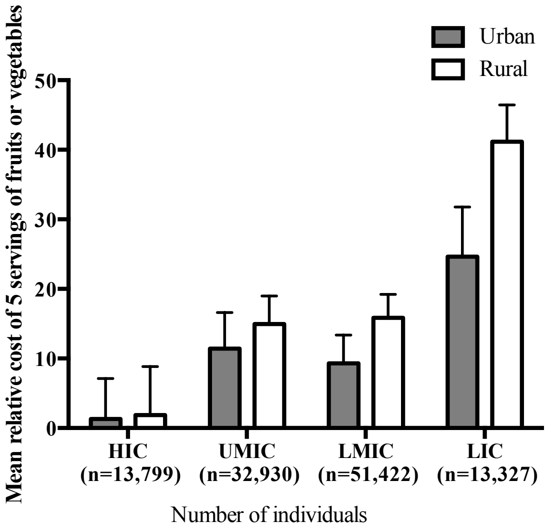


Figure S3B

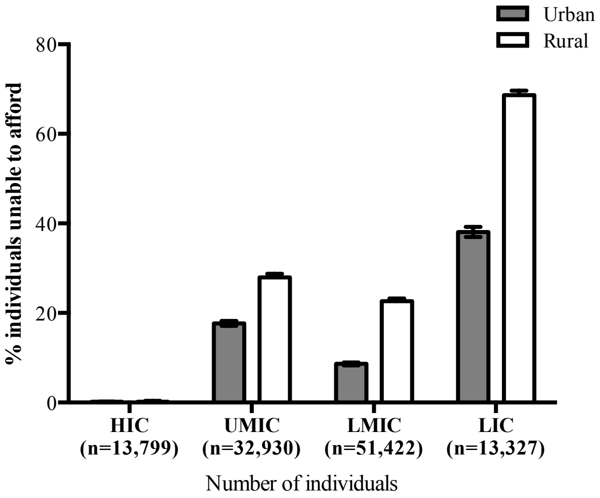


Figure S3.

Mean (95% CI) percent of income required to purchase 5 servings of vegetables or fruits per day (relative cost) (Figure 3A) and the percentage (95% CI) of individuals unable to afford 5 servings of vegetables or fruits per day (Figure 3B), by economic region and urban/rural location.

Figure S4

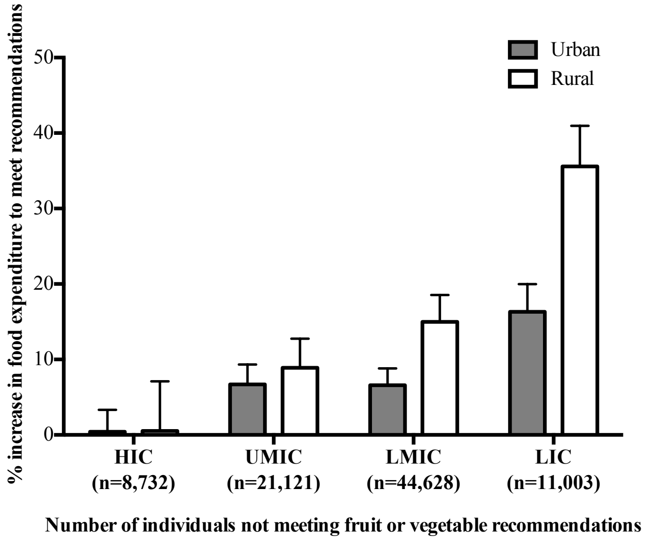


Figure S4.

Mean (95% CI) percent increase in food expenditure required to meet the recommendation of 2 servings of fruits and 3 servings per day in individuals not meeting the recommendation, by economic region and urban/rural location.

References for Supplementary Material

1. Teo K, Chow CK, Vaz M, Rangarajan S, Yusuf S; PURE Investigators-Writing Group. The Prospective Urban Rural Epidemiology (PURE) study: examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. Am Heart J 2009;158:1-7.e1.
2. Corsi DJ, Subramanian SV, Chow CK, McKee M, Chifamba J, Dagenais G, et al. Prospective Urban Rural Epidemiology (PURE) study: Baseline characteristics of the household sample and comparative analyses with national data in 17 countries. Am Heart J 2013 Oct;166(4):636-646.e4.
3. Kelemen LE, Anand SS, Vuksan V, Yi Q, Teo KK, Devanesen S, et al. Development and evaluation of cultural food frequency questionnaires for South Asians, Chinese, and Europeans in North America. J Am Diet Assoc 2003 Sep;103(9):1178-1184.
4. Bharathi AV, Kurpad AV, Thomas T, Yusuf S, Saraswathi G, Vaz M. Development of food frequency questionnaires and a nutrient database for the Prospective Urban and Rural Epidemiological (PURE) pilot study in South India: methodological issues. Asia Pac J Clin Nutr 2008;17(1):178-185.
5. Dehghan M, Martinez S, Zhang X, Seron P, Lanas F, Islam S, et al. Relative validity of an FFQ to estimate daily food and nutrient intakes for Chilean adults. Public Health Nutr 2013 Oct;16(10):1782-1788.
6. Dehghan M, del Cerro S, Zhang X, Cuneo JM, Linetzky B, Diaz R, et al. Validation of a semi-quantitative Food Frequency Questionnaire for Argentinean adults. PLoS One 2012;7(5):e37958.
7. Dehghan M, Ilow R, Zatonska K, Szuba A, Zhang X, Mente A, et al. Development, reproducibility and validity of the food frequency questionnaire in the Poland arm of the

Prospective Urban and Rural Epidemiological (PURE) study. J Hum Nutr Diet 2012 Jun;25(3):225-232.

8. Dehghan M, Lopez Jaramillo P, Duenas R, Anaya LL, Garcia RG, Zhang X, et al. Development and validation of a quantitative food frequency questionnaire among rural- and urban-dwelling adults in Colombia. J Nutr Educ Behav 2012 Nov-Dec;44(6):609-613.
9. Dehghan M, Al Hamad N, Yusufali A, Nusrath F, Yusuf S, Merchant AT. Development of a semi-quantitative food frequency questionnaire for use in United Arab Emirates and Kuwait based on local foods. Nutr J 2005 May 27;4:18.
10. Merchant AT, Dehghan M, Chifamba J, Terera G, Yusuf S. Nutrient estimation from an FFQ developed for a Black Zimbabwean population. Nutr J 2005 Dec 13;4:37.
11. Mikkelsen TB, Olsen SF, Rasmussen SE, Osler M. Relative validity of fruit and vegetable intake estimated by the food frequency questionnaire used in the Danish National Birth Cohort. Scand J Public Health. 2007;35(2):172-9.
12. Mohammadifard N, Omidvar N, Houshiarrad A, Neyestani T, Naderi GA, Soleymani B. Validity and reproducibility of a questionnaire for assessment of fruit and vegetable intake in Iranian adults. J Res Med Sci. 2011 Oct;16(10):1286-97.
13. Landais E, Gartner A, Bour A, McCullough F, Delpeuch F, Holdsworth M. Reproducibility and relative validity of a brief quantitative food frequency questionnaire for assessing fruit and vegetable intake in North-African women. J Hum Nutr Diet. 2014 Apr;27 Suppl 2:152-9.
14. Johansson I, Hallmans G, Wikman A, Biessy C, Riboli E, Kaaks R. Validation and calibration of frood-frequency questionnaire measurements in the Northern Sweden Health and Disease cohort. Public Health Nutr. 2002 Jun;5(3):487-96.

15. Marks GC, Hughes MC, van der Pols JC. Relative validity of food intake estimates using a food frequency questionnaire is associated with sex, age, and other personal characteristics. J Nutr. 2006 Feb;136(2):459-65.
16. Salvini S, Hunter DJ, Sampson L, Stampfer MJ, Colditz GA, Rosner B, Willett WC. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. Int J Epidemiol. 1989 Dec;18(4):858-67.
17. Chow CK, Lock K, Madhavan M, Corsi DJ, Gilmore AB, Subramanian SV, et al. Environmental Profile of a Community's Health (EPOCH): an instrument to measure environmental determinants of cardiovascular health in five countries. PLoS One 2010 Dec 10;5(12):e14294.
18. Corsi DJ, Subramanian SV, McKee M, Li W, Swaminathan S, Lopez-Jaramillo P, et al. Environmental Profile of a Community's Health (EPOCH): an ecometric assessment of measures of the community environment based on individual perception. PLoS One 2012;7(9):e44410.
19. Jones HA, Charlton KE. A Cross-sectional analysis of the cost and affordability of achieving recommended intakes of non-starchy fruits and vegetables in the capital of Vanuatu. BMC Public Health 2015 Mar;15:301.
20. Pollard CM, Landrigan TJ, Ellies PL, Kerr DA, Lester ML, Goodchild SE. Geographical factors as determinants of food security: a Western Australian food pricing and quality study. Asia Pac J Clin Nutr 2014;23(4):703-13.
21. Yuba TY, Sarti FM, Campino AC, Carmo HC. Evolution of the relative prices of food groups between 1939 and 2010 in the city of Sao Paulo, Southeastern Brazil. Rev Saude Publica 2013 Jun;47(3):549-59.

22. Claro RM, Monteiro CA. Family income, food prices and household purchases of fruits and vegetables in Brazil. Rev Saude Publica 2010 Dec;44(6):1014-20.
23. WHO Europe. Using price policies to promote healthier diets. Copenhagen: World Health Organization Europe; 2015.
24. Swinburg B, Sacks G, Vandavevijvere S, Kumanyika S, Lobstein T, et al. INFORMAS (International Network for Food and Obesity/non-communicable disease Research, Monitoring and Action Support): overview and key principles. Obes Rev 2013;14(S1)1-12.
25. FAO, IFAD and WFP. The state of food insecurity in the world 2013. Rome: Food and Agriculture Organization; 2013.
26. World Bank. GNI per capita, PPP. Washington: World Bank; 2008.
27. Micha R, Khatibzadeh S, Shi P, Andrews KG, Engell RE, Mozaffarian D. Global, regional and national consumption of major food groups in 1990 and 2010: a systematic analysis including 266 country-specific nutrition surveys worldwide. BMJ Open. 2015 Sep 24;5(9):e008705.
28. Thompson FE, Kipnis V, Subar AF, Krebs-Smith SM, Kahle LL, Midthune D, Potischman N, Schatzkin A. Evaluation of 2 brief instruments and a food-frequency questionnaire to estimate daily number of servings of fruit and vegetables. Am J Clin Nutr. 2000 Jun;71(6):1503-10.
29. Kristal AR, Shattuck AL, Henry HJ, Fowler AS. Rapid assessment of dietary intake of fat, fiber, and saturated fat: validity of an instrument suitable for community intervention research and nutritional surveillance. Am J Health Promot 1990;4:288–95.

30. Behavioral Risk Factor Surveillance System, 1998. Survey data, National Center for Chronic Disease Prevention and Health Promotion Centers for Disease Control and Prevention, US Department of Health and Human Services.
31. Thompson B, Demark-Wahnefried W, Taylor G, McClelland JW, Stables G, Havas S, et al. Baseline fruit and vegetable intake among adults in seven 5 A Day study centers located in diverse geographic areas. J Am Diet Assoc 1999;99:1241–48.
32. Kim DJ, Holowaty EJ. Brief, validated survey instruments for the measurement of fruit and vegetable intakes in adults: a review. Prev Med. 2003;36:440–447.
33. Smith AM, Smith C. Dietary intake and lifestyle patterns: correlates with socioeconomic, demographic and environmental factors. J of Hum Nutr Diet. 1994;7:283-294.
34. Thompson RL, Margetts BM, Speller VM, McVey D. The Health Education Authority's health and lifestyle survey 1993: who are the low fruit and vegetable consumers. J Epidemiol Community Health. 1999 May;53(5):294-9.
35. Wardle J, Parmenter K, Waller J. Nutrition knowledge and food intake. Appetite. 2000 Jun;34(3):269-75.
36. Baker AH, Wardle J. Sex differences in fruit and vegetable intake in older adults. Appetite. 2003 Jun;40(3):269-75.
37. Satheannoppakao W, Aekplakorn W, Pradipasen M. Fruit and vegetable consumption and its recommended intake associated with sociodemographic factors: Thailand National Health Examination Survey III. Public Health Nutr. 2009 Nov;12(11):2192-8.
38. Dehghan M, Akhtar-Danesh N, Merchant AT. Factors associated with fruit and vegetable consumption among adults. J Hum Nutr Diet. 2011 Apr;24(2):128-34.

39. Figueiredo IC, Jaime PC, Monteiro CA. Factors associated with fruit and vegetable intake among adults of the city of Sao Paulo, Southeastern Brazil. Rev Saude Publica. 2008 Oct;42(5):777-85.
40. Neutzling MB, Rombaldi AJ, Azevedo MR, Hallal PC. Factors associated with fruit and vegetable intake among adults in a southern Brazilian city. Cad Saude Publica. 2009 Nov;25(11):2365-74.
41. Lallukka T, Laaksonen M, Rahkonen O, Roos E, Lahelma E. Multiple socio-economic circumstances and healthy food habits. Eur J Clin Nutr. 2007;61:701–10.
42. Rohrmann S, Giovannucci E, Willett WC, Platz EA. Fruit and vegetable consumption, intake of micronutrients, and benign prostatic hyperplasia in US men. Am J Clin Nutr. 2007 Feb;85(2):523-9.
43. Corley J, Kyle JA, Starr JM, McNeill G, Deary IJ. Dietary factors and biomarkers of systemic inflammation in older people: the Lothian Birth Cohort 1936. Br J Nutr. 2015 Oct 14;114(7):1088-98.

Funding
The main PURE study and its components are funded by the Population Health Research Institute, the Canadian Institutes of Health Research, Heart and Stroke Foundation of Ontario and through unrestricted grants from several pharmaceutical companies [with major contributions from Boehringer Ingelheim (Germany & Canada), Astra Zeneca (Canada), Sanofi-Aventis (France and Canada), Servier, and GSK], and additional contributions from Novartis and King Pharma and from various national or local organizations in participating countries. These include: **Bangladesh:** Independent University, Bangladesh, Mitra and Associates; **Brazil:** Unilever Health Institute, Brazil; **Canada:** Public Health Agency of Canada and Champlain Cardiovascular Disease Prevention Network; **Chile:** Universidad de la Frontera; China: National Centre for Cardiovascular Diseases; **Colombia:** Colciencias, Grant number:6566-04-18062; **India:** Indian Council of Medical Research; **Malaysia:** Ministry of Science, Technology and Innovation of Malaysia Grant Nbr 100 - IRDC / BIOTEK 16/6/21 (13/2007), Grant Number 07-05-IFN-MEB010, Ministry of Higher Education of Malaysia Grant Nbr 600 - RMI/LRGS/5/3 (2/2011), Universiti Teknologi MARA, Universiti Kebangsaan Malaysia (UKM-Hejim-Komuniti-15-2010); **occupied Palestinian territory:** the United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA), occupied Palestinian territory; International Development Research Centre (IDRC), Canada; **Poland:** Polish Ministry of Science and Higher Education grant Nr 290/W-PURE/2008/0, Wroclaw Medical University; **South Africa:** The North-West University, SANPAD (SA and Netherlands Programme for Alternative Development), National Research Foundation, Medical Research Council of SA, The SA Sugar Association (SASA), Faculty of Community and Health Sciences (UWC); **Sweden**: AFA Insurance, Swedish Council for Working Life and Social Research, Swedish Research

Council for Environment, Agricultural Sciences and Spatial Planning, Swedish Heart and Lung Foundation, Swedish Research Council, Grant from the Swedish State under (LäkarUtbildningsAvtalet), Agreement, Grant from the Västra Götaland Region (FOUU); **TURKEY**: Metabolic Syndrome Society, Astra Zeneca, Turkey, Sanofi Aventis, Turkey; **UAE**: Sheikh Hamdan Bin Rashid Al Maktoum Award For Medical Sciences, Dubai Health Authority, Dubai UAE.

Clara Chow is supported by an Australian NHMRC Career development Fellowship co-funded by the National Heart Foundation.

Acknowledgments

Study Investigators:

Project office (Population Health Research Institute, Hamilton Health Sciences and McMaster University, Hamilton, Canada): S Yusuf* (Principal Investigator).

S Rangarajan (Project Manager); K K Teo, C K Chow, M O'Donnell, A Mente, D Leong, ASmyth, P Joseph, S Islam (Statistician), M Zhang (Statistician), W Hu (Statistician), G Wong (Statistician), ShunFu Lee (Statistician), A Casanova, L Dyal, M Dehghan (Nutritionist), J DeJesus, G Lewis, P Mackie, S Chin, L Farago, M Zarate, I Kay, D Agapay, R Solano, S Ramacham, N Kandy, J Rimac, S Trottier, W ElSheikh, M Mustaha, T Tongana, N Aoucheva, J Swallow, E Ramezani, N Aoucheva, A Aliberti, J Lindeman

ARGENTINA: R Diaz*; A Orlandini, B Linetsky, S Toscanelli, G Casaccia, JM Maini Cuneo; BANGLADESH: O Rahman*, R Yusuf, AK Azad, KA Rabbani, HM Cherry, A Mannan, I Hassan, AT Talukdar, RB Tooheen, MU Khan, M Sintaha, T Choudhury, R Haque, S Parvin; BRAZIL: A Avezum*, GB Oliveira, CS Marcilio, AC Mattos; CANADA: K Teo*, S Yusuf*, J Dejesus, D Agapay, T Tongana, R Solano, I Kay, S Trottier, J Rimac, W Elsheikh, L Heldman, E Ramezani, G Dagenais, P Poirier, G Turbide, D Auger, A LeBlanc De Bluts, MC Proulx, M Cayer, N Bonneville, S Lear, D Gasevic, E Corber, V de Jong, I Vukmirovich, A Wielgosz, G Fodor, A Pipe, A Shane; CHILE: F Lanas*, P Seron, S Martinez, A Valdebenito, M Oliveros; CHINA: Li Wei*, Liu Lisheng*, Chen Chunming, Wang Xingyu, Zhao Wenhua, Zhang Hongye, JiaXuan, Hu Bo, Sun Yi, Bo Jian, Zhao Xiuwen, Chang Xiaohong, Chen Tao, Chen Hui, Chang Xiaohong, Deng Qing, Cheng Xiaoru, Deng Qing, He Xinye, Hu Bo, JiaXuan, Li

Jian, Li Juan,Liu Xu, Ren Bing, Sun Yi, Wang Wei, Wang Yang, Yang Jun, Zhai Yi, Zhang Hongye, Zhao Xiuwen,Zhu Manlu, Lu Fanghong, Wu Jianfang, Li Yindong, Hou Yan, Zhang Liangqing, Guo Baoxia, Liao Xiaoyang, Zhang Shiying, BianRongwen, TianXiuzhen, Li Dong, Chen Di, Wu Jianguo, Xiao Yize, Liu Tianlu, Zhang Peng, Dong Changlin, Li Ning, Ma Xiaolan, Yang Yuqing, Lei Rensheng, Fu Minfan, He Jing, Liu Yu, Xing Xiaojie, Zhou Qiang; COLOMBIA: P Lopez-Jaramillo*, PACamacho Lopez, R Garcia, LJA Jurado, D Gómez Arbeláez, JF Arguello, R Dueñas, S Silva, LP Pradilla, F Ramirez, DI Molina, C Cure-Cure, M Perez, E Hernandez, E Arcos, S Fernandez, C Narvaez, A Sotomayor, H Garcia, G Sanchez, T David, A Rico; INDIA: P Mony *, M Vaz*, A V Bharathi, S Swaminathan, K Shankar AV Kurpad, KG Jayachitra, N Kumar, HAL Hospital, V Mohan, M Deepa, K Parthiban, M Anitha, S Hemavathy, T Rahulashankiruthiyayan, D Anitha, K Sridevi, R Gupta, RB Panwar, I Mohan, P Rastogi, S Rastogi, R Bhargava, R Kumar, J S Thakur, B Patro, PVM Lakshmi, R Mahajan, P Chaudary, V Raman Kutty, K Vijayakumar, K Ajayan, G Rajasree, AR Renjini, A Deepu, B Sandhya, S Asha, HS Soumya; IRAN: R Kelishadi*, A Bahonar, N Mohammadifard, H Heidari; MALAYSIA: K Yusoff*, TST Ismail, KK Ng, A Devi, NM Nasir, MM Yasin, M Miskan, EA Rahman, MKM Arsad, F Ariffin, SA Razak, FA Majid, NA Bakar, MY Yacob, N Zainon, R Salleh, MKA Ramli, NA Halim, SR Norlizan, NM Ghazali, MN Arshad, R Razali, S Ali, HR Othman, CWJCW Hafar, A Pit, N Danuri, F Basir, SNA Zahari, H Abdullah, MA Arippin, NA Zakaria, I Noorhassim, MJ Hasni, MT Azmi, MI Zaleha, KY Hazdi, AR Rizam, W Sazman, A Azman; OCCUPIED PALESTINIAN TERRITORY: R Khatib*, U Khammash, A Khatib, R Giacaman; PAKISTAN: R Iqbal*, A Afridi, R Khawaja, A Raza, K Kazmi; POLAND: W Zatonski*, A Szuba, K Zatonska, R Ilow#, M Ferus, B Regulska-Ilow, D Różańska, M Wolyniec; SOUTH AFRICA: A Kruger*, H H Voster, A E Schutte, E Wentzel-

Viljoen, FC Eloff, H de Ridder, H Moss, J Potgieter, AA Roux, M Watson, G de Wet, A Olckers, JC Jerling, M Pieters, T Hoekstra, T Puoane, E Igumbor, L Tsolekile, D Sanders, P Naidoo, N Steyn, N Peer, B Mayosi, B Rayner, V Lambert, N Levitt, T Kolbe-Alexander, L Ntyintyane, G Hughes, R Swart, J Fourie, M Muzigaba, S Xapa, N Gobile , K Ndayi, B Jwili, K Ndibaza, B Egbujie; SWEDEN: A Rosengren*, K Bengtsson Boström, U Lindblad, P Langkilde, A Gustavsson, M Andreasson, M Snällman, L Wirdemann, K Pettersson, E Moberg ; TURKEY: A Oguz*, AAK Akalin, KBT Calik, N Imeryuz, A Temizhan, E Alphan, E Gunes, H Sur, K Karsidag, S Gulec, Y Altuntas; UNITED ARAB EMIRATES: AM Yusufali*, W Almahmeed, H Swidan, EA Darwish, ARA Hashemi, N Al-Khaja, JM Muscat-Baron, SH Ahmed, TM Mamdouh, WM Darwish, MHS Abdelmotagali, SA Omer Awed, GA Movahedi, F Hussain, H Al Shaibani, RIM Gharabou, DF Youssef, AZS Nawati, ZAR Abu Salah, RFE Abdalla, SM Al Shuwaihi, MA Al Omairi, OD Cadigal; R.S. Alejandrino; ZIMBABWE: J Chifamba*, L Gwaunza, V Mapara, C Mahachi, P Mrambiwa, T Machiweni, R Mapanga, B Ncube.

*National Coordinator

Deceased

Chapter 2

Fruit, vegetable, and legume intake, and cardiovascular disease and deaths in 18 countries (PURE): a prospective cohort study

Reprinted from *The Lancet;* 390(10107); Victoria Miller, Andrew Mente, Mahshid Dehghan, Sumathy Rangarajan, Xiaohe Zhang, Sumathi Swaminathan, Gilles Dagenais, Rajeev Gupta, Viswanathan Mohan, Scott Lear, Shrikant I Bangdiwala, Aletta E Schutte, Edelweiss Wentzel-Viljoen, Alvaro Avezum, Yuksel Altuntas, Khalid Yusoff, Noorhassim Ismail, Nasheeta Peer, Jephat Chifamba, Rafael Diaz, Omar Rahman, Noushin Mohammadifard, Fernando Lana, Katarzyna Zatonska, Andreas Wielgosz, Afzalhussein Yusufali, Romaina Iqbal, Patricio Lopez-Jaramillo, Rasha Khatib, Annika Rosengren, V Raman Kutty, Wei Li, Jiankang Liu, Xiaoyun Liu, Lu Yin, Koon Teo, Sonia Anand, and Salim Yusuf; 2037-2049; Copyright 2017; with permission from Elsevier. doi:10.1016/ S0140-6736(17)32253-5

Fruit, vegetable, and legume intake, and cardiovascular disease and deaths in 18 countries (PURE): a prospective cohort study

*Victoria Miller, Andrew Mente, Mahshid Dehghan, Sumathy Rangarajan, Xiaohe Zhang, Sumathi Swaminathan, Gilles Dagenais, Rajeev Gupta, Viswanathan Mohan, Scott Lear, Shrikant I Bangdiwala, Aletta E Schutte, Edelweiss Wentzel-Viljoen, Alvaro Avezum, Yuksel Altuntas, Khalid Yusoff, Noorhassim Ismail, Nasheeta Peer, Jephat Chifamba, Rafael Diaz, Omar Rahman, Noushin Mohammadifard, Fernando Lana, Katarzyna Zatonska, Andreas Wielgosz, Afzalhussein Yusufali, Romaina Iqbal, Patricio Lopez-Jaramillo, Rasha Khatib, Annika Rosengren, V Raman Kutty, Wei Li, Jiankang Liu, Xiaoyun Liu, Lu Yin, Koon Teo, Sonia Anand, Salim Yusuf, on behalf of the Prospective Urban Rural Epidemiology (PURE) study investigators**

Summary

Background The association between intake of fruits, vegetables, and legumes with cardiovascular disease and deaths has been investigated extensively in Europe, the USA, Japan, and China, but little or no data are available from the Middle East, South America, Africa, or south Asia.

Methods We did a prospective cohort study (Prospective Urban Rural Epidemiology [PURE] in 135 335 individuals aged 35 to 70 years without cardiovascular disease from 613 communities in 18 low-income, middle-income, and high-income countries in seven geographical regions: North America and Europe, South America, the Middle East, south Asia, China, southeast Asia, and Africa. We documented their diet using country-specific food frequency questionnaires at baseline. Standardised questionnaires were used to collect information about demographic factors, socioeconomic status (education, income, and employment), lifestyle (smoking, physical activity, and alcohol intake), health history and medication use, and family history of cardiovascular disease. The follow-up period varied based on the date when recruitment began at each site or country. The main clinical outcomes were major cardiovascular disease (defined as death from cardiovascular causes and non-fatal myocardial infarction, stroke, and heart failure), fatal and non-fatal myocardial infarction, fatal and non-fatal strokes, cardiovascular mortality, non-cardiovascular mortality, and total mortality. Cox frailty models with random effects were used to assess associations between fruit, vegetable, and legume consumption with risk of cardiovascular disease events and mortality.

Findings Participants were enrolled into the study between Jan 1, 2003, and March 31, 2013. For the current analysis, we included all unrefuted outcome events in the PURE study database through March 31, 2017. Overall, combined mean fruit, vegetable and legume intake was 3·91 (SD 2·77) servings per day. During a median 7·4 years (5·5–9·3) of follow-up, 4784 major cardiovascular disease events, 1649 cardiovascular deaths, and 5796 total deaths were documented. Higher total fruit, vegetable, and legume intake was inversely associated with major cardiovascular disease, myocardial infarction, cardiovascular mortality, non-cardiovascular mortality, and total mortality in the models adjusted for age, sex, and centre (random effect). The estimates were substantially attenuated in the multivariable adjusted models for major cardiovascular disease (hazard ratio [HR] 0·90, 95% CI 0·74–1·10, p_{trend}=0·1301), myocardial infarction (0·99, 0·74–1·31; p_{trend}=0·2033), stroke (0·92, 0·67–1·25; p_{trend}=0·7092), cardiovascular mortality (0·73, 0·53–1·02; p_{trend}=0·0568), non-cardiovascular mortality (0·84, 0·68–1·04; p_{trend} =0·0038), and total mortality (0·81, 0·68–0·96; p_{trend}<0·0001). The HR for total mortality was lowest for three to four servings per day (0·78, 95% CI 0·69–0·88) compared with the reference group, with no further apparent decrease in HR with higher consumption. When examined separately, fruit intake was associated with lower risk of cardiovascular, non-cardiovascular, and total mortality, while legume intake was inversely associated with non-cardiovascular death and total mortality (in fully adjusted models). For vegetables, raw vegetable intake was strongly associated with a lower risk of total mortality, whereas cooked vegetable intake showed a modest benefit against mortality.

Interpretation Higher fruit, vegetable, and legume consumption was associated with a lower risk of non-cardiovascular, and total mortality. Benefits appear to be maximum for both non-cardiovascular mortality and total mortality at three to four servings per day (equivalent to 375–500 g/day).

Funding Full funding sources listed at the end of the paper (see Acknowledgments).

See **Online/Comment**
http://dx.doi.org/10.1016/S0140-6736(17)32251-1

*Investigators listed in the appendix

Population Health Research Institute, McMaster University, Hamilton, ON, Canada (V Miller BSc, A Mente PhD, M Dehghan PhD, S Rangarajan MSc, X Zhang MSc, S I Bangdiwala PhD, Prof K Teo MD, Prof S Anand MD, Prof S Yusuf DPhil); **St John's Research Institute, Bangalore, India** (S Swaminathan PhD); **Institut universitaire de cardiologie et pneumologie de Quebec, Université Laval, QC, Canada** (G Dagenais MD); **Eternal Heart Care Centre and Research Institute, Jaipur, India** (Prof R Gupta PhD); **Madras Diabetes Research Foundation, Dr Mohan's Diabetes Specialties Centre, Chennai, India** (Prof V Mohan MD); **Faculty of Health Sciences, Simon Fraser University, Vancouver, BC, Canada** (Prof S Lear PhD); **South African Medical Research Council/North-West University, Potchefstroom, South Africa** (A E Schutte PhD, E Wentzel-Viljoen PhD); **Dante Pazzanese Institute of Cardiology, University Santo Amaro, São Paulo, Brazil** (Prof A Avezum MD); **Istanbul Sisli Hamidiye Etfal Health Training and Research Center, Istanbul, Turkey** (Y Altuntas PhD); **Universiti Teknologi MARA, Selayang, Selangor, Malaysia and UCSI University, Cheras, Kuala Lumpur, Malaysia** (K Yusoff MBBS); **Department of**

Introduction

Several guidelines recommend the consumption of five or more servings per day of fruits, vegetables, and legumes.[1,2] This recommendation is largely based on observational data from Europe and the USA and a few studies from Japan and China. Consumption of these foods is higher in

Community Health, Faculty of Medicine, University of Kebangsaan Malaysia, Kuala Lumpur, Malaysia (N Ismail MD); **Non-Communicable Diseases Research Unit, South African Medical Research Council, Durban, South Africa** (N Peer PhD); **Physiology Department, University of Zimbabwe College of Health Sciences, Harare, Zimbabwe** (J Chifamba DPhil); **Estudios Clinicos Latinoamerica ECLA, Rosaria, Argentina** (R Diaz MD); **Independent University, Dhaka, Bangladesh** (O Rahman MD); **Isfahan Cardiovascular Research Center, Cardiovascular Research Institute, Isfahan University of Medical Sciences, Isfahan, Iran** (N Mohammadifard PhD); **Universidad de La Frontera, Temuco, Chile** (F Lana MD); **Department of Social Medicine, Medical University of Wroclaw, Wroclaw, Poland** (K Zatonska MD); **University of Ottawa Department of Medicine, Ottawa, ON, Canada** (Prof A Wielgosz MD); **Hatta Hospital, Dubai Health Authority and Dubai Medical College, Dubai, United Arab Emirates** (A Yusufali MD); **Aga Khan University, Karachi, Pakistan** (R Iqbal PhD); **Research Institute, FOSCAL and MASIRA Institute, Medical Studies, UDES, Bucaramanga, Colombia** (Prof P Lopez-Jaramillo MD); **Public Health Sciences, Stritch School of Medicine, IL, USA** (R Khatib PhD); **Department of Molecular and Clinical Medicine, Sahlgrenska Academy, University of Gothenburg and Sahlgrenska University Hospital/Őstra, Gőteborg, Sweden** (Prof A Rosengren MD); **Health Action by People, Trivandrum, India** (Prof V R Kutty MD); **State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Disease, Peking Union Medical College & Chinese Academy of Medical Sciences, Beijing, China** (Prof W Li PhD, X Liu PhD, L Yin PhD); **and Jianshe Road Community Health Center, Chengdu, Sichuan Province, China** (J Liu MSc)

Correspondence to:
Ms Victoria Miller, Population Health Research Institute, DBCVS Research Institute, McMaster University, Hamilton, ON, L8L 2X2, Canada.
victoria.miller@phri.ca

Research in context

Evidence before this study

We searched PubMed for articles published between Jan 1, 1960, and May 1, 2017, using the terms “fruit” OR “vegetable” OR “legume” OR “dietary pulse” OR “produce” OR “food” OR “diet” AND “cardiovascular” OR “coronary heart disease” OR “ischemic” OR “myocardial” OR “stroke” OR “death” OR “mortality” OR “sudden cardiac death”. We used search terms in English but did not apply any language restrictions. We screened papers by title and abstract to identify full-text reports that were relevant to the study aims. We also screened citation lists for these full-text reports to identify other relevant articles. Articles were considered relevant if they reported the relation between fruit, vegetable, or legume intake and cardiovascular events or total mortality. Numerous prospective cohort studies have reported inverse associations between fruit, vegetable and legume intake and cardiovascular events and total mortality, but existing evidence was limited to studies predominantly from European countries, the USA, Japan, and China, with little data available from other regions of the world. Additionally, previous studies have shown that the consumption of fruits, vegetables, and legumes is low among many populations, predominantly those in countries outside Europe or the USA. It is unclear whether these food items are beneficially associated with cardiovascular risk in non-Western countries.

Added value of this study

In our analysis of 135 335 participants from 18 countries, we assessed relationships across a broad range of intakes including very low consumption of fruits and vegetables, and high consumption of legumes. To our knowledge, this is the only global study relating fruit, vegetable, and legume intake to cardiovascular disease events and mortality. The results showed that non-cardiovascular mortality and total mortality are decreased with high intake of fruits, vegetables, and legumes compared with low intake.

Implications of all the available evidence

Many dietary guidelines recommend a minimum of 400 g/day of fruits and vegetables, which might not be achievable globally since fruits and vegetables have previously been shown to be unaffordable in low-income and lower-middle income countries. Our findings that even three servings per day (375 g/day) show similar benefit against the risk of non-cardiovascular and total mortality as higher intakes indicates that optimal health benefits can be achieved with a more modest level of consumption, an approach that is likely to be more affordable in poor countries.

Europe and the USA than other populations[3] and little information is available on any potential associations of lower levels of consumption of fruit and vegetables with cardiovascular disease or deaths outside European or US populations (such as those from the Middle East, South America, Africa, and south Asia). Even in studies from Europe and the USA,[4] the apparent benefits of fruit and vegetable consumption vary substantially (from as large as a 40% relative risk reduction to no benefit) depending on the outcome (myocardial infarction, stroke, or death), exposures reported (fruits, vegetables, or both), or the extent of statistical adjustment for other variables that might relate to cardiovascular disease or death (eg, other dietary variables or socioeconomic status). Moreover, mean consumption of fruit and vegetables in most countries is lower than current recommendations,[3,5] and partly because of the relative unaffordability of fruits and vegetables in poorer countries.[5] Although fruits are mainly consumed fresh in most parts of the world, vegetables are eaten raw or cooked in Europe and the USA, and mostly cooked in Asia and in other parts of the world.

Currently, most dietary guidelines do not differentiate between raw and cooked vegetable intake, despite potential differences in nutritional composition and digestibility.[6,7] Therefore, guidelines that are largely based on European and US data might not necessarily apply to other regions of the world. In this study, we investigated the association of fruit, vegetable, and legume consumption with cardiovascular outcomes and total mortality in a prospective cohort study from 18 countries from seven geographical regions: North America and Europe, South America, the Middle East, south Asia, China, southeast Asia, and Africa. This allowed us to investigate relationships across a broad range of intakes including very low consumption of fruits and vegetables, and high consumption of legumes. Additionally, we examined the associations of raw and cooked vegetable intake independently from each other with cardiovascular disease events and total mortality.

Methods

Study design and sample selection

We did a prospective cohort study (Prospective Urban Rural Epidemiology [PURE]) in individuals aged 35–70 years without cardiovascular disease from 613 communities in 18 low-income, middle-income, and high-income countries (HIC) in seven geographical regions: North America and Europe, South America, the Middle East, south Asia, China, southeast Asia, and Africa. A detailed description of participant, community, and country selection has been published previously[8–11] and is summarised in the appendix (p 3). We considered the heterogeneity of socioeconomic factors and the feasibility of carrying out long-term follow-up when selecting the participating countries. We included three HICs (Canada, Sweden, and United Arab Emirates), seven upper-middle income (UMICs; Argentina, Brazil, Chile, Malaysia, Poland, South Africa, and Turkey), four lower-middle income (LMICs; China, Colombia, Iran, and the occupied Palestinian territory) and four low-income countries (LICs; Bangladesh, India, Pakistan, and Zimbabwe), based on gross national income per capita from the World Bank classification for 2006 when the study was initiated. The study was approved by relevant institutional research ethics boards at all sites.

Procedures

At baseline, participants completed dietary assessments using country-specific (region-specific in India), validated food frequency questionnaires (FFQs).[12–22] For countries that had a previously validated FFQ (Canada, China, India, Malaysia, South Africa, Sweden, and Turkey) we used the nutrient databases that were used for the FFQ validation. For countries where a validated FFQ was not available, we developed and validated FFQs using a standard method (appendix p 9). To develop and validate the FFQs, a subgroup of participants from each country completed 24-h dietary recalls for each season (to account for seasonal changes in diet) and a food list was compiled based on the most frequently reported food items. To convert food into nutrients, we constructed country-specific nutrient databases with information about 43 macronutrients and micronutrients that were mainly based on the US Department of Agriculture (USDA) food-composition database (release 18 and 21), modified appropriately with reference to local food composition tables, and supplemented with recipes of locally eaten mixed dishes. The FFQ was administered to the same subgroup of participants in each country and we used Pearson correlation coefficients, using energy adjusted and deattenuated correlations, and weighted κ to validate the FFQs measured against 24-h dietary recalls. Our validation studies showed reasonable agreement between the FFQs and 24-h recalls for fruits (r_s =0·23–0·66) and vegetables (r_s =0·30–0·81) and the concordance rates of classification into the same quartiles ranged from 70% to 74% for fruit and 62% to 79% for vegetables.

See Online for appendix

Potatoes, other tubers, and legumes were not included as vegetables. Fruit and vegetable juices were excluded.

	<1 per day (n=9082)	≥1 to <2 per day (n=19 036)	≥2 to <3 per day (n=35 128)	≥3 to <4 per day (n=24 485)	≥4 to <5 per day (n=14 849)	≥5 to <6 per day (n=9790)	≥6 to <7 per day (n=6945)	≥7 to <8 per day (n=4857)	≥8 per day (n=11 163)
Age (year)	49·0 (41·0–58·0)	49·0 (40·0–58·0)	50·0 (42·0–58·0)	50·0 (42·0–57·0)	50·0 (42·0–58·0)	50·0 (42·0–58·0)	50·0 (43·0–58·0)	50·0 (43·0–58·0)	51·0 (44·0–59·0)
Female sex	5303 (58%)	11 218 (59%)	20 260 (58%)	14 156 (58%)	8592 (58%)	5618 (57%)	4048 (58%)	2862 (59%)	6856 (61%)
Urban location	2901 (32%)	7771 (41%)	16 447 (47%)	13 988 (57%)	9313 (63%)	5976 (61%)	4260 (61%)	3227 (66%)	7417 (66%)
Education level									
Less than graduation from high-school	6027/9031 (66%)	10 514/18 953 (55%)	14 625/35 033 (42%)	8838/24 436 (36%)	5393/14 825 (36%)	3809/9769 (39%)	2693/6933 (39%)	1769/4852 (36%)	3770/11 149 (34%)
High-school graduate	2502/9031 (28%)	6476/18 953 (34%)	15 737/35 033 (45%)	10 791/24 436 (44%)	5795/14 825 (39%)	3315/9769 (34%)	2206/6933 (32%)	1513/4852 (31%)	3395/11 149 (30%)
Some college or more	502/9031 (6%)	1963/18 953 (10%)	4671/35 033 (13%)	4807/24 436 (20%)	3637/14 825 (24%)	2645/9769 (27%)	2034/6933 (29%)	1570/4852 (32%)	3984/11 149 (36%)
Currently a smoker	2676/9031 (29%)	4480/18 953 (24%)	7836/35 033 (22%)	4982/24 436 (20%)	2883/14 825 (19%)	1761/9769 (18%)	1280/6933 (18%)	814/4852 (17%)	1698/11 149 (15%)
High physical activity level*	3253/7388 (36%)	7947/16 738 (42%)	13 817/33 335 (39%)	10 115/23 398 (41%)	5954/14 162 (40%)	4200/9272 (43%)	3168/6560 (46%)	2274/4606 (47%)	5346/10 486 (48%)
Waist-to-hip ratio	0·859 (0·086)	0·865 (0·090)	0·868 (0·081)	0·871 (0·082)	0·877 (0·086)	0·882 (0·086)	0·882 (0·088)	0·881 (0·088)	0·878 (0·088)
Energy intake (kcal/day)†	1442 (1077–1906)	1698 (1322–2205)	1847 (1461–2338)	2017 (1618–2517)	2160 (1726–2698)	2254 (1817–2767)	2363 (1923–2903)	2498 (2036–3054)	2869 (2305–3559)
Vegetable intake (servings per day)	0·27 (0·23)	0·76 (0·39)	1·60 (0·55)	1·86 (0·54)	2·18 (0·75)	2·67 (1·01)	3·13 (1·27)	3·63 (1·49)	4·91 (2·67)
Fruit intake (servings per day)	0·19 (0·22)	0·43 (0·35)	0·62 (0·35)	1·23 (0·60)	1·82 (0·83)	2·27 (1·13)	2·76 (1·40)	3·24 (1·58)	5·21 (3·03)
Legume intake (servings per day)	0·14 (0·16)	0·35 (0·33)	0·30 (0·36)	0·37 (0·41)	0·45 (0·49)	0·53 (0·57)	0·57 (0·60)	0·59 (0·60)	0·71 (0·77)
Starch intake (g/day)	604·8 (465·4)	659·6 (478·3)	507·7 (340·8)	491·8 (318·8)	510·8 (325·1)	506·4 (314·2)	505·0 (306·2)	506·9 (301·0)	544·5 (331·8)
Red meat intake (g/day)	32·8 (58·2)	37·7 (62·4)	57·2 (67·8)	71·6 (73·8)	75·3 (73·0)	83·2 (76·5)	88·0 (74·6)	89·7 (72·3)	93·8 (76·9)
White meat intake (g/day)	16·2 (30·0)	21·3 (33·2)	18·5 (30·6)	24·3 (36·4)	36·7 (47·8)	38·5 (44·9)	42·9 (44·8)	47·9 (48·7)	57·1 (62·2)

Total number of participants=135 335. Data are median (IQR), n (%), n/N (%), or mean (SD). *Defined as ≥3000 metabolic equivalent of task-min per week. †Plausible energy intake between ≥500 kcal/day and ≤5000 kcal/day.

***Table 1*: Participants' characteristics divided by number of fruit, vegetable, and legume servings per day**

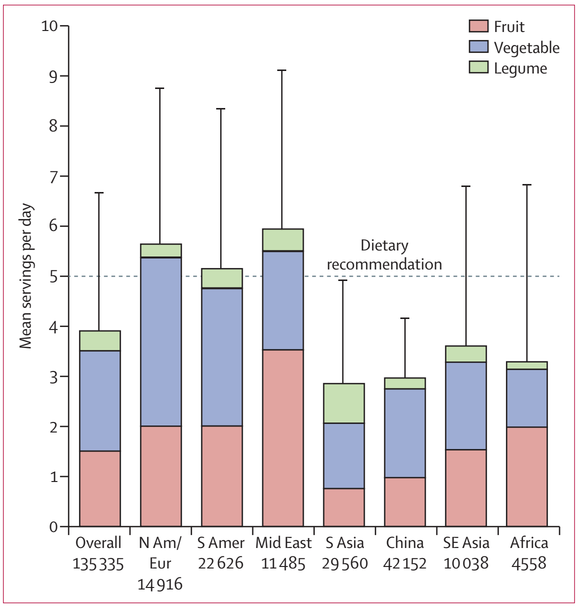


***Figure 1*: Mean fruit, vegetable, and legume intake overall and by geographical region**

Data are from 135 335 individuals. N Am/Eur=North America and Europe: Canada, Poland, and Sweden. S Amer=South America: Argentina, Brazil, Chile, and Colombia. Mid East=Middle East: Iran, occupied Palestinian territory, Turkey, and United Arab Emirates. S Asia=south Asia: Bangladesh, India, and Pakistan. SE Asia=southeast Asia: Malaysia. Africa=South Africa and Zimbabwe.

Legumes included beans, black beans, lentils, peas, chickpeas, and black-eyed peas. Because of inconsistent classification of legumes as a subset of vegetables,[23,24] we grouped total fruit, vegetable, and legume intake into a single food group for our primary analysis (presented in this report) and also present total fruits and vegetables (without legumes) as a separate food group in a secondary analysis (appendix). Although many nutritional qualities of legumes differ from those of vegetables (eg, variable amounts of starch and energy), both legumes and vegetables are good sources of plant protein, fibre, and isoflavones.[23]

We used standardised questionnaires to obtain information about demographic factors, socioeconomic status (education, income, and employment), lifestyle (smoking, physical activity, and alcohol intake), health history and medication use, and family history of cardiovascular disease. Physical assessments included standardised measurements of weight, height, waist and hip circumferences, and blood pressure. Case-report forms, death certificates, medical records and verbal autopsies were used to capture data about major cardiovascular events, and death during follow-up, which were adjudicated centrally in each country by trained physicians using predefined definitions. The follow-up varied based on the date when recruitment began at each site or country. During the follow-up, contact was made with every participant at least every 3 years either by telephone or by a face-to-face visit by the local research team. The median duration of follow-up by country is in the appendix (p 6).

Outcomes

The main clinical outcomes were major cardiovascular disease (defined as death from cardiovascular causes and non-fatal myocardial infarction, stroke, and heart failure), fatal and non-fatal myocardial infarction, fatal and non-fatal strokes, cardiovascular mortality, non-cardiovascular mortality, and total mortality.

Statistical analysis

We computed mean and median estimated fruit, vegetable, and legume intakes in servings overall and by geographical region. One serving was defined as 125 g of fruits or vegetables and 150 g of cooked legumes in accordance with USDA serving sizes. Participants were grouped into categories based on intake values for each dietary exposure (ranging from servings per month to servings per day depending on the dietary exposure). We calculated hazard ratios (HR) using multivariable Cox frailty analysis with random intercepts to account for the correlation of observations within centres (which therefore also accounted for clustering at region and country levels). In a minimally adjusted model, we adjusted for age, sex, and centre as a random effect. The primary model adjusted for age, sex, energy intake, current smoking status, urban or rural location, physical activity, baseline diabetes, education, and other dietary variables (white meat, red meat, bread, and cereal intake), and study centre (as a random effect). For the analyses of fruit intake, we adjusted for vegetable intake, and conversely, analyses of vegetable intake were adjusted for fruit intake. We adjusted for variations in socioeconomic status using education level and household income or wealth index; these produced similar results. The primary analysis did not adjust for obesity, hypertension, or hypercholesterolaemia because these factors might mediate the effects of fruits, vegetables, and legumes on the risk of cardiovascular disease and mortality. We did separate analyses that adjusted for these factors (appendix pp 57–68) and the results were largely similar to the primary model. To test for linear trends, categories of fruit, vegetable, and legume intake were replaced with continuous intake in the Cox frailty regression models. We did interaction tests for fruit, vegetable, and legume intake, and geographical region. Data were analysed with SAS version 9.4.

Role of the funding sources

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author (VMi), senior authors (SA and SY), and several co-authors (AM, XZ) had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Between Jan 1, 2003, and March 31, 2013, 148 723 participants completed the FFQ, of which 143 934 people had

	<1 per day (n=9082)	≥1 to <2 per day (n=19 036)	≥2 to <3 per day (n=35 128)	≥3 to <4 per day (n=24 485)	≥4 to <5 per day (n=14 849)	≥5 to <6 per day (n=9790)	≥6 to <7 per day (n=6945)	≥7 to <8 per day (n=4857)	≥8 per day (n=11 163)	p_{trend}
Median (IQR) fruit, vegetable, and legume servings per day	0·64 (0·41–0·83)	1·56 (1·30–1·79)	2·52 (2·29–2·75)	3·43 (3·21–3·70)	4·43 (4·19–4·70)	5·46 (5·22–5·72)	6·45 (6·22–6·71)	7·46 (7·22–7·71)	9·99 (8·82–12·15)	NA
Major cardiovascular disease events (n=4784)	373 (4%)	743 (4%)	1391 (4%)	882 (4%)	534 (4%)	267 (3%)	180 (3%)	131 (3%)	283(3%)	NA
Adjusted for age, sex, and centre (random effect)	1 (ref)	1·00 (0·88–1·14)	1·06 (0·94–1·20)	1·00 (0·87–1·14)	1·09 (0·94–1·25)	0·86 (0·73–1·02)	0·84 (0·70–1·02)	0·90 (0·73–1·10)	0·83 (0·70–0·99)	0·0015
Multivariable adjusted	1 (ref)	1·03 (0·89–1·18)	1·09 (0·96–1·25)	1·06 (0·92–1·22)	1·20 (1·02–1·40)	0·95 (0·79–1·14)	0·93 (0·76–1·14)	0·97 (0·77–1·21)	0·90 (0·74–1·10)	0·1301
Myocardial infarction events (n=2143)	164 (2%)	391 (2%)	565 (2%)	375 (2%)	254 (2%)	114 (1%)	77 (1%)	60 (1%)	143 (1%)	NA
Adjusted for age, sex, and centre (random effect)	1 (ref)	1·06 (0·88–1·28)	1·16 (0·97–1·40)	1·14 (0·94–1·38)	1·22 (0·99–1·50)	0·83 (0·65–1·07)	0·80 (0·61–1·07)	0·89 (0·65–1·21)	0·95 (0·65–1·21)	0·0403
Multivariable adjusted	1 (ref)	1·04 (0·85–1·27)	1·15 (0·95–1·40)	1·15 (0·93–1·43)	1·28 (1·02–1·61)	0·87 (0·66–1·14)	0·84 (0·62–1·13)	0·91 (0·65–1·27)	0·99 (0·74–1·31)	0·2033
Stroke events (n=2234)	157 (2%)	283 (1%)	727 (2%)	463 (2%)	233 (2%)	127 (1%)	81 (1%)	59 (1%)	104 (<1%)	NA
Adjusted for age, sex, and centre (random effect)	1 (ref)	1·00 (0·82–1·22)	1·03 (0·85–1·24)	0·99 (0·82–1·21)	1·07 (0·86–1·33)	1·03 (0·80–1·32)	1·00 (0·75–1·33)	1·11 (0·81–1·52)	0·85 (0·65–1·13)	0·5947
Multivariable adjusted	1 (ref)	1·02 (0·82–1·27)	1·05 (0·86–1·29)	1·04 (0·84–1·29)	1·16 (0·92–1·47)	1·12 (0·85–1·47)	1·11 (0·82–1·50)	1·19 (0·84–1·67)	0·92 (0·67–1·25)	0·7092
Cardiovascular death events (n=1649)	215 (2%)	361 (2%)	418 (1%)	245 (1%)	161 (1%)	68 (<1%)	57 (<1%)	43 (<1%)	81 (<1%)	NA
Adjusted for age, sex, and centre (random effect)	1 (ref)	0·81 (0·68–0·97)	0·81 (0·68–0·97)	0·73 (0·60–0·89)	0·77 (0·62–0·96)	0·51 (0·38–0·68)	0·65 (0·48–0·88)	0·73 (0·52–1·03)	0·59 (0·45–0·79)	<0·0001
Multivariable adjusted	1 (ref)	0·81 (0·71–1·06)	0·90 (0·73–1·10)	0·81 (0·65–1·02)	0·91 (0·71–1·17)	0·58 (0·42–0·80)	0·80 (0·57–1·10)	0·90 (0·62–1·31)	0·73 (0·53–1·02)	0·0568
Non-cardiovascular death events (n=3809)	486 (5%)	918 (5%)	1023 (3%)	485 (2%)	284 (2%)	199 (2%)	130 (2%)	80 (2%)	204 (2%)	NA
Adjusted for age, sex, and centre (random effect)	1 (ref)	0·95 (0·85–1·06)	0·80 (0·71–0·89)	0·61 (0·53–0·69)	0·60 (0·52–0·70)	0·65 (0·55–0·78)	0·62 (0·51–0·77)	0·55 (0·43–0·71)	0·56 (0·47–0·67)	<0·0001
Multivariable adjusted	1 (ref)	1·05 (0·93–1·19)	0·91 (0·80–1·03)	0·77 (0·66–0·89)	0·80 (0·68–0·95)	0·87 (0·71–1·05)	0·87 (0·70–1·09)	0·80 (0·62–1·05)	0·84 (0·68–1·04)	0·0038
Mortality events (n=5796)	736 (8%)	1371 (7%)	1529 (4%)	772 (3%)	468 (3%)	286 (3%)	198 (3%)	131 (3%)	305 (3%)	NA
Adjusted for age, sex, and centre (random effect)	1 (ref)	0·92 (0·84–1·01)	0·81 (0·74–0·89)	0·65 (0·58–0·74)	0·65 (0·58–0·74)	0·62 (0·53–0·71)	0·63 (0·54–0·75)	0·61 (0·50–0·74)	0·58 (0·50–0·74)	<0·0001
Multivariable adjusted	1 (ref)	1·01 (0·91–1·12)	0·91 (0·82–1·01)	0·78 (0·69–0·88)	0·83 (0·72–0·95)	0·78 (0·66–0·91)	0·84 (0·70–1·00)	0·83 (0·67–1·02)	0·81 (0·68–0·96)	0·0001

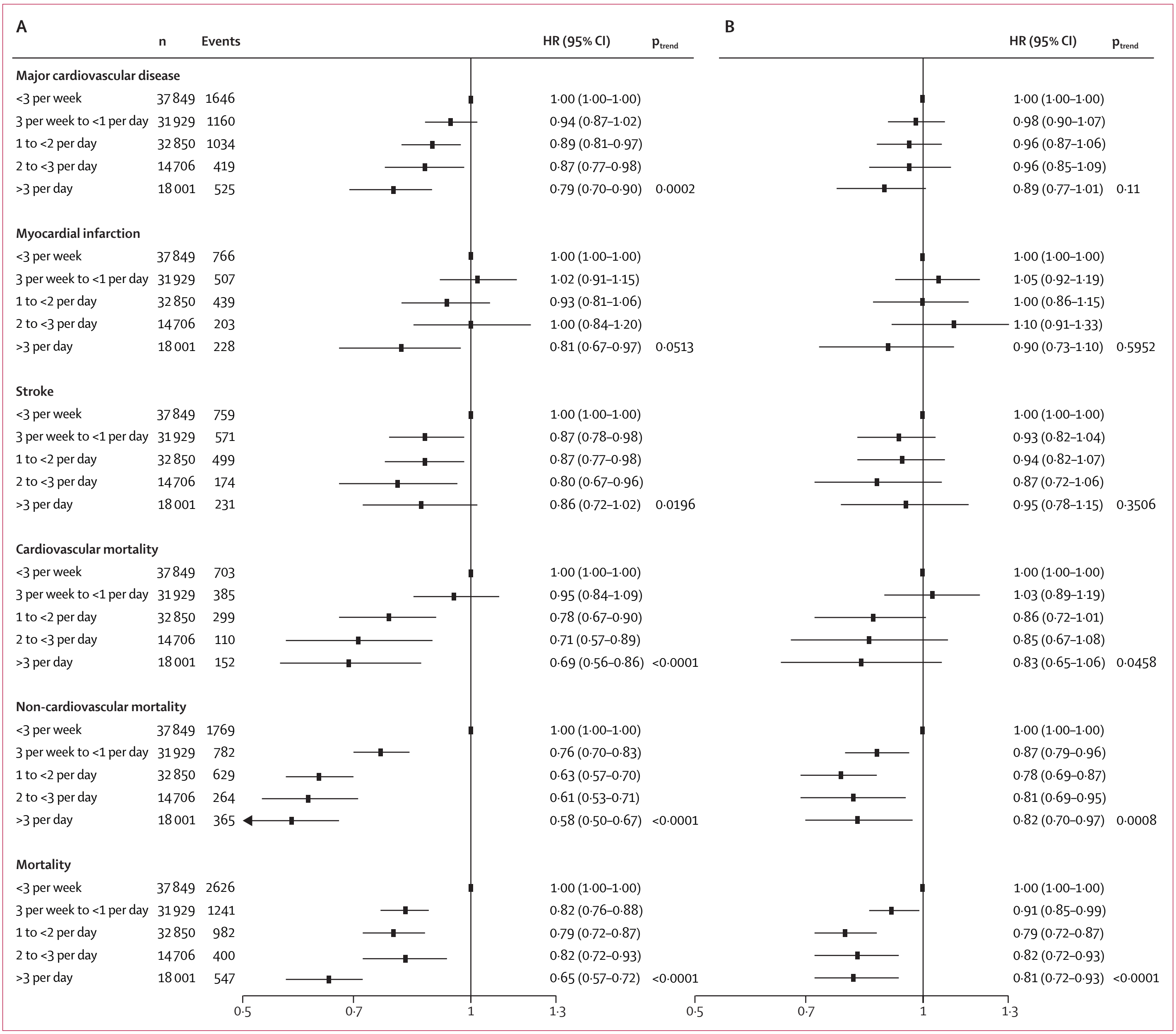
Total number of participants is 135 335. Data are n (%) or hazard ratio (95% CI) except where otherwise stated. Crude event rates are shown. For this analysis, the group with intake <1 serving per day was used as the reference (ref) group that all other groups were compared with. Additional sensitivity analyses with waist-to-hip ratio, hypertension status, and statin medication used in the model did not substantially change estimates of association (appendix). NA=not applicable. Major cardiovascular disease events=death from cardiovascular causes and non-fatal myocardial infarction, stroke, and heart failure. Multivariable adjusted=adjusted for age, sex, centre (random effect), energy intake, current smoker, diabetes, urban or rural location, physical activity, education level, and tertiles of white meat, red meat, breads, and cereals intake.

***Table 2*: Association of number of fruit, vegetable, and legume servings per day with cardiovascular outcomes and mortality**

plausible energy intake (500–5000 kcal/day) and were not missing information about their age and sex. We excluded 7369 people with a history of cardiovascular disease at baseline and 1230 people for whom follow-up information was not available. The remaining 135 335 individuals were included in this analysis (appendix p 7). For the current analysis, we included all unrefuted outcome events in the PURE study database through March 31, 2017.

People who consumed more fruits, vegetables, and legumes had higher education, higher levels of physical activity, lower rates of smoking, and higher energy, red meat and white meat intake, and were more likely to live in urban areas (table 1). Overall, mean fruit, vegetable, and legume intakes were 1·51 (SD 1·77), 2·01 (1·55), and 0·40 (0·48) servings per day, respectively. Combined mean fruit, vegetable, and legume intake was 3·91 (2·77) servings per day (figure 1).

During a median follow-up of 7·4 years (IQR 5·5–9·3), there were 4784 major cardiovascular disease events recorded (table 2). Higher total fruit, vegetable, and legume intake was inversely associated with major cardiovascular disease, myocardial infarction, cardiovascular mortality, non-cardiovascular mortality, and total mortality in the models adjusted for age, sex, and centre (as a random effect; table 2). Following multivariable adjustment, the associations were

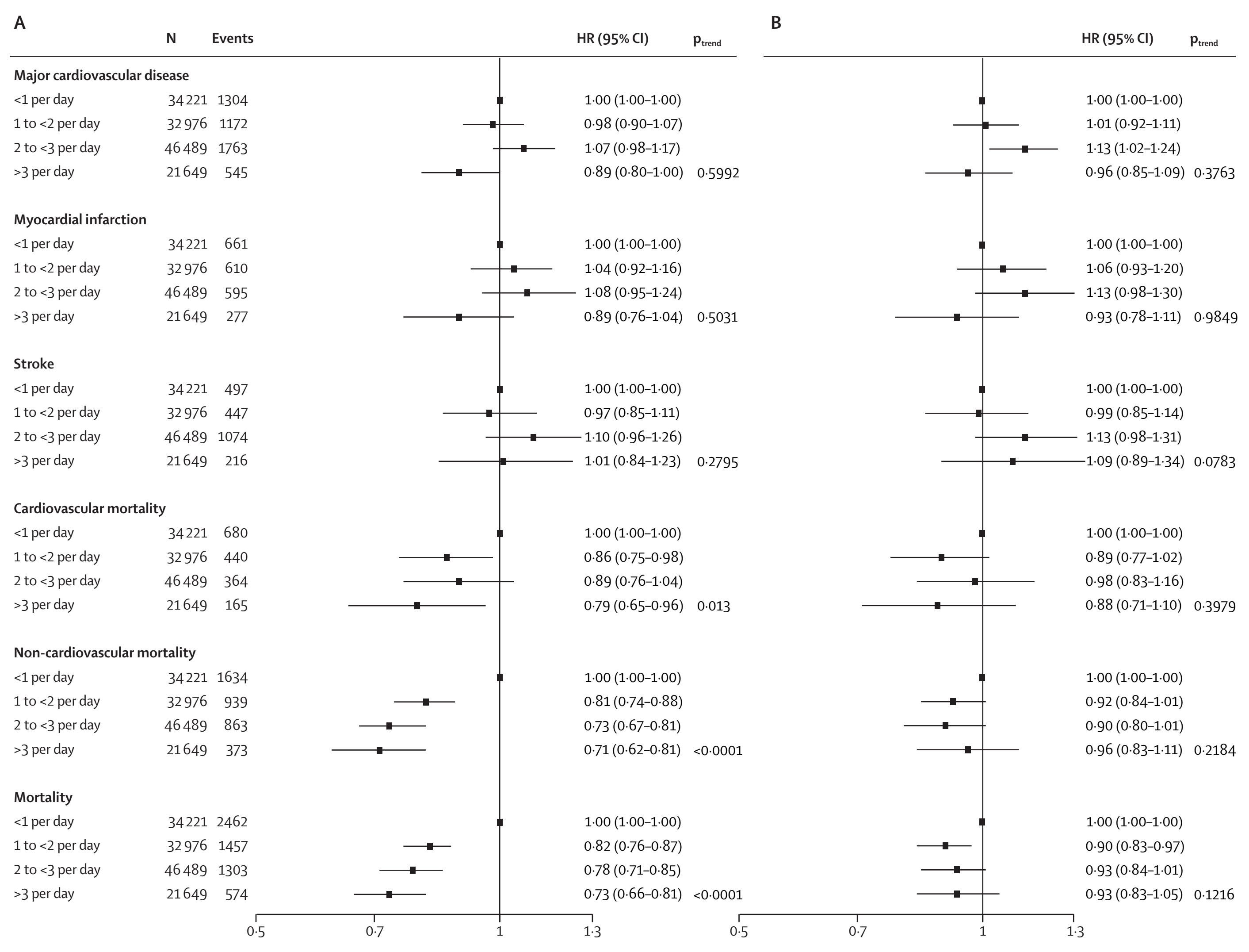


	n	Events	A: HR (95% CI)	A: p_{trend}	B: HR (95% CI)	B: p_{trend}
Major cardiovascular disease						
<3 per week	37 849	1646	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
3 per week to <1 per day	31 929	1160	0·94 (0·87–1·02)		0·98 (0·90–1·07)	
1 to <2 per day	32 850	1034	0·89 (0·81–0·97)		0·96 (0·87–1·06)	
2 to <3 per day	14 706	419	0·87 (0·77–0·98)		0·96 (0·85–1·09)	
>3 per day	18 001	525	0·79 (0·70–0·90)	0·0002	0·89 (0·77–1·01)	0·11
Myocardial infarction						
<3 per week	37 849	766	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
3 per week to <1 per day	31 929	507	1·02 (0·91–1·15)		1·05 (0·92–1·19)	
1 to <2 per day	32 850	439	0·93 (0·81–1·06)		1·00 (0·86–1·15)	
2 to <3 per day	14 706	203	1·00 (0·84–1·20)		1·10 (0·91–1·33)	
>3 per day	18 001	228	0·81 (0·67–0·97)	0·0513	0·90 (0·73–1·10)	0·5952
Stroke						
<3 per week	37 849	759	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
3 per week to <1 per day	31 929	571	0·87 (0·78–0·98)		0·93 (0·82–1·04)	
1 to <2 per day	32 850	499	0·87 (0·77–0·98)		0·94 (0·82–1·07)	
2 to <3 per day	14 706	174	0·80 (0·67–0·96)		0·87 (0·72–1·06)	
>3 per day	18 001	231	0·86 (0·72–1·02)	0·0196	0·95 (0·78–1·15)	0·3506
Cardiovascular mortality						
<3 per week	37 849	703	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
3 per week to <1 per day	31 929	385	0·95 (0·84–1·09)		1·03 (0·89–1·19)	
1 to <2 per day	32 850	299	0·78 (0·67–0·90)		0·86 (0·72–1·01)	
2 to <3 per day	14 706	110	0·71 (0·57–0·89)		0·85 (0·67–1·08)	
>3 per day	18 001	152	0·69 (0·56–0·86)	<0·0001	0·83 (0·65–1·06)	0·0458
Non-cardiovascular mortality						
<3 per week	37 849	1769	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
3 per week to <1 per day	31 929	782	0·76 (0·70–0·83)		0·87 (0·79–0·96)	
1 to <2 per day	32 850	629	0·63 (0·57–0·70)		0·78 (0·69–0·87)	
2 to <3 per day	14 706	264	0·61 (0·53–0·71)		0·81 (0·69–0·95)	
>3 per day	18 001	365	0·58 (0·50–0·67)	<0·0001	0·82 (0·70–0·97)	0·0008
Mortality						
<3 per week	37 849	2626	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
3 per week to <1 per day	31 929	1241	0·82 (0·76–0·88)		0·91 (0·85–0·99)	
1 to <2 per day	32 850	982	0·79 (0·72–0·87)		0·79 (0·72–0·87)	
2 to <3 per day	14 706	400	0·82 (0·72–0·93)		0·82 (0·72–0·93)	
>3 per day	18 001	547	0·65 (0·57–0·72)	<0·0001	0·81 (0·72–0·93)	<0·0001

***Figure 2*: Association of fruit intake with cardiovascular outcomes and mortality**

(A) Adjusted for age, sex, and centre (random effect). (B) Adjusted for age, sex, centre (random effect), energy intake, current smoker, diabetes, urban or rural location, physical activity, education level, and tertiles of white meat, red meat, and intake of breads, cereals, and vegetables. Crude event rates are shown. Additional sensitivity analyses with waist-to-hip ratio, hypertension status, and statin medication used in the model did not substantially change estimates of association (appendix). HR=hazard ratio. Major cardiovascular disease events=death from cardiovascular causes and non-fatal myocardial infarction, stroke, and heart failure.

markedly attenuated and only non-cardiovascular mortality and total mortality remained significant, with a non-significant trend for cardiovascular mortality (table 2). The HR for total mortality was lowest for three to four servings per day (0·78, 95% CI 0·69–0·88) compared with the reference group, with no further apparent decrease in HR with higher consumption. Total fruit, vegetable, and legume intake was inversely associated with total mortality in most geographical regions (south Asia, China, North America and Europe, the Middle East, and South America; appendix pp 35–36). Similarly, total fruit and vegetable intake was associated with major cardiovascular disease, myocardial infarction, cardiovascular mortality, non-cardiovascular mortality and total mortality in the age-adjusted and sex-adjusted models, and with lower non-cardiovascular mortality and total mortality in the fully adjusted models.



	N	Events	A: HR (95% CI)	A: ptrend	B: HR (95% CI)	B: ptrend
Major cardiovascular disease						
<1 per day	34 221	1304	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 to <2 per day	32 976	1172	0·98 (0·90–1·07)		1·01 (0·92–1·11)	
2 to <3 per day	46 489	1763	1·07 (0·98–1·17)		1·13 (1·02–1·24)	
>3 per day	21 649	545	0·89 (0·80–1·00)	0·5992	0·96 (0·85–1·09)	0·3763
Myocardial infarction						
<1 per day	34 221	661	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 to <2 per day	32 976	610	1·04 (0·92–1·16)		1·06 (0·93–1·20)	
2 to <3 per day	46 489	595	1·08 (0·95–1·24)		1·13 (0·98–1·30)	
>3 per day	21 649	277	0·89 (0·76–1·04)	0·5031	0·93 (0·78–1·11)	0·9849
Stroke						
<1 per day	34 221	497	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 to <2 per day	32 976	447	0·97 (0·85–1·11)		0·99 (0·85–1·14)	
2 to <3 per day	46 489	1074	1·10 (0·96–1·26)		1·13 (0·98–1·31)	
>3 per day	21 649	216	1·01 (0·84–1·23)	0·2795	1·09 (0·89–1·34)	0·0783
Cardiovascular mortality						
<1 per day	34 221	680	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 to <2 per day	32 976	440	0·86 (0·75–0·98)		0·89 (0·77–1·02)	
2 to <3 per day	46 489	364	0·89 (0·76–1·04)		0·98 (0·83–1·16)	
>3 per day	21 649	165	0·79 (0·65–0·96)	0·013	0·88 (0·71–1·10)	0·3979
Non-cardiovascular mortality						
<1 per day	34 221	1634	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 to <2 per day	32 976	939	0·81 (0·74–0·88)		0·92 (0·84–1·01)	
2 to <3 per day	46 489	863	0·73 (0·67–0·81)		0·90 (0·80–1·01)	
>3 per day	21 649	373	0·71 (0·62–0·81)	<0·0001	0·96 (0·83–1·11)	0·2184
Mortality						
<1 per day	34 221	2462	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 to <2 per day	32 976	1457	0·82 (0·76–0·87)		0·90 (0·83–0·97)	
2 to <3 per day	46 489	1303	0·78 (0·71–0·85)		0·93 (0·84–1·01)	
>3 per day	21 649	574	0·73 (0·66–0·81)	<0·0001	0·93 (0·83–1·05)	0·1216

***Figure 3*: Association of vegetable intake with cardiovascular outcomes and mortality**

(A) Adjusted for age, sex, and centre (random effect). (B) Adjusted for age, sex, centre (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, and tertiles of white meat, red meat, and intake of breads, cereals, and fruit. Crude event rates are shown. Additional sensitivity analyses with waist-to-hip ratio, hypertension status, and statin medication used in the model did not substantially change estimates of association (appendix). HR=hazard ratio. Major cardiovascular disease events=death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure.

After adjusting for age and sex, fruit consumption was inversely associated with the risk of major cardiovascular disease, stroke, cardiovascular mortality, non-cardiovascular mortality, and total mortality (figure 2). The associations were markedly attenuated after adjusting for additional lifestyle and dietary factors and only cardiovascular mortality, non-cardiovascular mortality, and total mortality remained significant (figure 2). An inverse association between higher fruit intake and the risk of total mortality was observed in south Asia, China, the Middle East, and South America (appendix pp 37–38).

In the age-adjusted and sex-adjusted models, vegetable intake was inversely associated with cardiovascular mortality, non-cardiovascular mortality, and total mortality (figure 3). After adjustment for additional covariates, vegetable intake was not significantly associated with these outcomes (figure 3). Similarly, no significant association was observed between vegetable intake and total mortality in most geographical regions (China, southeast Asia, Africa, the Middle East, and South America), but a beneficial association was shown in south Asia, and North America and Europe (appendix pp 39–40).

Legume consumption was inversely associated with cardiovascular mortality, non-cardiovascular mortality, and total mortality in the minimally adjusted models and with non-cardiovascular mortality and total mortality in the fully adjusted models (figure 4).

The percentage of total vegetable intake consumed as raw vegetables was low in south Asia, Africa, and southeast

	<1 serving per month (n=24 054)	1 serving per month to <1 per week (n=9086)	1 to <3 servings per week (n=15 410)	3 servings per week to <1 per day (n=16 263)	1 to 2 servings per day (n=15 818)	>2 servings per day (n=12 552)	p_{trend}
Median (IQR) raw vegetable servings per day	0·00 (0·00–0·00)	0·08 (0·05–0·11)	0·25 (0·20–0·33)	0·67 (0·52–0·82)	1·42 (1·19–1·66)	2·82 (2·35–3·61)	NA
Major cardiovascular disease events (n=3085)	1147 (5%)	313 (3%)	459 (3%)	477 (3%)	422 (3%)	267 (2%)	NA
Adjusted for age, sex, and centre (random effect)	1 (ref)	0·97 (0·84–1·12)	0·89 (0·77–1·03)	0·93 (0·79–1·09)	0·88 (0·74–1·05)	0·71 (0·58–0·87)	0·0056
Multivariable adjusted	1 (ref)	0·96 (0·82–1·13)	0·86 (0·73–1·01)	0·92 (0·77–1·11)	0·92 (0·76–1·12)	0·79 (0·63–1·00)	0·1303
Mortality events (n=4640)	2329 (10%)	480 (5%)	670 (4%)	536 (3%)	396 (3%)	229 (2%)	NA
Adjusted for age, sex, and centre (random effect)	1 (ref)	0·75 (0·68–0·84)	0·66 (0·59–0·74)	0·63 (0·55–0·72)	0·59 (0·50–0·70)	0·46 (0·38–0·56)	<0·0001
Multivariable adjusted	1 (ref)	0·86 (0·75–0·97)	0·77 (0·68–0·89)	0·76 (0·65–0·89)	0·81 (0·68–0·97)	0·69 (0·55–0·85)	0·0009

Total number of participants is 93 183. Data are n (%) or hazard ratio (95% CI) unless otherwise stated. Crude event rates are shown. For this analysis, the group with intake <1 serving per day was used as the reference (ref) group that all other groups were compared with. Sensitivity analyses adjusting for waist-to-hip ratio, hypertension status, and statin medication, and including the participants from China (by assuming that total vegetable intake was consumed as cooked vegetables) did not substantially change estimates of association (appendix). NA=not applicable. Major cardiovascular disease events=death from cardiovascular causes and non-fatal myocardial infarction, stroke, and heart failure. Multivariable adjusted=adjusted for age, sex, centre (random effect), energy intake, current smoker, diabetes, urban or rural location, physical activity, education level, and tertiles of white meat, red meat, breads, cereals, and fruit intake. Does not include data from China (n=42 152) because the food frequency questionnaires used in this country did not differentiate between raw and cooked vegetables.

***Table 3*: Association of raw vegetable intake with cardiovascular disease and mortality**

	<3 servings per week (n=20 890)	3 servings per week to <1 per day (n=33 395)	1 to 2 servings per day (n=25 813)	>2 servings per day (n=13 085)	p_{trend}
Median (IQR) cooked vegetable servings per day	0·26 (0·16–0·34)	0·68 (0·55–0·83)	1·34 (1·15–1·61)	2·66 (2·27–3·33)	NA
Major cardiovascular disease events (n=3085)	685 (3%)	1083 (3%)	811 (3%)	506 (4%)	NA
Adjusted for age, sex, and centre (random effect)	1 (ref)	1·02 (0·92–1·13)	0·98 (0·88–1·10)	1·06 (0·93–1·21)	0·6180
Multivariable adjusted	1 (ref)	1·10 (0·98–1·22)	1·07 (0·94–1·21)	1·17 (1·01–1·36)	0·0853
Mortality events (n=4640)	661 (2%)	610 (2%)	595 (1%)	277 (1%)	NA
Adjusted for age, sex, and centre (random effect)	1 (ref)	0·88 (0·82–0·95)	0·74 (0·67–0·80)	0·72 (0·65–0·80)	<0·0001
Multivariable adjusted	1 (ref)	0·99 (0·91–1·08)	0·86 (0·77–0·95)	0·91 (0·80–1·03)	0·0110

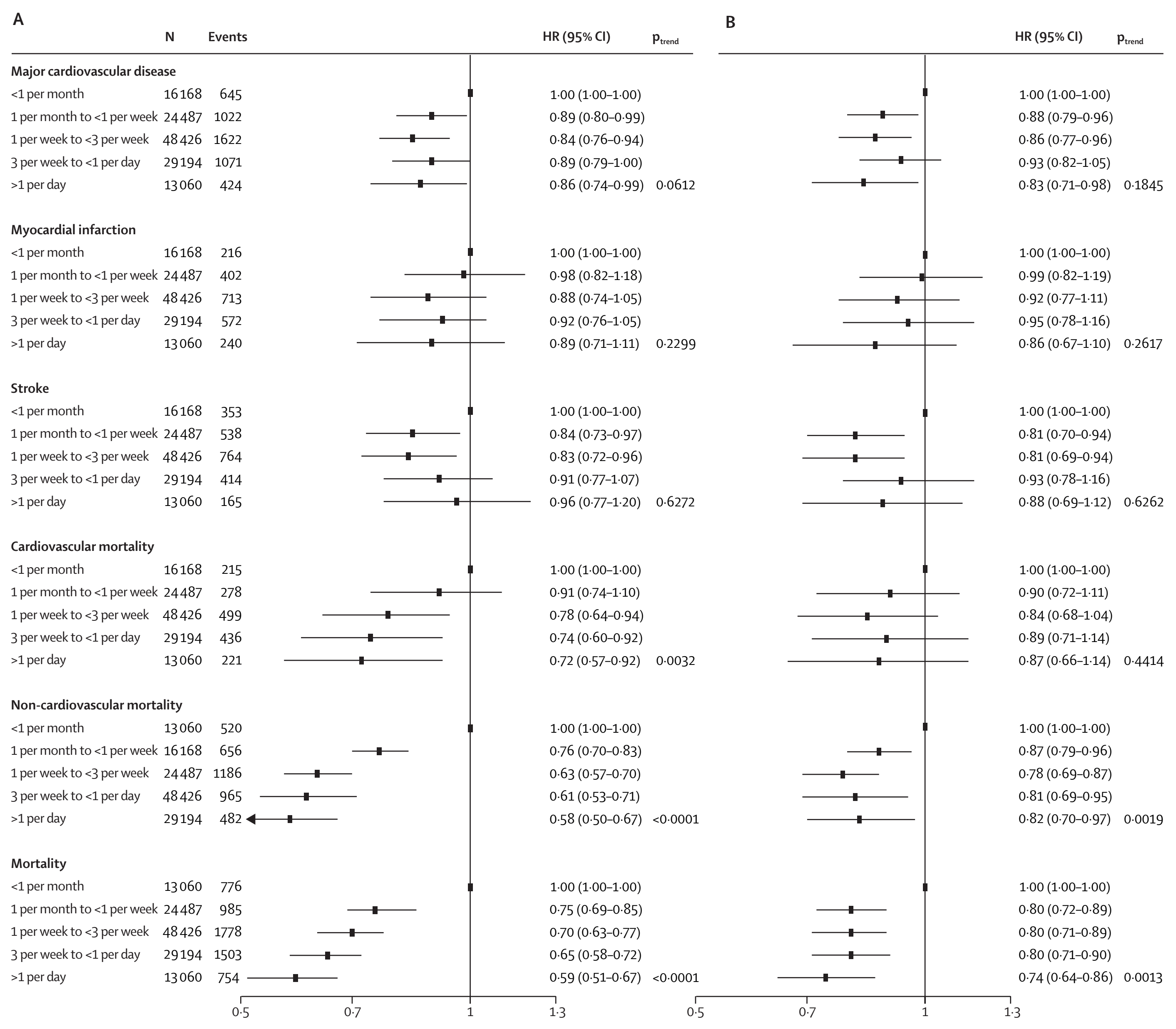
Total number of participants is 93 183. Data are n (%) or hazard ratio (95% CI) unless otherwise stated. Crude event rates are shown. For this analysis, the group with intake <1 serving per day was used as the reference (ref) group that all other groups were compared with. Sensitivity analyses adjusting for waist-to-hip ratio, hypertension status, and statin medication, and including the participants from China (by assuming that total vegetable intake was consumed as cooked vegetables) did not substantially change estimates of association (appendix). NA=not applicable. Major cardiovascular disease events=death from cardiovascular causes and non-fatal myocardial infarction, stroke, and heart failure. Multivariable adjusted=adjusted for age, sex, centre (random effect), energy intake, current smoker, diabetes, urban or rural location, physical activity, education level, and tertiles of white meat, red meat, breads, cereals, and fruit intake. Does not include data from China (n=42 152) because the food frequency questionnaires used in this country did not differentiate between raw and cooked vegetables.

***Table 4*: Association of cooked vegetable intake with cardiovascular disease and mortality**

Asia (figure 5). In the fully adjusted models, both raw and cooked vegetable intakes were inversely associated with total mortality (tables 3, 4). The risk of major cardiovascular disease was inversely associated with the level of raw vegetable intake, but not cooked vegetable intake (tables 3, 4). A non-significant inverse trend for the level of raw vegetable intake and risk of major cardiovascular disease was observed, but no association was shown for cooked vegetable intake.

Discussion

In this large, international prospective cohort study, we showed that greater fruit, vegetable, and legume intake was associated with a lower risk of major cardiovascular disease, myocardial infarction, cardiovascular mortality, non-cardiovascular mortality, and total mortality in the analyses adjusted for age and sex. With multivariable adjustment (demographic characteristics, lifestyle, health history, and dietary factors), the associations remained significant for non-cardiovascular mortality and total mortality, with a non-significant trend for lower cardiovascular mortality. Furthermore, an intake of three to four servings per day (equivalent to 375–500 g/day) was as beneficial as higher amounts of intake in reducing total mortality. Our findings indicate that even relatively moderate intakes of fruits, vegetables and legumes might lower the risk of death. A meta-analysis[4] reported the steepest reduction in risk of total mortality with up to 400 g



	N	Events	A: HR (95% CI)	A: p_{trend}	B: HR (95% CI)	B: p_{trend}
Major cardiovascular disease						
<1 per month	16 168	645	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 per month to <1 per week	24 487	1022	0·89 (0·80–0·99)		0·88 (0·79–0·96)	
1 per week to <3 per week	48 426	1622	0·84 (0·76–0·94)		0·86 (0·77–0·96)	
3 per week to <1 per day	29 194	1071	0·89 (0·79–1·00)		0·93 (0·82–1·05)	
>1 per day	13 060	424	0·86 (0·74–0·99)	0·0612	0·83 (0·71–0·98)	0·1845
Myocardial infarction						
<1 per month	16 168	216	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 per month to <1 per week	24 487	402	0·98 (0·82–1·18)		0·99 (0·82–1·19)	
1 per week to <3 per week	48 426	713	0·88 (0·74–1·05)		0·92 (0·77–1·11)	
3 per week to <1 per day	29 194	572	0·92 (0·76–1·05)		0·95 (0·78–1·16)	
>1 per day	13 060	240	0·89 (0·71–1·11)	0·2299	0·86 (0·67–1·10)	0·2617
Stroke						
<1 per month	16 168	353	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 per month to <1 per week	24 487	538	0·84 (0·73–0·97)		0·81 (0·70–0·94)	
1 per week to <3 per week	48 426	764	0·83 (0·72–0·96)		0·81 (0·69–0·94)	
3 per week to <1 per day	29 194	414	0·91 (0·77–1·07)		0·93 (0·78–1·16)	
>1 per day	13 060	165	0·96 (0·77–1·20)	0·6272	0·88 (0·69–1·12)	0·6262
Cardiovascular mortality						
<1 per month	16 168	215	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 per month to <1 per week	24 487	278	0·91 (0·74–1·10)		0·90 (0·72–1·11)	
1 per week to <3 per week	48 426	499	0·78 (0·64–0·94)		0·84 (0·68–1·04)	
3 per week to <1 per day	29 194	436	0·74 (0·60–0·92)		0·89 (0·71–1·14)	
>1 per day	13 060	221	0·72 (0·57–0·92)	0·0032	0·87 (0·66–1·14)	0·4414
Non-cardiovascular mortality						
<1 per month	13 060	520	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 per month to <1 per week	16 168	656	0·76 (0·70–0·83)		0·87 (0·79–0·96)	
1 per week to <3 per week	24 487	1186	0·63 (0·57–0·70)		0·78 (0·69–0·87)	
3 per week to <1 per day	48 426	965	0·61 (0·53–0·71)		0·81 (0·69–0·95)	
>1 per day	29 194	482	0·58 (0·50–0·67)	<0·0001	0·82 (0·70–0·97)	0·0019
Mortality						
<1 per month	13 060	776	1·00 (1·00–1·00)		1·00 (1·00–1·00)	
1 per month to <1 per week	24 487	985	0·75 (0·69–0·85)		0·80 (0·72–0·89)	
1 per week to <3 per week	48 426	1778	0·70 (0·63–0·77)		0·80 (0·71–0·89)	
3 per week to <1 per day	29 194	1503	0·65 (0·58–0·72)		0·80 (0·71–0·90)	
>1 per day	13 060	754	0·59 (0·51–0·67)	<0·0001	0·74 (0·64–0·86)	0·0013

***Figure 4*: Association of legume intake with cardiovascular outcomes and mortality**

(A) Adjusted for age, sex, and centre (random effect). (B) Adjusted for age, sex, centre (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, and tertiles of white meat, red meat, and intake of breads and cereals. Crude event rates are shown. Additional sensitivity analyses with waist-to-hip ratio, hypertension status, and statin medication used in the model did not substantially change estimates of association (appendix). HR=hazard ratio. Major cardiovascular disease events=death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure.

of fruits and vegetables per day (equivalent to 3·2 servings in our study), with modest additional benefit for intake above this level (relative risk [RR] 0·76, 95% CI 0·73–0·79 for 400 g/day and RR 0·69, 0·66–0·73 for 800 g/day), which is consistent with our findings.

We showed an 11% lower risk of major cardiovascular disease for the highest fruit intake category compared with the lowest intake category. Our findings are consistent with those from a meta-analysis[4] which showed a beneficial effect (RR 0·87, 95% CI 0·82–0·92) for the prevention of cardiovascular disease. Our results suggest a more modest effect on stroke than some, but not all, previous studies.[4,25] Fewer studies have investigated the association between fruit intake and stroke compared with cardiovascular disease or coronary heart disease. Not all studies of cardiovascular disease reported on both stroke and myocardial infarction, and some reported on only one outcome. This might be partly due to the varying

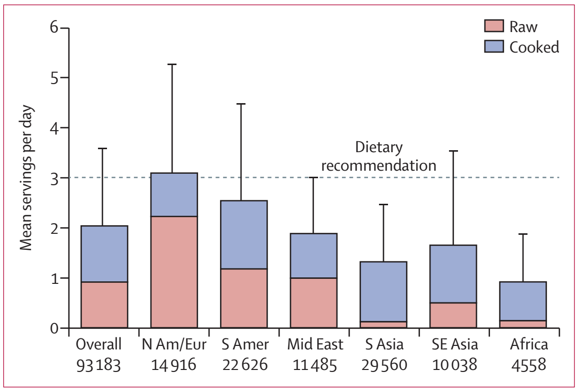


***Figure 5*: Mean raw and cooked vegetable intake overall and by geographical region**

Data are from 93 183 participants, not including those from China (n=42 152) because the food frequency questionnaires used in this country did not differentiate between raw and cooked vegetables. N Am/Eur=North America and Europe: Canada, Poland, and Sweden. S Amer=South America: Argentina, Brazil, Chile, and Colombia. Mid East=Middle East: Iran, occupied Palestinian territory, Turkey, and United Arab Emirates. S Asia=south Asia: Bangladesh, India, and Pakistan. SE Asia=southeast Asia: Malaysia. Africa=South Africa and Zimbabwe.

goals of different studies, but could also be due to a data-derived emphasis about which outcome to report (with the potential that the outcome and exposure with the greater effect size might have been selectively reported or emphasised). Among studies with a similar range of fruit intake as our PURE study, most reported a nominal reduction in the risk of stroke including the Health Professionals Follow-up study,[26] the Monitoring Project on Risk Factors and Chronic Diseases in the Netherlands study,[7] and the Multiethnic Cohort study.[27] Our findings for fruit and cardiovascular mortality and total mortality are in keeping with results from most previous prospective studies.[4,25] We showed a 14% and 21% lower risk of these outcomes when one to two daily servings (equivalent to 125–250 g/day) of fruit were consumed, and intake above this level had little added benefit on risk lowering.

Our finding of a minimal decrease in the risk of major cardiovascular disease and total mortality with higher vegetable intake is consistent with many earlier studies, although substantial heterogeneity in the results of different studies has been reported in a meta-analysis;[4] and, only three of the 18 cohorts included in the meta-analysis reported a robust, inverse association between vegetable intake and cardiovascular disease.[28–30] Among studies with similar daily vegetable intake to the PURE study, the effect estimates for cardiovascular disease, stroke, and total mortality were similar to those in our study.[26–32] Collectively, these data indicate that the association of total vegetable intake and these outcomes is much more modest than generally assumed.

Vegetables might be consumed raw or cooked and the cooking process might alter the bioavailability of nutrients (such as phytochemicals, vitamins, minerals, and fibre), and digestibility. Some evidence suggests that cooking vegetables can degrade nutrient and enzyme content and possibly create harmful byproducts.[33] However, for some nutrients such as lycopene and β carotene, their bioavailability might be enhanced by cooking.[34,35] Little information is available on the effect of raw vegetable intake on cardiovascular disease and total mortality in low-income and middle-income countries. In a cohort of 451 151 participants from ten European countries,[31] both raw and cooked vegetable intakes were inversely associated with risk of mortality, but a stronger beneficial effect was noted for raw vegetable consumption. Additionally, beneficial associations of raw vegetable intake with total mortality, ischaemic heart disease,[36] and ischaemic stroke[7] have been reported. Our findings of a trend towards greater beneficial associations for raw vegetable intake and major cardiovascular disease and total mortality are consistent with some previous reports,[31] but differ from a meta-analysis that showed similar beneficial effects for raw and cooked vegetable intake and the risk of mortality (RR 0·88, 95% CI 0·79–0·98 for raw vegetables and 0·87, 0·80–0·94 for cooked).[4] However, a key limitation of this meta-analysis was that few studies examined raw and cooked vegetable intakes separately within the same study, which can lead to additional confounding (two studies on both raw and cooked vegetables and three on cooked vegetables only).

We found that higher legume consumption was associated with a lower risk of major cardiovascular disease, cardiovascular mortality, non-cardiovascular mortality and total mortality, although only the associations with non-cardiovascular mortality and total mortality were significant after full adjustment. A meta-analysis showed that legume consumption reduced the risk of ischaemic heart disease, but not stroke.[37] However, this meta-analysis included data from only six cohorts, predominantly in North America and Europe, with little data from other regions. Furthermore, the consumption of legumes was low in these studies (eg, <25 g/day) compared with our study (overall mean legume intake of 60 g/day). Legumes contribute a substantial amount of energy and protein for many populations in south Asia, Africa, and Latin America,[38] which are included in PURE. In Costa Rica, higher legume intake (>86 g/day *vs* <86 g/month) was associated with a 38% lower risk of myocardial infarction,[39] which is consistent with our findings. Legumes are frequently consumed as an alternative to meat, and higher glycaemic grains and starches (eg, pasta and white bread), and might be beneficial as a replacement for these foods.[37,40,41] Overall, our findings suggest that higher legume consumption is associated with lower mortality in populations.

Several mechanisms have been proposed to explain the lower risk of cardiovascular disease with higher consumption of fruits, vegetables, and legumes. One explanation is that antioxidants and polyphenols in fruits and vegetables, such as vitamin C, vitamin E, and carotenoids, might prevent lipid oxidation in arterial vessel walls,[42] lower blood pressure,[43,44] and improve endothelial function.[45] Several studies have reported inverse associations between potassium[46] and magnesium[47] with blood pressure. Fruits and vegetables are good sources of dietary fibre, which has been shown to reduce the insulin response to carbohydrates,[48] and decrease total and LDL cholesterol.[49] Similarly, legumes contain fibre and phytochemicals, and legume consumption has been shown to reduce blood pressure, total and LDL cholesterol, and triglycerides.[50]

Our study included data from geographical regions such as the Middle East, South America, Africa, south Asia, and southeast Asia from which little or no data are available on the associations between fruits, vegetables, and legume intake and cardiovascular disease or deaths. Additional strengths of this study included the prospective design, the large sample size, the use of validated, country-specific FFQs to estimate intake, the broad range of intake of fruits, vegetables, and legumes (0 to >1000 g/day), and standardised methods to collect and adjudicate events. Our study also had some limitations. First, fruit, vegetable, and legume consumption was measured using validated country-specific FFQs at baseline and their consumption was assumed to remain unchanged during follow-up. Consequently, measurement errors might have occurred that would probably have resulted in an underestimation of the relationship between dietary intake and cardiovascular disease and mortality. Second, we did not examine the associations of the different types of fruits and vegetables versus cardiovascular disease or mortality; the power to detect these associations was low since such data were not available from China (which removed about 40% of the study population), and the numbers of events per type of fruits and vegetables was relatively low. Furthermore, there were additional confounders as the consumption of different types of vegetables and fruits varied by region, and those consuming low amounts of one vegetable or fruit might have consumed large amounts of another fruit or vegetable. Controlling for such factors requires even more events and therefore the current results are not robust. Third, environmental factors (use of pesticides and herbicides, and water contamination) that might affect the nutritional quality of fruits, vegetables, and legumes were not measured and might have contributed to the differences between our study and previous studies. Fourth, the methods of cultivation, types of fruits and vegetables commonly consumed, and cooking methods (eg, frying *vs* other methods) might have varied across countries. Fifth, in observational studies, the possibility of residual confounding from unmeasured or imprecise measurement of covariates cannot be completely ruled out. This is exemplified by the near halving of apparent associations in the fully adjusted model compared with the minimally adjusted model, and suggests that part of the apparent large benefits in some previous studies might have been due to incomplete adjustment. Sixth, some event misclassification cannot be ruled out. However, the number of misclassified events was probably minimal, as most events were ascertained using supporting documents and adjudicated using standardised definitions. Last, because of the low number of events in some geographical regions, our findings within the geographical regions of the Middle East, Southeast Asia, and North America and Europe are presently not robust. However, we intend to re-examine these relationships in the future when more participants have been enrolled (the study has now nearly enrolled 200 000 people) and longer follow-up will be available. Nevertheless, to our knowledge, this is the first study to report on the associations of fruit, vegetable, and legume intake with cardiovascular risk in countries at varying economic levels and from different regions.

In summary, our results show that higher fruit, vegetable, and legume intake is associated with a lower risk of non-cardiovascular mortality and total mortality, with a non-significant trend for cardiovascular mortality, in a global population. Previous research[4] and many dietary guidelines in North America and Europe[2] recommended intake ranging from 400 to 800 g/day, but these targets are unaffordable for most individuals in LIC and LMIC.[5] Even a small reduction in the recommendation from 400 to 375 g/day may have important implications on household spending and food security in poorer countries. Our findings that even three servings per day (375 g/day) show similar benefit against non-cardiovascular and total mortality as higher intakes and trends towards benefit for cardiovascular mortality, indicate that optimal health benefits can be achieved with a more modest level of consumption, an approach that is likely to be much more affordable.

Contributors

VMi and AM designed the study, were involved in data management and statistical analysis, and wrote the first and subsequent drafts of the report. MD developed and validated the country-specific food frequency questionnaires, supervised the collection of dietary information, and commented on drafts of the report. SR coordinated the worldwide Prospective Urban Rural Epidemiology (PURE) study and reviewed and commented on drafts of the report. XZ and SIB contributed to the statistical analysis and reviewed and commented on drafts of the report. KT was the co-principal investigator of PURE and reviewed and commented on drafts of the report. SY designed the study, conceived and initiated PURE, supervised its conduct and data analysis, and provided comments on all drafts. All other authors coordinated the study, collected data in their respective countries, and provided comments on drafts of the report.

Declaration of interests

We declare no competing interests.

Acknowledgments

SY is supported by the Heart & Stroke Foundation/Marion W. Burke Chair in Cardiovascular Disease. The PURE Study is an investigator-initiated study funded by the Population Health Research Institute, the Canadian Institutes of Health Research (CIHR), Heart and Stroke Foundation of Ontario, support from CIHR's Strategy for Patient Oriented Research (SPOR) through the Ontario SPOR Support Unit, as well as the Ontario Ministry of Health and Long-Term Care and through unrestricted grants from several pharmaceutical companies, with major contributions from AstraZeneca (Canada), Sanofi-Aventis

(France and Canada), Boehringer Ingelheim (Germany and Canada), Servier, and GlaxoSmithkline, and additional contributions from Novartis and King Pharma and from various national or local organisations in participating countries; these include: Argentina: Fundacion ECLA; Bangladesh: Independent University, Bangladesh and Mitra and Associates; Brazil: Unilever Health Institute, Brazil; Canada: Public Health Agency of Canada and Champlain Cardiovascular Disease Prevention Network; Chile: Universidad de la Frontera; China: National Center for Cardiovascular Diseases; Colombia: Colciencias, grant number 6566-04-18062; India: Indian Council of Medical Research; Malaysia: Ministry of Science, Technology and Innovation of Malaysia, grant numbers 100 - IRDC/BIOTEK 16/6/21 (13/2007) and 07-05-IFN-BPH 010, Ministry of Higher Education of Malaysia grant number 600 - RMI/LRGS/5/3 (2/2011), Universiti Teknologi MARA, Universiti Kebangsaan Malaysia (UKM-Hejim-Komuniti-15-2010); occupied Palestinian territory: the UN Relief and Works Agency for Palestine Refugees in the Near East, occupied Palestinian territory; International Development Research Centre, Canada; Philippines: Philippine Council for Health Research & Development; Poland: Polish Ministry of Science and Higher Education grant number 290/W-PURE/2008/0, Wroclaw Medical University; Saudi Arabia: the Deanship of Scientific Research at King Saud University, Riyadh, Saudi Arabia (research group number RG -1436-013); South Africa: the North-West University, SANPAD (SA and Netherlands Programme for Alternative Development), National Research Foundation, Medical Research Council of SA, The SA Sugar Association (SASA), Faculty of Community and Health Sciences (UWC); Sweden: grants from the Swedish state under the Agreement concerning research and education of doctors; the Swedish Heart and Lung Foundation; the Swedish Research Council; the Swedish Council for Health, Working Life and Welfare, King Gustaf V's and Queen Victoria Freemasons Foundation, AFA Insurance, Swedish Council for Working Life and Social Research, Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning, grant from the Swedish State under the Läkar Utbildnings Avtalet agreement, and grant from the Västra Götaland Region; Turkey: Metabolic Syndrome Society, AstraZeneca, Turkey, Sanofi Aventis, Turkey; United Arab Emirates (UAE): Sheikh Hamdan Bin Rashid Al Maktoum Award For Medical Sciences and Dubai Health Authority, Dubai UAE.

References

1 WHO. Diet, nutrition and the prevention of chronic diseases. World Health Organization Technical Report Series, 1990. http://apps.who.int/iris/bitstream/10665/39426/1/WHO_TRS_797_(part1).pdf (accessed Jan 20, 2017).

2 National Nutrition Council. Dietary recommendations to improve public health and prevent chronic diseases. Methodology and scientific basis. National Nutrition Council 2011. https://helsedirektoratet.no/publikasjoner/kostrad-for-a-fremme-folkehelsen-og-forebygge-kroniske-sykdommer-metodologi-og-vitenskapelig-kunnskapsgrunnlag (accessed July 4, 2017).

3 Micha R, Khatibzadeh S, Shi P, Andrews KG, Engell RE, Mozaffarian D. Global, regional and national consumption of major food groups in 1990 and 2010: a systematic analysis including 266 country-specific nutrition surveys worldwide. *BMJ Open* 2015; **5:** e008705.

4 Aune D, Giovannucci E, Boffetta P, et al. Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality-a systematic review and dose-response meta-analysis of prospective studies. *Int J Epidemiol* 2017; published on Feb 22. DOI:10.1093/ije/dyw319.

5 Miller V, Yusuf S, Chow CK, et al. Availability, affordability, and consumption of fruits and vegetables in 18 countries across income levels: findings from the Prospective Urban Rural Epidemiology (PURE) study. *Lancet Glob Health* 2016; **4:** e695–703.

6 Eichholzer M, Luthy J, Gutzwiller F, Stahelin HB. The role of folate, antioxidant vitamins and other constituents in fruit and vegetables in the prevention of cardiovascular disease: the epidemiological evidence. *Int J Vitam Nutr Res* 2001; **71:** 5–17.

7 Oude Griep LM, Verschuren WM, Kromhout D, Ocke MC, Geleijnse JM. Raw and processed fruit and vegetable consumption and 10-year stroke incidence in a population-based cohort study in the Netherlands. *Eur J Clin Nutr* 2011; **65:** 791–99.

8 Chow CK, Teo KK, Rangarajan S, et al. Prevalence, awareness, treatment, and control of hypertension in rural and urban communities in high-, middle-, and low-income countries. *JAMA* 2013; **310:** 959–68.

9 Corsi DJ, Subramanian SV, Chow CK, et al. Prospective Urban Rural Epidemiology (PURE) study: baseline characteristics of the household sample and comparative analyses with national data in 17 countries. *Am Heart J* 2013; **166:** 636–46.

10 Teo K, Chow CK, Vaz M, Rangarajan S, Yusuf S. The Prospective Urban Rural Epidemiology (PURE) study: examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. *Am Heart J* 2009; **158:** 1–7.

11 Yusuf S, Islam S, Chow CK, et al. Use of secondary prevention drugs for cardiovascular disease in the community in high-income, middle-income, and low-income countries (the PURE Study): a prospective epidemiological survey. *Lancet* 2011; **378:** 1231–43.

12 Dehghan M, Al Hamad N, Yusufali A, Nusrath F, Yusuf S, Merchant AT. Development of a semi-quantitative food frequency questionnaire for use in United Arab Emirates and Kuwait based on local foods. *Nutr J* 2005; **4:** 18.

13 Dehghan M, del Cerro S, Zhang X, et al. Validation of a semi-quantitative Food Frequency Questionnaire for Argentinean adults. *PLoS One* 2012; **7:** e37958.

14 Dehghan M, Ilow R, Zatonska K, et al. Development, reproducibility and validity of the food frequency questionnaire in the Poland arm of the Prospective Urban and Rural Epidemiological (PURE) study. *J Hum Nutr Diet* 2012; **25:** 225–32.

15 Dehghan M, Lopez Jaramillo P, Duenas R, et al. Development and validation of a quantitative food frequency questionnaire among rural- and urban-dwelling adults in Colombia. *J Nutr Educ Behav* 2012; **44:** 609–13.

16 Dehghan M, Martinez S, Zhang X, et al. Relative validity of an FFQ to estimate daily food and nutrient intakes for Chilean adults. *Public Health Nutr* 2013; **16:** 1782–88.

17 Merchant AT, Dehghan M, Chifamba J, Terera G, Yusuf S. Nutrient estimation from an FFQ developed for a Black Zimbabwean population. *Nutr J* 2005; **4:** 37.

18 Bharathi AV, Kurpad AV, Thomas T, Yusuf S, Saraswathi G, Vaz M. Development of food frequency questionnaires and a nutrient database for the Prospective Urban and Rural Epidemiological (PURE) pilot study in south India: methodological issues. *Asia Pac J Clin Nutr* 2008; **17:** 178–85.

19 Gunes FE, Imeryuz N, Akalin A, et al. Development and validation of a semi-quantitative food frequency questionnaire to assess dietary intake in Turkish adults. *J Pak Med Assoc* 2015; **65:** 756–63.

20 Iqbal R, Ajayan K, Bharathi AV, et al. Refinement and validation of an FFQ developed to estimate macro- and micronutrient intakes in a south Indian population. *Public Health Nutr* 2009; **12:** 12–18.

21 Mahajan R, Malik M, Bharathi AV, et al. Reproducibility and validity of a quantitative food frequency questionnaire in an urban and rural area of northern India. *Natl Med J India* 2013; **26:** 266–72.

22 Dehghan M, Al-Hamad N, McMillan CR, Prakash P, Merchant AT. Comparison of a semi-quantitative food frequency questionnaire with 24-hour dietary recalls to assess dietary intake of adult Kuwaitis. *Saudi Med J* 2009; **30:** 159–61.

23 Agudo A. Measuring intake of fruit and vegetables. Joint FAO/WHO workshop on fruits and vegetables for health, 2004. http://www.who.int/dietphysicalactivity/publications/f&v_intake_measurement.pdf (accessed Jan 20, 2017).

24 Painter J, Rah JH, Lee YK. Comparison of international food guide pictorial representations. *J Am Diet Assoc* 2002; **102:** 483–89.

25 Wang X, Ouyang Y, Liu J, et al. Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies. *BMJ* 2014; **349:** g4490.

26 Joshipura KJ, Ascherio A, Manson JE, et al. Fruit and vegetable intake in relation to risk of ischemic stroke. *JAMA* 1999; **282:** 1233–39.

27 Sharma S, Cruickshank JK, Green DM, Vik S, Tome A, Kolonel LN. Impact of diet on mortality from stroke: results from the US multiethnic cohort study. *J Am Coll Nutr* 2013; **32:** 151–59.

28 Larsson SC, Virtamo J, Wolk A. Total and specific fruit and vegetable consumption and risk of stroke: a prospective study. *Atherosclerosis* 2013; **227:** 147–52.

29 Zhang X, Shu XO, Xiang YB, et al. Cruciferous vegetable consumption is associated with a reduced risk of total and cardiovascular disease mortality. *Am J Clin Nutr* 2011; **94:** 240–46.

30 Buil-Cosiales P, Toledo E, Salas-Salvado J, et al. Association between dietary fibre intake and fruit, vegetable or whole-grain consumption and the risk of CVD: results from the PREvencion con DIeta MEDiterranea (PREDIMED) trial. *Br J Nutr* 2016; **116:** 534–46.

31 Leenders M, Sluijs I, Ros MM, et al. Fruit and vegetable consumption and mortality: European prospective investigation into cancer and nutrition. *Am J Epidemiol* 2013; **178:** 590–602.

32 Nguyen B, Bauman A, Gale J, Banks E, Kritharides L, Ding D. Fruit and vegetable consumption and all-cause mortality: evidence from a large Australian cohort study. *Int J Behav Nutr Phys Act* 2016; **13:** 9.

33 Link LB, Potter JD. Raw versus cooked vegetables and cancer risk. *Cancer Epidemiol Biomarkers Prev* 2004; **13:** 1422–35.

34 Dewanto V, Wu X, Adom KK, Liu RH. Thermal processing enhances the nutritional value of tomatoes by increasing total antioxidant activity. *J Agric Food Chem* 2002; **50:** 3010–14.

35 Rock CL, Lovalvo JL, Emenhiser C, Ruffin MT, Flatt SW, Schwartz SJ. Bioavailability of β-carotene is lower in raw than in processed carrots and spinach in women. *J Nutr* 1998; **128:** 913–16.

36 Key TJ, Thorogood M, Appleby PN, Burr ML. Dietary habits and mortality in 11,000 vegetarians and health conscious people: results of a 17 year follow up. *BMJ* 1996; **313:** 775–79.

37 Afshin A, Micha R, Khatibzadeh S, Mozaffarian D. Consumption of nuts and legumes and risk of incident ischemic heart disease, stroke, and diabetes: a systematic review and meta-analysis. *Am J Clin Nutr* 2014; **100:** 278–88.

38 Akibode S, Maredia M. Global and regional trends in production, trade and consumption of food legume crops. SPIA: Michigan State University, Department of Agricultural Fare, 2011.

39 Kabagambe EK, Baylin AF, Ruiz-Narvarez E, Fau S X, Campos H. Decreased consumption of dried mature beans is positively associated with urbanization and nonfatal acute myocardial infarction. *J Nutr* 2005; **135:** 1770–75.

40 Micha R, Wallace SK, Mozaffarian D. Red and processed meat consumption and risk of incident coronary heart disease, stroke, and diabetes mellitus: a systematic review and meta-analysis. *Circulation* 2010; **121:** 2271–83.

41 Mozaffarian D, Kamineni A, Carnethon M, Djousse L, Mukamal KJ, Siscovick D. Lifestyle risk factors and new-onset diabetes mellitus in older adults: the cardiovascular health study. *Arch Intern Med* 2009; **169:** 798–807.

42 Asplund K. Antioxidant vitamins in the prevention of cardiovascular disease: a systematic review. *J Intern Med* 2002; **251:** 372–92.

43 Appel LJ, Moore TJ, Obarzanek E, et al. A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. *N Engl J Med* 1997; **336:** 1117–24.

44 John JH, Ziebland S, Yudkin P, Roe LS, Neil HA. Effects of fruit and vegetable consumption on plasma antioxidant concentrations and blood pressure: a randomised controlled trial. *Lancet* 2002; **359:** 1969–74.

45 Lefer AM. Prostacyclin, high density lipoproteins, and myocardial ischemia. *Circulation* 1990; **81:** 2013–15.

46 Whelton PK, He J, Cutler JA, et al. Effects of oral potassium on blood pressure. Meta-analysis of randomized controlled clinical trials. *JAMA* 1997; **277:** 1624–32.

47 Jee SH, Miller ER 3rd, Guallar E, Singh VK, Appel LJ, Klag MJ. The effect of magnesium supplementation on blood pressure: a meta-analysis of randomized clinical trials. *Am J Hypertens* 2002; **15:** 691–96.

48 Anderson JW, O'Neal DS, Riddell-Mason S, Floore TL, Dillon DW, Oeltgen PR. Postprandial serum glucose, insulin, and lipoprotein responses to high- and low-fiber diets. *Metab Clin Exper* 1995; **44:** 848–54.

49 Ripsin CM, Keenan JM, Jacobs DR Jr, et al. Oat products and lipid lowering. a meta-analysis. *JAMA* 1992; **267:** 3317–25.

50 Anderson JW, Major AW. Pulses and lipaemia, short- and long-term effect: potential in the prevention of cardiovascular disease. *Br J Nutr* 2002; **88** (suppl 3): S263–71.

THE LANCET

Supplementary appendix

Supplement to: Miller V, Mente A, Dehghan M, et al, on behalf of the Prospective Urban Rural Epidemiology (PURE) study investigator. Fruit, vegetable, and legume intake, and cardiovascular disease and deaths in 18 countries (Prospective Urban Rural Epidemiology [PURE]): a prospective cohort study. *Lancet* 2017; published online Aug 29. http://dx.doi.org/10.1016/S0140-6736(17)32253-5.

Supplementary methods

Tables

Figures

Supplementary Methods: PURE Study Participant Selection Methodology as Excerpted from Teo et al(1).

Selection of Countries

The choice and number of countries selected in PURE reflects a balance between involving a large number of communities in countries at different economic levels, with substantial heterogeneity in social and economic circumstances and policies, and the feasibility of centers to successfully achieve long-term follow-up. Thus, PURE included sites in which investigators are committed to collecting good-quality data for a low-budget study over the planned 10-year follow-up period and did not aim for a strict proportionate sampling of the entire world.

Selection of Communities

Within each country, urban and rural communities were selected based on broad guidelines. A common definition for "community" that is applicable globally is difficult to establish (2). In PURE, a community was defined as a group of people who have common characteristics and reside in a defined geographic area. A city or large town was not usually considered to be a single community, rather communities from low-, middle-, and high-income areas were selected from sections of the city and the community area defined according to a geographical measure (e.g., a set of contiguous postal code areas or a group of streets or a village). The primary sampling unit for rural areas in many countries was the village. The reason for inclusion of both urban and rural communities is that for many countries, urban and rural 9 environments exhibit distinct characteristics in social and physical environment, and hence, by sampling both, we ensured considerable variation in societal factors across PURE communities. The number of communities selected in each country varied, with the aim to recruit communities with substantial heterogeneity in social and economic circumstances balanced against the capacity of local investigators to maintain follow-up. In some countries (e.g., India, China, Canada, and Colombia), communities from several states/provinces were included to capture regional diversity, in policy, socioeconomic status, culture, and physical environment. In other countries (e.g., Iran, Poland, Sweden, and Zimbabwe), fewer communities were selected.

Selections of Households and Individuals

Within each community, sampling was designed to achieve a broadly representative sample of that community of adults aged between 35 and 70 years. The choice of sampling frame within each center was based on both "representativeness" and feasibility of long-term follow-up, following broad study guidelines. Once a community was identified, where possible, common and standardized approaches were applied to the enumeration of households, identification of individuals, recruitment procedures, and data collection. The method of approaching households differed between regions. For example, in rural areas of India and China, a community announcement was made to the village through contact of a community leader, followed by in-person door-to-door visits of all households. In contrast in Canada, initial contact was by mail followed by telephone inviting members of the households to a central clinic. Households were eligible if at least 1 member of the household was between the 10 ages of 35 and 70 years and the household members intended to continue living in their current home for a further 4 years. For each approach, at least 3 attempts at contact were made. All individuals within these households between 35 and 70 years providing written informed consent were enrolled. When an eligible household or eligible individual in a household refused to participate, demographics and self-reported data about CVD risk factors, education, and history of CVD, cancers and deaths in the households within the two previous years were recorded. To ensure standardization and high data quality, we used a comprehensive operations manual, training workshops, DVDs, regular communication with study personnel and standardized report forms. We entered all data in a customized database programmed with range and consistency checks which was transmitted electronically to the Population Health Research Institute in Hamilton (Ontario, Canada) where further quality checks were implemented.

Supplementary Methods: Representativeness of the PURE Cohort as Excerpted from Yusuf et al(3).

a) Are the countries included in the PURE cohort atypical?

We compared the countries participating in PURE with those participating in MONICA (4), the largest previous study of this nature conducted. We plotted national income (GDP/capita) against research output (number of publications recorded in SCOPUS between 1996 and 2010 per 100,000 population). These data are taken from a previous study undertaken by one of the investigators (M. McKee) on global health research capacity to inform policy discussions within WHO (5). These graphs demonstrate that PURE has captured the full diversity of countries on these two dimensions, in marked contrast to MONICA, which was concentrated in high income countries with substantial research capacity.

b) Are PURE populations representative of the countries in which they are situated?

The PURE household population compared to national statistics had more women (sex ratio 95.1 men per 100 women vs 100.3) and was older (33.1 years vs 27.3), although age had a positive linear relationship between the two data sources (Pearson's r = 0.92). PURE was 59.3% urban compared to an average of 63.1% in participating countries. The distribution of education was less than 7% different for each category, although PURE households typically had higher levels of education. For example, 37.8% of PURE household members had completed secondary education compared to 31.3% in the national data. However, age-adjusted annual mortality rates showed positive correlation for men (r = 0.91) and women (r = 0.92) but were lower in PURE compared to national statistics (7.9 per 1000 vs 8.7 for men; 6.7 vs 8.1 for women). These findings indicate that modest differences exist between the PURE household population 12 and national data for the indicators studied. These differences, however, are unlikely to have much influence on exposure-disease (or health systems assessments vs outcome) associations derived in PURE. In addition, mortality rates reported for the two years prior to enrolment in the PURE study were closely correlated with the mortality rates observed in the study participants during follow up (r=0.87 and 0.85 respectively). Further, incidence estimates from PURE, adjusted or stratified according to sex and/or urban/rural location, will enable valid comparisons of the relative rates of various cardiovascular outcomes across countries.

Supplementary Methods: Collection of Demographics, Risk Factors and Outcome Events

We collected data at national, community, household, and individual levels with standardized questionnaires. 1 Questions about age, sex, education, smoking status, hypertension, diabetes, and obesity were identical to those in the INTERHEART and INTERSTROKE studies (6,7). We obtained BP measurements in individuals and so hypertension was defined as those with a BP >140/>90 or those who were already on treatment. Fasting glucose was available in most individuals (76%) and so diabetes was defined as those who were reported as having diabetes and those with a fasting glucose >7.0 mmol/L. (Sensitivity analyses indicate a very high correlation between self report of diabetes alone versus self-report and fasting glucose >7.0 in the 110,000 people with both measures, and so self-report is a reasonable surrogate for the prevalence of diabetes) Total cholesterol was available in 122,640 individuals and a value of >5.2 mmol/L was considered to be elevated. In most of the LIC and MIC there was no central system of death or event registration. We therefore; 1) obtained information on prior medical illness and medically certified cause of death where available, 2) captured best available information from reliable sources in those instances where medical information was not available in order to be able to arrive at a probable diagnosis or cause of death. Event documentation was based on information from household interviews and medical records, death certificates and other sources. We also used Verbal Autopsies to ascertain cause of death in addition to medical records which were reviewed by a health professional. This approach has been used in several studies conducted in LIC and MIC (8,9) To ensure a standard approach and accuracy for classification of events across all countries and over time, the first 100 CVD events (deaths, MI, strokes, heart failure or cancers) for China and India, and 50 cases for other countries were adjudicated both locally and also by the adjudication chair, and if necessary further training was provided. Thereafter, every year, 50 cases for China and India and 25 cases for each of the remaining countries were adjudicated as above.

Table S1. PURE baseline response rates (%) by country and urban/rural location.

PURE Baseline Response Rates			
Region	**Country**	**Urban**	**Rural**
High-income countries	Canada	68	72
	Sweden	48	54
	UAE	63	83
Middle-income countries	Argentina	64	86
	Brazil	60	77
	Chile	79	89
	China	78	80
	Colombia	70	71
	Iran	76	90
	Malaysia	72	84
	Poland	74	70
	South Africa	69	50
	Turkey	78	84
	Occupied Palestinian territory	82	85
Low-income countries	Bangladesh	69	95
	India	57	60
	Pakistan	80	81
	Zimbabwe	75	87

Table S2. Median duration of follow-up in years by country

Country	Median
UAE	5·83
Canada	6·81
Sweden	8·92
Poland	7·22
Argentina	7·82
Chile	8·60
Malaysia	6·34
Turkey	6·17
Iran	6·94
Occupied Palestinian territory	3·20
Brazil	3·74
South Africa	6·71
Colombia	5·72
China	8·10
India	9·91
Pakistan	5·20
Bangladesh	7·38
Zimbabwe	9·02
Total	7·36

Supplementary Figure 1: Flowchart of participants in the PURE Study.

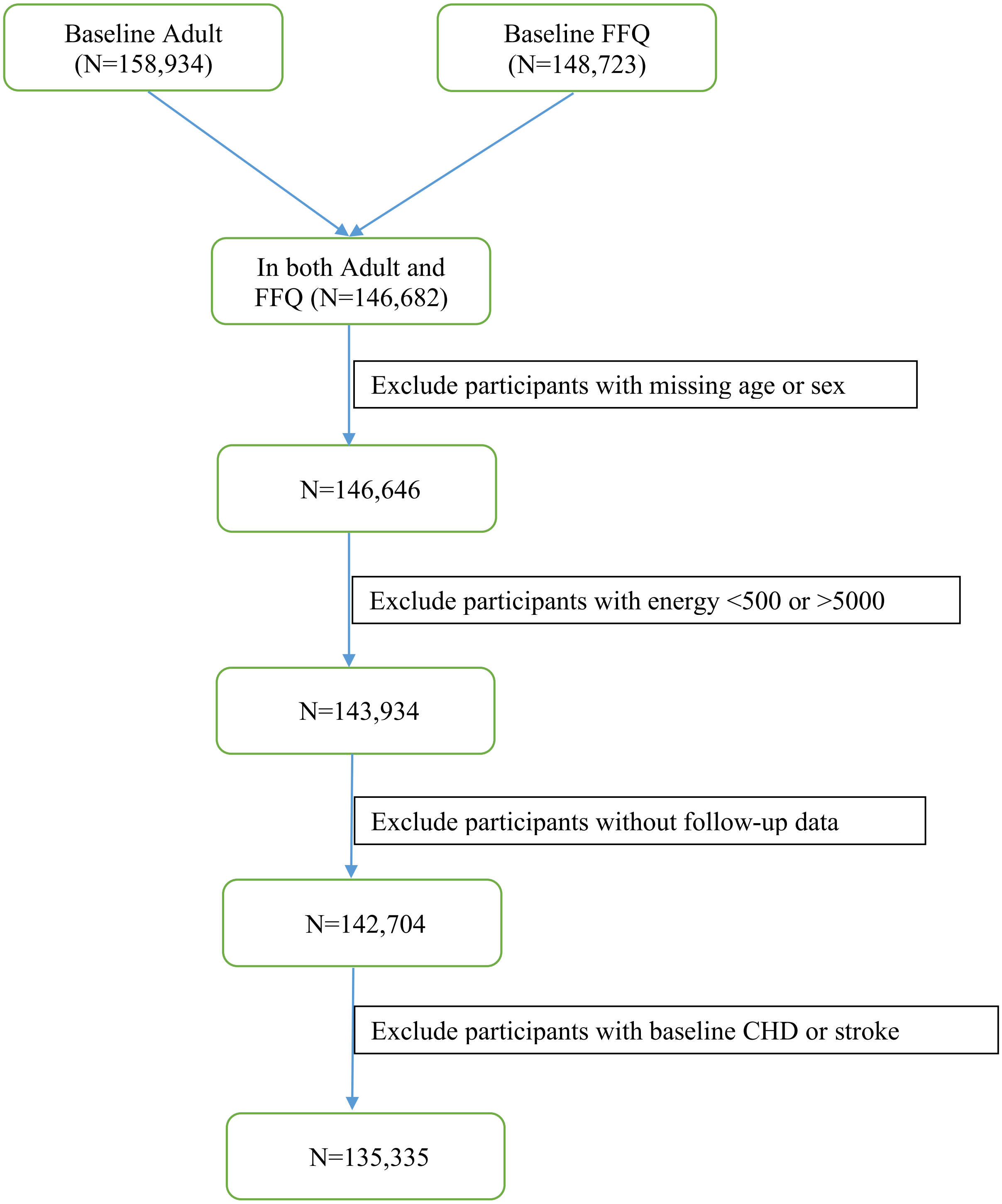


Table S3: Classification of PURE countries into geographic regions

Region	**Country**	**N**
North America and Europe (N=14,916)	Canada	9,114
	Poland	1,869
	Sweden	3,933
South America (N=22,626)	Argentina	7,039
	Brazil	5,615
	Chile	3,400
	Colombia	6,572
Middle East (N=11,485)	Iran	5,166
	Occupied Palestinian territory	1,400
	Turkey	3,636
	United Arab Emirates	1,283
South Asia (N=29,560)	Bangladesh	2,759
	India	25,324
	Pakistan	1,477
China (N=42,152)	China	42,152
South East Asia (N=10,038)	Malaysia	10,038
Africa (N=4,558)	South Africa	3,466
	Zimbabwe	1,092

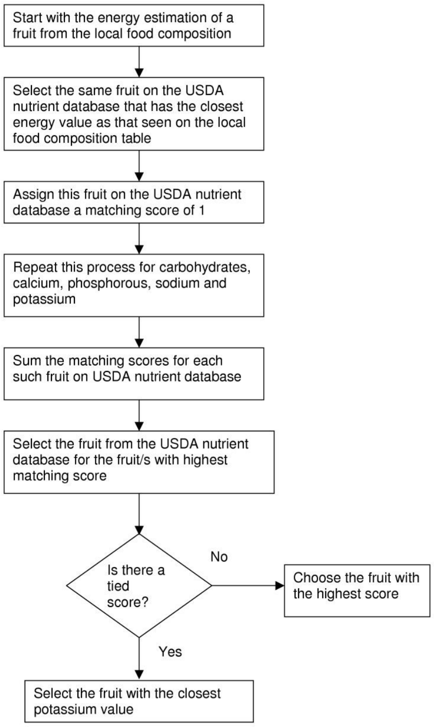
Supplementary Methods: Methodology Used for Estimating Dietary Intake

Participants' habitual food intake was recorded using country-specific (region specific in India) validated food frequency questionnaires (FFQs) at the baseline. For countries where a validated FFQ was not available, we developed and validated FFQs using a standard method (Table S1 in the Supplementary Appendix) (10-21). To convert food into nutrients, country specific nutrient databases were constructed with information on 43 macro and micro-nutrients. The nutrient database is primarily based on the United State Department of Agriculture (USDA) food composition database (release 18 and 21), modified with reference to local food composition tables, and supplemented with recipes of local mixed dishes (22). However, for Canada, China, India, Malaysia, South Africa, Sweden, and Turkey we used the nutrient databases that were used for FFQs validation. The FFQ was administered to the same subgroup of participants and Pearson correlation coefficients, using energy-adjusted and de-attenuated correlations, and weighted kappa were used to validate the FFQs measures against 24-hour dietary recalls. Our validation studies demonstrated reasonable agreement between the FFQs and 24-hour dietary recalls for fruits (r_s=0.23-0.66) and vegetables (r_s=0.30-0.81) which is in keeping with previous FFQ validation studies (23-27). The concordance rates of classification into the same quartiles ranged from 70% to 74% for fruit and 62% to 79% for vegetables, which is consistent with previous studies which reported cross classification (23-27).

Supplementary Methods: Food Composition Database Compilation as Excerpted from Merchant & Dehghan (28).

We used the USDA nutrient database as the primary nutrient data source for the PURE study because it is regularly updated, comprehensive and, the data are freely available. To ensure that the nutrient content of the foods were appropriate for the local countries, we referred to other sources such as the INFOODS food composition tables, or local food composition tables. As there are many entries for a single food (18 types of rice for instance) we developed the following algorithm to select the food from the USDA nutrient database that most closely matched the local food. To match the foods we considered total energy content and the following nutrients (macronutrients and minerals) for fruits and vegetables: energy, carbohydrates, calcium, phosphorous, sodium and potassium; dairy: energy, protein, fat, calcium, phosphorous; cereals: energy, carbohydrates, calcium, and phosphorous; and meats and eggs: energy, protein, fat, and iron, because these nutrients were likely to be present in those food groups. We only used macronutrients and minerals for matching because the assays for these nutrients have high within laboratory agreements (29). We did not include vitamins in the matching process because their estimation is sensitive to the assay, method of food preparation, and storage (30-33).

To select the food most similar to the local food we started with the estimated nutrient intake from the local food composition table. We first compared total energy intake for 100 g of that food estimated from the local food composition table. The food with the most similar total energy intake per 100 g of that food on the USDA nutrient database was given a matching score of 1. We repeated this process with the next nutrient and so on, until all the entries for that particular food group were exhausted. The food in the USDA nutrient database with the highest total watching score was considered as being the closest to the local food. In case of a tie we considered the closest match with potassium for fruits and vegetables, total fat for dairy and meats, protein for eggs, and carbohydrates for cereals. The algorithm to select fruits is described in the figure below.



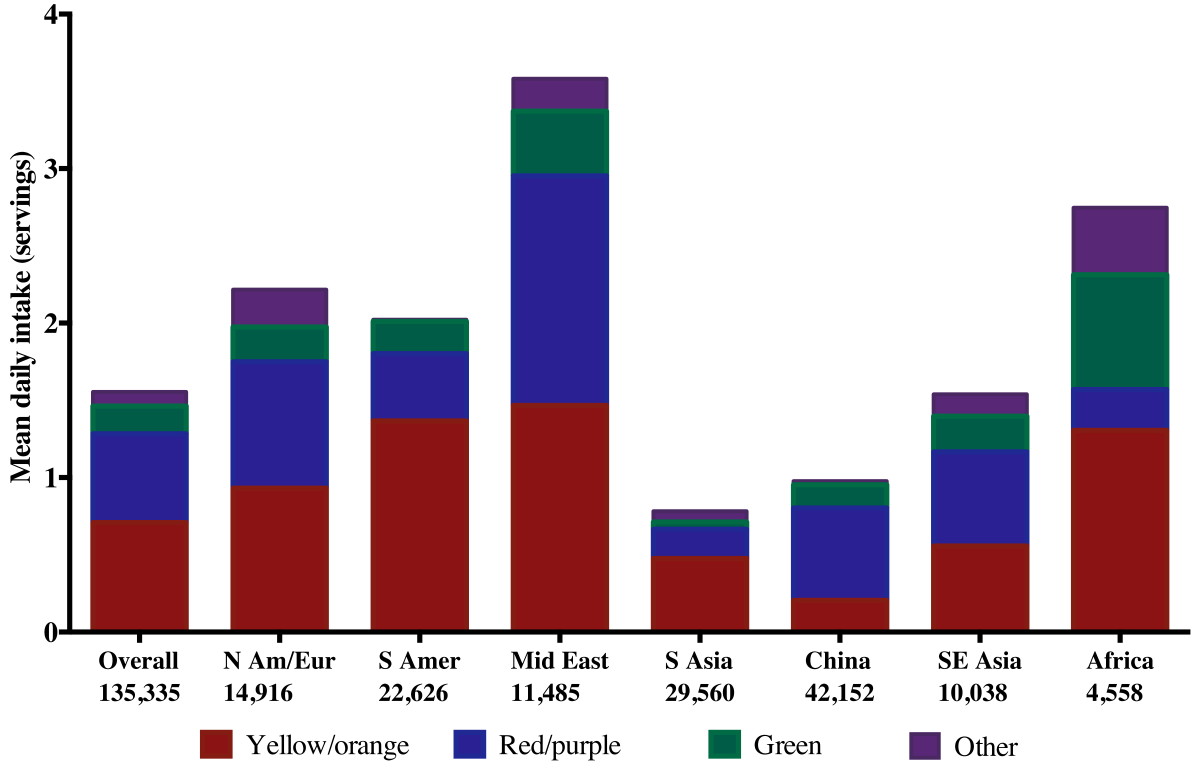


Figure S2: Mean contribution of fruit type to total fruit intake by region

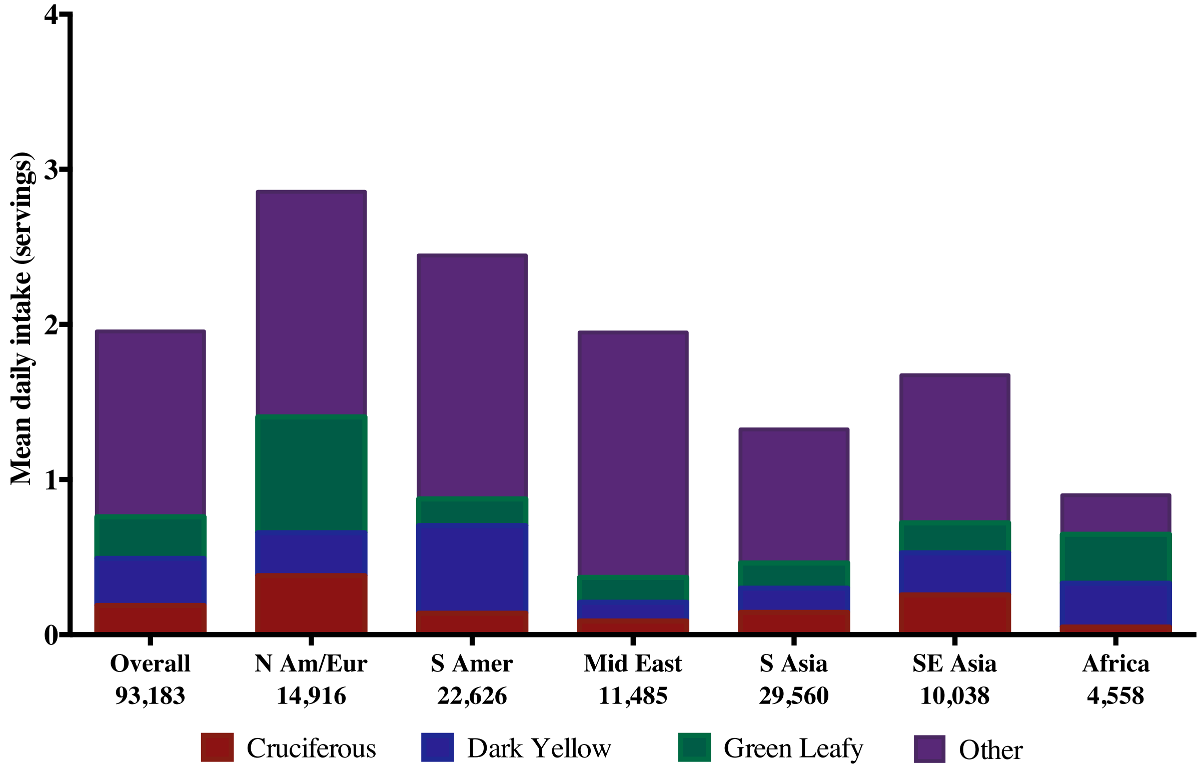


Figure S3: Mean contribution of vegetable type to total vegetable intake by region, not including China (N=42,152)

Supplementary Methods: Prospective Follow-up for Cardiovascular Events and Mortality

History of disease was collected at baseline from every participant with standardized questionnaires. Follow-up was initiated in all sites by 2008 and completed by March, 2017. Up to three attempts were made to interview all households to document events.

Information on specific events (death, myocardial infarction, stroke, heart failure, cancer, hospitalizations, new diabetes, injury, tuberculosis, human immunodeficiency viral infections, malaria, pneumonia, asthma, chronic obstructive pulmonary disease) were obtained from participants or their family members. This information was adjudicated centrally in each country by trained physicians using standardized definitions. Because the PURE study involves urban and rural areas from middle- and low-income countries, supporting documents to confirm cause of death and/or event varied in degrees of completion and availability. In most of middle- and low-income countries there was no central system of death or event registration. Therefore, information was obtained about prior medical illness and medically certified cause of death where available, and, second, best available information was captured from reliable sources in those instances where medical information was not available in order to be able to arrive at a probable diagnosis or cause of death. Event documentation was based on information from death certificates (available in 100% of deaths), medical records (MI: 49·4%, stroke 80·8% and heart failure: 76·2%), household interviews and other sources. Verbal Autopsies were also used to ascertain cause of death in addition to medical records which were reviewed by a health professional. This approach has been used in several studies conducted in middle- and low-income countries.

To ensure a standard approach and accuracy for classification of events across all countries and over time, the first 100 CVD events (deaths, MI, strokes, heart failure or cancers) for China and India, and 50 cases for other countries were adjudicated both locally and also by the adjudication chair, and if necessary further training was provided. Thereafter, every year, 50 cases for China and India and 25 cases for each of the remaining countries were adjudicated as above.

Supplementary Methods: Event Definitions

FATAL EVENTS

Cardiovascular Death – Definitions

01.00 DEATH DUE TO CARDIOVASCULAR EVENTS

01.10 Sudden unexpected Cardiovascular Death (SCVD)

Without evidence of other cause of death, death that occurred suddenly and unexpectedly (examples: witnessed collapse, persons resuscitated from cardiac arrest who later died) or persons seen alive less than 12 hours prior to discovery of death (example persons found dead in his/her bed).

- SCVD is either definite, probable or possible according to the following characteristics:

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
01.11: Definite	One of the following in persons with: • known cardiovascular disease, or • diabetes with an additional risk factor such as hypertension, smoking, dyslipidemia, micro albuminuria, serum creatinine 50% above upper limit of normal, or • 3 of the above risk factors, or • 2 of the above risk factors in men aged 60 and more and women aged 65 and more	No ICD-10 Code
01.12: Probable	One of the following in persons with: • diabetes, or • 2 of the above risk factors in men aged less than 60 and in women less than 65, or • one of the above risk factor in men aged 60 and more and in women aged 65 and more, or • typical of chest pain or sudden severe dyspnea of less than 20-minute duration preceding the event	
01.13: Possible	In persons without risk factor	
*For **SCVD**, the patient was well or had a stable CVD (example stable angina) when last seen alive. The event of a sudden death occurring during the hospitalization of MI is considered a fatal MI and not sudden death.*		

01.30 Fatal Myocardial Infarction

Symptoms of Myocardial Infarction:

Typical symptoms or suggestive symptoms of MI according to physician are characterized by severe anterior chest pain as tightness, crushing, burning, lasting at least 20 minutes, occurring at rest, or on exertion, that may radiate to the arms or neck or jaw and may be associated with dyspnea, diaphoresis and nausea. However, death associated with nausea and vomiting with or without chest pain not due to another cause may be considered as possible MI if ECG and cardiac markers are not done. These symptoms may have occurred the last month before death.

Fatal myocardial infarction is either definite, probable or possible according to the following characteristics:

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
01.31: Definite	1. Autopsy demonstrating fresh myocardial infarction and/or recent coronary occlusion, **or** 2. ECG showing new and definite sign of MI (Minnesota code 1-1-1) **or** 3. Symptoms typical or atypical or inadequately described but attributed to cardiac origin lasting at least 20 minutes and by troponin or cardiac enzymes (CKMB, CK, SGOT, SLDH) above center laboratory ULN 4. ECG with new ischemic changes (new ST elevation/depression or T wave inversion ≥ 2 mm) and by troponin or cardiac enzymes (CKMB, CK, SGOT, SLDH) above center laboratory ULN	I21- I22
01.32: Probable	1. ECG with sign of probable MI (Minnesota code 1-2-1), **or** 2. Typical symptoms lasting at least 20 minutes considered of cardiac origin, with only new ST-T changes (new ST elevation/depression or T wave inversion ≥ 1 but < 2mm) without documented increased cardiac markers or enzyme as in PURE definition 1.31 (above), **or** 3. Increased cardiac enzymes as in PURE definition 1.31 (above) showing a typical pattern of MI as above without symptoms or significant ECG changes	
01.33: Possible	1. ECG with sign of possible MI (Minnesota code 1-3-1) **or** 2. Typical symptoms or symptoms suggestive of MI according to the physician lasting at least 20 minutes without documented ECG or cardiac marker.	

The **Minnesota codes** for MI is taken from Rose and Blackburn and published in their book “Evaluation Methods of Cardiovascular Disease WHO 1969”.

- **Definite MI** is Q/R ratio ≥1/3 and Q duration ≥ 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-1-1)
- **Probable MI** is Q/R ratio ≥1/3 and Q duration between 0.02 and 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-2-1)
- **Possible MI** is Q/R ratio between 1/5 and 1/3 and Q duration between 0.02 and 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-3-1)

01.40 Fatal Stroke

Fatal stroke is either definite or possible according to the following characteristics:

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
01.41: Definite	Stroke death is defined as death within 30 days from an acute focal neurological deficit ***diagnosed by a physician*** and thought to be of vascular origin (without other cause such as brain tumor) with signs and symptoms lasting >= 24 hrs. Stroke death is also considered if death occurred within 24 hrs. of onset of persisting signs and symptoms, or if there is evidence of a recent stroke on autopsy. N.B. • **In a subject with a stroke <= 30 days**: If death occurred with a pneumonia due to possible aspiration, death will be considered to be due to stroke. • **In a subject with a stroke > 30 days**: If death occurred with a pneumonia due to possible aspiration, the adjudicator will make a decision according to his/her clinical judgment if death is related to stroke or not. • Subarachnoid hemorrhage death manifested by sudden onset headache with/without focal signs and imaging (CT or MRI) evidence of bleeding primarily in the subarachnoid space is considered a fatal stroke in absence of trauma or brain tumor or malformation • Subdural hematoma death is not considered as a stroke death and may be related to previous trauma or other cause.	I60- I64, I69
01.43: Possible	Death in a participant with a history of sudden onset of focal neurological deficit of one or more limbs, loss of vision or slurred speech lasting about 24 hours.	

01.50 Fatal Congestive Heart Failure

Fatal congestive heart failure is either definite or possible according to the following characteristics:

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
01.51: <u>Definite</u>	The diagnosis of congestive heart failure may be an autopsy finding in absence of other cause or requires signs (rales, increased jugular venous pressure or ankle edema) or symptoms (nocturnal paroxysmal dyspnea, dyspnea at rest or ankle edema) of congestive heart failure **<u>and</u>** one or both of the following: • radiological signs of pulmonary congestion, • treatment of heart failure with diuretics *If sudden death occurred in a patient with chronic severe heart failure, it should be adjudicated as fatal congestive heart failure.*	I50
01.52: <u>Probable</u>	Progressive shortness of breath on lying down or at night, improving on sitting up AND any of the following signs or symptoms: swelling of feet, distension of abdomen, progressive cough in a person with known hypertension or a history of previous MI/angina or other heart disease	
01.53: <u>Possible</u>	Progressive shortness of breath on lying down or at night, improving on sitting up AND any of the following signs or symptoms: swelling of feet, distension of abdomen, progressive cough	

01.60 Death Due to Other Cardiovascular Deaths *(other causes [1.10 to 1.50 above] having been excluded)*

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
01.61	Arterial rupture of aneurysm	I71- I72
01.62	Pulmonary embolism *NOTE: Death associated with pulmonary embolism occurring* ***<u>within 2 weeks</u>*** *after a fracture such as hip, femur should attributed to death due to injury. Refer to Injury, Section 6.0*	I26
01.63	Arrhythmic death (A-V block, sustained ventricular tachycardia in absence of other causes)	I44- I45, I47- I49
01.64	Death after invasive cardiovascular intervention: a perioperative death extending to 30 days after coronary or arterial surgical revascularization and to 7 days after a coronary or arterial percutaneous dilatation (angioplasty) with or without a stent or an invasive diagnostic procedure.	I97
01.65	Congenital heart disease	Q20-Q28
01.66	Heart valve disease (including rheumatic heart disease)	I01, I05- I09, I34- I37
01.67	Endocarditis	I33, I38
01.68	Myocarditis	I40
01.69	Tamponade (pericarditis)	I30, I31, I32
01.70	Other cardiovascular events *(Excluding* ***1.61 to 1.69*** *above)* ***Valid*** *ICD-10 codes would include the following:* *I11, I12, I13, I23, I24, I25, I27, I28, I42, I51, I52, I65-I68, I73, I74, I96, I98, I99* *(Refer to ICD-10 Listing for associated definitions for each code)*	Any valid 'I' (Cardiovascular) ICD-10 code that can be classified as underlying CoD, not specified above

NON-FATAL EVENTS
Cardiovascular Events – Definitions
10.00 NON-FATAL CARDIOVASCULAR EVENTS
10.10 Non-Periprocedural Myocardial Infarction (MI)

MI is considered either definite, probable or possible according to the following characteristics:

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
10.11: Definite	5. ECG showing new and definite sign of MI (Minnesota code 1-1-1) **or** 6. Symptoms typical or atypical or inadequately described but attributed to cardiac origin lasting at least 20 minutes and by troponin or cardiac enzymes (CKMB, CK, SGOT, SLDH) above center laboratory ULN 7. ECG with new ischemic changes (new ST elevation/depression or T wave inversion ≥ 2 mm) and by troponin or cardiac enzymes (CKMB, CK, SGOT, SLDH) above center laboratory ULN Please note that increased markers may occur in trauma (CK, AST, myoglobin and CK MB to a lesser degree); renal insufficiency, heart failure, pulmonary embolism…. (troponin), cardioversion (all)	I21-I22
10.12: Probable	4. ECG with new and probable sign of MI (Minnesota code 1-2-1), **or** 5. Typical symptoms lasting at least 20 minutes considered of cardiac origin, with only new ST-T changes (new ST elevation/depression or T wave inversion ≥ 1 but < 2mm) without documented increased cardiac markers as in PURE definition 10.11 (above), **or** 6. Increased cardiac enzymes showing a typical pattern of MI as above without symptoms or significant ECG changes.	
10.13: Possible	1. ECG with new and possible sign of MI (Minnesota code 1-3-1), **or** 2. Typical symptoms lasting 20 minutes and more considered to be of cardiac origin without documented ECG or cardiac marker.	

10.20 Periprocedural Myocardial Infarction

PURE Adjudication Code	**Event Type**	**Acceptable ICD-10 codes**
10.21: <u>Definite</u>	1. ECG showing new and definite sign of MI (Minnesota code 1-1-1), or 2. Increased cardiac markers within 48 hours of procedure: • percutaneous coronary intervention: CKMB should be ≥ 5 X ULN or troponin ≥ 5 X above lower level of necrosis **OR** > 20% increase in cardiac markers if elevated at the beginning of the procedure in a patient with symptoms suggestive of myocardial ischemia • Coronary surgery: Increased cardiac markers CKMB should be ≥ 10X ULN or troponin ≥ 10X above lower limit of necrosis.	I21-I22

The **Minnesota codes** for MI is taken from Rose and Blackburn and published in their book “Evaluation Methods of Cardiovascular Disease WHO 1969”.

- **Definite MI** is Q/R ratio ≥1/3 and Q duration ≥ 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-1-1)
- **Probable MI** is Q/R ratio ≥1/3 and Q duration between 0.02 and 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-2-1)
- **Possible MI** is Q/R ratio between 1/5 and 1/3 and Q duration between 0.02 and 0.03 second in one of the following leads: I, II, V2, 3, 4, 5, 6. (code 1-3-1)

10.30 Stroke/Transient Ischemic Attack (TIA)

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
10.31: <u>Definite</u>	Stroke is defined as an acute focal neurological deficit ***diagnosed by a physician*** and thought to be of vascular origin (without other case such as brain tumor) with signs and symptoms lasting >= 24 hrs. N.B. • Subarachnoid hemorrhage manifested by sudden onset headache with/without focal signs and imaging (CT or MRI or lumbar puncture) showing evidence of bleeding primarily in the subarachnoid space is considered a stroke in absence of trauma or brain tumor or malformation • Subdural hematoma is <u>not</u> considered as a stroke and may be related to previous trauma or other cause.	I60-I64, I69
10.33: <u>Possible</u>	Stroke is possible if there is a history of sudden onset of focal neurological deficit of one or more limbs, loss of vision or slurred speech lasting about 24 hours or more	
10.34: <u>TIA</u>	The diagnosis of TIA requires the presence of acute focal neurological deficit thought to be of vascular origin with signs and symptoms lasting less than 24 hours	G45

10.40 Congestive Heart Failure

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
10.41: <u>Definite</u>	The diagnosis of congestive heart failure requires signs (rales, increased jugular venous pressure or ankle edema) or symptoms (nocturnal paroxysmal dyspnea, dyspnea at rest or ankle edema) of congestive heart failure **<u>and</u>** one or both of the following: • radiological signs of pulmonary congestion, • Treatment of heart failure with diuretics.	I50
10.42: <u>Probable</u>	Progressive shortness of breath on lying down or at night, improving on sitting up AND any of the following signs or symptoms: swelling of feet, distension of abdomen, progressive cough in a person with known hypertension or a history of previous MI/angina or other heart disease	
10.43: <u>Possible</u>	Congestive heart failure is considered possible when there is progressive shortness of breath on lying down or at night, improving on sitting up AND any of the following signs or symptoms: swelling of feet, distension of abdomen, progressive cough	

10.50 Effort Angina with documented Ischemia

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
10.51: Definite	Stress test with ECG with new ST depression >1 mm or positive imaging (ECHO, Scan) compatible with ischemia	I20
10.52: Probable	Typical effort angina (i.e. Squeezing, pressure or burning type pain touching the sternum occurring on exertion and relieved by rest or nitroglycerin)	

10.60 Unstable Angina

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
10.61: <u>Definite</u>	1. Hospitalization for typical symptoms with new ECG changes (T wave inversion < 2mm) **<u>OR</u>** 2. Coronary revascularization within one week of admission, and treated as unstable angina.	I20
10.62: <u>Probable</u>	1. Hospitalization for typical symptoms lasting at least 10 but less than 20 minutes without ECG or cardiac enzyme elevation	
10.63: <u>Possible</u>	Not hospitalized for typical symptoms of angina occurring at rest and treated as unstable angina: rest, anti-platelets, nitrates, beta blockers and/or calcium channel blockers.	

10.80 Other Non-Fatal Cardiovascular Events

PURE Adjudication Code	Event Type	Acceptable ICD-10 codes
10.81	Rheumatic Heart Disease	I01, I05-I09
10.82	Other valvular heart diseases (excluding Rheumatic Heart Disease)	I34-I37
10.83	Pericarditis	I30, I31, I32
10.84	Endocarditis	I33, I38
10.85	Myocarditis	I40
10.86	Congenital heart disease	Q20-Q28
10.87	Atrial fibrillation/flutter Atrial Fibrillation and Flutter are due to an abnormal cardiac rhythm at the atrial level, and the diagnosis is made on the electrocardiographic (ECG) tracing or monitor.	I48
10.88	Peripheral artery disease (lower limb iliac to popliteal and carotid)	I73
10.90	Pulmonary embolism	I26
10.91	Other cardiac or arterial diseases: *Specific details should be provided on the corresponding Adjudication Form*	Any ‘I’ (Cardiovascular) ICD-10 code not classified above

Table S4: Ranking of cause-specific deaths by region.

Region	Cardiovascular disease	Cancer	Respiratory	Injury
Overall	1	2	3	4
South Asia	1	2	3	4
China	2	1	4	3
Southeast Asia	1	2	3	4
Africa	1	3	2	4
North America/Europe	2	1	3	4
Middle East	1	2	4	3
South America	2	1	3	4

Table S5: Association of total fruit and vegetable intake with cardiovascular outcomes and mortality

All participants (N=135,335)	**<1/day (N=16,512)**	**1 to <2/day (N=20,478)**	**2 to <3/day (N=36,036)**	**3 to <4/day (N=22,134)**	**4 to <5/day (N=12,604)**	**5 to <6/day (N=8,575)**	**6 to 7/day (N=5,841)**	**7 to 8/day (N=4,068)**	**>8/day (N=9,078)**	**P-trend**
Median fruit and vegetable intake, servings	0·61 (0·38-0·81) per day	1·53 (1·27-1·77) per day	2·47 (2·25-2·71) per day	3·43 (3·20-3·70) per day	4·44 (4·20-4·70) per day	5·45 (5·22-5·72) per day	6·45 (6·22-6·71) per day	7·45 (7·22-7·71) per day	9·98 (8·82-12·12) per day	
Major CVD										
Major CVD events (N=4,784)	712 (4·31)	746 (3·64)	1,414 (3·92)	782 (3·53)	409 (3·25)	234 (2·73)	153 (2·62)	97 (2·38)	237 (2·61)	
Age and sex adjusted	1·00	0·93 (0·84-1·04)	1·02 (0·92-1·13)	0·99 (0·88-1·11)	0·98 (0·86-1·12)	0·84 (0·71-0·98)	0·81 (0·67-0·98)	0·76 (0·60-0·95)	0·81 (0·69-0·96)	0·0010
Multivariable adjusted	1·00	0·94 (0·83-1·05)	1·04 (0·93-1·17)	1·05 (0·92-1·19)	1·06 (0·91-1·23)	0·89 (0·75-1·06)	0·90 (0·73-1·10)	0·81 (0·64-1·03)	0·88 (0·73-1·07)	0·1045
MI										
MI events (N=2,143)	361 (2·19)	396 (1·93)	541 (1·50)	329 (1·49)	186 (1·48)	104 (1·21)	64 (1·10)	47 (1·16)	115 (1·27)	
Age and sex adjusted	1·00	0·97 (0·84-1·12)	1·10 (0·95-1·28)	1·09 (0·92-1·29)	1·01 (0·83-1·23)	0·83 (0·66-1·06)	0·74 (0·56-0·98)	0·78 (0·57-1·08)	0·89 (0·70-1·13)	0·0345
Multivariable adjusted	1·00	0·96 (0·82-1·13)	1·10 (0·94-1·30)	1·13 (0·94-1·36)	1·07 (0·87-1·32)	0·86 (0·67-1·11)	0·80 (0·59-1·08)	0·87 (0·62-1·22)	0·95 (0·72-1·25)	0·2400
Stroke										
Stroke events (N=2,234)	283 (1·71)	274 (1·34)	785 (2·18)	406 (1·83)	176 (1·40)	110 (1·28)	70 (1·20)	39 (0·96)	91 (1·00)	
Age and sex adjusted	1·00	0·89 (0·75-1·05)	0·98 (0·83-1·15)	0·97 (0·82-1·16)	0·96 (0·78-1·18)	0·96 (0·76-1·23)	0·95 (0·71-1·26)	0·82 (0·64-1·10)	0·84 (0·64-1·10)	0·3790
Multivariable adjusted	1·00	0·87 (0·72-1·05)	0·98 (0·82-1·17)	1·02 (0·84-1·23)	1·01 (0·81-1·27)	1·02 (0·79-1·33)	1·04 (0·77-1·41)	0·80 (0·54-1·17)	0·90 (0·67-1·23)	0·9628
CV mortality										
CV mortality events (N=1,649)	409 (2·48)	350 (1·71)	367 (1·02)	195 (0·88)	123 (0·98)	67 (0·78)	49 (0·84)	23 (0·57)	66 (0·73)	
Age and sex adjusted	1·00	0·83 (0·72-0·97)	0·84 (0·71-0·98)	0·72 (0·60-0·87)	0·80 (0·64-0·99)	0·66 (0·50-0·87)	0·75 (0·55-1·03)	0·51 (0·33-0·79)	0·65 (0·49-0·87)	<0·0001
Multivariable adjusted	1·00	0·88 (0·74-1·04)	0·87 (0·73-1·05)	0·80 (0·65-0·99)	0·89 (0·70-1·14)	0·72 (0·53-0·98)	0·90 (0·64-1·27)	0·60 (0·38-0·95)	0·78 (0·55-1·09)	0·0574
Non-CV mortality										
Non-CV mortality events	1,044 (6·32)	772 (3·77)	875 (2·43)	384 (1·73)	236 (1·87)	157 (1·83)	100 (1·71)	66 (1·62)	175 (1·93)	

(N=3,809)										
Age and sex adjusted	1·00	0·80 (0·73-0·89)	0·71 (0·63-0·79)	0·55 (0·48-0·63)	0·60 (0·51-0·70)	0·57 (0·48-0·69)	0·54 (0·43-0·67)	0·50 (0·38-0·65)	0·55 (0·46-0·66)	<0·0001
Multivariable adjusted	1·00	0·89 (0·80-0·99)	0·84 (0·74-0·95)	0·71 (0·61-0·82)	0·81 (0·68-0·96)	0·77 (0·63-0·94)	0·79 (0·62-1·00)	0·73 (0·55-0·96)	0·84 (0·67-1·04)	0·0142
Mortality										
Mortality events (N=5,796)	1,547 (9·37)	1,195 (5·84)	1,313 (3·64)	614 (2·77)	380 (3·01)	235 (2·74)	159 (2·72)	96 (2·36)	257 (2·83)	
Age and sex adjusted	1·00	0·81 (0·75-0·88)	0·74 (0·68-0·81)	0·60 (0·54-0·74)	0·65 (0·58-0·74)	0·59 (0·51-0·68)	0·60 (0·50-0·71)	0·51 (0·41-0·64)	0·58 (0·50-0·68)	<0·0001
Multivariable adjusted	1·00	0·88 (0·80-0·96)	0·84 (0·76-0·93)	0·73 (0·65-0·82)	0·82 (0·72-0·94)	0·73 (0·62-0·86)	0·81 (0·67-0·97)	0·69 (0·55-0·86)	0·81 (0·68-0·96)	0·0006

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure MI: myocardial infarction; CV mortality: cardiovascular mortality; Non-CV mortality: non-cardiovascular mortality. Crude event rates shown. Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, bread and cereals and vegetable intake. Additional sensitivity analyses with waist-to-hip ratio, hypertension status and statin medication used in the model did not materially alter estimates of association (Shown in the Supplementary Appendix).

Table S6: Association of fruit intake with cardiovascular outcomes and mortality.

All participants (N=135,335)	**<3/week (N=37,849)**	**3/week to <1/day (N=31,929)**	**1 to <2/day (N=32,850)**	**2 to <3/day (N=14,706)**	**>3/day (N=18,001)**	**P-trend**
Median fruit intake, servings	0·20 (0·08-0·31) per day	0·68 (0·55-0·83) per day	1·41 (1·19-1·68) per day	2·41 (2·19-2·66) per day	4·28 (3·52-5·81) per day	
Major CVD						
Major CVD events (N=4,784)	1,646 (4·35)	1,160 (3·63)	1,034 (3·15)	419 (2·85)	525 (2·92)	
Age and sex adjusted	1·00	0·94 (0·87-1·02)	0·89 (0·81-0·97)	0·87 (0·77-0·98)	0·79 (0·70-0·90)	0·0002
Multivariable adjusted	1·00	0·98 (0·90-1·07)	0·96 (0·87-1·06)	0·96 (0·85-1·09)	0·89 (0·77-1·01)	0·1100
MI						
MI events (N=2,143)	766 (2·02)	507 (1·59)	439 (1·34)	203 (1·38)	228 (1·27)	
Age and sex adjusted	1·00	1·02 (0·91-1·15)	0·93 (0·81-1·06)	1·00 (0·84-1·20)	0·81 (0·67-0·97)	<0·0001
Multivariable adjusted	1·00	1·05 (0·92-1·19)	1·00 (0·86-1·15)	1·10 (0·91-1·33)	0·90 (0·73-1·10)	0·5952
Stroke						
Stroke events (N=2,234)	759 (2·01)	571 (1·79)	499 (1·52)	174 (1·18)	231 (1·28)	
Age and sex adjusted	1·00	0·87 (0·78-0·98)	0·87 (0·77-0·98)	0·80 (0·67-0·96)	0·86 (0·72-1·02)	0·0196
Multivariable adjusted	1·00	0·93 (0·82-1·04)	0·94 (0·82-1·04)	0·87 (0·72-1·05)	0·95 (0·78-1·15)	0·3506
CV mortality						
CV death events (N=1,649)	703 (1·87)	385 (1·21)	299 (0·91)	110 (0·75)	152 (0·84)	
Age and sex adjusted	1·00	0·95 (0·84-1·09)	0·78 (0·67-0·90)	0·71 (0·57-0·89)	0·69 (0·56-0·86)	<0·0001
Multivariable adjusted	1·00	1·03 (0·89-1·19)	0·86 (0·72-1·01)	0·85 (0·67-1·08)	0·83 (0·65-1·06)	0·0458
Non-CV mortality						
Non-CV death events (N=3,809)	1,769 (4·67)	782 (2·45)	629 (1·91)	264 (1·80)	365 (2·03)	
Age and sex	1·00	0·76 (0·70-0·83)	0·63 (0·57-0·70)	0·61 (0·53-0·71)	0·58 (0·50-0·67)	<0·0001

adjusted						
Multivariable adjusted	1·00	0·87 (0·79-0·96)	0·78 (0·69-0·87)	0·81 (0·69-0·95)	0·82 (0·70-0·97)	0·0008
Mortality						
Mortality events (N=5,796)	2,626 (6·94)	1,241 (3·89)	982 (2·99)	400 (2·72)	547 (3·04)	
Age and sex adjusted	1·00	0·82 (0·76-0·88)	0·79 (0·72-0·87)	0·82 (0·72-0·93)	0·65 (0·57-0·72)	<0·0001
Multivariable adjusted	1·00	0·91 (0·85-0·99)	0·79 (0·72-0·87)	0·82 (0·72-0·93)	0·81 (0·72-0·93)	<0·0001

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. MI: myocardial infarction; CV mortality: cardiovascular mortality; Non-CV mortality: non-cardiovascular mortality. Crude event rates shown. Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals and vegetable intake. Additional sensitivity analyses with waist-to-hip ratio, hypertension status and statin medication used in the model did not materially alter estimates of association (Shown in the Supplementary Appendix).

Table S7: Association of vegetable intake with cardiovascular outcomes and mortality.

All participants (N=135,335)	**<1/day (N=34,221)**	**1 to <2/day (N=32,976)**	**2 to <3/day (N=46,489)**	**>3/day (N=21,649)**	**P-trend**
Median vegetable intake, servings	0·55 (0·31-0·78) per day	1·53 (1·25-1·70) per day	2·06 (2·01-2·23) per day	4·13 (3·48-5·31) per day	
Major CVD					
Major CVD events (N=4,784)	1,304 (3·81)	1,172 (3·55)	1,763 (3·79)	545 (2·52)	
Age and sex adjusted	1·00	0·98 (0·90-1·07)	1·07 (0·98-1·17)	0·89 (0·80-1·00)	0·5992
Multivariable adjusted	1·00	1·01 (0·92-1·11)	1·13 (1·02-1·24)	0·96 (0·85-1·09)	0·3763
MI					
MI events (N=2,143)	661 (1·93)	610 (1·85)	595 (1·28)	277 (1·28)	
Age and sex adjusted	1·00	1·04 (0·92-1·16)	1·08 (0·95-1·24)	0·89 (0·76-1·04)	0·5031
Multivariable adjusted	1·00	1·06 (0·93-1·20)	1·13 (0·98-1·30)	0·93 (0·78-1·11)	0·9849
Stroke					
Stroke events (N=2,234)	497 (1·45)	447 (1·36)	1,074 (2·31)	216 (1·00)	
Age and sex adjusted	1·00	0·97 (0·85-1·11)	1·10 (0·96-1·26)	1·01 (0·84-1·23)	0·2795
Multivariable adjusted	1·00	0·99 (0·85-1·14)	1·13 (0·98-1·31)	1·09 (0·89-1·34)	0·0783
CV death					
CV death events (N=1,649)	680 (1·99)	440 (1·33)	364 (0·78)	165 (0·76)	
Age and sex adjusted	1·00	0·86 (0·75-0·98)	0·89 (0·76-1·04)	0·79 (0·65-0·96)	0·0130
Multivariable adjusted	1·00	0·89 (0·77-1·02)	0·98 (0·83-1·16)	0·88 (0·71-1·10)	0·3979
Non-CV death					
Non-CV death events (N=3,809)	1,634 (4·77)	939 (2·85)	863 (1·86)	373 (1·72)	
Age and sex adjusted	1·00	0·81 (0·74-0·88)	0·73 (0·67-0·81)	0·71 (0·62-0·81)	<0·0001
Multivariable adjusted	1·00	0·92 (0·84-1·01)	0·90 (0·80-1·01)	0·96 (0·83-1·11)	0·2184
Mortality					
Mortality events (N=5,796)	2,462 (7·19)	1,457 (4·42)	1,303 (2·80)	574 (2·65)	
Age and sex adjusted	1·00	0·82 (0·76-0·87)	0·78 (0·71-0·85)	0·73 (0·66-0·81)	<0·0001
Multivariable adjusted	1·00	0·90 (0·83-0·97)	0·93 (0·84-1·01)	0·93 (0·83-1·05)	0·1216

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. MI: myocardial infarction; CV mortality: cardiovascular mortality; Non-CV mortality: non-cardiovascular mortality. Crude event rates shown. Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals and fruit intake. Additional sensitivity analyses with waist-to-hip ratio, hypertension status and statin medication used in the model did not materially alter estimates of association (Shown in the Supplementary Appendix).

Table S8: Association of legume intake with cardiovascular outcomes and mortality.

All participants (N=135,335)	**<1/month (N=37,849)**	**1/week to <1/week (N=24,487)**	**1 to <3/week (N=48,426)**	**3/week to <1/day (N=29,194)**	**>1/day (N=13,060)**	**P-trend**
Median legume intake, servings	0·00 (0·00-0·009) per day	0·087 (0·065-0·11) per day	0·26 (0·19-0·33) per day	0·61 (0·52-0·75) per day	1·42 (1·14-1·59) per day	
Major CVD						
Major CVD events (N=4,784)	645 (3·99)	1,022 (3·59)	1,622 (3·35)	1,071 (3·67)	424 (3·25)	
Age and sex adjusted	1·00	0·89 (0·80-0·99)	0·84 (0·76-0·94)	0·89 (0·79-1·00)	0·86 (0·74-0·99)	0·0612
Multivariable adjusted	1·00	0·88 (0·79-0·98)	0·86 (0·77-0·96)	0·93 (0·82-1·05)	0·83 (0·71-0·98)	0·1845
MI						
MI events (N=2,143)	216 (1·34)	402 (1·41)	713 (1·47)	572 (1·96)	240 (1·84)	
Age and sex adjusted	1·00	0·98 (0·82-1·18)	0·88 (0·74-1·05)	0·92 (0·76-1·05)	0·89 (0·71-1·11)	0·2299
Multivariable adjusted	1·00	0·99 (0·92-1·19)	0·92 (0·77-1·11)	0·95 (0·78-1·16)	0·86 (0·67-1·10)	0·2617
Stroke						
Stroke events (N=2,234)	353 (2·18)	538 (1·89)	764 (1·58)	414 (1·42)	165 (1·26)	
Age and sex adjusted	1·00	0·84 (0·73-0·97)	0·83 (0·72-0·96)	0·91 (0·77-1·07)	0·96 (0·77-1·20)	0·6272
Multivariable adjusted	1·00	0·81 (0·70-0·94)	0·81 (0·69-0·94)	0·93 (0·78-1·16)	0·88 (0·69-1·12)	0·6262
CV death						
CV death events (N=1,649)	215 (1·33)	278 (0·98)	499 (1·03)	436 (1·49)	221 (1·69)	
Age and sex adjusted	1·00	0·91 (0·74-1·10)	0·78 (0·64-0·94)	0·74 (0·60-0·92)	0·72 (0·57-0·92)	0·0032
Multivariable adjusted	1·00	0·90 (0·72-1·11)	0·84 (0·68-1·04)	0·89 (0·71-1·14)	0·87 (0·66-1·14)	0·4414
Non-CV death						
Non-CV death events (N=3,809)	520 (3·22)	656 (2·30)	1,186 (2·45)	965 (3·31)	482 (3·69)	
Age and sex	1·00	0·76 (0·70-0·83)	0·63 (0·57-0·70)	0·61 (0·53-0·71)	0·58 (0·50-0·67)	<0·0001

adjusted						
Multivariable adjusted	1·00	0·87 (0·79-0·96)	0·78 (0·69-0·87)	0·81 (0·69-0·95)	0·82 (0·70-0·97)	0·0019
Mortality						
Mortality events (N=5,796)	776 (4·80)	985 (3·46)	1,778 (3·67)	1,503 (5·15)	754 (5·77)	
Age and sex adjusted	1·00	0·76 (0·69-0·85)	0·70 (0·63-0·77)	0·65 (0·58-0·72)	0·59 (0·51-0·67)	<0·0001
Multivariable adjusted	1·00	0·80 (0·72-0·89)	0·80 (0·71-0·89)	0·80 (0·71-0·90)	0·74 (0·64-0·86)	0·0013

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. MI: myocardial infarction; CV death: cardiovascular mortality; Non-CV death: non-cardiovascular mortality. Crude event rates shown. Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals intake. Additional sensitivity analyses with waist-to-hip ratio, hypertension status and statin medication used in the model did not materially alter estimates of association (Shown in the Supplementary Appendix).

Table S9: Association of total fruit, vegetable and legume intake with total mortality by region using region-specific tertiles.

All participants (N=135,335)	**T1**	**T2**	**T3**
South Asia (N=29,560)			
N	9,755	9,755	10,050
Mortality events (N=2,527)	1,130 (11·58)	832 (8·53)	565 (5·62)
Median fruit, vegetable and legume intake, servings	1·21 (0·88-1·48) per day	2·30 (2·00-2·65) per day	4·37 (3·57-5·77) per day
Age and sex adjusted	1·00	0·81 (0·74-0·89)	0·60 (0·54-0·67)
Multivariable adjusted	1·00	0·84 (0·76-0·93)	0·68 (0·60-0·78)
China (N=42,152)			
N	13,911	13,909	14,332
Mortality events (N=1,156)	535 (3·85)	562 (2·60)	259 (1·81)
Median fruit, vegetable and legume intake, servings	2·13 (1·46-2·35) per day	2·81 (2·65-2·98) per day	3·88 (3·48-4·59) per day
Age and sex adjusted	1·00	0·77 (0·67-0·89)	0·56 (0·48-0·66)
Multivariable adjusted	1·00	0·95 (0·82-1·09)	0·83 (0·69-0·98)
Southeast Asia (N=10,038)			
N	3,315	3,310	3,413
Mortality events (N=327)	99 (2·99)	114 (3·44)	114 (3·34)
Median fruit, vegetable and legume intake, servings	1·27 (0·77-1·68) per day	2·68 (2·28-3·27) per day	5·38 (4·37-7·72) per day
Age and sex adjusted	1·00	1·22 (0·93-1·60)	1·13 (0·86-1·48)
Multivariable adjusted	1·00	1·35 (1·00-1·82)	1·24 (0·89-1·73)
Africa (N=4,558)			
N	1,504	1,505	1549
Mortality events (N=451)	154 (10·24)	124 (8·24)	173 (11·17)
Median fruit, vegetable and legume intake, servings	0·66 (0·43-0·90) per day	2·12 (1·69-2·59) per day	5·38 (4·06-8·65) per day
Age and sex adjusted	1·00	0·94 (0·64-1·27)	0·83 (0·60-1·16)
Multivariable adjusted	1·00	1·30 (0·72-2·35)	1·42 (0·76-2·65)
North America/Europe (N=14,916)			
N	4,923	4,922	5,071
Mortality events (n=328)	129 (2·62)	116 (2·36)	83 (1·64)
Median fruit, vegetable and legume intake, servings	2·76 (2·05-3·35) per day	5·00 (4·43-5·65) per day	8·35 (7·25-10·23) per day
Age and sex adjusted	1·00	0·94 (0·64-1·27)	0·83 (0·60-1·16)
Multivariable adjusted	1·00	1·10 (0·84-1·44)	0·82 (0·59-1·14)
Middle East (N=11,485)			
N	3,790	3,791	3,904
Mortality events (N=208)	74 (1·95)	62 (1·64)	72 (1·84)
Median fruit, vegetable and legume intake,	3·15 (2·49-3·63) per day	5·18 (4·62-5·78) per day	8·62 (7·30-10·90) per day

servings			
Age and sex adjusted	1·00	0·90 (0·64-1·27)	0·83 (0·60-1·16)
Multivariable adjusted	1·00	0·90 (0·63-1·29)	0·78 (0·50-1·18)
South America (N=22,626)			
N	7,466	7,467	7,693
Mortality events (n=799)	335 (4·49)	267 (3·58)	197 (2·56)
Median fruit, vegetable and legume intake, servings	2·19 (1·52-2·76) per day	4·55 (3·91-5·21) per day	7·82 (6·74-9·71) per day
Age and sex adjusted	1·00	0·89 (0·75-1·05)	0·79 (0·65-0·96)
Multivariable adjusted	1·00	0·92 (0·77-1·09)	0·84 (0·67-1·04)

Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals intake. P for heterogeneity=0·0174.

Table S10: Association of fruit intake with total mortality by region using region-specific tertiles.

All participants (N=135,335)	**T1**	**T2**	**T3**
South Asia (N=29,560)			
N	9,756	9,753	10,051
Mortality events (N=2,527)	1,170 (11·99)	794 (8·14)	563 (5·60)
Median fruit intake, servings	0·07 (0·024-0·13) per day	0·36 (0·27-0·48) per day	1·22 (0·87-2·00) per day
Age and sex adjusted	1·00	0·78 (0·72-0·86)	0·61 (0·55-0·68)
Multivariable adjusted	1·00	0·88 (0·80-0·98)	0·80 (0·70-0·91)
China (N=42,152)			
N	13,911	13,907	14,334
Mortality events (N=1,156)	537 (3·86)	346 (2·49)	273 (1·90)
Median fruit intake, servings	0·23 (0·14-0·33) per day	0·67 (0·54-0·83) per day	1·65 (1·28-2·35) per day
Age and sex adjusted	1·00	0·75 (0·65-0·86)	0·57 (0·49-0·87)
Multivariable adjusted	1·00	0·88 (0·77-1·02)	0·78 (0·66-0·93)
Southeast Asia (N=10,038)			
N	3,313	3,341	3,384
Mortality events (N=327)	117 (3·53)	97 (2·90)	113 (3·34)
Median fruit intake, servings	0·37 (0·17-0·57) per day	1·15 (0·93-1·24) per day	2·28 (1·96-3·31) per day
Age and sex adjusted	1·00	0·81 (0·61-1·07)	0·95 (0·72-1·24)
Multivariable adjusted	1·00	0·83 (0·61-1·13)	1·02 (0·73-1·43)
Africa (N=4,558)			
N	1,504	1,504	1,550
Mortality events (N=451)	147 (9·77)	136 (9·04)	168 (10·84)
Median fruit intake, servings	0·12 (0·00-0·25) per day	0·89 (0·61-1·14) per day	3·18 (2·21-6·45) per day
Age and sex adjusted	1·00	1·15 (0·90-1·48)	0·92 (0·67-1·26)
Multivariable adjusted	1·00	1·64 (0·94-2·86)	1·14 (0·63-2·07)
North America/Europe (N=14,916)			
N	4,922	4,923	5,071
Mortality events (n=328)	117 (2·38)	116 (2·36)	95 (1·87)
Median fruit intake, servings	0·72 (0·44-0·96) per day	0·67 (1·43-1·94) per day	3·21 (2·64-4·07) per day
Age and sex adjusted	1·00	0·92 (0·71-1·19)	0·75 (0·57-0·99)
Multivariable adjusted	1·00	1·05 (0·80-1·38)	1·12 (0·85-1·47)
Middle East (N=11,485)			
N	3,790	3,791	3,904

Mortality events (N=208)	68 (1·79)	59 (1·56)	81 (2·07)
Median fruit intake, servings	1·36 (0·96-1·63) per day	2·68 (2·29-3·19) per day	5·78 (4·55-7·75) per day
Age and sex adjusted	1·00	0·82 (0·58-1·17)	0·87 (0·61-1·23)
Multivariable adjusted	1·00	0·86 (0·59-1·25)	0·88 (0·57-1·35)
South America (N=22,626)			
N	7,466	7,467	7,693
Mortality events (n=799)	283 (3·79)	264 (3·54)	252 (3·28)
Median fruit intake, servings	0·56 (0·29-0·77) per day	1·50 (1·23-1·79) per day	3·35 (2·63-4·59) per day
Age and sex adjusted	1·00	0·82 (0·69-0·97)	0·73 (0·61-0·87)
Multivariable adjusted	1·00	0·88 (0·73-1·05)	0·78 (0·64-0·95)

Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals and vegetable intake. P for heterogeneity=0·1514.

Table S11: Association of vegetable intake with total mortality by region using region-specific tertiles.

All participants (N=135,335)	**T1**	**T2**	**T3**
South Asia (N=29,560)			
N	9,754	9,748	10,058
Mortality events (N=2,527)	1,087 (11·14)	843 (8·65)	597 (5·94)
Median vegetable intake, servings	0·37 (0·24-0·51) per day	0·96 (0·80-1·17) per day	2·16 (1·74-2·91) per day
Age and sex adjusted	1·00	0·85 (0·77-0·94)	0·65 (0·58-0·72)
Multivariable adjusted	1·00	0·95 (0·86-1·06)	0·81 (0·71-0·93)
China (N=42,152)			
N	10,943	17,367	13,842
Mortality events (N=1,156)	383 (3·50)	484 (2·79)	289 (2·09)
Median vegetable intake, servings	1·20 (0·31-1·60) per day	2·00 (2·00-2·01) per day	2·10 (2·06-2·20) per day
Age and sex adjusted	1·00	0·98 (0·84-1·13)	0·77 (0·63-0·94)
Multivariable adjusted	1·00	1·20 (1·03-1·40)	1·08 (0·90-1·29)
Southeast Asia (N=10,038)			
N	3,315	3,311	3,412
Mortality events (N=327)	102 (3·08)	114 (3·44)	111 (3·25)
Median vegetable intake, servings	0·51 (0·27-0·69) per day	1·14 (0·91-1·50) per day	2·75 (1·98-4·09) per day
Age and sex adjusted	1·00	1·17 (0·89-1·13)	1·09 (0·67-0·94)
Multivariable adjusted	1·00	1·14 (0·85-1·54)	1·11 (0·80-1·55)
Africa (N=4,558)			
N	1,502	1,506	1,550
Mortality events (N=451)	155 (10·32)	152 (10·09)	144 (9·29)
Median vegetable intake, servings	0·37 (0·23-0·48) per day	0·88 (0·72-1·05) per day	1·93 (1·54-2·56) per day
Age and sex adjusted	1·00	1·03 (0·81-1·31)	0·99 (0·74-1·31)
Multivariable adjusted	1·00	1·09 (0·69-1·71)	1·06 (0·67-1·69)
North America/Europe (N=14,916)			
N	4,922	4,922	5,072
Mortality events (n=328)	140 (2·84)	107 (2·17)	81 (1·60)
Median vegetable intake, servings	1·42 (1·02-1·78) per day	2·87 (2·49-3·30) per day	5·21 (4·40-6·53) per day
Age and sex adjusted	1·00	0·85 (0·65-1·09)	0·62 (0·47-1·09)
Multivariable adjusted	1·00	0·94 (0·71-1·23)	0·72 (0·52-1·00)
Middle East (N=11,485)			
N	3,791	3,789	3,905

Mortality events (N=208)	79 (2·08)	69 (1·82)	60 (1·54)
Median vegetable intake, servings	0·96 (0·68-1·19) per day	1·79 (1·58-2·00) per day	2·85 (2·49-3·46) per day
Age and sex adjusted	1·00	1·02 (0·73-1·42)	0·97 (0·68-1·40)
Multivariable adjusted	1·00	1·03 (0·73-1·44)	1·01 (0·67-1·51)
South America (N=22,626)			
N	7,466	7,468	7,692
Mortality events (n=799)	346 (4·63)	275 (3·68)	178 (2·31)
Median vegetable intake, servings	1·02 (0·70-1·30) per day	2·28 (1·90-2·75) per day	4·33 (3·71-5·41) per day
Age and sex adjusted	1·00	0·98 (0·83-1·15)	0·83 (0·67-1·03)
Multivariable adjusted	1·00	1·03 (0·86-1·20)	0·96 (0·76-1·21)

Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals and fruit intake. P for heterogeneity=0·0434.

Table S12: Association of legume intake with total mortality by region using region-specific tertiles.

All participants (N=135,335)	**T1**	**T2**	**T3**
South Asia (N=29,560)			
N	9,755	9,752	10,053
Mortality events (N=2,527)	859 (8·81)	834 (8·55)	834 (8·30)
Median legume intake, servings	0·28 (0·17-0·36) per day	0·64 (0·54-0·74) per day	1·24 (1·03-1·62) per day
Age and sex adjusted	1·00	0·86 (0·78-0·95)	0·81 (0·73-0·89)
Multivariable adjusted	1·00	0·98 (0·88-1·09)	0·96 (0·85-1·08)
China (N=42,152)			
N	13,494	14,303	14,355
Mortality events (N=1,156)	481 (3·56)	328 (2·29)	347 (2·42)
Median legume intake, servings	0·033 (0·00-0·059) per day	0·15 (0·11-0·19) per day	0·38 (0·29-0·52) per day
Age and sex adjusted	1·00	0·75 (0·66-0·86)	0·66 (0·55-0·80)
Multivariable adjusted	1·00	0·83 (0·71-0·97)	0·88 (0·75-1·04)
Southeast Asia (N=10,038)			
N	3,267	3,367	3,3402
Mortality events (N=327)	120 (3·67)	84 (2·49)	123 (3·62)
Median legume intake, servings	0·00 (0·00-0·037) per day	0·22 (0·16-0·23) per day	0·50 (0·42-0·87) per day
Age and sex adjusted	1·00	0·68 (0·50-0·92)	0·91 (0·67-1·23)
Multivariable adjusted	1·00	0·73 (0·53-1·00)	0·86 (0·61-1·22)
Africa (N=4,558)			
N	1,910	2,648	N/A
Mortality events (N=451)	170 8·90)	281 (10·61)	
Median legume intake, servings	0·00 (0·00-0·00) per day	0·17 (0·090-0·35) per day	
Age and sex adjusted	1·00	0·90 (0·72-1·12)	
Multivariable adjusted	1·00	0·89 (0·60-1·32)	
North America/Europe (N=14,916)			
N	4,921	4,931	5,064
Mortality events (n=328)	124 (2·52)	95 (1·93)	109 (2·15)
Median legume intake, servings	0·059 (0·034-0·087) per day	0·18 (0·15-0·22) per day	0·43 (0·33-0·61) per day
Age and sex adjusted	1·00	0·69 (0·53-0·90)	0·78 (0·60-1·01)
Multivariable adjusted	1·00	0·75 (0·57-1·00)	0·93 (0·70-1·24)
Middle East (N=11,485)			
N	3,807	3,773	3,905

Mortality events (N=208)	58 (1·52)	78 (2·07)	72 (1·84)
Median legume intake, servings	0·16 (0·10-0·21) per day	0·35 (0·35-0·41) per day	0·68 (0·57-0·87) per day
Age and sex adjusted	1·00	1·36 (0·96-1·93)	1·19 (0·82-1·71)
Multivariable adjusted	1·00	1·40 (0·97-2·02)	1·20 (0·80-1·79)
South America (N=22,626)			
N	7,702	7,275	7,649
Mortality events (n=799)	341 (4·43)	278 (3·82)	180 (2·35)
Median legume intake, servings	0·038 (0·00-0·087) per day	0·19 (0·19-0·25) per day	0·66 (0·53-1·45) per day
Age and sex adjusted	1·00	1·09 (0·91-1·30)	1·00 (0·80-1·30)
Multivariable adjusted	1·00	1·13 (0·94-1·36)	1·09 (0·80-1·26)

Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals intake. P for heterogeneity=0·1794.

Table S13: Association of raw vegetable intake with total mortality by region using region-specific tertiles.

All participants (N=135,335)	**T1**	**T2**	**T3**
South Asia (N=29,560)			
N	15,949	13,611	N/A
Mortality events (N=2,527)	1,783 (11·18)	744 (5·47)	
Median raw vegetable intake, servings	0·00 (0·00-0·00) per day	0·10 (0·036-0·24) per day	
Age and sex adjusted	1·00	0·67 (0·61-0·74)	
Multivariable adjusted	1·00	0·88 (0·77-1·00)	
Southeast Asia (N=10,038)			
N	3,320	3,294	3,422
Mortality events (N=327)	107 (3·22)	114 (3·46)	106 (3·10)
Median raw vegetable intake, servings	0·051 (0·00-0·077) per day	0·23 (0·17-0·33) per day	0·77 (0·51-1·15) per day
Age and sex adjusted	1·00	1·09 (0·83-1·42)	0·98 (0·75-1·28)
Multivariable adjusted	1·00	1·01 (0·76-1·34)	0·95 (0·70-1·30)
Africa (N=4,558)			
N	2,846	1,712	N/A
Mortality events (N=451)	235 (8·26)	216 12·62)	
Median raw vegetable intake, servings	0·00 (0·00-0·00) per day	0·23 (0·069-0·50) per day	
Age and sex adjusted	1·00	0·95 (0·71-1·27)	
Multivariable adjusted	1·00	1·28 (0·71-2·31)	
North America/Europe (N=14,916)			
N	4,922	4,921	5,073
Mortality events (n=328)	143 (2·91)	98 (1·99)	87 (1·71)
Median raw vegetable intake, servings	0·84 (0·56-1·07) per day	1·87 (1·59-2·17) per day	3·47 (2·92-4·41) per day
Age and sex adjusted	1·00	0·75 (0·58-0·98)	0·67 (0·51-0·88)
Multivariable adjusted	1·00	0·85 (0·65-1·12)	0·81 (0·59-1·10)
Middle East (N=11,485)			
N	3,790	3789	3,906
Mortality events (N=208)	89(2·35)	54 (1·43)	65 (1·43)
Median raw vegetable intake, servings	0·31 (0·22-0·46) per day	0·84 (0·72-0·98) per day	1·53 (1·32-1·91) per day
Age and sex adjusted	1·00	0·66 (0·46-0·94)	0·73 (0·52-1·03)
Multivariable adjusted	1·00	0·70 (0·48-1·01)	0·79 (0·54-1·15)
South America (N=22,626)			
N	5,705	4,489	12,432
Mortality events (n=799)	279 (4·89)	182 (4·05)	338 (2·72)
Median raw vegetable	0·84 (0·56-	1·87 (1·59-	3·47 (2·92-4·41)

intake, servings	1·07) per day	2·17) per day	per day
Age and sex adjusted	1·00	0·99 (0·83-1·19)	0·84 (0·66-1·07)
Multivariable adjusted	1·00	1·05 (0·87-1·27)	1·01 (0·78-1·31)

Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals and fruit intake. P for heterogeneity=0·1055.

Table S14: Association of cooked vegetable intake with total mortality by region using region-specific tertiles.

All participants (N=135,335)	**T1**	**T2**	**T3**
South Asia (N=29,560)			
N	9,755	9,755	10,050
Mortality events (N=2,527)	1,043 (10·69)	854 (8·75)	630 (6·27)
Median cooked vegetable intake, servings	0·33 (0·22-0·46) per day	0·88 (0·73-1·05) per day	2·00 (1·59-2·70) per day
Age and sex adjusted	1·00	0·83 (0·76-0·92)	0·66 (0·59-0·73)
Multivariable adjusted	1·00	0·93 (0·84-1·04)	0·82 (0·71-0·94)
Southeast Asia (N=10,038)			
N	3,367	3,529	3,142
Mortality events (N=327)	109 (3·24)	98 (2·78)	120 (3·82)
Median cooked vegetable intake, servings	0·33 (0·17-0·46) per day	0·78 (0·62-0·97) per day	1·87 (1·24-2·82) per day
Age and sex adjusted	1·00	0·99 (0·75-1·31)	1·19 (0·93-1·54)
Multivariable adjusted	1·00	0·95 (0·70-1·30)	1·19 (0·87-1·62)
Africa (N=4,558)			
N	1,505	1,503	1,550
Mortality events (N=451)	154 (10·23)	137 (9·12)	160 (10·32)
Median cooked vegetable intake, servings	0·29 (0·17-0·37) per day	0·60 (0·52-0·72) per day	1·23 (1·00-1·65) per day
Age and sex adjusted	1·00	0·82 (0·65-1·03)	0·94 (0·74-1·20)
Multivariable adjusted	1·00	0·82 (0·56-1·21)	0·93 (0·64-1·36)
North America/Europe (N=14,916)			
N	4,922	4,922	5,072
Mortality events (n=328)	135 (2·74)	110 (2·23)	83 (1·64)
Median cooked vegetable intake, servings	0·32 (0·21-0·41) per day	0·69 (0·59-0·80) per day	1·38 (1·12-1·83) per day
Age and sex adjusted	1·00	0·82 (0·64-1·06)	0·65 (0·49-0·86)
Multivariable adjusted	1·00	0·90 (0·63-1·29)	0·76 (0·50-1·18)
Middle East (N=11,485)			
N	3,792	3,789	3,904
Mortality events (N=208)	85 (2·24)	56 (1·48)	67 (1·72)
Median cooked vegetable intake, servings	0·27 (0·16-0·40) per day	0·79 (0·65-0·893per day	1·45 (1·12-1·80) per day
Age and sex adjusted	1·00	0·97 (0·66-1·41)	1·16 (0·81-1·14)
Multivariable adjusted	1·00	0·95 (0·64-1·41)	1·16 (0·74-1·82)
South America (N=22,626)			
N	5,705	4,489	12,432
Mortality events (n=799)	279 (4·89)	182 (4·05)	338 (2·72)
Median cooked vegetable	0·32 (0·21-	0·69 (0·59-	1·38 (1·12-1·83)

intake, servings	0·41) per day	0·80) per day	per day
Age and sex adjusted	1·00	1·00 (0·81-1·14)	0·92 (0·77-1·11)
Multivariable adjusted	1·00	1·04 (0·87-1·24)	1·07 (0·88-1·31)

Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals and fruit intake. P for heterogeneity=0·0364.

Table S15: Association of total fruit, vegetable, and legume intake with total mortality by region using overall tertiles.

All participants (N=135,335)	**T1**	**T2**	**T3**
Fruit, vegetable and legume intake cutpoint, servings	<2·50 per day	2·50 to 4·09 per day	>4·09 per day
South Asia (N=29,560)			
N	16,091	7,690	5,779
Mortality events (N=2,527)	1713 (10·65)	520 (6·76)	294 (5·09)
Age and sex adjusted	1·00	0·71 (0·64-0·79)	0·60 (0·53-0·68)
Multivariable adjusted	1·00	0·78 (0·69-0·87)	0·70 (0·60-0·82)
China (N=42,152)			
N	13,662	22,706	5,784
Mortality events (N=1,156)	522 (3·82)	530 (2·33)	104 (1·80)
Age and sex adjusted	1·00	0·72 (0·63-0·82)	0·57 (0·46-0·72)
Multivariable adjusted	1·00	0·92 (0·80-1·06)	0·90 (0·71-1·15)
Southeast Asia (N=10,038)			
N	4,695	2,131	3,212
Mortality events (N=327)	141 (3·00)	78 (3·66)	108 (3·36)
Age and sex adjusted	1·00	1·30 (0·99-1·72)	1·12 (0·87-1·44)
Multivariable adjusted	1·00	1·22 (0·85-1·77)	1·26 (0·89-1·77)
Africa (N=4,558)			
N	2,555	856	1,147
Mortality events (N=451)	249 (9·75)	65 (7·59)	137 (11·94)
Age and sex adjusted	1·00	0·99 (0·73-1·34)	0·92 (0·68-1·26)
Multivariable adjusted	1·00	0·90 (0·53-1·53)	0·78 (0·43-1·40)
North America/Europe (N=14,916)			
N	1,969	3,382	9,565
Mortality events (n=328)	63 (3·20)	75 (2·22)	190 (1·99)
Age and sex adjusted	1·00	0·68 (0·49-0·95)	0·66 (0·49-0·88)
Multivariable adjusted	1·00	0·89 (0·62-1·27)	0·95 (0·68-1·33)
Middle East (N=11,485)			
N	960	2,777	7,748
Mortality events (N=208)	20 (2·08)	54 (1·94)	134 (1·73)
Age and sex adjusted	1·00	1·03 (0·61-1·72)	0·86 (0·54-1·39)
Multivariable adjusted	1·00	1·12 (0·65-1·95)	0·96 (0·55-1·67)
South America (N=22,626)			

N	4,726	5,119	12,778
Mortality events (n=799)	214 (4·53)	215 (4·20)	370 (2·90)
Age and sex adjusted	1·00	0·94 (0·77-1·13)	0·83 (0·69-1·00)
Multivariable adjusted	1·00	1·00 (0·82-1·22)	0·94 (0·76-1·15)

Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals intake.

Table S16: Association of total fruit intake with total mortality by region using overall tertiles.

All participants (N=135,335)	**T1**	**T2**	**T3**
Fruit intake cutpoint, servings	<0·53 per day	0·53 to 1·50 per day	>1·50 per day
South Asia (N=29,560)			
N	17,903	7,872	3,785
Mortality events (N=2,527)	1866 (10·42)	496 (6·30)	165 (4·36)
Age and sex adjusted	1·00	0·73 (0·66-0·81)	0·52 (0·44-0·62)
Multivariable adjusted	1·00	0·86 (0·77-0·97)	0·70 (0·58-0·85)
China (N=42,152)			
N	17,245	16,560	8,347
Mortality events (N=1,156)	622 (3·61)	384 (2·32)	150 (1·80)
Age and sex adjusted	1·00	0·73 (0·64-0·84)	0·58 (0·48-0·70)
Multivariable adjusted	1·00	0·91 (0·79-1·04)	0·79 (0·64-0·97)
Southeast Asia (N=10,038)			
N	2,292	4,241	3,505
Mortality events (N=327)	82 (3·58)	129 (3·04)	116 (3·31)
Age and sex adjusted	1·00	0·89 (0·67-1·19)	0·96 (0·71-1·29)
Multivariable adjusted	1·00	0·99 (0·73-1·34)	1·07 (0·75-1·52)
Africa (N=4,558)			
N	1,751	1,247	1,560
Mortality events (N=451)	168 (9·59)	115 (9·22)	168 (10·77)
Age and sex adjusted	1·00	1·18 (0·91-1·52)	0·91 (0·67-1·24)
Multivariable adjusted	1·00	1·11 (0·70-1·78)	0·82 (0·50-1·34)
North America/Europe (N=14,916)			
N	1,602	4,920	8,394
Mortality events (n=328)	38 (2·37)	121 (2·46)	169 (2·01)
Age and sex adjusted	1·00	0·90 (0·62-1·30)	0·74 (0·52-1·05)
Multivariable adjusted	1·00	1·06 (0·72-1·56)	1·05 (0·70-1·56)
Middle East (N=11,485)			
N	298	2,145	9,042
Mortality events (N=208)	9 (3·02)	42 (1·96)	157 (1·74)
Age and sex adjusted	1·00	0·86 (0·41-1·80)	0·66 (0·33-1·33)
Multivariable adjusted	1·00	0·98 (0·45-2·14)	0·76 (0·34-1·63)
South America (N=22,626)			
N	3,571	7,674	11,381

Mortality events (n=799)	140 (3·92)	269 (3·51)	390 (3·43)
Age and sex adjusted	1·00	0·89 (0·72-1·09)	0·75 (0·62-0·92)
Multivariable adjusted	1·00	0·92 (0·75-1·14)	0·84 (0·68-1·05)

Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals and vegetable intake.

Table S17: Association of vegetable intake with total mortality by region using overall tertiles.

All participants (N=135,335)	**T1**	**T2**	**T3**
Vegetable intake cutpoint, servings	<1·32 per day	1·32 to 2·06 per day	>2·06 per day
South Asia (N=29,560)			
N	18,574	5,381	5,605
Mortality events (N=2,527)	1865 (10·04)	358 (6·65)	304 (5·42)
Age and sex adjusted	1·00	0·75 (0·66-0·84)	0·67 (0·59-0·76)
Multivariable adjusted	1·00	0·85 (0·74-0·96)	0·82 (0·71-0·95)
China (N=42,152)			
N	6,293	25,463	10,396
Mortality events (N=1,156)	229 (3·64)	710 (2·79)	217 (2·09)
Age and sex adjusted	1·00	0·88 (0·75-1·04)	0·77 (0·63-0·94)
Multivariable adjusted	1·00	1·11 (0·94-1·32)	1·04 (0·84-1·29)
Southeast Asia (N=10,038)			
N	5,415	2,371	2,252
Mortality events (N=327)	164 (3·03)	84 (3·54)	79 (3·51)
Age and sex adjusted	1·00	1·21 (0·93-1·58)	1·16 (0·89-1·52)
Multivariable adjusted	1·00	1·23 (0·90-1·69)	1·18 (0·86-1·62)
Africa (N=4,558)			
N	3,123	753	682
Mortality events (N=451)	314 (10·05)	65 (8·63)	72 (10·56)
Age and sex adjusted	1·00	0·93 (0·69-1·24)	1·17 (0·88-1·56)
Multivariable adjusted	1·00	0·91 (0·63-1·30)	1·39 (0·96-2·02)
North America/Europe (N=14,916)			
N	2,130	2,575	10,211
Mortality events (n=328)	73 (3·43)	62 (2·41)	193 (1·89)
Age and sex adjusted	1·00	0·75 (0·54-1·05)	0·64 (0·48-0·85)
Multivariable adjusted	1·00	0·90 (0·63-1·28)	0·79 (0·58-1·09)
Middle East (N=11,485)			
N	3,420	3,486	4,579
Mortality events (N=208)	73 (2·13)	67 (1·49)	68 (1·49)
Age and sex adjusted	1·00	1·06 (0·75-1·49)	0·89 (0·62-1·28)
Multivariable adjusted	1·00	1·10 (0·78-1·56)	0·93 (0·63-1·39)
South America (N=22,626)			
N	5,705	4,489	12,432
Mortality events (n=799)	279 (4·89)	182 (4·05)	338 (2·72)

Age and sex adjusted	1·00	0·90 (0·74-1·09)	0·87 (0·73-1·05)
Multivariable adjusted	1·00	0·94 (0·77-1·14)	0·97 (0·79-1·18)

Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals and fruit intake.

Table S18: Association of legume intake with total mortality by region using overall tertiles.

All participants (N=135,335)	**T1**	**T2**	**T3**
Legume intake cutpoint, servings	<0·14 per day	0·14 to 0·40 per day	>0·40 per day
South Asia (N=29,560)			
N	2,127	6,100	21,333
Mortality events (N=2,527)	167 (7·85)	562 (9·21)	1,798 (8·43)
Age and sex adjusted	1·00	0·91 (0·76-1·10)	0·74 (0·562-0·88)
Multivariable adjusted	1·00	0·96 (0·80-1·16)	0·88 (0·73-1·07)
China (N=42,152)			
N	20,348	15,396	6,408
Mortality events (N=1,156)	630 (3·10)	376 (2·44)	150 (2·34)
Age and sex adjusted	1·00	0·75 (0·66-0·86)	0·66 (0·55-0·80)
Multivariable adjusted	1·00	0·96 (0·83-1·11)	0·94 (0·77-1·15)
Southeast Asia (N=10,038)			
N	3,267	3,978	2,793
Mortality events (N=327)	120 (3·67)	114 (2·87)	93 (3·33)
Age and sex adjusted	1·00	0·75 (0·56-1·00)	0·81 (0·59-1·12)
Multivariable adjusted	1·00	0·78 (0·57-1·05)	0·79 (0·56-1·10)
Africa (N=4,558)			
N	3,072	957	529
Mortality events (N=451)	309 (10·06)	99 (10·34)	43 (8·13)
Age and sex adjusted	1·00	0·91 (0·72-1·15)	0·68 (0·49-0·95)
Multivariable adjusted	1·00	1·03 (0·75-1·40)	0·63 (0·41-0·98)
North America/Europe (N=14,916)			
N	6,040	6,032	2,844
Mortality events (n=328)	146 (2·42)	129 (2·14)	53 (1·86)
Age and sex adjusted	1·00	0·78 (0·62-0·99)	0·75 (0·55-1·03)
Multivariable adjusted	1·00	0·84 (0·65-1·08)	0·91 (0·65-1·28)
Middle East (N=11,485)			
N	1,551	4,834	5,100
Mortality events (N=208)	29 (1·87)	81 (1·68)	98 (1·92)
Age and sex adjusted	1·00	0·81 (0·53-1·25)	0·94 (0·61-1·45)
Multivariable adjusted	1·00	1·87 (0·56-1·36)	0·96 (0·60-1·55)
South America (N=22,626)			
N	8,251	7,368	7,007
Mortality events (n=799)	360 (4·36)	279 (3·79)	160 (2·28)

Age and sex adjusted	1·00	1·07 (0·90-1·27)	0·97 (0·77-1·21)
Multivariable adjusted	1·00	1·11 (0·93-1·33)	1·08 (0·85-1·36)

Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals intake.

Supplementary Methods: Association of raw and cooked vegetables and major CVD and mortality including China.

We conducted a sensitivity analysis on raw and cooked vegetable intake and major CVD and mortality with China included whereby we assumed that total vegetable intake was consumed as cooked vegetables. This sensitivity analysis includes the 42,152 participants from China.

Table S19A: Association of raw vegetable intake with major cardiovascular disease and mortality.

All participants (N=93,183	**<1/month (N=66,206)**	**1/month to <1/week (N=9,086)**	**1 to <3/week (N=15,410)**	**3/week to <1/day (N=16,263)**	**1 to 2/day (N=15,818)**	**>2/day (N=12,552)**	**P-trend**
Median raw vegetable intake, servings	0·00 (0·00-0·00) per day	0·080 (0·051-0·11) per day	0·25 (0·20-0·33) per day	0·67 (0·52-0·82) per day	1·42 (1·19-1·66) per day	2·82 (2·35-3·61) per day	
Major CVD							
Major CVD events (N=4,784)	2,846 (4·30)	313 (3·44)	459 (2·98)	477 (2·93)	422 (2·67)	267 (2·13)	
Multivariable adjusted	1·00	0·99 (0·85-1·16)	0·89 (0·76-1·04)	0·96 (0·81-1·14)	0·96 (0·79-1·16)	0·82 (0·65-1·02)	0·1803
Mortality							
Mortality events (N=4,640)	2,329 (9·68)	480 (5·28)	670 (4·35)	536 (3·30)	396 (2·50)	229 (1·82)	
Multivariable adjusted	1·00	0·93 (0·82-1·05)	0·85 (0·74-0·97)	0·85 (0·73-0·99)	0·92 (0·77-1·09)	0·79 (0·64-0·98)	0·0451

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. Adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals, and fruit intake. Assumes total vegetable intake was consumed cooked in China (N=42,152) and does not exclude the participants from this region. P for heterogeneity=0.60 with cooked vegetable intake for major CVD and P for heterogeneity=0.35 with cooked vegetable intake for total mortality.

Table S19B: Association of cooked vegetable intake with major cardiovascular disease and mortality.

All participants (N=93,183	**<3/week (N=24,147)**	**3/week to <1/day (N=35,327)**	**1 to 2/day (N=31,567)**	**>2/day (N=44,294)**	**P-trend**
Median cooked vegetable intake, servings	0·24 (0·13-0·34) per day	0·68 (0·55-0·83) per day	1·40 (1·18-1·62) per day	2·06 (2·01-2·23) per day	
Major CVD					
Major CVD events (N=4,784)	833 (3·45)	1,146 (3·24)	1,069 (3·39)	1,736 (3·92)	
Multivariable adjusted	1·00	1·08 (0·97-1·19)	1·06 (0·96-1·19)	1·18 (1·05-1·32)	0·0855
Mortality					
Mortality events (N=4,640)	661 (1·93)	610 (1·85)	595 (1·28)	277 (1·28)	
Multivariable adjusted	1·00	1·02 (0·94-1·10)	0·90 (0·82-0·99)	0·96 (0·86-1·06)	0·1020

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. Adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals, and fruit intake. Assumes total vegetable intake was consumed cooked in China (N=42,152) and does not exclude the participants from this region. P for heterogeneity=0.60 with raw vegetable intake for major CVD and P for heterogeneity=0.35 with raw vegetable intake for total mortality.

Supplementary Methods: Association of fruits, vegetables and legumes adjusting for waist-to-hip ratio, hypertension status and cholesterol medication use.

Table S20: Association of total fruit, vegetable and legume intake with cardiovascular disease and mortality.

All participants (N=135,335)	**<1/day (N=9,082)**	**1 to <2/day (N=19,036)**	**2 to <3/day (N=35,128)**	**3 to <4/day (N=24,485)**	**4 to <5/day (N=14,849)**	**5 to <6/day (N=9,790)**	**6 to 7/day (N=6,945)**	**7 to 8/day (N=4,857)**	**>8/day (N=11,163)**	**P-trend**
Median fruit, vegetable and legume intake, servings	0·64 (0·41-0·83) per day	1·56 (1·30-1·79) per day	2·52 (2·29-2·75) per day	3·43 (3·21-3·70) per day	4·43 (4·19-4·70) per day	5·46 (5·22-5·72) per day	6·45 (6·22-6·71) per day	7·46 (7·22-7·71) per day	9·99 (8·82-12·15) per day	
Major CVD										
Major CVD events (N=4,784)	373 (4·11)	743 (3·90)	1391 (3·96)	882 (3·60)	534 (3·60)	267 (2·73)	180 (2·459	131 (2·70)	283(2·54)	
Multivariable adjusted	1·00	1·03 (0·89-1·20)	1·07 (0·93-1·24)	1·04 (0·90-1·21)	1·19 (1·01-1·40)	0·94 (0·78-1·14)	0·93 (0·75-1·15)	0·97 (0·79-1·26)	0·90 (0·73-1·10)	0·1543
MI										
MI events (N=2,143)	164 (1·81)	391 (2·05)	565 (1·61)	375 (1·53)	254 (1·71)	114 (1·16)	77 (1·11)	60 (1·24)	143 (1·28)	
Multivariable adjusted	1·00	1·04 (0·84-1·28)	1·14 (0·93-1·40)	1·15 (0·92-1·43)	1·27 (1·00-1·61)	0·86 (0·65-1·14)	0·84 (0·62-1·15)	0·96 (0·68-1·35)	0·98 (0·73-1·31)	0·2440
Stroke										
Stroke events (N=2,234)	157 (1·73)	283 (1·49)	727 (2·07)	463 (1·89)	233 (1·57)	127 (1·30)	81 (1·17)	59 (1·21)	104 (0·93)	
Multivariable adjusted	1·00	1·04 (0·83-1·30)	1·03 (0·83-1·27)	1·02 (0·81-1·27)	1·16 (0·91-1·48)	1·10 (0·83-1·46)	1·11 (0·81-1·52)	1·22 (0·85-1·73)	0·92 (0·67-1·28)	0·6652
CV mortality										
CV mortality events (N=1,649)	215 (2·37)	361 (1·90)	418 (1·19)	245 (1·00)	161 (1·08)	68 (0·69)	57 (0·82)	43 (0·89)	81 (0·73)	
Multivariable adjusted	1·00	0·84 (0·68-1·04)	0·85 (0·69-1·05)	0·78 (0·62-0·99)	0·90 (0·69-1·17)	0·58 (0·42-0·81)	0·82 (0·58-1·16)	0·87 (0·58-1·29)	0·67 (0·46-0·95)	0·0565
Non-CV mortality										
Non-CV mortality events (N=3,809)	486 (5·35)	918 (4·82)	1,023 (2·91)	485 (1·98)	284 (1·91)	199 (2·03)	130 (2·03)	80 (1·65)	204 (1·83)	
Multivariable adjusted	1·00	1·05 (0·92-1·19)	0·90 (0·79-1·03)	0·74 (0·63-0·92)	0·77 (0·64-0·92)	0·86 (0·70-1·05)	0·81 (0·64-1·03)	0·78 (0·59-1·03)	0·79 (0·63-0·99)	0·0007
Mortality										
Mortality events	736 (810)	1,371 (7·20)	1,529 (4·35)	772 (3·15)	468 (3·15)	286 (2·92)	198 (2·85)	131 (2·70)	305 (2·73)	

(N=5,796)										
Multivariable adjusted	1·00	1·01 (0·90-1·12)	0·90 (0·80-1·00)	0·75 (0·66-0·85)	0·81 (0·71-0·94)	0·78 (0·66-0·93)	0·82 (0·68-0·99)	0·81 (0·65-1·01)	0·77 (0·64-0·92)	0·0001

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. MI: myocardial infarction; CV mortality: cardiovascular mortality; non-CV mortality: non-cardiovascular mortality. Crude event rates shown. Adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals, waist-to-hip ratio, hypertension status, and cholesterol medication use.

Table S21: Association of total fruit and vegetable intake with cardiovascular disease and mortality.

All participants (N=135,335)	**<1/day (N=16,512)**	**1 to <2/day (N=20,478)**	**2 to <3/day (N=36,036)**	**3 to <4/day (N=22,134)**	**4 to <5/day (N=12,604)**	**5 to <6/day (N=8,575)**	**6 to 7/day (N=5,841)**	**7 to 8/day (N=4,068)**	**>8/day (N=9,078)**	**P-trend**
Median fruit and vegetable intake, servings	0·61 (0·38-0·81) per day	1·53 (1·27-1·77) per day	2·47 (2·25-2·71) per day	3·43 (3·20-3·70) per day	4·44 (4·20-4·70) per day	5·45 (5·22-5·72) per day	6·45 (6·22-6·71) per day	7·45 (7·22-7·71) per day	9·98 (8·82-12·12) per day	
Major CVD										
Major CVD events (N=4,784)	712 (4·31)	746 (3·64)	1,414 (3·92)	782 (3·53)	409 (3·25)	234 (2·73)	153 (2·62)	97 (2·38)	237 (2·61)	
Multivariable adjusted	1·00	0·95 (0·80-1·12)	1·10 (0·93-1·30)	1·12 (0·92-1·35)	1·05 (0·84-1·31)	0·87 (0·69-1·13)	0·82 (0·60-1·12)	0·91 (0·65-1·29)	0·93 (0·70-1·24)	0·1304
MI										
MI events (N=2,143)	361 (2·19)	396 (1·93)	541 (1·50)	329 (1·49)	186 (1·48)	104 (1·21)	64 (1·10)	47 (1·16)	115 (1·27)	
Multivariable adjusted	1·00	0·96 (0·82-1·13)	1·10 (0·94-1·30)	1·13 (0·94-1·36)	1·07 (0·87-1·32)	0·86 (0·67-1·11)	0·80 (0·59-1·08)	0·87 (0·62-1·22)	0·95 (0·72-1·25)	0·3089
Stroke										
Stroke events (N=2,234)	283 (1·71)	274 (1·34)	785 (2·18)	406 (1·83)	176 (1·40)	110 (1·28)	70 (1·20)	39 (0·96)	91 (1·00)	
Multivariable adjusted	1·00	0·83 (0·68-1·01)	0·93 (0·78-1·10)	0·97 (0·79-1·18)	0·99 (0·75-1·28)	0·98 (0·75-1·29)	1·03 (0·76-1·40)	0·78 (0·53-1·17)	0·89 (0·65-1·21)	0·8972
CV mortality										
CV death events (N=1,649)	409 (2·48)	350 (1·71)	367 (1·02)	195 (0·88)	123 (0·98)	67 (0·78)	49 (0·84)	23 (0·57)	66 (0·73)	
Multivariable adjusted	1·00	0·83 (0·70-1·00)	0·81 (0·67-0·98)	0·77 (0·62-0·97)	0·89 (0·69-1·15)	0·75 (0·55-1·02)	0·91 (0·64-1·29)	0·56 (0·34-0·92)	0·69 (0·48-1·00)	0·0567
Non-CV mortality										
Non-CV death events (N=3,809)	1,044 (6·32)	772 (3·77)	875 (2·43)	384 (1·73)	236 (1·87)	157 (1·83)	100 (1·71)	66 (1·62)	175 (1·93)	
Multivariable adjusted	1·00	0·90 (0·80-1·00)	0·81 (0·72-0·92)	0·67 (0·58-0·78)	0·79 (0·66-0·94)	0·77 (0·63-0·95)	0·74 (0·57-0·94)	0·71 (0·53-0·95)	0·78 (0·62-0·99)	0·0030
Mortality										
Mortality events (N=5,796)	1,547 (9·37)	1,195 (5·84)	1,313 (3·64)	614 (2·77)	380 (3·01)	235 (2·74)	159 (2·72)	96 (2·36)	257 (2·83)	
Multivariable adjusted	1·00	0·87 (0·80-0·96)	0·81 (0·73-0·89)	0·70 (0·62-0·79)	0·81 (0·71-0·94)	0·75 (0·63-0·88)	0·78 (0·64-0·95)	0·67 (0·52-0·85)	0·76 (0·63-0·91)	0·0002

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. MI: myocardial infarction; CV mortality: cardiovascular mortality; non-CV mortality: non-cardiovascular mortality. Crude event rates shown. Adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals, waist-to-hip ratio, hypertension status, and cholesterol medication use.

Table S22: Association of total fruit intake with cardiovascular disease and mortality.

All participants (N=135,335)	**<3/week (N=37,849)**	**3/week to <1/day (N=31,929)**	**1 to <2/day (N=32,850)**	**2 to <3/day (N=14,706)**	**>3/day (N=18,001)**	**P-trend**
Median fruit intake, servings	0·20 (0·08-0·31) per day	0·68 (0·55-0·83) per day	1·41 (1·19-1·68) per day	2·41 (2·19-2·66) per day	4·28 (3·52-5·81) per day	
Major CVD						
Major CVD events (N=4,784)	1,646 (4·35)	1,160 (3·63)	1,034 (3·15)	419 (2·85)	525 (2·92)	
Multivariable adjusted	1·00	0·98 (0·90-1·07)	0·95 (0·86-1·04)	0·98 (0·86-1·12)	0·89 (0·77-1·02)	0·1384
MI						
MI events (N=2,143)	766 (2·02)	507 (1·59)	439 (1·34)	203 (1·38)	228 (1·27)	
Multivariable adjusted	1·00	1·06 (0·93-1·21)	0·98 (0·85-1·15)	1·13 (0·93-1·37)	0·91 (0·74-1·13)	0·7148
Stroke						
Stroke events (N=2,234)	759 (2·01)	571 (1·79)	499 (1·52)	174 (1·18)	231 (1·28)	
Multivariable adjusted	1·00	0·90 (0·80-1·02)	0·93 (0·81-1·06)	0·89 (0·74-1·08)	0·93 (0·76-1·14)	0·3568
CV mortality						
CV death events (N=1,649)	703 (1·87)	385 (1·21)	299 (0·91)	110 (0·75)	152 (0·84)	
Multivariable adjusted	1·00	1·02 (0·88-1·19)	0·85 (0·71-1·02)	0·89 (0·69-1·15)	0·84 (0·65-1·08)	0·0841
Non-CV mortality						
Non-CV death events (N=3,809)	1,769 (4·67)	782 (2·45)	629 (1·91)	264 (1·80)	365 (2·03)	
Multivariable adjusted	1·00	0·86 (0·78-0·96)	0·76 (0·68-0·86)	0·78 (0·67-0·92)	0·82 (0·70-0·97)	0·0005
Mortality						
Mortality events (N=5,796)	2,626 (6·94)	1,241 (3·89)	982 (2·99)	400 (2·72)	547 (3·04)	
Multivariable adjusted	1·00	0·91 (0·84-0·99)	0·78 (0·71-0·86)	0·82 (0·72-0·93)	0·82 (0·71-0·94)	<0·0001

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. MI: myocardial infarction; CV mortality: cardiovascular mortality; non-CV mortality: non-cardiovascular mortality. Crude event rates shown. Adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals, vegetable intake, waist-to-hip ratio, hypertension status, and cholesterol medication use.

Table S23: Association of total vegetable intake with cardiovascular disease and mortality.

All participants (N=135,335)	**<1/day (N=34,221)**	**1 to <2/day (N=32,976)**	**2 to <3/day (N=46,489)**	**>3/day (N=21,649)**	**P-trend**
Median vegetable intake, servings	0·55 (0·31-0·78) per day	1·53 (1·25-1·70) per day	2·06 (2·01-2·23) per day	4·13 (3·48-5·31) per day	
Major CVD					
Major CVD events (N=4,784)	1,304 (3·81)	1,172 (3·55)	1,763 (3·79)	545 (2·52)	
Multivariable adjusted	1·00	0·99 (0·90-1·09)	1·11 (1·00-1·22)	0·95 (0·83-1·08)	0·5803
MI					
MI events (N=2,143)	661 (1·93)	610 (1·85)	595 (1·28)	277 (1·28)	
Multivariable adjusted	1·00	1·06 (0·92-1·21)	1·10 (0·95-1·28)	0·91 (0·76-1·09)	0·7114
Stroke					
Stroke events (N=2,234)	497 (1·45)	447 (1·36)	1,074 (2·31)	216 (1·00)	
Multivariable adjusted	1·00	0·97 (0·83-1·12)	1·12 (0·97-1·30)	1·09 (0·88-1·35)	0·0853
CV mortality					
CV death events (N=1,649)	680 (1·99)	440 (1·33)	364 (0·78)	165 (0·76)	
Multivariable adjusted	1·00	0·88 (0·76-1·02)	0·96 (0·80-1·14)	0·87 (0·69-1·08)	0·2898
Non-CV mortality					
Non-CV death events (N=3,809)	1,634 (4·77)	939 (2·85)	863 (1·86)	373 (1·72)	
Multivariable adjusted	1·00	0·90 (0·81-0·99)	0·89 (0·79-1·00)	0·94 (0·80-1·18)	0·1414
Mortality					
Mortality events (N=5,796)	2,462 (7·19)	1,457 (4·42)	1,303 (2·80)	574 (2·65)	
Multivariable adjusted	1·00	0·88 (0·81-0·96)	0·91 (0·83-1·00)	0·91 (0·81-1·03)	0·0711

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. MI: myocardial infarction; CV mortality: cardiovascular mortality; non-CV mortality: non-cardiovascular mortality. Crude event rates shown. Adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals, fruit intake, waist-to-hip ratio, hypertension status, and cholesterol medication use.

Table S24: Association of legume intake with cardiovascular disease and mortality.

All participants (N=135,335)	**<1/month (N=37,849)**	**1/week to <1/week (N=24,487)**	**1 to <3/week (N=48,426)**	**3/week to <1/day (N=29,194)**	**>1/day (N=13,060)**	**P-trend**
Median legume intake, servings	0·00 (0·00-0·009) per day	0·087 (0·065-0·11) per day	0·26 (0·19-0·33) per day	0·61 (0·52-0·75) per day	1·42 (1·14-1·59) per day	
Major CVD						
Major CVD events (N=4,784)	645 (3·99)	1,022 (3·59)	1,622 (3·35)	1,071 (3·67)	424 (3·25)	
Multivariable adjusted	1·00	0·88 (0·78-0·98)	0·84 (0·75-0·94)	0·90 (0·79-1·03)	0·81 (0·68-0·95)	0·0585
MI						
MI events (N=2,143)	216 (1·34)	402 (1·41)	713 (1·47)	572 (1·96)	240 (1·84)	
Multivariable adjusted	1·00	0·98 (0·81-1·19)	0·90 (0·75-1·09)	0·92 (0·75-1·14)	0·83 (0·64-1·07)	0·1470
Stroke						
Stroke events (N=2,234)	353 (2·18)	538 (1·89)	764 (1·58)	414 (1·42)	165 (1·26)	
Multivariable adjusted	1·00	0·80 (0·69-0·93)	0·79 (0·68-0·92)	0·90 (0·75-1·08)	0·85 (0·66-1·09)	0·3624
CV mortality						
CV death events (N=1,649)	215 (1·33)	278 (0·98)	499 (1·03)	436 (1·49)	221 (1·69)	
Multivariable adjusted	1·00	0·93 (0·74-1·17)	0·84 (0·67-1·05)	0·88 (0·69-1·13)	0·87 (0·65-1·17)	0·3844
Non-CV mortality						
Non-CV death events (N=3,809)	520 (3·22)	656 (2·30)	1,186 (2·45)	965 (3·31)	482 (3·69)	
Multivariable adjusted	1·00	0·76 (0·66-0·87)	0·80 (0·70-0·91)	0·78 (0·66-0·91)	0·70 (0·58-0·85)	0·0054
Mortality						
Mortality events (N=5,796)	776 (4·80)	985 (3·46)	1,778 (3·67)	1,503 (5·15)	754 (5·77)	
Multivariable adjusted	1·00	0·80 (0·71-0·90)	0·80 (0·71-0·90)	0·81 (0·71-0·902	0·75 (0·64-0·87)	0·0039

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. MI: myocardial infarction; CV mortality: cardiovascular mortality; non-CV mortality: non-cardiovascular mortality. Crude event rates shown. Adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals, waist-to-hip ratio, hypertension status, and cholesterol medication use.

Table S25: Association of raw vegetable intake with cardiovascular disease and mortality.

All participants (N=93,183	**<1/month (N=24,054)**	**1/month to <1/week (N=9,086)**	**1 to <3/week (N=15,410)**	**3/week to <1/day (N=16,263)**	**1 to 2/day (N=15,818)**	**>2/day (N=12,552)**	**P-trend**
Median raw vegetable intake, servings	0·00 (0·00-0·00) per day	0·080 (0·051-0·11) per day	0·25 (0·20-0·33) per day	0·67 (0·52-0·82) per day	1·42 (1·19-1·66) per day	2·82 (2·35-3·61) per day	
Major CVD							
Major CVD events (N=3,085)	1,147 (4·77)	313 (3·44)	459 (2·98)	477 (2·93)	422 (2·67)	267 (2·13)	
Multivariable adjusted	1·00	0·94 (0·79-1·11)	0·81 (0·68-0·96)	0·89 (0·74-1·07)	0·89 (0·73-1·09)	0·78 (0·62-0·99)	0·1091
Mortality							
Mortality events (N=4,640)	2,329 (9·68)	480 (5·28)	670 (4·35)	536 (3·30)	396 (2·50)	229 (1·82)	
Multivariable adjusted	1·00	0·83 (0·73-0·95)	0·74 (0·64-0·85)	0·72 (0·61-0·85)	0·76 (0·63-0·92)	0·66 (0·53-0·83)	0·0002

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. MI: myocardial infarction; CV mortality: cardiovascular mortality; non-CV mortality: non-cardiovascular mortality. Crude event rates shown. Adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals, fruit intake, waist-to-hip ratio, hypertension status, and cholesterol medication use.

Table S26: Association of cooked vegetable intake with cardiovascular disease and mortality.

All participants (N=93,183)	**<3/week (N=20,890)**	**3/week to <1/day (N=33,395)**	**1 to 2/day (N=25,813)**	**>2/day (N=13,085)**	**P-trend**
Median cooked vegetable intake, servings	0·26 (0·16-0·34) per day	0·68 (0·55-0·83) per day	1·34 (1·15-1·61) per day	2·66 (2·27-3·33) per day	
Major CVD					
Major CVD events (N=3,085)	685 (3·28)	1,083 (3·24)	811 (3·14)	506 (3·87)	
Multivariable adjusted	1·00	1·08 (0·96-1·21)	1·04 (0·91-1·19)	1·13 (0·97-1·32)	0·2366
Mortality					
Mortality events (N=4,640)	661 (1·93)	610 (1·85)	595 (1·28)	277 (1·28)	
Multivariable adjusted	1·00	1·00 (0·91-1·09)	0·85 (0·76-0·96)	0·89 (0·78-1·02)	0·0610

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. MI: myocardial infarction; CV death: cardiovascular mortality; Non-CV death: non-cardiovascular mortality. Crude event rates shown. Adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals, fruit intake, waist-to-hip ratio, hypertension status, and cholesterol medication use.

Table S27: Association of total fruit, vegetable and legume intake with cardiovascular disease and mortality with >6 servings/day as the highest intake category.

All participants (N=135,335)	**<1/day (N=9,082)**	**1 to <2/day (N=19,036)**	**2 to <3/day (N=35,128)**	**3 to <4/day (N=24,485)**	**4 to <5/day (N=14,849)**	**5 to <6/day (N=9,790)**	**>6/day (N=22,965)**	**P-trend**
Median fruit, vegetable and legume intake, servings	0·64 (0·41-0·83) per day	1·56 (1·30-1·79) per day	2·52 (2·29-2·75) per day	3·43 (3·21-3·70) per day	4·43 (4·19-4·70) per day	5·46 (5·22-5·72) per day	7·92 (6·80-9·92) per day	
Major CVD								
Major CVD events (N=4,784)	373 (4·11)	745 3·90)	1,391 3·96)	882 3·60)	534 (3·60)	267 (2·73)	594 2·59)	
Age and sex adjusted	1·00	1·00 (0·88-1·14)	1·06 (0·94-1·20)	1·00 (0·87-1·14)	1·09 (0·94-1·25)	0·86 (0·73-1·02)	0·85 (0·73-0·98)	0·0037
Multivariable adjusted	1·00	1·03 (0·89-118)	1·09 (0·96-1·25)	1·06 (0·92-1·23)	1·20 (1·03-1·40)	0·95 (0·80-1·14)	0·93 (0·78-1·10)	02932
Mortality								
Mortality events (N=5,796)	736 (8·10)	1,371 (7·20)	1,529 (4·35)	772 (3·15)	468 (3·15)	286 (2·92)	198 (2·85)	
Age and sex adjusted	1·00	0·92 (0·84-1·01)	0·81 (0·74-0·89)	0·65 (0·58-0·72)	0·65 (0·58-0·74)	0·62 (0·54-0·71)	0·60 (0·53-0·68)	<0·0001
Multivariable adjusted	1·00	1·01 (0·91-1·12)	0·91 (0·82-1·01)	0·78 (0·69-0·88)	0·83 (0·72-0·95)	0·78 (0·66-0·91)	0·82 (0·71-0·95)	0·0001

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. Crude event rates shown. Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals intake.

Table S28: Association of total fruit and vegetable intake with cardiovascular disease and mortality with >6 servings/day as the highest intake category.

All participants (N=135,335)	**<1/day (N=16,512)**	**1 to <2/day (N=20,478)**	**2 to <3/day (N=36,036)**	**3 to <4/day (N=22,134)**	**4 to <5/day (N=12,604)**	**5 to <6/day (N=8,575)**	**>6/day (N=18,996)**	**P-trend**
Median fruit, vegetable intake, servings	0·64 (0·38-0·81) per day	1·53 (1·27-1·77) per day	2·27 (2·25-2·71) per day	3·43 (3·20-3·71) per day	4·44 (4·40-4·70) per day	5·45 (5·22-5·71) per day	7·88 (6·78-9·86) per day	
Major CVD								
Major CVD events (N=4,874)	712 (4·31)	746 (3·64)	1,414 (3·92)	782 (3·53)	409 (3·25)	234 (2·73)	487 (2·56)	
Age and sex adjusted	1·00	0·93 (0·84-1·04)	1·02 (0·92-1·13)	0·99 (0·88-1·11)	0·98 (0·86-1·12)	0·84 (0·71-0·98)	0·80 (0·70-0·98)	0·0018
Multivariable adjusted	1·00	0·93 (0·83-1·05)	1·04 (0·93-1·17)	1·05 (0·92-1·19)	1·06 (0·92-1·22)	0·89 (0·75-1·06)	0·87 (0·74-1·02)	0·1910
Mortality								
Mortality events (N=5,796)	1,547 (9·37)	1,195 (5·84)	1,313 (3·64)	614 (2·77)	380 (3·01)	235 (2·74)	159 (2·72)	
Age and sex adjusted	1·00	0·81 (0·75-0·88)	0·74 (0·68-0·81)	0·60 (0·54-0·66)	0·65 (0·58-0·74)	0·59 (0·51-0·68)	0·57 (0·51-0·64)	<0·0001
Multivariable adjusted	1·00	0·88 (0·80-0·96)	0·84 (0·76-0·93)	0·73 (0·65-0·82)	0·82 (0·72-0·94)	0·73 (0·62-0·86)	0·78 (0·68-0·90)	<0·0001

Major CVD: defined as death from cardiovascular causes and nonfatal myocardial infarction, stroke, and heart failure. Crude event rates shown. Age and sex adjusted: adjusted for age, sex, center (random effect), multivariable adjusted: adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals intake.

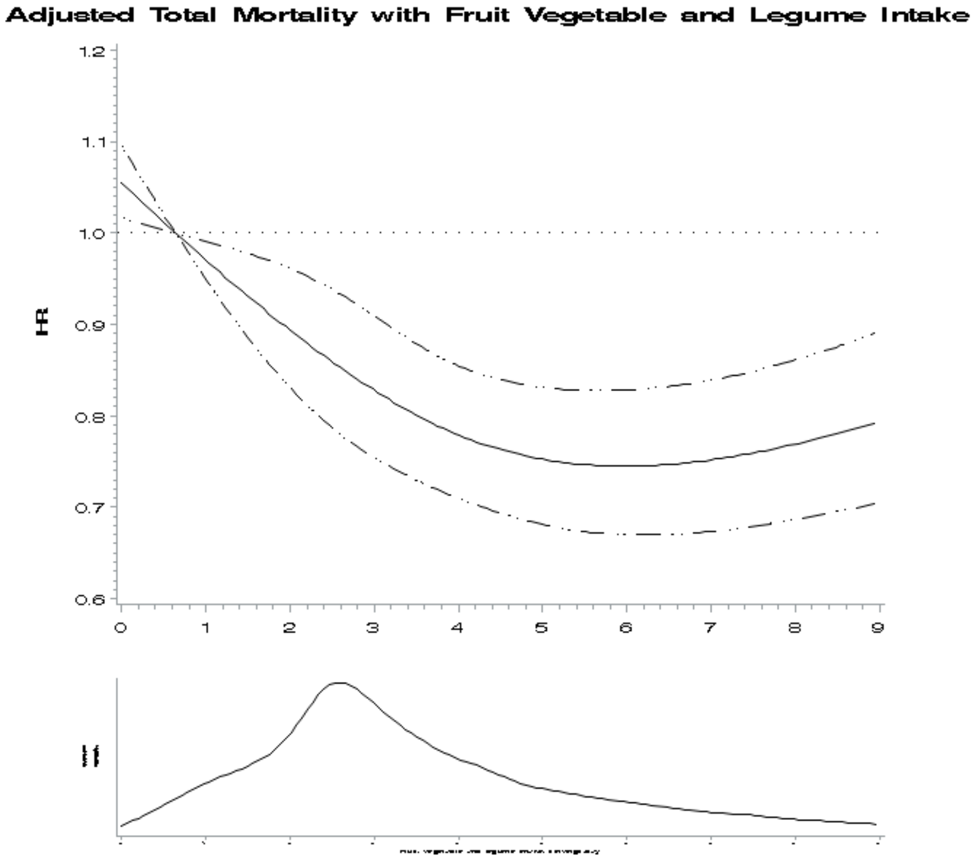


Figure S4. Total fruit, vegetable and legume intake versus mortality.
Cubic spline for the association between total fruit, vegetable and legume intake and total mortality (N=135,335). Adjusted for age, sex, center (random effect), energy intake, current smoker, diabetes, urban/rural location, physical activity, education level, tertiles of white meat, red meat, breads and cereals intake.

References:

1. Teo K, Chow CK, Vaz M, Rangarajan S, Yusuf S, Group PI-W. The Prospective Urban Rural Epidemiology (PURE) study: examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. Am Heart J 2009;158(1):1-7 e1.
2. MacQueen KM, McLellan E, Metzger DS, et al. What is community? An evidence-based definition for participatory public health. Am J Public Health 2001;91:1929-38.
3. Yusuf S, Rangarajan S, Teo K, Islam S, Liu L, et al. cardiovascular risk and events in 17 low-, middle-, and high-income countries. N Engl J Med 2014;371(9):818-27.
4. Tunstall-Pedoe H, Vanuzzo D, Hobbs M, Mahonen M, Cepaitis Z, Kuulasmaa K, et al. Estimation of contribution of changes in coronary care to improving survival, event rates, and coronary heart disease mortality across the WHO MONICA Project populations. Lancet 2000;355(9205):688-700.
5. Corsi DJ, Subramanian SV, Chow CK, McKee M, Chifamba J, Dagenais G, et al. Prospective Urban Rural Epidemiology (PURE) study: Baseline characteristics of the household sample and comparative analyses with national data in 17 countries. Am Heart J 2013;166(4):636-46 e4.
6. Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet 2004;364(9438):937-52.
7. O'Donnell MJ, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, et al. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. Lancet 2010;376(9735):112-23.
8. Gajalakshmi V, Peto R. Verbal autopsy of 80,000 adult deaths in Tamilnadu, South India. BMC Public Health 2004;4:47.
9. Jha P, Gajalakshmi V, Gupta PC, Kumar R, Mony P, Dhingra N, et al. Prospective study of one million deaths in India: rationale, design, and validation results. PLoS Med 2006;3(2):e18.

Bharathi AV, Kurpad AV, Thomas T, Yusuf S, Saraswathi G, Vaz M. Development of food frequency questionnaires and a nutrient database for the Prospective Urban and Rural Epidemiological (PURE) pilot study in South India: methodological issues. Asia Pac J Clin Nutr 2008;17(1):178-85.

11. Dehghan M, Al HN, Yusufali A, Nusrath F, Yusuf S, Merchant AT. Development of a semi-quantitative food frequency questionnaire for use in United Arab Emirates and Kuwait based on local foods. Nutr J 2005;4:18.
12. Dehghan M, Al-Hamad N, McMillan CR, Prakash P, Merchant AT. Comparison of a semi-quantitative food frequency questionnaire with 24-hour dietary recalls to assess dietary intake of adult Kuwaitis. Saudi Med J 2009 Jan;30(1):159-61.
13. Dehghan M, Lopez-Jaramillo P, Duenas R, Anaya LL, Garcia RG, Zhang X, et al. Development and validation of a quantitative food frequency questionnaire among rural- and urban-dwelling adults in Colombia. J Nutr Educ Behav 2012 Nov;44(6):609-13.
14. Dehghan M, Ilow R, Zatonska K, Szuba A, Zhang X, Mente A, et al. Development, reproducibility and validity of the food frequency questionnaire in the Poland arm of the Prospective Urban and Rural Epidemiological (PURE) study. J Hum Nutr Diet 2012 Jun;25(3):225-32.
15. Dehghan M, del CS, Zhang X, Cuneo JM, Linetzky B, Diaz R, et al. Validation of a semi-quantitative Food Frequency Questionnaire for Argentinean adults. PLoS One 2012;7(5):e37958.
16. Dehghan M, Martinez S, Zhang X, Seron P, Lanas F, Islam S, et al. Relative validity of an FFQ to estimate daily food and nutrient intakes for Chilean adults. Public Health Nutr 2012 Sep 21;1-7.
17. Dehghan M, Mente A, Teo KK, Gao P, Sleight P, Dagenais G, et al. Relationship between healthy diet and risk of cardiovascular disease among patients on drug therapies for secondary prevention: a prospective cohort study of 31 546 high-risk individuals from 40 countries. Circulation 2012 Dec 4;126(23):2705-12.
18. Gunes FE, Imeryuz N, Akalin A, Bekiroglu N, Alphan E, Oguz A, et al. Development and validation of a semi-quantitative food frequency questionnaire to assess dietary intake in Turkish adults. J Pak Med Assoc 2015 Jul;65(7):756-63.
19. Iqbal R, Ajayan K, Bharathi AV, Zhang X, Islam S, Soman CR, et al. Refinement and validation of an FFQ developed to estimate macro- and micronutrient intakes in a south Indian population. Public Health Nutr 2009 Jan;12(1):12-8.
20. Mahajan R, Malik M, Bharathi AV, Lakshmi PVM, Parto BK, Rana SK, et al. Reproducibility and validity of a quantitative food frequency questionnaire in an urban and rural area of northern India. The national medical journal of India 2013;26(5):266-72.
21. Merchant AT, Dehghan M. Food composition database development for between country comparisons. Nutr J 2006;5:2.

22. Mikkelsen TB, Olsen SF, Rasmussen SE, Osler M. Relative validity of fruit and vegetable intake estimated by the food frequency questionnaire used in the Danish National Birth Cohort. Scandinavian journal of public health. 2007;35(2):172-9.
23. Mohammadifard N, Omidvar N, Houshiarrad A, Neyestani T, Naderi GA, Soleymani B. Validity and reproducibility of a food frequency questionnaire for assessment of fruit and vegetable intake in Iranian adults(*). Journal of research in medical sciences : the official journal of Isfahan University of Medical Sciences. 2011;16(10):1286-97.
24. Landais E, Gartner A, Bour A, McCullough F, Delpeuch F, Holdsworth M. Reproducibility and relative validity of a brief quantitative food frequency questionnaire for assessing fruit and vegetable intakes in North-African women. Journal of human nutrition and dietetics : the official journal of the British Dietetic Association. 2014;27 Suppl 2:152-9.
25. Johansson I, Hallmans G, Wikman A, Biessy C, Riboli E, Kaaks R. Validation and calibration of food-frequency questionnaire measurements in the Northern Sweden Health and Disease cohort. Public health nutrition. 2002;5(3):487-96.
26. Marks GC, Hughes MC, van der Pols JC. Relative validity of food intake estimates using a food frequency questionnaire is associated with sex, age, and other personal characteristics. The Journal of nutrition. 2006;136(2):459-65.
27. Salvini S, Hunter DJ, Sampson L, Stampfer MJ, Colditz GA, Rosner B, et al. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. International journal of epidemiology. 1989;18(4):858-67.
28. Merchant AT, Dehghan M.Food coposition database development for between country comparisons. Nutrition Journal. 2006;5:2.
29. Cromwell GL, Calvert CC, Cline TR, Crenshaw JD, Crenshaw TD, Easter RA, Ewan RC, Hamilton CR, Hill GM, Lewis AJ, Mahan DC, Miller ER, Nelssen JL, Pettigrew JE, Tribble LF, Veum TL, Yen JT. Variability among sources and laboratories in nutrient analyses of corn and soybean meal. NCR-42 Committee on Swine Nutrition. North Central Regional-42. J Anim Sci. 1999, 77: 3262-3273.
30. Engler PP, Bowers JA. B-vitamin retention in meat during storage and preparation. A review. J Am Diet Assoc. 1976, 69: 253-257.
31. Nagy S, Smoot JM. Temperature and storage effects on percent retention and percent U.S. recommended dietary allowance of vitamin C in canned single-strength orange juice. J Agric Food Chem. 1976, 25: 135-138. 10.1021/jf60209a031.
32. Wen J, Morrissey PA, Buckley DJ, Sheehy PJ. Oxidative stability and alpha-tocopherol retention in turkey burgers during refrigerated and frozen storage as influenced by dietary alpha-tocopheryl acetate. Br Poult Sci. 1996, 37: 787-795.
33. Williams PG. Vitamin retention in cook/chill and cook/hot-hold hospital food-services. J Am Diet Assoc. 1996, 96: 490-498. 10.1016/S0002-8223(96)00135-6.

Funding/Support:
Dr S Yusuf is supported by the Mary W Burke endowed chair of the Heart and Stroke Foundation of Ontario.
The PURE Study is an investigator-initiated study that is funded by the Population Health Research Institute, the Canadian Institutes of Health Research, Heart and Stroke Foundation of Ontario, Support from CIHR's Strategy for Patient Oriented Research, through the Ontario SPOR Support Unit, as well as the Ontario Ministry of Health and Long-Term Care and through unrestricted grants from several pharmaceutical companies [with major contributions from Astra Zeneca (Canada), Sanofi-Aventis (France and Canada), Boehringer Ingelheim (Germany & Canada), Servier, and GSK], and additional contributions from Novartis and King Pharma and from various national or local organizations in participating countries.

These include: **Argentina:** Fundacion ECLA; **Bangladesh:** Independent University, Bangladesh and Mitra and Associates; **Brazil:** Unilever Health Institute, Brazil; **Canada:** Public Health Agency of Canada and Champlain Cardiovascular Disease Prevention Network; **Chile:** Universidad de la Frontera; **China:** National Center for Cardiovascular Diseases; **Colombia:** Colciencias, Grant number:6566-04-18062; **India:** Indian Council of Medical Research; **Malaysia:** Ministry of Science, Technology and Innovation of Malaysia Grant Nbr 100 - IRDC / BIOTEK 16/6/21 (13/2007), Grant Number 07-05-IFN-BPH 010, Ministry of Higher Education of Malaysia Grant Nbr 600 - RMI/LRGS/5/3 (2/2011), Universiti Teknologi MARA, Universiti Kebangsaan Malaysia (UKM-Hejim-Komuniti-15-2010); **occupied Palestinian territory:** the United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA), occupied Palestinian territory; International Development Research Centre (IDRC), Canada; **Philippines:** Philippine Council for Health Research & Development (PCHRD); **Poland:** Polish Ministry of Science and Higher Education grant Nr 290/W-PURE/2008/0, Wroclaw Medical University; **Saudi Arabia:** The Deanship of Scientific Research at King Saud University, Riyadh, Saudi Arabia (Research group number: RG -1436-013); **South Africa:** The North-West University, SANPAD (SA and Netherlands Programme for Alternative Development), National Research Foundation, Medical Research Council of SA, The SA Sugar Association (SASA), Faculty of Community and Health Sciences (UWC); **Sweden**: Grants from the Swedish state under the Agreement concerning research and education of doctors; the Swedish Heart and Lung Foundation; the Swedish Research Council; the Swedish Council for Health, Working Life and Welfare, King Gustaf V:s and Queen Victoria Freemasons Foundation, AFA Insurance, Swedish Council for Working Life and Social Research, Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning, Grant from the Swedish State under (LäkarUtbildningsAvtalet) Agreement, Grant from the Västra Götaland Region (FOUU); **TURKEY**: Metabolic Syndrome Society, Astra Zeneca, Turkey, Sanofi Aventis, Turkey; **UAE**: Sheikh Hamdan Bin Rashid Al Maktoum Award For Medical Sciences and Dubai Health Authority, Dubai UAE.

Role of Sponsor: The funders and sponsors had no role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; in the preparation, review, or approval of the manuscript; or in the decision to submit the manuscript for publication.

PURE Project Office Staff, National Coordinators, Investigators, and Key Staff:

Project office (Population Health Research Institute, Hamilton Health Sciences and McMaster University, Hamilton, Canada): S Yusuf* (Principal Investigator).
S Rangarajan (Project Manager); K K Teo, C K Chow, M O'Donnell, A Mente, D Leong, A Smyth, P Joseph, A Merchant, S Islam (Statistician), M Zhang (Statistician), W Hu (Statistician), C Ramasundarahettige (Statistician), G Wong (Statistician), S Bangdiwala, L Dyal, A Casanova, M Dehghan (Nutrition Epidemiologist), G Lewis, A Aliberti, A Arshad, A Reyes, A Zaki, B Lewis, B Zhang, D Agapay, D Hari, E Milazzo, E Ramezani, F Hussain, F Shifaly, G McAlpine, I Kay, J Lindeman, J Rimac, J Swallow, L Heldman, M(a) Mushtaha, M(o) Mushtaha, M Trottier, M Riggi, N Aoucheva, N Kandy, P Mackie, R Solano, S Chin, S Ramacham, S Shahrook, S Trottier, T Tongana, W ElSheikh, Y Iyengar.
Core Laboratories: M McQueen, K Hall, J Keys (Hamilton), X Wang (Beijing, China), J Keneth, A Devanath (Bangalore, India).
Argentina: R Diaz*; A Orlandini, B Linetsky, S Toscanelli, G Casaccia, JM Maini Cuneo; **Bangladesh:** O Rahman*, R Yusuf, AK Azad, KA Rabbani, HM Cherry, A Mannan, I Hassan, AT Talukdar, RB Tooheen, MU Khan, M Sintaha, T Choudhury, R Haque, S Parvin; **Brazil:** A Avezum*, GB Oliveira, CS Marcilio, AC Mattos; **Canada:** K Teo***,** S Yusuf***,** J Dejesus, D Agapay, T Tongana, R Solano, I Kay, S Trottier, J Rimac, W Elsheikh, L Heldman, E Ramezani, G Dagenais, P Poirier, G Turbide, D Auger, A LeBlanc De Bluts, MC Proulx, M Cayer, N Bonneville, S Lear, D Gasevic, E Corber, V de Jong, I Vukmirovich, A Wielgosz, G Fodor, A Pipe, A Shane; **CHILE:** F Lanas*, P Seron, S Martinez, A Valdebenito, M Oliveros; **CHINA:** Li Wei*, Liu Lisheng*, Chen Chunming, Wang Xingyu, Zhao Wenhua, Zhang Hongye, JiaXuan, Hu Bo, Sun Yi, Bo Jian, Zhao Xiuwen, Chang Xiaohong, Chen Tao, Chen Hui, Chang Xiaohong, Deng Qing, Cheng Xiaoru, Deng Qing, He Xinye, Hu Bo, JiaXuan, Li Jian, Li Juan,Liu Xu, Ren Bing, Sun Yi, Wang Wei, Wang Yang, Yang Jun, Zhai Yi, Zhang Hongye, Zhao Xiuwen,Zhu Manlu, Lu Fanghong, Wu Jianfang, Li Yindong, Hou Yan, Zhang Liangqing, Guo Baoxia, Liao Xiaoyang, Zhang Shiying, BianRongwen, TianXiuzhen, Li Dong, Chen Di, Wu Jianguo, Xiao Yize, Liu Tianlu, Zhang Peng, Dong Changlin, Li Ning, Ma Xiaolan, Yang Yuqing, Lei Rensheng, Fu Minfan, He Jing, Liu Yu, Xing Xiaojie, Zhou Qiang, ; **Colombia:** P Lopez-Jaramillo*, PA Camacho Lopez, R Garcia, LJA Jurado, D Gómez-Arbeláez, JF Arguello, R Dueñas, S Silva, LP Pradilla, F Ramirez, DI Molina, C Cure-Cure, M Perez, E Hernandez, E Arcos, S Fernandez, C Narvaez, J Paez, A Sotomayor, H Garcia, G Sanchez, T David, A Rico; **India:** P Mony *, M Vaz*, A V Bharathi, S Swaminathan, K Shankar AV Kurpad, KG Jayachitra, N Kumar, HAL Hospital, V Mohan, M Deepa, K Parthiban, M Anitha, S Hemavathy, T Rahulashankiruthiyayan, D Anitha, K Sridevi, R Gupta, RB Panwar, I Mohan, P Rastogi, S Rastogi, R Bhargava, R Kumar, J S Thakur, B Patro, PVM Lakshmi, R Mahajan, P Chaudary, V Raman Kutty, K Vijayakumar, K Ajayan, G Rajasree, AR Renjini, A Deepu, B Sandhya, S Asha, HS Soumya; **Iran:** R Kelishadi*, A Bahonar, N Mohammadifard, H Heidari; **Malaysia:** K Yusoff*, TST Ismail, KK Ng, A Devi, NM Nasir, MM Yasin, M Miskan, EA Rahman, MKM Arsad, F Ariffin, SA Razak, FA Majid, NA Bakar, MY Yacob, N Zainon, R Salleh, MKA Ramli, NA Halim, SR Norlizan, NM Ghazali, MN Arshad, R Razali, S Ali, HR Othman, CWJCW Hafar, A Pit, N Danuri, F Basir, SNA Zahari, H Abdullah, MA Arippin, NA Zakaria, I

Noorhassim, MJ Hasni, MT Azmi, MI Zaleha, KY Hazdi, AR Rizam, W Sazman, A Azman; **OCCUPIED PALESTINIAN TERRITORY:** R Khatib*, U Khammash, A Khatib, R Giacaman; **PAKISTAN:** R Iqbal*, A Afridi, R Khawaja, A Raza, K Kazmi; **PHILIPPINES:** A Dans*, HU Co, JT Sanchez, L Pudol, C Zamora-Pudol, LAM Palileo-Villanueva, MR Aquino, C Abaquin, SL Pudol, ML Cabral; **Poland:** W Zatonski*, A Szuba, K Zatonska, R Ilow#, M Ferus, B Regulska-Ilow, D Różańska, M Wolyniec; **SAUDI ARABIA:** KF AlHabib*, A Hersi, T Kashour, H Alfaleh, M Alshamiri, HB Altaradi, O Alnobani, A Bafart, N Alkamel, M Ali, M Abdulrahman, R Nouri; **South Africa:** A Kruger*, H H Voster, A E Schutte, E Wentzel-Viljoen, FC Eloff, H de Ridder, H Moss, J Potgieter, AA Roux, M Watson, G de Wet, A Olckers, JC Jerling, M Pieters, T Hoekstra, T Puoane, E Igumbor, L Tsolekile, D Sanders, P Naidoo, N Steyn, N Peer, B Mayosi, B Rayner, V Lambert, N Levitt, T Kolbe-Alexander, L Ntyintyane, G Hughes, R Swart, J Fourie, M Muzigaba, S Xapa, N Gobile , K Ndayi, B Jwili, K Ndibaza, B Egbujie; **SWEDEN:** A Rosengren*, K Bengtsson Boström, A Gustavsson, M Andreasson, M Snällman, L Wirdemann ; **TANZANIA:** K Yeates*, J Sleeth, K Kilonzo**;** **TURKEY:** A Oguz*, N Imeryuz, Y Altuntas, S Gulec, A Temizhan, K Karsidag, KBT Calik, AAK Akalin, OT Caklili, MV Keskinler, AN Erbakan; **UNITED ARAB EMIRATES:** AM Yusufali*, W Almahmeed, H Swidan, EA Darwish, ARA Hashemi, N Al-Khaja, JM Muscat-Baron, SH Ahmed, TM Mamdouh, WM Darwish, MHS Abdelmotagali, SA Omer Awed, GA Movahedi, F Hussain, H Al Shaibani, RIM Gharabou, DF Youssef, AZS Nawati, ZAR Abu Salah, RFE Abdalla, SM Al Shuwaihi, MA Al Omairi, OD Cadigal; R.S. Alejandrino; **Zimbabwe:** J Chifamba*, L Gwaunza, G Terera, C Mahachi, P Murambiwa, T Machiweni, R Mapanga.

*National Coordinator

Deceased

Chapter 3

Comparison of two methods for modeling nutrient replacement in an international population

Prepared for submission; Victoria Miller, Andrew Mente

Abstract

Background: Energy-adjustment analyses are widely used in nutrition epidemiologic studies to model the impact of replacing a single nutrient with another nutrient on risk of disease. This approach has been used predominantly in studies of North American and European populations and it is unknown whether it is applicable to a global population with a broader range of intake for most nutrients and non-linear associations (i.e., U-shape or threshold) between exposure and health outcomes. An alternative approach would be to examine the joint association of nutrients in combination (e.g., low-low, low-high, high-low, and high-high) on the risk of the outcome. Our aim was to perform exploratory analyses to compare the observed results from both methods separately to expectant results based on the observed association of each nutrient versus mortality risk (the primary result in a conventional nutrition epidemiology paper).

Methods: The Prospective Urban Rural Epidemiology (PURE) study is an international cohort of 135,607 individuals aged 35 to 70 years in 18 high, middle and low-income countries on 5 continents. Dietary information was collected using country-specific food frequency questionnaires at baseline. Standardized questionnaires were used to collect information about demographic factors, socioeconomic status (income, occupation, and education), lifestyle (physical activity, alcohol intake and smoking status), health history and medication use. Cox frailty models with random effects were used to assess associations between nutrient consumption with risk of mortality. Multivariate nutrient density models were used to calculate hazard ratios (HRs) for the isocaloric replacement of percent of energy from saturated fat with other nutrients. Using binary categorizes of intake, we estimated the joint effect of saturated fat with carbohydrate, protein, monounsaturated fat and polyunsaturated fat, separately. We compared the observed results for both the multivariable nutrient density and joint effect models separately with the expectant results and scored the level of agreement.

Findings: During 8.2 years (IQR 5.7-9.6) years of follow-up, 6,630 total deaths were documented. Higher carbohydrate intake was positively associated with an increased risk of mortality (Q5 vs. Q1: HR 1.28; 95% CI 1.13 to 1.44; p-trend <0.0001), while total fat (HR 0.77; 95% CI 0.68 to 0.87; p-trend<0.0001), saturated fat (HR 0.85; 95% CI 0.75 to 0.96; p-trend=0.0028), monounsaturated fat (HR 0.83; 95% CI 0.74 to 0.94; p-trend=0.0001), polyunsaturated fat (HR 0.82; 95% CI 0.74 to 0.91; p-trend=0.0002), and protein (HR 0.83; 95% CI 0.74 to 0.94; p-trend<0.0001) was inversely associated with death. In the multivariate nutrient density models, replacement of 3% of energy from saturated fat with polyunsaturated fat (HR 0.94; 95% CI 0.90 to 0.98; p-value=0.0075) or protein (HR 0.94; 95% CI 0.90 to 0.98; p-value=0.0039) was significantly associated with a lower risk of total mortality, whereas no significant association was observed for monounsaturated fat (HR 0.95; 95% CI 0.89 to 1.01; p-value=0.1012) or carbohydrate (HR 0.98; 95% CI 0.96 to 1.01; p-value=0.1725). In the joint effect analyses, higher saturated fat with monounsaturated fat, polyunsaturated fat, protein or carbohydrate, separately, was associated with lower mortality compared to the reference. The observed results from the energy-adjustment analysis did not agree with the expectant results for any of the nutrient replacement analyses (0 of 4 nutrients), while the joint effect approach had markedly better agreement (3 of 4 nutrients).

Interpretation: Use of conventional energy-adjustment may misinform the effect of nutrients on mortality in diverse populations. Our findings highlight the need for new novel methods to examine the relationship between nutrients in combination and disease risk, particularly when the population of interest has a wide range of nutrient intake.

Introduction

Current global dietary guidelines recommend that adults consume fewer than 30% of total energy intake from total fats with less than 10% of energy from saturated fats[1]. The origin of this recommendation has been traced to the Seven Countries Study which was the first study to show that saturated fat and cholesterol increased the risk of coronary heart disease[2,3]. In subsequent years, however, prospective cohort studies[4], largely done in North America and Europe, found no significant association between higher saturated fat intake and risk of cardiovascular disease and mortality. Macronutrient consumption varies substantially across geographic regions with individuals in North America and Europe consuming the greatest percent of energy from total fat and saturated fat, and China and Africa consuming the lowest[5,6]. Until recently, little information was available on any potential associations of macronutrients with death outside of North American and Europe populations. In our recent work within 18 high, middle and low-income countries in the PURE study, we found that higher carbohydrate intake was associated with an increased risk of total mortality, while total fat and each type of fat (saturated, monounsaturated, polyunsaturated) were associated with lower risk of total mortality[5]. Collectively, these data question the scientific basis of low dietary fat and saturated fat guidelines.

Nutrients are not consumed in isolation but rather are consumed concomitantly in foods. Assessing associations of individual nutrients with health outcomes associations may be impacted by potential interactions between the effects of different nutrients, the high degree of collinearity between nutrients, or the replacement of one nutrient for another. In randomized controlled trials, the effect of replacing saturated fat with other nutrients (e.g., polyunsaturated fat or carbohydrates) is conducted directly using an experimental design whereby one nutrient is replaced for another in the two diets that are being compared. Conversely, in observational studies, the replacement of one nutrient for another nutrient is modeled using a multivariable regression model known as the energy-adjustment approach. This has come to be accepted as a conventional approach for studying the replacement of one nutrient for another in observational studies. However, a key assumption of this method is that there is a linear relationship between each of the nutrients and the outcome event. This may not be appropriate in a global population with a broader range of intake for most nutrients than observed in previous studies conducted mostly in Western populations with a narrower range of intake.

The joint effect approach is commonly used in many areas of epidemiology to examine the effect of two or more independent variables in combination in relation to risk of outcome events. For example, studies have investigated the joint effect of genes and lifestyle behaviors on the risk of type 2 diabetes[7], low birthweight and lifestyle behaviors on the risk of hypertension[8] and television watching and physical activity on the risk of type 2 diabetes[9]. In nutritional epidemiology, previous joint effect analyses have mostly been used to assess the joint effect of foods or dietary patterns on disease risk (e.g., coffee and tea at baseline and follow-up on diabetes risk[10] and ‘Prudent’ and ‘Western’ diet scores versus coronary artery disease risk[11]). However, similar methodology can be used to investigate the joint association of two or more nutrients. With this approach, each nutrient is divided into categories of intake (e.g., halves or tertiles) to form mutually exclusive groups and the association is modeled using multivariable regression. Thus, the joint effect approach demonstrates the effect of various levels of nutrients jointly on the outcome of interest.

In the present analysis from the PURE study, our aim was to perform exploratory analyses to compare the observed results from both methods separately (energy-adjusted and joint effect) to expectant results for the associations of saturated fat intake with other nutrients on the risk of total mortality.

Methods

Study design and sample selection

A detailed description of participant, community and country selection has been described previously[12-15]. Briefly, we enrolled adults between the ages of 35 and 70 years from 18 low, middle and high-income countries on five continents between January 1, 2003 and March 31, 2013. We considered the heterogeneity of socioeconomic factors and the feasibility of carrying out long-term follow up when selecting the participating countries. We included: three high-income (HICs: Canada, Sweden and United Arab Emirates), seven upper-middle income (UMICs: Argentina, Brazil, Chile, Malaysia, Poland, South Africa, and Turkey), four lower-middle income (LMICs: China, Colombia, Iran and the occupied Palestinian territory) and four lower-income countries (LICs: Bangladesh, India, Pakistan and Zimbabwe), based on gross national income per capita from the World Bank classification for 2006 when the study was initiated.

Procedures

Habitual dietary intake was measured using country-specific (region-specific in India) validated food frequency questionnaires (FFQ) at baseline. Previously validated FFQs and the corresponding nutrient database were used for Canada, China, India, Malaysia, South Africa, Sweden, and Turkey. For countries where a validated FFQ was not available, a standard method was used to develop and validate the country-specific FFQ[16-26]. Within these countries, a subset of participants (60-250 participants) completed 24-hour dietary recalls for each season and the most frequently recalled food items were compiled. To convert food into nutrients, country-specific nutrient databases were developed with information on 43 macro and micronutrients. A master international nutrient database was created primarily based on the United States Department of Agriculture (USDA) food composition database (release 18 and 21), modified appropriately with reference to local food composition tables, and supplemented with recipes of locally eaten mixed dishes.

Standardized questionnaires were used to collect information about demographic factors, socio-economic status (education, income and employment), lifestyle (smoking, physical activity and alcohol intake), health history and medication use, and family history of cardiovascular disease. Physical assessment included standardized weight, height, and waist and hip circumference and blood pressure measurements. Standardized case-report forms were used to capture data on major cardiovascular events, and death during follow-up, which were adjudicated centrally in each country by trained physicians using pre-defined definitions. For the current analysis, we included all outcome events known to us until October 2017.

Outcome

The main clinical outcome was total mortality (defined as death from cardiovascular and non-cardiovascular causes, not including death from injury).

Models of energy-adjustment

The independent effect of a nutrient on risk of disease can be evaluated in several ways: the standard multivariate model, nutrient residual model, energy-partition model and multivariate nutrient density model[27-33]. Since similar associations have been reported from the four energy-adjustment models when the nutrients are in continuous form[27,28,34], we present the results of the multivariate nutrient density model only in this paper. For the sake of simplicity, the model shown below does not include confounders other than energy intake.

The Multivariate Nutrient Density Model

Disease risk= m1*carbohydrate + m2*monounsaturated fat + m3*polyunsaturated fat + m4*trans-fat + m5*protein + m6*total energy intake

where m1-m6 are regression coefficients and one macronutrient is excluded from the model (saturated fat)

The multivariate nutrient density method models the nutrients as a percentage of energy, total energy intake and excludes one macronutrient from the model. The regression coefficients have isocaloric interpretation and represent the substitution of equal energy.

Joint effect modeling

The joint effect of two nutrients can be investigated by dividing the nutrients into categories of intake (halves, tertiles, quartiles, etc.). For each nutrient, we visually inspected the change in slope observed from the restricted cubic splines and used the corresponding nutrient intake value as the cut point to define binary categories of intake. We performed sensitivity analyses using the median nutrient intake value as the binary cut point (appendix).

The Joint Effect Model

Group 1= % energy from saturated fat <7 and % energy from monounsaturated fat <6
Group 2= % energy from saturated fat ≥7 and % energy from monounsaturated fat <6
Group 3= % energy from saturated fat <7 and % energy from monounsaturated fat ≥6
Group 4= % energy from saturated fat ≥7 and % energy from monounsaturated fat ≥6

Disease risk= m1*group 2 + m2*group 3 + m3*group 4 + m4*total energy intake

where m1-m4 are regression coefficients and group 1 is the reference group

Statistical Analysis

Means (SDs) were calculated to summarize continuous variables and percentages for categorical variables. Location was categorized by urban or rural community. Physical activity was categorized into three levels based on the metabolic equivalent of task (MET) per minutes per week: low (<600 MET minutes per week), moderate (600-3000 MET minutes per week) and high (>3000 MET minutes per week). Education was categorized as less than graduation from high-school, high-school graduate, and some college or more. Smoking status was categorized as never, former or current smoker. Waist-to-hip ratio was used as a continuous variable and computed as waist circumference (cm) divided by hip circumference (cm). When appropriate, energy intake was used as a continuous variable. We computed percent energy provided by nutrients by dividing energy provided by the nutrient by total energy intake. Participants were grouped into quintiles of intake values for each dietary exposure (percent energy from carbohydrate, protein, total fat, saturated fatty acid, monounsaturated fatty acid and polyunsaturated fatty acid). We calculated hazard ratios (HR) using multivariable Cox frailty analysis with random intercepts to account for clustering which therefore also accounted for clustering at region and country levels. In a minimally adjusted model, we adjusted for age, sex and center as a random effect. The fully adjusted model adjusted for age, sex, urban or rural location, physical activity, education level, current smoking status, waist-to-hip ratio, history of diabetes and study center (as a random effect). We used restricted cubic splines with four knots to assess the shape of associations between nutrients and total mortality.

We used the findings from the fully adjusted Cox frailty analysis of quintiles of each nutrient and the risk of mortality to hypothesize the expected results of both replacing percent energy from saturated fats with other nutrients and the joint effect of saturated fats with other nutrients. We estimated the effect of isocaloric replacement (3% of energy) of saturated fat with other nutrients using multivariable nutrient density models. In this modeling, the percent of energy from carbohydrate, monounsaturated fat, polyunsaturated fat, protein and other were included as exposures, and total energy as a covariate. In these models, the coefficients indicated the change in total mortality risk by replacement of saturated fat for a 3% increase in the other energy-contributing nutrients.

We estimated the joint effect of saturated fat with other nutrients (carbohydrate, protein, monounsaturated fat, and polyunsaturated fat) using binary categorizes of intake. In these models, the coefficients represent the combined effect of saturated fat with each nutrient on the risk of mortality. We compared the observed results for both the multivariable nutrient density and joint effect models with the expectant results and scored the level of agreement. A binary score of 0 or 1 was allocated for each comparison. For all analyses, the criterion for statistical significance was α=0.05. Statistical analyses were done with SAS software, version 9.4. Spline curves were generated with SAS LGTPHCURV9 Macro.

Results

Data collection occurred between January 1, 2003 and March 31, 2013. A total of 153,485 participants completed the FFQ of which 143,974 people had plausible energy intake (500–5000 kcal per day) and were not missing information on age and sex. We excluded 7,412 people with a history of CVD at baseline and 955 people because follow-up information was not available. The remaining 135,607 individuals are included in this analysis. For the current analysis, we included all unrefuted cases of total mortality in the PURE study database through October 1, 2017.

The characteristics of the participants are shown in Table 1. Across the study population, mean percent of energy from carbohydrate intake was 61.2% (SD 11.6), percent of energy from protein was 15.2% (3.6), and percent energy from total fat was 23.5% (9.3). Percent energy from total fat consisted of 7.6% (SD 4.1) from unsaturated fat, 7.7% (3.5) from monounsaturated fat, and 5.0%

(2.9) from polyunsaturated fat. People who consumed more saturated fat had higher waist-to-hip ratio, higher education, higher levels of physical activity, higher prevalence of diabetes at baseline, higher energy intake, higher percent energy from protein, total fat, monounsaturated fat, and polyunsaturated fat intake, and lower percent energy from carbohydrate intake (Table 1). During a median follow-up of 8.2 (IQR 5.7-9.6) years, 6,630 cases of total mortality occurred.

Associations between nutrients and mortality (conventional multivariable modeling, reported in Dehghan et al. 2017)

Higher carbohydrate intake was positively associated with total mortality in both the age and sex-adjusted (Q5 vs. Q1: HR 1.39; 95% CI 1.26 to 1.54; p-trend <0.0001) and fully adjusted models (Q5 vs. Q1: HR 1.28; 95% CI 1.13 to 1.44; p-trend <0.0001)(Table 2). Conversely, in comparisons between quintile 5 and quintile 1 higher total fat (HR 0.77; 95% CI 0.68 to 0.87; p-trend<0.0001), saturated fat (HR 0.85; 95% CI 0.75 to 0.96; p-trend=0.0028), monounsaturated fat (HR 0.83; 95% CI 0.74 to 0.94; p-trend=0.0001), polyunsaturated fat (HR 0.82; 95% CI 0.74 to 0.91; p-trend=0.0002), and protein (HR 0.83; 95% CI 0.74 to 0.94; p-trend<0.0001) was inversely associated with risk of total mortality (Table 2).

Expected results of a conventional nutrient replacement analyses, based on the associations of each nutrient with mortality

The direction and magnitude of associations from the Cox frailty analyses were similar for saturated, monounsaturated and polyunsaturated fat and protein. From this information, it would be expected that the replacement of saturated fat with these other nutrients would not alter the risk of death. Conversely, since higher carbohydrate intake was associated with *increased* the risk of death (unlike saturated fat which was associated with *lower* risk of death), we would expect that replacing saturated fat with carbohydrate would be associated with *higher* risk of mortality. For the joint effect analysis, we expect that the lowest risk of mortality would be observed in the groups with both higher saturated fat and higher monounsaturated and polyunsaturated fat and protein, separately. Lastly, we expect that the group with higher saturated fat and lower carbohydrate intake would have the strongest inverse association with mortality.

Restricted multivariable cubic spline plots for total mortality using four knots are shown in Figure 1. Multivariable splines for carbohydrate had a non-linear increasing trend in total mortality (p for non-linearity =0.0250). However, multivariable splines for total fat (p for non-linearity=0.3844), saturated fat (p for non-linearity =0.0314), monounsaturated fat (p for non-linearity =0.0095), polyunsaturated fat (p for non-linearity =0.3092), and protein (p for non-linearity =0.0163) showed non-linear decreasing trends in total mortality.

Conventional nutrient replacement analysis (Method 1)

In the nutrient replacement analyses, isocaloric replacement of 3% of energy from saturated fat with polyunsaturated fat (HR 0.94; 95% CI 0.90 to 0.98; p-value=0.0075) or protein (HR 0.94; 95% CI 0.90 to 0.98; p-value=0.0039) was significantly associated with a lower risk of total mortality, whereas no association was observed for carbohydrate or monounsaturated fat (Table 3). As shown in Table 4, there was disagreement between the expectant results and the observed results for all nutrients.

Joint effects analysis (Method 2)

Tables 5 to 8 show the joint effect of saturated fat with monounsaturated fat, polyunsaturated fat, carbohydrate and protein, separately. Compared to the saturated fat <7%E and monounsaturated fat intake <6%E group, the group with ≥7%E saturated fat and ≥6%E monounsaturated fat was associated with a 14% decrease in the risk of total mortality (HR 0.86; 95% CI 0.77 to 0.97; p-value=0.0103). For saturated fat and polyunsaturated fat, the greatest reduction in risk of total mortality was observed for the ≥7%%E saturated fat and ≥5%E polyunsaturated fat group compared to the reference (<7%E saturated fat and <5%E monounsaturated fat; HR 0.87; 95% CI 0.77 to 0.98; p-value=0.0171). Compared to the saturated fat <7%E and carbohydrate ≥60%E group, the ≥7%E saturated fat and ≥60%E carbohydrate group was inversely associated with death (HR 0.87; 95% CI 0.78 to 0.96; p-value=0.0055). Lastly, a 15% decrease in the risk of total mortality was observed for the ≥7%E saturated fat and ≥13%E protein group; HR 0.85; 95% CI 0.75 to 0.96; p-trend=0.0111), compared to the reference group.

As shown in Table 9, there was agreement between the expectant results and the observed results for all nutrients except carbohydrate.

Comparison of the two methods

The observed results from the conventional nutrient replacement analysis had poor agreement with the expectant results (Agreement score=0/4), while joint effect approach had higher agreement (Agreement score=3/4)(Table 10).

Discussion

In this updated analysis of the paper by Dehghan et al.[5], we showed that greater carbohydrate intake was positively associated with total mortality, whereas total fat and individual types of fat (saturated fat, monounsaturated fat, and polyunsaturated fat) are inversely associated with risk of death. The spline plots showed significant non-linear associations between each type of fat and total mortality, and a trend for a non-linear association between carbohydrate intake and mortality.

Using multivariate nutrient density models, we replaced 3% of energy from saturated fat for other nutrients. No association was observed for the isocaloric replacement with carbohydrate or monounsaturated fat, and an inverse association was observed for the substitution of polyunsaturated fat or protein. The energy-adjustment modeling assumes linearity of associations, but our data suggests that across a wide range of intake, the relationship between nutrients and mortality may not be linear. For example, the effect of replacing 3% of energy from saturated fat among individuals consuming 15% of energy from saturated fat may be different than among those consuming 5% of energy from saturated fat.

Numerous studies have examined the replacement of saturated fat with other macronutrients on the risk of mortality[4]. In a pooled analysis of 11 cohorts in America and Europe, the substitution of 5% of energy from saturated fat with polyunsaturated fat decreased the risk of coronary deaths by 26% (HR=0.74; 95% CI 0.61 to 0.89), but no association was observed for the replacement of saturated fat with carbohydrate or monounsaturated fat[35]. Several foods high in polyunsaturated fats are low in saturated fat including plant-based oils (corn, safflower, soybean and sunflower oil) and walnuts, while many major saturated fat food sources contain little to no polyunsaturated fat (coconut oil, palm oil, egg, butter, cheese, whole milk, and red meat)[36]. Conversely, saturated and monounsaturated fat are moderately correlated (r=0.76; p-value<0.0001) due to shared food sources such as eggs, red meat, margarine, butter, whole milk and cheese[36]. In real-life settings, individuals can easily substitute polyunsaturated fat for saturated fat, demonstrating a feasible dietary intervention to improve mortality risk, while replacing saturated fat for monounsaturated may be less biologically attainable. This is exemplified by the relatively small number of individuals who were classified as high intake of saturated fat and low monounsaturated fat intake or vice versa in the joint effect models. Most studies, including this present study, did not consider type of carbohydrates in our energy replacement models. The substitution of carbohydrates for fat intake may vary by the type of carbohydrate consumed including dietary fiber content[37], level of processing (whole and refined grain)[38], and the glycemic index[35,39]. For example, Jakobsen et al.[40] demonstrated that replacing saturated fat with low glycemic carbohydrates (fruits, vegetables, pulses and grains) decreased the risk of cardiovascular disease and replacement with high glycemic carbohydrate increased the risk.

As an alternative to the energy-adjustment model, we modeled the joint effect of saturated fat with other nutrients. The risk of mortality was reduced for all categories of saturated fat with monounsaturated fat, polyunsaturated fat and protein, separately, compared to the reference of low intake of both nutrients. This finding is consistent with the expectant results and suggests that increasing the percent energy from these nutrients in place of saturated fat does not offer substantial benefit against premature death. Conversely, for the joint effect of saturated fat and carbohydrate intake an inverse association was found for higher saturated fat intake regardless of carbohydrate intake which was inconsistent with the expectant results. Overall, the findings of the joint effect analysis appear to have relatively good agreement with the expectant results. Importantly, there are several limitations to the joint effect method. First, this method may be overly simplistic when used for macronutrients because it does not consider the joint effect of all energy-contributing nutrients. To do so, a single model that includes the joint effect of all macronutrients (saturated, monounsaturated and polyunsaturated fat, carbohydrate, and protein) is required. However, this approach is infeasible due to the number of cells and scarcity of individuals within the cells. Alternatively, we modeled the joint effect of saturated fat and carbohydrate with monounsaturated fat or polyunsaturated fat or protein, separately (appendix) and the findings were generally consistent with primary models which adjusted for carbohydrate intake. Second, the categorization of nutrient intake into binary categories was based on visual inspection of the cubic splines and the effect estimates may differ based on the threshold selected (e.g., <7%E vs. <6%E from saturated fat). However, we observed similar associations when median nutrient intake was used as the threshold (appendix). Third, the FFQs used to measure dietary intake are suitable for classifying individuals into categories of intake and not estimating absolute intake. The categorization of nutrient intake via absolute intake may be subject to misclassification.

Our study had some limitations. We measured dietary intake using country-specific validated food frequency questionnaires at baseline only and we did not examine changes in dietary intake which may have varied during follow-up. Epidemiological studies measuring dietary intake are often subject to recall and measurement error due to imprecise measurement tools[41,42]. The use of self-reported dietary questionnaires may result in measurement error which may distort true associations. Second, we were unable to measure trans-fat intake, but we included a variable for other energy-contributing nutrients in the energy adjustment models to minimize the effect this may have on our results. Third, the possibility of residual confounding cannot be completely ruled out and is exemplified by the attenuation of apparent effects in the fully adjusted model compared to the age and sex-adjusted model. Fourth, the energy adjustment models simulate macronutrient replacement and are less robust than direct observations of interventions in randomized controlled trials[43]. Last, we were not able to establish a causal relationship between macronutrient intake and total mortality given the nature of an observational study.

In conclusion, we found that higher total fat, fat subtypes and protein intake was beneficially associated with the risk of mortality, whereas high carbohydrate intake was associated with increased the risk of death. From these results, it was expected that carbohydrates would provide the least favorable health effect in replacement analyses, but the conventional replacement analysis was not consistent with expected results based on direct associations of each nutrient versus mortality. Specifically, the conventional method did not show any agreement with expected results for any of the 4 nutrients. By contrast, the findings of the joint effect analyses showed agreement with expected results for 3 of 4 nutrients. Given the complexity of modeling the combined effect of nutrients on disease risk, caution should be taken when interpreting findings from energy-adjustment and joint effect analyses, particularly when the population of interest is diverse with a wide range of nutrient intake. Our findings highlight the need for new novel methods to examine the relationship between nutrients in combination and disease risk.

Table 1. Characteristics of the study population at baseline by quintiles of percent energy of saturated fatty acid intake

	Overall (N=135,607)	**Q1 (N=27,339)**	**Q2 (N=27,024)**	**Q3 (N=26,739)**	**Q4 (N=27,056)**	**Q5 (N=27,449)**
Age*	50.0 (42.0-58.0)	49.0 (41.0-57.0)	50.0 (42.0-58.0)	50.0 (42.0-58.0)	50.0 (42.0-58.0)	50.0 (42.0-58.0)
Sex- F (%)	79,071 (58.3)	15,393 (56.3)	15,509 (57.4)	15,836 (59.2)	15,671 (57.9)	16,662 (60.7)
Urban location	71,427 (52.7)	73,72 (27.0)	13,083 (48.4)	17,578 (65.7)	17,347 (64.1)	16,047 (41.5)
Waist-to-hip ratio	0.872 ± 0.0852	0.856 ± 0.0812	0.866 ± 0.0797	0.875 ± 0.0854	0.882 ± 0.0877	0.879 ± 0.0882
Education						
Pre-secondary school	57,572/135,254 (42.6)	15,571/27,212 (57.2)	11,255/26,941 (41.8)	9,443/26,659 (35.4)	10,235/27,011 (37.9)	11,068/27,431 (40.4)
Secondary school	51,829/135,254 (38.3)	10,198/27,2712 (37.5)	11,902/26,941 (44.2)	10,654/26,659 (40.0)	9,405/27,011 (34.8)	9,670/27,431 (35.3)
Post-secondary school	25,853/135,254 (19.1)	1,443/27,212 (5.2)	3,784/26,941 (14.1)	6,562/26.659 (24.6)	7,371/27,011 (27.3)	6,693/27,431 (24.4)
Physical activity						
Low	22,080/126,202 (17.5)	5,493/26,120 (21.0)	5,243/25,549 (20.5)	4,221/25,267 (16.7)	3,789/25,363 (14.9)	3,334/23,858 (14.0)
Moderate	47,932/126,202 (38.0)	9,142/26,120 (35.0)	9,989/25,594 (39.0)	10,315/25,267 (40.8)	9,914/25,363 (39.1)	8,572/23,858 (35.9)
High	56,190/126,202 (44.5)	11,485/26.120 (44.0)	10,362/25,594 (40.5)	10,731/25,267 (42.5)	11,660/25,363 (46.0)	11,952/23,858 (50.1)
Current smoker	28,483 (21.0)	6,882 (25.2)	5,660 (20.9)	5,026 (18.8)	5,031 (18.6)	5,884 (21.4)
Diabetes	9,667/135,390 (7.1)	1,244/27,280 (4.6)	1,734/27,024 (6.4)	2,161/26,679 (8.1)	2,208/27,021 (8.2)	2,320/27,439 (8.5)
Energy intake*	2,011 (1,556-2,596)	1,811 (1,428-2,355)	1,962 (1,519-2,542)	2,005 (1,563-2,564)	2,113 (1,651-2,681)	2,208 (1,685-2,812)
Energy from carbohydrate (%)	61.2 ± 11.6	74.8 ± 8.9	65.6 ± 8.5	60.8 ± 6.9	55.4 ± 6.4	49.7 ± 7.9
Energy from protein (%)	15.2 ± 3.6	12.6 ± 2.6	15.2 ± 2.9	15.9 ± 3.3	16.2 ± 3.4	16.2 ± 4.1
Energy from total fat (%)	23.5 ± 9.3	12.5 ± 7.0	19.5 ± 6.5	23.7 ± 5.1	28.4 ± 4.9	33.6 ± 5.5
Energy from saturated fatty acids (%)	7.6 ± 4.1	2.7 ± 0.9	5.0 ± 0.6	7.1 ± 0.6	9.5 ± 0.8	13.9 ± 2.7
Energy from monounsaturated fatty acids (%)	7.7 ± 3.5	3.4 ± 1.3	6.2 ± 1.7	8.3 ± 2.3	9.9 ± 2.7	10.4 ± 3.5
Energy from polyunsaturated fatty acids (%)	5.0 ± 2.9	1.7 ± 1.4	4.3 ± 2.1	5.6 ± 2.7	6.2 ± 3.1	6.2 ± 3.0

Data are mean ± standard deviation, median (IQR)*, or n/N (%), MET=metabolic equivalents.
Low physical activity: <600 MET min per week, moderate physical activity: 600-3000 MET min per week, high physical activity: >3000 MET min per week
Missing: Education: 353 participants, Physical activity: 9,405 participants, Diabetes: 217 participants, WHR: 7,564 participants

Table 2. Association between quintiles of percent energy from macronutrients and total mortality (N=135,607)

All participants (N=135,607)	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**P-trend**
Percent energy from carbohydrate						
Mortality events (N=6,630)	1,096 (4.00)	1,087 (4.00)	1,135 (4.21)	1,397 (4.87)	2,005 (7.36)	
Median intake (%)	46.46 (42.65-49.08)	54.70 (52.97-56.32)	60.85 (59.36-62.38)	67.67 (65.77-69.71)	77.18 (74.42-80.69)	
Age and sex adjusted	1.00	1.00 (0.91-1.09)	0.92 (0.84-1.01)	1.11 (1.01-1.22)	1.39 (1.26-1.54)	<0.0001
Fully adjusted	1.00	1.08 (0.97-1.19)	1.07 (0.96-1.19)	1.17 (1.05-1.32)	1.28 (1.13-1.44)	<0.0001
Percent energy from total fat						
Mortality events (N=6,630)	1,887 (6.95)	1,342 (5.01)	1,172 (4.35)	1,107 (4.06)	1,122 (4.08)	
Median intake (%)	10.58 (8.07-12.63)	17.99 (16.30-19.63)	21.12 (22.74-25.39)	29.04 (27.81-30.26)	35.25 (33.27-38.24)	
Age and sex adjusted	1.00	0.79 (0.73-0.86)	0.65 (0.59-0.70)	0.64 (0.58-0.70)	0.63 (0.57-0.69)	<0.0001
Fully adjusted	1.00	0.90 (0.83-0.98)	0.81 (0.74-0.90)	0.82 (0.73-0.91)	0.77 (0.68-0.87)	<0.0001
Percent energy from saturated fatty acids						
Mortality events (N=6,630)	1,872 (6.85)	1,311 (4.85)	1,065 (3.98)	1,039 (3.84)	1,343 (4.89)	
Median intake (%)	2.80 (2.02-3.41)	4.95 (4.44-5.48)	7.10 (6.55-7.65)	9.48 (8.83-10.16)	13.16 (11.90-15.10)	
Age and sex adjusted	1.00	0.86 (0.80-0.93)	0.71 (0.65-0.77)	0.67 (0.61-0.73)	0.70 (0.63-0.78)	<0.0001
Fully adjusted	1.00	0.96 (0.89-1.05)	0.89 (0.81-0.98)	0.85 (0.77-0.95)	0.85 (0.75-0.96)	0.0028
Percent energy from monounsaturated fatty acids						
Mortality events (N=6,630)	2,050 (7.48)	1,488 (5.49)	1,198 (4.45)	973 (3.61)	921 (3.39)	
Median intake (%)	3.37 (2.46-4.01)	5.43 (5.00-5.87)	7.29 (6.79-7.82)	9.45 (8.88-10.08)	12.47 (11.54-13.82)	
Age and sex adjusted	1.00	0.90 (0.83-0.96)	0.77 (0.71-0.83)	0.66 (0.60-0.72)	0.67 (0.60-0.74)	<0.0001
Fully adjusted	1.00	1.03 (0.95-1.12)	0.92 (0.84-1.02)	0.82 (0.73-0.91)	0.83 (0.74-0.94)	0.0001
Percent energy from polyunsaturated fatty acids						
Mortality events (N=6,630)	1,641 (6.08)	1,241 (4.59)	1,225 (4.52)	1,230 (4.52)	1,293 (4.75)	
Median intake (%)	2.07 (1.63-2.46)	3.34 (3.08-3.60)	4.40 (4.12-4.70)	5.74 (5.35-6.19)	8.52 (7.49-10.39)	
Age and sex adjusted	1.00	0.81 (0.75-0.88)	0.77 (0.71-0.84)	0.71 (0.66-0.78)	0.64 (0.59-0.70)	<0.0001

Fully adjusted	1.00	0.90 (0.83-0.98)	0.91 (0.83-0.99)	0.86 (0.78-0.94)	0.82 (0.74-0.91)	0.0002
Percent energy from protein						
Mortality events (N=6,630)	2,442 (8.86)	1,464 (5.43)	1,015 (3.76)	802 (2.96)	907 (3.36)	
Median intake (%)	10.79 (9.96-11.50)	13.13 (12.63-13.62)	15.01 (14.56-15.46)	16.87 (16.37-17.43)	19.75 (18.78-21.41)	
Age and sex adjusted	1.00	0.93 (0.86-1.01)	0.83 (0.75-0.90)	0.71 (0.64-0.79)	0.74 (0.66-0.82)	<0.0001
Fully adjusted	1.00	1.03 (0.94-1.12)	0.93 (0.84-1.03)	0.83 (0.74-0.93)	0.83 (0.74-0.94)	<0.0001

Age and sex adjusted: adjusted for age, sex and center (random effect).

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline.

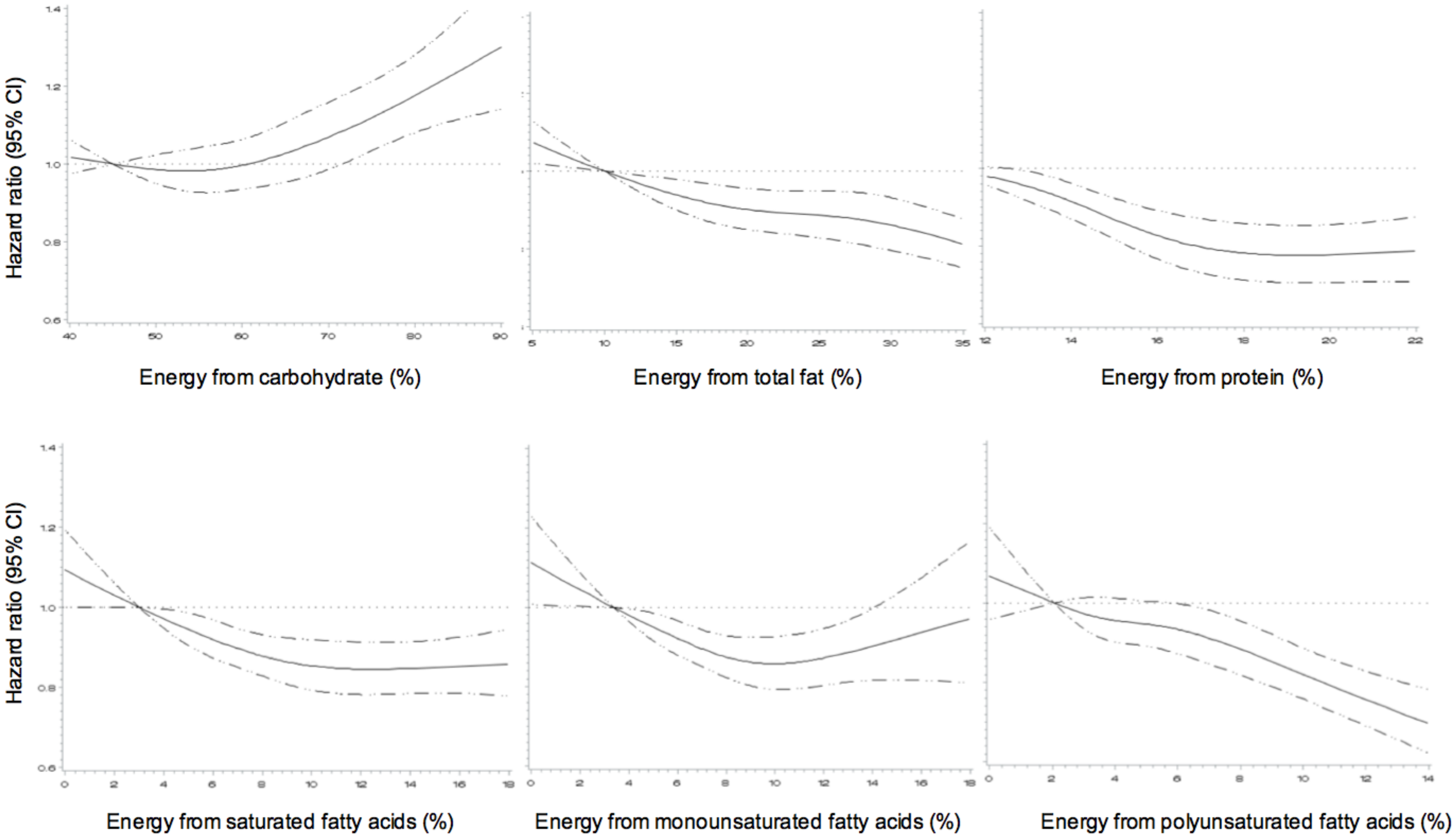


Figure 1. Association between estimated percentage energy from nutrients and total mortality (N=135,607)

Adjusted for age, sex, geographic region, education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline.

Table 3. Risk of mortality associated with isocaloric (3% of energy) replacement of saturated fat with other nutrients (N=135,607)

Nutrient	**Replaced with**	**HR (95% CI)**	**P-value**
Saturated Fat	Monounsaturated fat	0.95 (0.89-1.01)	0.1012
	Polyunsaturated fat	0.94 (0.90-0.98)	0.0075
	Carbohydrate	0.98 (0.96-1.01)	0.1725
	Protein	0.94 (0.90-0.98)	0.0039
	Other	0.99 (0.94-1.05)	0.7368

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline.

Table 4. Comparison of expectant result of conventional nutrient replacement analysis with observed result of energy-adjustment analysis

Nutrient	**Replaced with**	**Expectant result of energy-adjustment analysis based on single nutrient analyses**	**Observed result of energy-adjustment analysis**
Saturated Fat	Monounsaturated fat	Neutral	Beneficial (directionally)
	Polyunsaturated fat	Neutral	Beneficial
	Carbohydrate	Harmful	Beneficial (directionally)
	Protein	Neutral	Beneficial

Table 5. Joint association of saturated and monounsaturated fat intake with risk of mortality (N=135,607)

N	**Death**	**Saturated fat (%E)**	**Monounsaturated fat (%E)**	**HR (95% CI)**	**P-value**
52,168	3,100 (5.94)	<7%	<6%	1.00	
11,988	888 (7.41)	≥7%	<6%	0.95 (0.83-1.09)	0.4741
14,392	615 (4.27)	<7%	≥6%	0.94 (0.84-1.05)	0.2612
57,059	2,027 (3.55)	≥7%	≥6%	0.86 (0.77-0.97)	0.0103

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline, percent energy of carbohydrate. P for interaction=0.5994.

Table 6. Joint association of saturated and polyunsaturated fat intake with risk of mortality (N=135,607)

N	**Death**	**Saturated fat (%E)**	**Polyunsaturated fat (%E)**	**HR (95% CI)**	**P-value**
52,420	2,881 (5.50)	<7%	<5%	1.00	
28,582	1,214 (4.25)	≥7%	<5%	0.89 (0.80-1.00)	0.0418
14,140	834 (5.90)	<7%	≥5%	0.93 (0.84-1.03)	0.1580
40,465	1,701 (4.20)	≥7%	≥5%	0.87 (0.77-0.98)	0.0171

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline, percent energy of carbohydrate. P for interaction=0.3612.

Table 7. Joint association of percent energy from saturated fat and carbohydrate intake with risk of mortality (N=135,607)

N	**Death**	**Saturated fat (%E)**	**Carbohydrate (%E)**	**HR (95% CI)**	**P-value**
54,708	3,194 (5.84)	<7%	≥60%	1.00	
16,688	837 (5.02)	≥7%	≥60%	0.87 (0.78-0.96)	0.0055
52,359	2,078 (3.97)	≥7%	<60%	0.92 (0.84-1.01)	0.0866
11,852	521 (4.40)	<7%	<60%	1.01 (0.87-1.18)	0.8560

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline, percent energy of polyunsaturated fat. P for interaction=0.7898.

Table 8. Joint association of saturated fat and protein intake with risk of mortality (N=135,607)

N	**Death**	**Saturated fat (%E)**	**Protein (%E)**	**HR (95% CI)**	**P-value**
24,522	2,139 (8.72)	<7%	<13%	1.00	
14,673	1,009 (6.88)	≥7%	<13%	0.93 (0.82-1.06)	0.2713
42,038	1,576 (3.75)	<7%	≥13%	0.94 (0.85-1.04)	0.2591
54,374	1,906 (3.51)	≥7%	≥13%	0.85 (0.75-0.96)	0.0111

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline, percent energy of carbohydrate. P for interaction=0.4735.

Table 9. Comparison of expectant and observed results from joint effect analysis

Nutrient	**Replaced with**	**Expectant joint effect analysis result based on single nutrient analyses**	**Observed joint effect analysis result**
Saturated Fat	Monounsaturated fat	Neutral	Neutral
	Polyunsaturated fat	Neutral	Neutral
	Carbohydrate	Harmful	Neutral
	Protein	Neutral	Neutral

Table 10. Comparison of agreement scores between Method 1 (energy-adjustment analysis) and Method 2 (joint effects analysis)

Nutrient	**Replaced with**	**Method 1 agreement between expected versus observed?**	**Method 2 agreement between expected versus observed?**
Saturated fat	Monounsaturated fat	No	Yes
	Polyunsaturated fat	No	Yes
	Carbohydrate	No	No
	Protein	No	Yes
Agreement score		**0 of 4**	**3 of 4**

References

1. WHO. World Health Organization healthy diet fact sheet number 391. 2017.
2. Teicholz N. The Big Fat Surprise: Why Butter Meat & Cheese Belong in a Healthy Diet. New York: Simon & Schuster; 2014.
3. Coronary heart disease in seven countries. Summary. *Circulation* 1970; **41**(4 Suppl): I186-95.
4. de Souza RJ, Mente A, Maroleanu A, et al. Intake of saturated and trans unsaturated fatty acids and risk of all cause mortality, cardiovascular disease, and type 2 diabetes: systematic review and meta-analysis of observational studies. *BMJ (Clinical research ed)* 2015; **351**: h3978.
5. Dehghan M, Mente A, Zhang X, et al. Associations of fats and carbohydrate intake with cardiovascular disease and mortality in 18 countries from five continents (PURE): a prospective cohort study. *Lancet (London, England)* 2017; **390**(10107): 2050-62.
6. Mente A, Dehghan M, Rangarajan S, et al. Association of dietary nutrients with blood lipids and blood pressure in 18 countries: a cross-sectional analysis from the PURE study. *The lancet Diabetes & endocrinology* 2017; **5**(10): 774-87.
7. Cornelis MC, Qi L, Zhang C, et al. Joint effects of common genetic variants on the risk for type 2 diabetes in U.S. men and women of European ancestry. *Annals of internal medicine* 2009; **150**(8): 541-50.
8. Li Y, Ley SH, VanderWeele TJ, et al. Joint association between birth weight at term and later life adherence to a healthy lifestyle with risk of hypertension: a prospective cohort study. *BMC medicine* 2015; **13**: 175.
9. Hu FB, Li TY, Colditz GA, Willett WC, Manson JE. Television watching and other sedentary behaviors in relation to risk of obesity and type 2 diabetes mellitus in women. *Jama* 2003; **289**(14): 1785-91.
10. Bhupathiraju SN, Pan A, Manson JE, Willett WC, van Dam RM, Hu FB. Changes in coffee intake and subsequent risk of type 2 diabetes: three large cohorts of US men and women. *Diabetologia* 2014; **57**(7): 1346-54.
11. Hu FB, Rimm EB, Stampfer MJ, Ascherio A, Spiegelman D, Willett WC. Prospective study of major dietary patterns and risk of coronary heart disease in men. *The American journal of clinical nutrition* 2000; **72**(4): 912-21.
12. Chow CK, Teo KK, Rangarajan S, et al. Prevalence, awareness, treatment, and control of hypertension in rural and urban communities in high-, middle-, and low-income countries. *Jama* 2013; **310**(9): 959-68.
13. Corsi DJ, Subramanian SV, Chow CK, et al. Prospective Urban Rural Epidemiology (PURE) study: Baseline characteristics of the household sample and comparative analyses with national data in 17 countries. *American heart journal* 2013; **166**(4): 636-46.e4.
14. Teo K, Chow CK, Vaz M, Rangarajan S, Yusuf S. The Prospective Urban Rural Epidemiology (PURE) study: examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. *American heart journal* 2009; **158**(1): 1-7.e1.
15. Yusuf S, Islam S, Chow CK, et al. Use of secondary prevention drugs for cardiovascular disease in the community in high-income, middle-income, and low-income countries (the PURE Study): a prospective epidemiological survey. *Lancet (London, England)* 2011; **378**(9798): 1231-43.

16. Dehghan M, Al Hamad N, Yusufali A, Nusrath F, Yusuf S, Merchant AT. Development of a semi-quantitative food frequency questionnaire for use in United Arab Emirates and Kuwait based on local foods. *Nutrition journal* 2005; **4**: 18.
17. Dehghan M, Al-Hamad N, McMillan CR, Prakash P, Merchant AT. Comparison of a semi-quantitative food frequency questionnaire with 24-hour dietary recalls to assess dietary intake of adult Kuwaitis. *Saudi medical journal* 2009; **30**(1): 159-61.
18. Dehghan M, del Cerro S, Zhang X, et al. Validation of a semi-quantitative Food Frequency Questionnaire for Argentinean adults. *PloS one* 2012; **7**(5): e37958.
19. Dehghan M, Ilow R, Zatonska K, et al. Development, reproducibility and validity of the food frequency questionnaire in the Poland arm of the Prospective Urban and Rural Epidemiological (PURE) study. *Journal of human nutrition and dietetics : the official journal of the British Dietetic Association* 2012; **25**(3): 225-32.
20. Dehghan M, Lopez Jaramillo P, Duenas R, et al. Development and validation of a quantitative food frequency questionnaire among rural- and urban-dwelling adults in Colombia. *Journal of nutrition education and behavior* 2012; **44**(6): 609-13.
21. Dehghan M, Martinez S, Zhang X, et al. Relative validity of an FFQ to estimate daily food and nutrient intakes for Chilean adults. *Public health nutrition* 2013; **16**(10): 1782-8.
22. Gunes FE, Imeryuz N, Akalin A, et al. Development and validation of a semi-quantitative food frequency questionnaire to assess dietary intake in Turkish adults. *JPMA The Journal of the Pakistan Medical Association* 2015; **65**(7): 756-63.
23. Merchant AT, Dehghan M, Chifamba J, Terera G, Yusuf S. Nutrient estimation from an FFQ developed for a Black Zimbabwean population. *Nutrition journal* 2005; **4**: 37.
24. Bharathi AV, Kurpad AV, Thomas T, Yusuf S, Saraswathi G, Vaz M. Development of food frequency questionnaires and a nutrient database for the Prospective Urban and Rural Epidemiological (PURE) pilot study in South India: methodological issues. *Asia Pacific journal of clinical nutrition* 2008; **17**(1): 178-85.
25. Iqbal R, Ajayan K, Bharathi AV, et al. Refinement and validation of an FFQ developed to estimate macro- and micronutrient intakes in a south Indian population. *Public health nutrition* 2009; **12**(1): 12-8.
26. Mahajan R, Malik M, Bharathi AV, et al. Reproducibility and validity of a quantitative food frequency questionnaire in an urban and rural area of northern India. *The National medical journal of India* 2013; **26**(5): 266-72.
27. Bellach B, Kohlmeier L. Energy adjustment does not control for differential recall bias in nutritional epidemiology. *Journal of clinical epidemiology* 1998; **51**(5): 393-8.
28. Hu FB, Stampfer MJ, Rimm E, et al. Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. *American journal of epidemiology* 1999; **149**(6): 531-40.
29. Willett W. Nutritional Epidemiology. USA: Oxford University Press; 2013.
30. Mackerras D. Energy adjustment: the concepts underlying the debate. *Journal of clinical epidemiology* 1996; **49**(9): 957-62.
31. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. *American journal of epidemiology* 1986; **124**(1): 17-27.
32. Pike MC, Bernstein L, Peters RK. Re: "Total energy intake: implications for epidemiologic analyses". *American journal of epidemiology* 1989; **129**(6): 1312-5.

33. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. *The American journal of clinical nutrition* 1997; **65**(4 Suppl): 1220S-8S; discussion 9S-31S.
34. Brown CC, Kipnis V, Freedman LS, Hartman AM, Schatzkin A, Wacholder S. Energy adjustment methods for nutritional epidemiology: the effect of categorization. *American journal of epidemiology* 1994; **139**(3): 323-38.
35. Jakobsen MU, O'Reilly EJ, Heitmann BL, et al. Major types of dietary fat and risk of coronary heart disease: a pooled analysis of 11 cohort studies. *The American journal of clinical nutrition* 2009; **89**(5): 1425-32.
36. Thompson J. Nutrition: an applied approach. New York: Pearson; 2017.
37. Pereira MA, O'Reilly E, Augustsson K, et al. Dietary fiber and risk of coronary heart disease: a pooled analysis of cohort studies. *Archives of internal medicine* 2004; **164**(4): 370-6.
38. Jensen MK, Koh-Banerjee P, Hu FB, et al. Intakes of whole grains, bran, and germ and the risk of coronary heart disease in men. *The American journal of clinical nutrition* 2004; **80**(6): 1492-9.
39. Ludwig DS. The glycemic index: physiological mechanisms relating to obesity, diabetes, and cardiovascular disease. *Jama* 2002; **287**(18): 2414-23.
40. Jakobsen MU, Dethlefsen C, Joensen AM, et al. Intake of carbohydrates compared with intake of saturated fatty acids and risk of myocardial infarction: importance of the glycemic index. *The American journal of clinical nutrition* 2010; **91**(6): 1764-8.
41. Paeratakul S, Popkin BM, Kohlmeier L, Hertz-Picciotto I, Guo X, Edwards LJ. Measurement error in dietary data: implications for the epidemiologic study of the diet-disease relationship. *European journal of clinical nutrition* 1998; **52**(10): 722-7.
42. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. *Journal of the National Cancer Institute* 2011; **103**(14): 1086-92.
43. Ramsden CE, Zamora D, Majchrzak-Hong S, et al. Re-evaluation of the traditional diet-heart hypothesis: analysis of recovered data from Minnesota Coronary Experiment (1968-73). *BMJ (Clinical research ed)* 2016; **353**: i1246.

Supplementary appendix

Supplement to: Miller V, Mente A. Comparision of two methods for modeling nutrient replacement in an international population.

Appendix material:

Supplementary methods

Tables

Energy-adjustment Models[1-7]

The Standard Multivariate Model

Disease risk= s1* carbohydrate + s2*monounsaturated fat + s3*polyunsaturated fat + s4*trans-fat + s5*protein + s6*total energy intake

where s1-s6 are regression coefficients and one macronutrient is excluded from the model (saturated fat)

The standard multivariate method models the absolute intake of nutrients and total energy intake while excluding one macronutrient from the model. The regression coefficients are interpreted as the effect of substituting a certain amount of each nutrient for the same amount of nutrient excluded from the model while holding constant the intakes of energy and the other macronutrients.

The Nutrient Residual Model

Disease risk= n1* carbohydrate residual + n2*monounsaturated fat residual + n3*polyunsaturated fat residual + n4*trans-fat residual + n5*protein residual + n6*total energy intake

where n1-n6 are regression coefficients and one macronutrient is excluded from the model (saturated fat)

The nutrient residual method regresses nutrient intake on total energy intake to obtain energy-adjusted nutrient intakes as the residuals and includes the residuals and total energy intake in the model. Similar to the standard multivariate method, the macronutrient being substituted for is not included in the model.

The Energy Partition Model

Disease risk= e1* saturated fat + e2*monounsaturated fat + e3*polyunsaturated fat + e4*trans-fat + e5*protein + e6*carbohydrate

where e1-e6 are regression coefficients and all energy contributing nutrients are included in the model

The energy-partition method is not isocaloric and models the absolute intake of all energy-bearing nutrients simultaneously. The regression coefficient for each nutrient is interpreted as the effect of increasing each nutrient while holding the other nutrients constant.

The Multivariate Nutrient Density Model

Disease risk= m1* carbohydrate + m2*monounsaturated fat + m3*polyunsaturated fat + m4*trans-fat + m5*protein + m6*total energy intake

where m1-m6 are regression coefficients and one macronutrient is excluded from the model (saturated fat)

The multivariate nutrient density method models the nutrients as a percentage of energy, total energy intake and excludes one macronutrient from the model. The regression coefficients have isocaloric interpretation and represent the substitution of equal energy.

Table S1. Joint association of saturated and monounsaturated fat intake with risk of mortality using median cutpoints (N=135,607)

N	**Death**	**Saturated fat (%E)**	**Monounsaturated fat (%E)**	**HR (95% CI)**	**P-value**
54,778	3,314 (6.05)	<7.1%	<7.3%	1.00	
13,170	996 (7.56)	≥7.1%	<7.3%	0.96 (0.84-1.11)	0.6186
12,878	537 (4.17)	<7.1%	≥7.3%	0.94 (0.83-1.07)	0.3226
54,509	1,783 (3.27)	≥7.1%	≥7.3%	0.87 (0.77-0.98)	0.0237

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline, percent energy of carbohydrate.

Table S2. Joint association of saturated and polyunsaturated fat intake with risk of mortality using median cutpoints (N=135,607)

N	**Death**	**Saturated fat (%E)**	**Polyunsaturated fat (%E)**	**HR (95% CI)**	**P-value**
47,689	2,485 (5.21)	<7.1%	<4.4%	1.00	
19,697	780 (3.96)	≥7.1%	<4.4%	0.90 (0.79-1.03)	0.1244
19,967	1,064 (5.33)	<7.1%	≥4.4%	0.91 (0.83-1.01)	0.0713
47,982	2,301 (4.80)	≥7.1%	≥4.4%	0.85 (0.75-0.96)	0.0103

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline, percent energy of carbohydrate.

Table S3. Joint association of percent energy from saturated fat and carbohydrate intake with risk of mortality using median cutpoints (N=135,607)

N	**Death**	**Saturated fat (%E)**	**Carbohydrate (%E)**	**HR (95% CI)**	**P-value**
54,173	3,082 (5.69)	<7.1%	≥60.8%	1.00	
13,478	697 (5.17)	≥7.1%	≥60.8%	0.88 (0.79-0.99)	0.0318
54,201	2,179 (4.02)	≥7.1%	<60.8%	0.94 (0.85-1.05)	0.2782
13,483	672 (4.98)	<7.1%	<60.8%	1.03 (0.89-1.21)	0.6664

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline, percent energy of polyunsaturated fat.

Table S4. Joint association of saturated fat and protein intake with risk of mortality using median cutpoints (N=135,607)

N	**Death**	**Saturated fat (%E)**	**Protein (%E)**	**HR (95% CI)**	**P-value**
41,338	3,063 (7.41)	<7.1%	<15.0%	1.00	
26,323	1,579 (6.04)	≥7.1%	<15.0%	0.90 (0.81-1.00)	0.0405
26,318	850 (3.23)	<7.1%	≥15.0%	0.88 (0.79-0.98)	0.0196
41,356	1,138 (2.75)	≥7.1%	≥15.0%	0.79 (0.71-0.89)	<0.0001

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline, percent energy of carbohydrate.

Table S5. Joint association of saturated, monounsaturated fat and carbohydrate intake with risk of mortality (N=135,607)

N	N Death (%)	Saturated fat (%E)	Monounsaturated fat (%E)	Carbohydrate (%E)	HR (95% CI)	P-value
45,551	2,794 (6.13)	Low	Low	High	1.00	
6,617	306 (4.62)	Low	Low	Low	1.00 (0.82-1.21)	0.0007
4,051	365 (9.01)	High	Low	Low	1.07 (0.90-1.28)	0.4377
5,235	215 (4.11)	Low	High	Low	0.90 (0.75-1.09)	0.2911
48,308	1,713 (3.55)	High	High	Low	0.82 (0.75-0.91)	<0.0001
7,937	523 (6.59)	High	Low	High	0.88 (0.76-1.01)	0.0691
9,157	400 (4.37)	Low	High	High	0.92 (0.81-1.04)	0.1766
8,751	314 (3.59)	High	High	High	0.84 (0.74-0.97)	0.0136

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline.

Table S6. Joint association of saturated, polyunsaturated fat and carbohydrate intake with risk of mortality (N=135,607)

N	N Death (%)	Saturated fat (%E)	Polyunsaturated fat (%E)	Carbohydrate (%E)	HR (95% CI)	P-value
45,259	2,598 (5.74)	Low	Low	High	1.00	
7,161	283 (3.95)	Low	Low	Low	0.91 (0.75-1.10)	0.3340
18,657	734 (3.93)	High	Low	Low	0.84 (0.75-0.95)	0.0040
4,691	238 (5.07)	Low	High	Low	0.97 (0.81-1.17)	0.7356
33,702	1,344 (3.99)	High	High	Low	0.84 (0.76-0.93)	0.0006
9,925	480 (4.84)	High	Low	High	0.88 (0.78-0.99)	0.0382
9,449	596 (6.31)	Low	High	High	0.88 (0.79-0.98)	0.0175
6,763	357 (5.28)	High	High	High	0.76 (0.66-0.87)	0.0001

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline.

Table S7. Joint association of saturated, protein and carbohydrate intake with risk of mortality (N=135,607)

N	**N Death (%)**	**Saturated fat (%E)**	**Protein (%E)**	**Carbohydrate (%E)**	**HR (95% CI)**	**P-value**
23,546	2,017 (8.57)	Low	Low	High	1.00	
976	122 (12.50)	Low	Low	Low	1.36 (1.03-1.81)	0.0323
7,944	531 (6.68)	High	Low	Low	0.90 (0.78-1.04)	0.1620
10,876	399 (3.67)	Low	High	Low	0.86 (0.72-1.02)	0.0762
44,415	1,547 (3.48)	High	High	Low	0.84 (0.75-0.93)	0.0013
6,729	478 (7.10)	High	Low	High	0.92 (0.80-1.05)	0.2126
31,162	1,177 (378)	Low	High	High	0.95 (0.86-1.05)	0.3226
9,959	359 (3.60)	High	High	High	0.76 (0.66-0.88)	0.0002

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline.

Table S8. Joint association of saturated, monounsaturated fat and protein intake with risk of mortality (N=135,607)

N	**N Death (%)**	**Saturated fat (%E)**	**Monounsaturated fat (%E)**	**Protein (%E)**	**HR (95% CI)**	**P-value**
29,961	1,156 (3.86)	Low	Low	High	1.00	
22,207	1,944 (8.75)	Low	Low	Low	1.07 (0.96-1.18)	0.2209
7,021	556 (7.92)	High	Low	Low	0.98 (0.83-1.15)	0.8018
2,315	195 (8.42)	Low	High	Low	1.05 (0.86-1.27)	0.6331
7,652	453 (5.92)	High	High	Low	0.95 (0.81-1.11)	0.4960
4,967	332 (6.68)	High	Low	High	0.97 (0.82-1.15)	0.7041
12,077	420 (3.48)	Low	High	High	0.90 (0.79-1.03)	0.1123
49,407	1,574 (3.19)	High	High	High	0.82 (0.74-0.91)	0.0002

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline.

Table S9. Joint association of saturated, polyunsaturated fat and protein intake with risk of mortality (N=135,607)

N	N Death (%)	Saturated fat (%E)	Polyunsaturated fat (%E)	Protein (%E)	HR (95% CI)	P-value
32,745	1,201 (6.41)	Low	Low	High	1.00	
19,675	1,680 (8.54)	Low	Low	Low	1.10 (0.99-1.22)	0.0717
3,052	264 (8.65)	High	Low	Low	1.17 (0.98-1.40)	0.0760
4,847	459 (9.47)	Low	High	Low	0.97 (0.83-1.12)	0.6425
11,621	745 (6.41)	High	High	Low	0.87 (0.76-1.00)	0.0566
25,530	950 (3.72)	High	Low	High	0.84 (0.76-0.94)	0.0025
9,293	375 (4.04)	Low	High	High	0.96 (0.84-1.10)	0.5973
28,844	956 (3.31)	High	High	High	0.87 (0.78-0.97)	0.0141

Fully adjusted: adjusted for age, sex, center (random effect), education level, urban/rural location, physical activity, waist-to-hip ratio, current smoker, energy intake, diabetes at baseline.

References

1. Hu FB, Stampfer MJ, Rimm E, et al. Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. *American journal of epidemiology* 1999; **149**(6): 531-40.
2. Bhupathiraju SN, Pan A, Manson JE, Willett WC, van Dam RM, Hu FB. Changes in coffee intake and subsequent risk of type 2 diabetes: three large cohorts of US men and women. *Diabetologia* 2014; **57**(7): 1346-54.
3. Willett W. Nutritional Epidemiology. USA: Oxford University Press; 2013.
4. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. *American journal of epidemiology* 1986; **124**(1): 17-27.
5. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. *The American journal of clinical nutrition* 1997; **65**(4 Suppl): 1220S-8S; discussion 9S-31S.
6. Mackerras D. Energy adjustment: the concepts underlying the debate. *Journal of clinical epidemiology* 1996; **49**(9): 957-62.
7. Pike MC, Bernstein L, Peters RK. Re: "Total energy intake: implications for epidemiologic analyses". *American journal of epidemiology* 1989; **129**(6): 1312-5.

DISCUSSION

4.1 Overview of main findings

Most country-specific and global dietary guidelines are based on evidence from North American and European populations. However, it is not unclear whether these findings can be extrapolated to other geographic regions where cultural factors, food availability and affordability, cooking methods, dietary patterns, lifestyle factors, and disease prevalence differ markedly.

The findings from chapter 1 showed that intake of fruits and vegetables are low worldwide, particularly in LICs and MICs. At the community level, the number of different types of fruits and vegetables available for sale was greatest in HICs, intermediate in UMICs, lower in LMICs, and lowest in LICs. The cost of purchasing 2 servings of fruit and 3 servings of vegetables per day accounted for 52% of household income in LICs, 18% in LMICs, 16% in UMICs, and 2% in HICs. Fruit and vegetable recommendations were considered affordable if their daily cost did not exceed 20% of daily income per household member. Using this threshold, fruits and vegetables were not affordable for 58%, 25%, 18%, and <1% of individuals in LICs, LMICs, UMICs and HICs, respectively. Our results show that increased costs of fruits and vegetables relative to income were associated with reduced consumption, highlighting the need for policies that expand availability and affordability of these foods, which may improve the diet quality of many populations, especially in LICs and LMICs.

Chapter 2 presented the association of fruit, vegetable, and legume intake and risk of cardiovascular disease and mortality in 18 countries enrolled in the PURE study. Higher total fruit, vegetable, and legume consumption was inversely associated with cardiovascular, non-cardiovascular, and total mortality, respectively, in the fully adjusted models. Maximum benefit for both non-cardiovascular and total mortality appears at three to four servings per day (equivalent to 375-500 g/day). When examined separately, fruit consumption was inversely associated with cardiovascular, non-cardiovascular, and total mortality, while legume intake was associated with lower risk of non-cardiovascular death and total mortality. For vegetables, raw vegetable consumption was inversely associated with major cardiovascular disease and total mortality, while cooked vegetable intake was associated with total mortality. The findings from this large, international prospective cohort study provide evidence that fruits, vegetables, and legumes are important dietary factors for the prevention of cardiovascular disease and premature death.

The PURE study previously demonstrated that carbohydrate consumption was positively associated with the risk of total mortality, while protein, total fat and fat subtypes (saturated, monounsaturated and polyunsaturated fats) were inversely associated with death. In Chapter 3, we modeled the both the replacement and combined effect of saturated fat with other nutrients on mortality risk. Results from the multivariate nutrient density analysis indicated that replacing percent energy from saturated fat with polyunsaturated fat or protein was significantly associated with a decreased risk of mortality, and no association was observed for replacement with monounsaturated fat or carbohydrate. In the joint effect analyses, higher saturated fat with monounsaturated fat, polyunsaturated fat, protein or carbohydrate, separately, was associated with lower mortality compared to the reference of low intake of both nutrients. The conventional replacement method did not show any agreement with expected results for any of the 4 nutrients. By contrast, the findings of the joint effect analyses showed agreement with expected results for 3 of 4 nutrients. Given the complexity of modeling the combined effect of nutrients on disease risk, caution should be taken when interpreting findings from energy-adjustment and joint effect analyses, particularly when the population of interest is diverse with a wide range of nutrient intake. Our findings highlight the need for new novel methods to examine the relationship between nutrients in combination and disease risk.

4.2 Methodological issues

First, this section discusses the internal validity of the findings reported in this thesis by considering participant selection, measurement of exposure and outcome variables, and the methods used to account for confounding. Second, this section discusses the extent to which the findings are generalizable. Last, issues related to effect modification and statistical analyses are addressed.

Internal validity

Threats to internal validity may occur due to selection, information and confounding bias.

Selection bias

Selection bias refers to a difference in probabilities of individuals being selected and retained in the study[24]. Systematic differences in the characteristics of the participants and the target population, particularly the exposure and the outcome of interest, may result in error in the estimated effect[25]. Selection bias includes volunteer bias (differences in study volunteers and non-volunteers), and loss to follow-up bias (participants remaining in the study differ from those lost to follow-up)[25-27].

The PURE study selected countries and communities to reflect a balance between including countries and communities with heterogeneity in social and economic circumstances and policies, and the feasibility of conducting long-term follow-up. Within each community, the aim was to achieve a broadly representative sample of adults aged 35 to 70 years. The sampling frames used to sample households within urban and rural communities differed between countries and regions. For example, in Canada initial contact was made via a mail out and followed by telephone invitations, while in rural India and China, a community leader made an announcement and subsequent follow-up visits occurred through in-person door-to-door visits of all households. Information on demographic characteristics, health history and socioeconomic status was recorded for households and individuals who refused to participate. Standardization of sampling methods was achieved through in-person and video training workshops, a comprehensive operations manual, standardized report forms and regular communication with study personnel.

Among eligible individuals, participation rates were high (ranged from 45% to 95% by country and urban/rural location), with generally higher response rates in rural areas and middle-income countries. An earlier publication from the PURE study compared characteristics of the study participants with national and World Health Organization data and found similar age, sex and education distributions, and rates of mortality[28]. These findings demonstrate that the PURE participants are a reasonable representation of the enrolled countries. Additionally, rates of follow-up were high (>90% at the 3 and 6-year visits) which minimizes the risk of loss to follow-up bias.

Information bias

Information bias refers to measurement error or misclassification of the exposure and/or outcome variables resulting in distortion of the estimated effect[25,29]. Misclassification of the exposure and/or outcome can be non-differential or differential. Non-differential misclassification occurs when the misclassification is unrelated to the outcome/exposure and the frequency of errors is approximately the same in both groups.[30] Alternatively, differential misclassification occurs when the frequency of misclassification is different for those with and without the disease/exposure[30].

In the PURE study, non-differential misclassification may have occurred during the collection of the participants' diet history. Participants' habitual food intake was recorded using country-specific (region specific in India) validated FFQs at baseline. Semi-quantitative FFQs measure dietary intake through participant recall during the past year and recall errors may occur. However, these errors are unlikely to differ systematically between the groups due to the prospective study design. Our validation studies demonstrated reasonable agreement between the FFQs and 24-hour dietary recalls for fruits (r_s=0.23-0.66) and vegetables (r_s=0.30-0.81) which is in keeping with previous FFQ validation studies[31-36]. Additionally, the de-attenuated correlation coefficients exceeded 0.4 for most nutrients and foods demonstrating moderate to high agreement.

Information on fruit and vegetable availability and cost, and household income was collected systematically. For fruit and vegetable availability and cost, trained research staff conducted a 1 km observation walk within each community and used a binary checklist of types of fruits and vegetables to assess availability. For costs, the non-sale price of 7 types of fruits and vegetables were collected from grocery stores or market places within the walk zone. The assessment was completed independently of information on dietary intake and any error in assessing availability and cost is likely to be non-differential. Household income (measured as income from wage and salary) may be inaccurately reported due to nonresponse, lack of knowledge or recall difficulties, sensitivity, and failure to account for non-monetary sources, property ownership or government assistance programs[37]. Our comparisons between the median income collected for each country and the World Bank per capita gross national income estimates demonstrated high agreement (r=0.91, p<0.001).

Non-differential misclassification of the outcomes may have occurred. However, the number of misclassified events was probably minimal, as most events were ascertained using supporting documents and adjudicated using standardized definitions. Further, outcome ascertainment and adjudication were conducted independently of exposure status and it is unlikely that misclassification differed systematically between groups.

Confounding bias

Confounding bias refers to a distorted estimate of the exposure effect due to the effect of an extraneous variable that is independently associated with both the exposure and outcome and is not an intermediate in the causal chain[25,30,38]. Identifying and statistically adjusting for confounding variables may address confounding bias.

In the analyses conducted, we adjusted for demographic, lifestyle, health history, socioeconomic and dietary history factors including age, sex, urban/rural location, physical activity, smoking status, waist-to-hip ratio, diabetes at baseline, hypertension status, statin use, education level, and energy, red meat, white meat, starch and fruit and vegetable intake. Additionally, we included center as a random effect in the statistical models to account for clustering at the region, country and center levels. In Chapters 2 and 3, we were unable to adjust for environmental factors (use of pesticides and herbicides, and water contamination), methods of cultivation and cooking methods (e.g., frying vs. other methods) that might affect the nutritional quality of foods and nutrients. The possibility of residual confounding from unmeasured or imprecise measurement of covariates cannot be ruled out. This is exemplified by the attenuation of apparent associations in the fully adjusted model compared to the minimally adjusted model for the analyses on fruits, vegetables, and legumes (Chapter 2), and macronutrients (Chapter 3).

External validity

External validity refers to the generalization of an observed causal relationship across different populations, settings, measures and times[39].

The PURE study did not aim for a strict proportionate sampling of the entire world[40]. The number of choice of countries and communities reflected a balance between the feasibility of conducting long-term follow-up and heterogeneity in social and economic conditions. Restricting enrolment to countries with previous collaborations between researchers may limit the generalizability of findings. However, the countries were selected from 7 geographic regions (North America and Europe, South America, the Middle East, South Asia, China, Southeast Asia, and Africa) representing a broad range of countries at various levels of economic development. Since limited data is available for many low-income and middle-income countries, results from the PURE study offer the greatest estimate of the true relationship between the exposures and outcomes examined.

Effect modification

Effect modification is present when the magnitude of the association differs by another variable[41]. In assessing the associations between the exposures and outcomes, most analyses were stratified by either economic (HIC, UMIC, LMIC, and LIC) or geographic region (North America and Europe, South America, the Middle East, South Asia, China, Southeast Asia, and Africa). The interaction terms for these analyses were not statistically significant. This finding may reflect an absence of a true difference in the association across these subgroups, or inadequate power given the relatively small number of events in each subgroup.

Conclusions and Epilogue

Dietary guidelines are based largely on data from North America and Europe and may not be appropriate for non-Western populations. We described the relationship between dietary factors and risk of cardiovascular disease and mortality in 18 countries. Our findings indicate that total fruits, vegetables and legumes, protein and fat intake are inversely associated with the risk of mortality, while carbohydrate intake is positively associated with premature death. Additionally, we found that current recommendations for essential foods such as fruits and vegetables are unaffordable for the majority of individuals in low-income countries. Interventions to improve diet and nutrition recommendations should be tailored to the geographic setting to ensure balance in the feasibility of achieving recommending intake and minimizing disease risk.

Despite the growing emphasis on a healthy diet to prevent the risk of chronic diseases the evidence to support nutritional recommendations in many regions of the world is sparse. The PURE study is unique as it is the only large study from 5 continents (18 high, middle- and low-income countries) examining the association of diet to a range of conditions in middle age and older age. Large studies which include individuals from different regions of the world with prolonged follow-up are needed to accrue sufficient numbers of cases of common conditions to inform global dietary guidelines. Future region and country-specific analyses of diet-disease relationships in the PURE study will complement the global analyses by considering contextual and cultural factors unique to each geographic region and country. Additionally, well designed qualitative studies are needed to understand the food environment and food purchasing behaviors. Collectively, these research efforts may yield important information necessary to derive tailored dietary interventions and recommendations.