Cardiovascular disease risk and dietary fibre intake in the United Kingdom Women's Cohort Study

Diane Erin Threapleton

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Food Science and Nutrition

January 2014

Declaration

I confirm that the work submitted is my own, except where work which has formed part of jointly authored publications has been included. My contribution and the contribution of the other authors to this work has been explicitly indicated below. I confirm that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 2, relating to a systematic literature review and meta-analysis of dietary fibre intake and coronary heart disease, stroke and cardiovascular disease risk incorporates a large part of two jointly authored publications:

- Threapleton, D.E. Greenwood, D.C. Evans, C.E. Cleghorn, C.L. Nykjaer, C. Woodhead, C. Cade, J.E. Gale, C.P & Burley, V.J. 2013. Dietary fiber intake and risk of first stroke: a systematic review and meta-analysis. *Stroke,* 44**,** 1360-8.
- Threapleton, D.E. Greenwood, D.C. Evans, C.E. Cleghorn, C.L. Nykjaer, C. Woodhead, C. Cade, J.E. Gale, C.P & Burley, V.J. 2013. Dietary fibre intake and cardiovascular disease: A systematic review and meta-analysis. *British Medical Journal*, 374: f6879.

A large systematic review relating to all dietary carbohydrates and cardio-metabolic health outcomes was commissioned by the Scientific Advisory Committee on Nutrition (SACN) to update reports published by the Committee on Medical Aspects of Food Policy (COMA, 1991). This work was undertaken within the Nutritional Epidemiology Group at the University of Leeds. I was involved in this project from the start and made academic contributions at each stage of the work. I mention this previous work here as I have conducted an update of part of this review, specifically focusing on dietary fibre intake and cardiovascular disease events.

The meta-analyses presented in Chapter 2 and publications include studies identified both during the initial systematic review (conducted prior to the work presented in this thesis) and also update searches (conducted for the purpose of this thesis). Meta-analyses were conducted by Dr Darren Greenwood and I was not directly involved with combining risk estimates or producing forest plots or restricted cubic spline graphs for the combined study data.

Dr Victoria Burley was the project lead for the main systematic review which examined dietary carbohydrates and cardiometabolic health outcomes. Victoria Burley, Christine Cleghorn and I searched databases for the main systematic review. I conducted update searches. Chris Gale helped develop search strategy terms relating to cardiovascular disease outcomes for the main

review. Article screening was undertaken by Victoria Burley, Christine Cleghorn, Charlotte Evans, Camilla Nykjaer and I. Data extraction was carried out by Victoria Burley, Darren Greenwood, Christine Cleghorn, Charlotte Evans, Camilla Nykjaer and I. Quality of data extraction and checking was carried out by Camilla Nykjaer, Christine Cleghorn, Charlotte Woodhead and I. Statistical analysis was undertaken by Darren Greenwood. I was the lead author responsible for writing and submitting manuscripts and incorporated comments from the above co-authors.

Chapter 5, relating to dietary fibre intake and cardiovascular disease mortality was largely based on a jointly authored publication:

- Threapleton, D.E. Greenwood, D.C. Burley, V.J. Aldwairji, M & Cade, J.E. 2013. Dietary fibre and cardiovascular disease mortality in the UK Women's Cohort Study. *European Journal of Epidemiology,* 28**,** 335-46.

I was jointly responsible for the design of the analysis, along with Janet Cade, Victoria Burley and Darren Greenwood, and I undertook the analysis and interpretation. Darren Greenwood provided statistical advice. Maryam Aldwairji calculated Association of Official Analytical Chemist fibre values for food items. I was the lead author responsible for writing and submitting the manuscript, and incorporated comments from the above co-authors.

Acknowledgments

I owe thanks to many people for their invaluable help, support and advice during the course of this work.

I sincerely thank Professor Janet Cade, Dr Victoria Burley and Dr Darren Greenwood for their trust and belief in my ability and continual encouragement which has kept me motivated month to month. I feel very privileged to have such wonderful supervisors who are both experts in their field and have allowed me the freedom to independently direct my work. Additional thanks are owed to Darren Greenwood for his statistical advice on many methods that were unfamiliar to me and for his significant contribution to the results of meta-analyses, presented in Chapter 2.

Maryam Aldiwarji sourced Association of Official Analytical Chemist fibre values for the foods used in the Food Frequency Questionnaire and I thank her for sharing this information with me. I thank Neil Hancock for his superb ability with Microsoft Access and for helping me to compact complex datasets into a manageable format, thus saving me much time. I also thank Neil for helping to process fibre from key food sources reported in food diaries. I am grateful to Dr Chris Gale for his help in categorising cardiovascular event types from the MINAP and HES datasets.

I acknowledge the time and effort contributed by Undergraduate and Postgraduate students who collectively processed many hundreds of food diaries and thank them for their hard work which allowed me to have access to this information. I also thank all those who have previously contributed to the initiation, data collection, management and processing of information collated for the UK Women's Cohort Study, notably James Thomas for his data management and for creating systems which allowed me to derive fibre values from different food sources.

I am grateful to the thousands of women who participated in this study and who gave up their time to complete questionnaires, in particular the four day food diaries. I thank the participants for allowing this study to trace important disease outcomes and for their contribution to the important data I have been able to access.

Significant thanks are owed to Kellogg's Sales and Marketing UK Ltd who sponsored my studentship. I am very grateful for both their generous contribution, which has permitted this research to be undertaken, and for the independent association between us. I thank them for the ability to present independent findings, since Kellogg's have had no influence on the direction of my work or presentation of results.

I have loved working alongside my colleagues and friends in the Nutritional Epidemiology Group and I thank them all for their support and for the good words of advice and encouragement over many cups of tea during the past three years.

I owe huge thanks to my family for their inexhaustible support during difficulties and to James for his patience, encouragement and confidence in my abilities. I also thank my family, friends and James for all the pep talks, which have elevated my spirits along the way.

The work presented in Chapter 2 has been carried out by a team which has included Dr Victoria Burley, Dr Darren Greenwood, Dr Charlotte Evans, Dr Chris Gale, Christine Cleghorn, Camilla Nykjaer and Charlotte Woodhead. My own contributions are fully and explicitly indicated in the thesis and included: developing search strategies for the initial review and managing update searches, screening articles, data extraction, data extraction quality checking and interpretation of meta-analysis results. The other members of the group and their contributions have been as follows: Victoria Burley was the project lead for the main systematic review concerning dietary carbohydrates and cardiometabolic health outcomes. Victoria Burley and Christine Cleghorn searched databases for the main systematic review. Chris Gale helped develop search strategy terms relating to cardiovascular disease outcomes for the main review. Article screening was undertaken by Victoria Burley, Christine Cleghorn, Charlotte Evans, Camilla Nykjaer and I. Data extraction was carried out by Victoria Burley, Darren Greenwood, Christine Cleghorn, Charlotte Evans, Camilla Nykjaer and I. Quality of data extraction and checking was carried out by Camilla Nykjaer, Christine Cleghorn, Charlotte Woodhead and I. Statistical analysis was undertaken by Darren Greenwood.

Abstract

Background: Dietary fibre has been associated with risk of cardiovascular disease (CVD) in some cohort studies around the world. Key health messages may be created by examining the types or sources of fibre in the diet and associated risk of different CVD events but few studies have explored this.

Methods: I conducted a systematic literature review of published studies investigating dietary fibre intake and CVD. Associations were explored using dose-response meta-analysis in addition to potential non-linear associations. CVD event data for the UK Women's Cohort Study were obtained from death records, hospital episode statistics (HES) and the Myocardial Ischaemia National Audit Project (MINAP). Capture-recapture methods were then applied to estimate the potential for missing cases.

Survival analyses for coronary heart disease (CHD), stroke and total CVD risk in association with total fibre intake and fibre from key food sources were conducted using a cohort approach for food frequency data and case-cohort methods were used for analyses with food diary data.

Results: Meta-analyses broadly supported inverse associations between CVD and fibre intake. Combined data from 9 studies indicate lower CVD risk per 7g/day greater intake in total fibre, relative risk 0.91 (95% confidence intervals (CI) 0.88, 0.94).

After 14 years, 821 CHD and 388 stroke cases were observed. Total fibre, soluble, insoluble and fibre from cereals assessed using FFQs were associated with lower risk of stroke. With each 6g/day higher total fibre intake, hazard ratio (HR) 0.89 (95% CI: 0.81, 0.99). Higher fibre density was associated with lower risk of myocardial infarction, for every 2g/1000kcal/day higher intake, HR 0.89 (95% CI: 0.81, 0.98). Higher cereal fibre intake, calculated using food diaries, was associated with lower risk of acute coronary events HR 0.76 (95% CI: 0.58, 1.00).

Conclusion: Fibre intake is inversely associated with CVD risk in a dose response relationship after accounting for other potentially confounding influences. Associations were stronger for stroke risk, when the source of fibre was cereals and in those without hypertension.

Contents

List of figures 16

List of tables 18

Abbreviations 22

Chapter 1 Introduction and Objectives 24

1.1 Overview of Cardiovascular Disease 24

1.1.1 Cardiovascular disease in the UK 24

1.1.2 Cardiovascular disease types and pathogenesis 25

1.1.3 Women and cardiovascular disease 26

1.1.4 Risk factors for cardiovascular disease 27

1.1.4.1 Diet as a risk factor 29

1.2 Dietary fibre 30

1.2.1 Constituents of dietary fibre 30

Non-digestible oligosaccharides 30

Non-starch polysaccharides 31

Resistant starch 31

Lignin 32

1.2.2 Dietary fibre chemical analysis methods 32

1.2.3 UK and international definitions of 'dietary fibre' 33

1.2.4 UK and international intakes of dietary fibre and recommendations 34

1.3 Potential mechanisms for fibre and CVD risk factors 35

1.3.1 Circulating lipid levels 35

1.3.2 Overweight 36

1.3.3 Blood pressure 37

1.3.4 Glucose metabolism and diabetes 37

1.3.5 Inflammation 38

1.4 Gaps in current epidemiological research 38

1.5 Aims 39
1.6 Summary 40
Chapter 2 Systematic literature review and meta-analysis 41
2.1 Chapter overview 41
2.2 Background 42
2.3 Methods 44
2.3.1 Literature searches 44
Main review: 44
Update: 45
2.3.2 Screening articles 45
Main review: 45
2.3.3 Data Extraction 46
2.3.4 Statistical analysis 46
2.4 Results 48
2.4.1 Included articles 48
2.4.2 Meta-analyses and comparison across similar studies 56
Total dietary fibre 57
Insoluble fibre 60
Soluble fibre 61
Cereal fibre 64
Vegetable fibre 65
Fruit fibre 66
Other sources of fibre 67
Meta-regression 68
2.5 Discussion 71
2.5.1 Result summary 71
2.5.2 Findings in context, other published reviews 72
2.5.3 Limitations 73
2.5.4 Strengths 75

2.6 Summary 75

Chapter 3 Methods, the UK Women's Cohort Study and dietary fibre intake 77

3.1 Chapter overview 77

3.2 Method 77

3.2.1 Study population 77

3.2.2 Dietary assessment: Food Frequency Questionnaire 79

FFQ calculations for fibre from food sources 80

Correlation of baseline FFQ fibre intakes, by type 81

Characteristics of low and high fibre consumers as estimated from FFQ 83

Exploring linear dose-response associations with CVD for each fibre exposure 87

3.2.3 Dietary assessment: Food Diary 88

Characteristics of responders to follow-up data collection 88

Diary data quality and cleaning 89

Food diary calculations for fibre from food sources 90

Correlation of food diary fibre intakes 91

Characteristics of low and high fibre consumers as estimated from food diaries 93

Exploring linear dose-response associations with CVD, for each fibre exposure 98

3.3 FFQ values compared to food diaries 99

3.4 Discussion 100

3.4.1 How FFQs compare to diary methods 100

3.4.2 Fibre intakes in UKWCS compared to other study populations 100

3.4.3 Limitations and strengths of dietary assessment methods 102

3.5 Summary 104

Chapter 4 Sources of cardiovascular event data and case ascertainment rates 105

4.1 Chapter overview 105

4.2 Background 105

4.2.1 Ethical considerations and approval for access to data 106

4.2.2 Data security and anonymity 108

4.3 Methods 109

4.3.1 Overview of the three sources of outcome data 109
4.3.2 Mortality data 110
Source summary 110
Case definitions 110
Data summary 110
4.3.3 Myocardial Ischaemia National Audit Project 111
Source summary 111
Case definitions 111
Data cleaning and summary 112
4.3.4 Hospital Episode Statistics 113
Source summary 113
Case definitions 113
Data cleaning and summary 117
4.3.5 Capture-recapture methodology 121
Three-source approach 121
Two-source approach 123
4.4 Results 124
4.4.1 Case cross-matches 124
4.4.2 Estimation of total CHD cases 126
4.4.3 Estimation of total stroke cases 129
4.4.4 Estimation of total ACS cases from HES and MINAP 129
4.5 Discussion 132
4.5.1 Summary of findings 132
4.5.2 Strengths and limitations 134
Model assumptions 134
Sources of data 135
4.6 Summary 136
Chapter 5 Dietary fibre and fatal ischaemic heart disease, stroke and cardiovascular disease 138
5.1 Chapter overview 138

5.2 Background 139
5.3 Method 140
5.3.1 Dietary data 140
5.3.2 Mortality data 141
5.3.3 Exclusions 141
5.3.4 Testing dose-response and non-linear associations 141
5.3.5 Descriptive statistics 142
5.3.6 Survival analyses 142
5.3.7 Checking proportional hazards 142
5.3.8 Confounder adjustment 144
5.3.9 Subgroup analyses 147
5.3.10 Statistical significance 148
5.4 Results 148
5.4.1 Descriptive statistics 148
Case numbers 152
5.4.2 Survival analysis 152
Full sample results-overview 153
Modelling risk in high and low intakes compared to the mid-intake group 153
Total NSP and AOAC fibre intake and fibre density 154
Soluble and insoluble fibre 154
Fibre from food sources 154
Subgroup analyses 155
5.5 Discussion 161
5.5.1 Result summary 161
5.5.2 Comparison with other published studies 162
5.5.3 Strengths and limitations 163
5.6 Summary 165
Chapter 6 Dietary fibre intake and risk of total CVD, non-fatal CVD and CVD subtypes 166

6.1 Chapter overview 166
6.2 Background 167
6.3 Method 168
6.3.1 Dietary data 168
6.3.2 Mortality data 169
6.3.3 HES records 169
6.3.4 MINAP records 169
6.3.5 Censor date 169
6.3.6 Exclusions 170
6.3.7 Testing dose-response and non-linear associations 170
6.3.8 Survival analyses 171
6.3.9 Confounder adjustment 172
6.3.10 Cohort subgroup analyses 172
6.3.11 Case subtype analyses 173
6.4 Results 173
6.4.1 Sample and case numbers 173
6.4.2 Descriptive statistics 177
6.4.3 Survival analysis 183
Total CHD, stroke and CVD incidence (fatal plus non-fatal events) 183
Non-fatal CHD, non-fatal stroke and non-fatal CVD 188
Haemorrhagic, ischaemic and unspecified stroke incidence (fatal plus non-fatal events) 193
Risk of MI, ACS and chronic heart disease (fatal plus non-fatal events) 200
6.5 Discussion 204
6.5.1 Total CVD, CHD and stroke 204
6.5.2 Non-fatal cardiovascular disease 207
6.5.3 Haemorrhagic vs. ischaemic stroke 208
6.5.4 Acute vs. chronic heart disease 210

6.5.5 Strengths and limitations....210
6.6 Summary....212
Chapter 7 Dietary fibre intake and risk of cardiovascular disease, a case-cohort approach....213
7.1 Chapter overview....213
7.2 Background....214
7.3 Methods....214
7.3.1 Dietary data....214
7.3.2 Mortality data....215
7.3.3 HES records....215
7.3.4 MINAP records....215
7.3.5 Censor date....216
7.3.6 Exclusions....216
7.3.7 Testing dose-response and non-linear associations....216
7.3.8 Case-cohort analysis and selection of sub-cohort participants....216
7.3.9 Survival analyses using the case-cohort method....217
7.3.10 Confounder adjustment....218
7.3.11 Cohort subgroup analyses....218
7.4 Results....218
7.4.1 Sample and case numbers....218
7.4.2 Descriptive statistics....219
7.4.3 Survival analysis....223
Incident MI or ACS....223
Fatal IHD, Stroke and CVD....223
7.5 Discussion....226
7.5.1 Result summary....226
7.5.2 Findings compared with previous chapters....227
7.5.3 Results compared to other studies....228

7.5.4 Strengths and limitations of the case-cohort approach 230

7.5.5 Strengths and limitations of dietary and covariate assessment 231

7.6 Summary 232

Chapter 8 Summary and Conclusions 234

8.1 What was already known about fibre and CVD risk? 234

8.2 What this work has added 234

8.3 Thesis result summary 235

8.4 Summary discussion 237

8.5 Strengths and limitations 240

8.6 Recommendations 242

8.6.1 In the cohort 242

8.6.2 Other studies 243

8.6.3 Public health messages 245

8.7 Conclusions 247

References 249

Appendices 266

Appendix I: Medline search strategy 266

Appendix II: Structured guidelines for article screening 268

Appendix III: Structured flow chart for article inclusion 269

Appendix IV: Individual study results for studies included in Chapter 2 systematic literature review and meta-analyses 271

Appendix V: FFQ items contributing to sources of fibre 291

Appendix VI: Research Ethics Committee approval 292

Appendix VII: National Research Ethics Committee Section 251 approval 295

Appendix VIII: Definition for final diagnosis field in the MINAP dataset 297

List of figures

Figure 1.1 Overview of Cardiovascular Disease Sub-types 25

Figure 2.1 Flow chart for update search publication identification and inclusion 49

Figure 2.2 Main plus update search flow diagram for publication identification and inclusion for CHD, stroke and CVD outcomes combined 50

Figure 2.3 Forest plots and restricted cubic spline figures for total fibre intake and CVD, CHD and stroke 59

Figure 2.4 Forest plots and restricted cubic spline figures for insoluble and soluble fibre 63

Figure 2.5 Forest plots and restricted cubic spline figures for cereal fibre 65

Figure 2.6 Forest plots and restricted cubic spline figures for fruit and vegetable fibre 67

Figure 3.1 Recruitment of UKWCS participants 78

Figure 3.2 Example section of baseline FFQ form 79

Figure 3.3 Correlation between NSP and AOAC fibre, estimated from FFQ 82

Figure 3.4 Correlation between NSP intake from fruit and cereal foods, estimated from FFQ .. 82

Figure 3.5 Percent contributions of food groups calculated from FFQ to total NSP intake in all participants and NSP intake quintiles (as presented in Table 3.2) 85

Figure 3.6 Correlation between NSP intake and NSP density estimated from food diaries 92

Figure 3.7 Correlation between fruit and vegetable fibre estimated from food diaries 92

Figure 3.8 Percent contributions of food groups calculated from food diary to total NSP intake (as presented in Table 3.7) 94

Figure 4.1 Process of obtaining ethical approval for access to and linkage with cardiovascular event data 108

Figure 4.2 Time periods covered by different case datasets 109

Figure 4.3 MINAP case frequency reported each quarter (Q) since 1st January 2003 to 30th June 2011, for English residents 113

Figure 4.4 Source overlap including only the main HES diagnosis field 125

Figure 4.5 Source overlap including all HES diagnosis fields 125

Figure 5.1 Data sources used in this chapter: Dietary data from FFQs and mortality (CHD, stroke and CVD) data spanning from study baseline..... 139

Figure 5.2 Survival function for IHD mortality in those consuming lower and higher NSP density diets..... 143

Figure 5.3 Example of simplified directed acyclic graph..... 144

Figure 5.4 Causal diagram for the variables associated with fibre and CVD 145

Figure 6.1 Data sources used in this chapter: Dietary data from FFQs, mortality data spanning from study baseline, HES data since 1998 and MINAP since 2003..... 167

Figure 7.1 Data sources used in this chapter: Dietary data from case and sub-cohort food diaries, mortality and HES cases identified after receipt of food diaries plus MINAP cases identified since 2003..... 213

List of tables

Table 2.1 Summary of outcomes and exposures reported in cohort studies identified during the systematic review and update searches 51

Table 2.2 Characteristics of cohort studies identified during the systematic review and update searches 53

Table 2.3 Total fibre intake and stroke risk by stroke subtype 60

Table 2.4 Study subgroup combined risk estimates for total fibre intake and CVD, CHD and stroke 70

Table 3.1 Correlation between fibre types and sources of fibre assessed using FFQ 81

Table 3.2 Characteristics of all participants, groups based on total NSP intake (FFQ) 84

Table 3.3 Characteristics of all participants, groups based on NSP density of the diet (FFQ) 86

Table 3.4 Median intakes and mean difference between categories for different dietary exposures estimated from the FFQ 87

Table 3.5 Baseline characteristics of participants that provided complete dietary and lifestyle information at phase 2, follow-up. 89

Table 3.6 Correlation between total and sources of fibre assessed using food diaries 91

Table 3.7 Characteristics of all case and sub-cohort participants who meet the UK dietary NSP recommended intake level, as assessed using food diaries 95

Table 3.8 Characteristics of all case and sub-cohort participants, grouped based on total NSP intake, as assessed using food diaries 96

Table 3.9 Characteristics of all case and sub-cohort participants, grouped based on NSP density of the diet, as assessed using food diaries 97

Table 3.10 Median intakes and mean difference between categories for different dietary exposures estimated from food diaries 98

Table 3.11 Energy and fibre intakes estimated from FFQs and food diaries 99

Table 4.1 Fatal IHD, stroke and CVD case numbers used in examining associations between dietary fibre and CVD mortality and CVD incidence 111

Table 4.2 Final diagnosis frequency and outcome classification in all MINAP records and for English participants only 112

Table 4.3 ICD10 codes used to identify relevant HES records 114

Table 4.4 Coronary heart disease classification codes and outcome grouping..... 115

Table 4.5 Stroke classification codes and outcome grouping..... 116

Table 4.6 Proportion of participants with no CVD-related HES records, by region..... 118

Table 4.7 Example dataset displaying multiple diagnoses per participant..... 119

Table 4.8 Row and case frequency since study baseline of different CVD outcomes using the primary and multiple diagnosis fields from HES..... 119

Table 4.9 Row and case frequency since study follow-up for each CVD outcome, calculated using the primary HES diagnosis field..... 120

Table 4.10 Three-way cross over table illustrating all potential combinations of case reporting from the three event data sources..... 122

Table 4.11 Eight models fitted for three-source data, from independent to dependant lists . 122

Table 4.12 CHD case capture and overlap within 3 sources..... 124

Table 4.13 Stroke case capture and overlap within 2 sources 125

Table 4.14 ACS case capture and overlap within HES and MINAP over different timeframes. 126

Table 4.15 Assessing best model fit for list dependency: Estimation of expected missing cases and total cases estimated from each of eight interaction models, using ACS cases identified from the main HES diagnosis field..... 128

Table 4.16 Assessing best model fit for list dependency: Estimation of expected missing cases and total cases estimated from each of eight interaction models, using ACS cases identified from all HES diagnosis fields 128

Table 4.17 Observed and expected cell counts and case ascertainment rates for CHD events reported by mortality records, MINAP and HES..... 130

Table 4.18 Observed and expected cell counts and case ascertainment rates for stroke events reported by mortality records and from the primary and all HES diagnosis fields 130

Table 4.19 Observed and expected cell counts and case ascertainment rates for ACS in HES and MINAP 131

Table 5.1 Correlation between continuous variables..... 146

Table 5.2 Baseline cross-sectional characteristics in those excluded because of history of stroke and heart-attacks, those subsequently suffering a fatal stoke or IHD and those women still alive at censor date or whose cause of death was not attributed to CVD 150

Table 5.3 Dietary and lifestyle characteristics across increasing NSP quintiles, after applying exclusions to the sample 151

Table 5.4 IHD, stroke and CVD case numbers in full sample and subgroup analyses 152

Table 5.5 Cardiovascular mortality risk assessed using both categorised fibre intake and by fitting a linear dose-response trend, using fibre as a continuous variable 156

Table 5.6 Risk of fatal IHD, stroke and CVD in postmenopausal women, using continuous fibre variables 159

Table 5.7 Risk of fatal IHD, stroke and CVD in those with healthy BMI at baseline (18.5-24.9kg/m^2), using continuous fibre variables 159

Table 5.8 Risk of fatal IHD, stroke and CVD in those with BMI classed as overweight or obese at baseline (≥25kg/m^2), using continuous fibre variables 160

Table 5.9 Risk of fatal IHD, stroke and CVD in those without history of hypertension or angina at baseline, using continuous fibre variables 160

Table 6.1 Cohort participant and total (fatal plus non-fatal) CHD, stroke and CVD case frequency (%) included in unadjusted and fully adjusted models 175

Table 6.2 Cohort participant and stroke sub-type case frequency (%) included in unadjusted and fully adjusted models 175

Table 6.3 Participant and CHD sub-type case frequency (%) included in unadjusted and fully adjusted models 176

Table 6.4 Participant and non-fatal CHD, stroke and CVD case frequency (%) included in unadjusted and fully adjusted models 176

Table 6.5 Baseline cross-sectional characteristics for fatal plus non-fatal stroke and stroke subtype cases, non-fatal stroke cases and non cases 178

Table 6.6 Baseline cross-sectional characteristics for fatal plus non-fatal CHD and CHD subtype cases, non-fatal CHD cases and non cases 180

Table 6.7 Baseline cross-sectional characteristics for non-fatal and total CVD (fatal plus non-fatal) cases in addition to non-cases 182

Table 6.8 Dietary fibre intake and associated risk for total CHD, total stroke and total CVD .. 185

Table 6.9 Dietary fibre intake and associated risk for non-fatal CHD, non-fatal stroke and non-fatal CVD 190

Table 6.10 Dietary fibre intake and associated risk for haemorrhagic, ischaemic and 'unspecified' stroke..... 196

Table 6.11 Dietary fibre intake and associated risk for 'mostly ischaemic' type stroke (ischaemic type plus unspecified stroke)..... 199

Table 6.12 Dietary fibre intake and associated risk for MI, ACS or chronic heart disease 201

Table 7.1 Frequency of available participants and mean follow-up duration in both sub-cohort and case groups 219

Table 7.2 Follow-up assessment cross-sectional characteristic for MI, ACS, IHD mortality, stroke mortality cases and randomly selected sub-cohort participants 221

Table 7.3 Total dietary fibre and fibre density of the diet and associated risk of total MI/ACS or fatal IHD, Stroke and CVD 224

Table 7.4 Fibre from cereals and breakfast cereals and associated risk of total MI/ACS or fatal IHD, Stroke and CVD 224

Table 7.5 Fibre from fruit and vegetables and associated risk of total MI/ACS or fatal IHD, Stroke and CVD 225

Table 7.6 Fibre from legumes and nuts/seeds and associated risk of MI/ACS or fatal IHD, Stroke and CVD..... 225

Abbreviations

ACS	Acute coronary syndrome
AIC	Akaike's information criterion
ANOVA	Analysis of variance
AOAC	Association of Official Analytical Chemists
ATBC	Alpha-tocopherol Beta-carotene Study
BMI	Body Mass Index
CHD	Coronary Heart Disease
CI	Confidence Interval
CVD	Cardiovascular Disease
DAAG	Data Access Advisory Group
DAG	Directed Acyclic Graph
ECC	Ethics and Confidentiality Committee
EFSA	European Food Safety Authority
EPIC	European Prospective Investigation into Cancer and Nutrition
EU	European Union
FFQ	Food Frequency Questionnaire
GI	Glycaemic Index
GP	General Practitioner
HDL-C	High-density lipoprotein cholesterol
HES	Hospital Episode Statistics
HPFS	Health Professionals Follow-up Study
HR	Hazard Ratio
ICD	International Classification of Disease codes
ID	Identification
IHD	Ischaemic Heart Disease
IQR	Interquartile range
LDL-C	Low-density lipoprotein cholesterol
MET	Metabolic Equivalent of Task
MI	Myocardial Infarction
MINAP	Myocardial Ischaemia National Audit Project
NDNS	National Diet and Nutrition Survey
NHANES	National Health and Nutrition Examination Survey
NHS	National Health Service
NHSIC	National Health Service Information Centre
NoSA	Notice of substantial Amendment
NIGB	National Information Governance Board
NIH-AARP	National Institutes of Health- American Association of Retired Persons
NRES	National Research Ethics Service
NSP	Non-starch Polysaccharide
NS-SEC	National Statistics Socio-Economic Classification
ONS	Office of National Statistics
RCT	Randomised Controlled Trial
REC	Research Ethics Committee
RR	Relative Risk

SACN	Scientific Advisory Committee on Nutrition
SCFA	Short-chain Fatty-acid
SCORE	Systematic Coronary Risk Evaluation
SD	Standard Deviation
SES	Socio-economic Status
UK	United Kingdom
UKWCS	United Kingdom Women's Cohort Study
US	United States
WCRF	World Cancer Research Fund

Chapter 1 Introduction and Objectives

1.1 Overview of Cardiovascular Disease

Cardiovascular disease (CVD) is a broad term for conditions that affect the heart and blood vessels such as coronary heart disease (CHD), cerebrovascular disease and peripheral arterial disease. Acute events are typically classed as either coronary or cerebrovascular (WHO, 2013) and vessels are either occluded by a process involving atherosclerotic plaque ruptures that cause a clotting cascade (described below) or vessels may rupture. Both occlusion and haemorrhage result in downstream tissue damage, with significant consequences. Atherosclerotic CVD is a chronic condition that develops throughout the life course and is normally advanced by the time symptoms present. CVD remains the primary cause of premature death within Europe but over 80% of total CVD mortality worldwide occurs in developing countries (Perk et al., 2012).

1.1.1 Cardiovascular disease in the UK

The past few decades have brought declining incidence and mortality rates for CVD in many developed countries (AHA, 2011, Allender et al., 2008, Roger et al., 2012, Pearson-Stuttard et al., 2012). For example, the age standardised death rates (for deaths under 65 years) in the United Kingdom (UK) have fallen from 143/100,000 in 1980 to 33/100,000 in 2009 in males and from 36/100,000 to 8/100,000 in females (Nichols et al., 2012). Prevalence rates in the UK have not followed the same trajectory of decline and have increased over the past few decades, peaking in the early to mid 2000's. Since 2002, prevalence rates have remained relatively stable in men and a slight decrease in the prevalence rate has been observed in women (Townsend et al., 2012). Today, there are an estimated 600,000 individuals of each sex in the UK who have suffered a stroke and for CHD the prevalence rates are even greater, with over 1.6 million males and over 1 million females having CHD (Townsend et al., 2012). CHD mortality rates vary greatly between countries and substantial changes over short periods of time reflect changing environmental rather than genetic factors. CVD mortality rates in the UK are not declining as fast as in some other developed countries (Capewell et al., 2008).

Despite impressive improvements in CVD incidence, it remains a significant social and financial burden with over 3 million people in the UK living with CVD (Townsend et al., 2012) and annual costs of CVD are estimated to exceed £30 billion in the UK (Allender et al., 2008). The burden of disease is set to increase over coming decades, with the growing size of the ageing

population in the UK (Capewell et al., 2008). Additionally, the burden of disease remains unequal across the socioeconomic gradient in the UK; for example the risk of myocardial infarction (MI) hospital admission in patients under 55 years in the most deprived quintile of the UK was double that in the least deprived quintile (Pearson-Stuttard et al., 2012). Uptake of treatments for CVD are however equitable across social groups and the disparity is therefore likely related to differences in major CVD risk factors (Pearson-Stuttard et al., 2012).

1.1.2 Cardiovascular disease types and pathogenesis

Strokes and acute coronary events occur when blood flow to cardiac or cerebral tissue is disrupted causing myocardial damage or neurological deficits. Coronary artery disease manifests as angina, silent ischaemia, unstable angina, myocardial infarction (MI), arrhythmias, heart failure and sudden death (Grech, 2003). Strokes can be grouped into two main types, the more common ischaemic type which is caused by a vessel blockage and the less common haemorrhagic type which results from a ruptured vessel (Frizzell, 2005) (Figure 1.1). Ischaemic strokes are either caused by atherosclerosis within vessels or from a thrombus (blood clot) that often originates from the heart region. It is estimated that approximately 20% of strokes are cardiothrombotic and occur after recent MI, as a result of atrial fibrillation or the thrombus may originate in the aortic arch or carotid arteries (Frizzell, 2005).

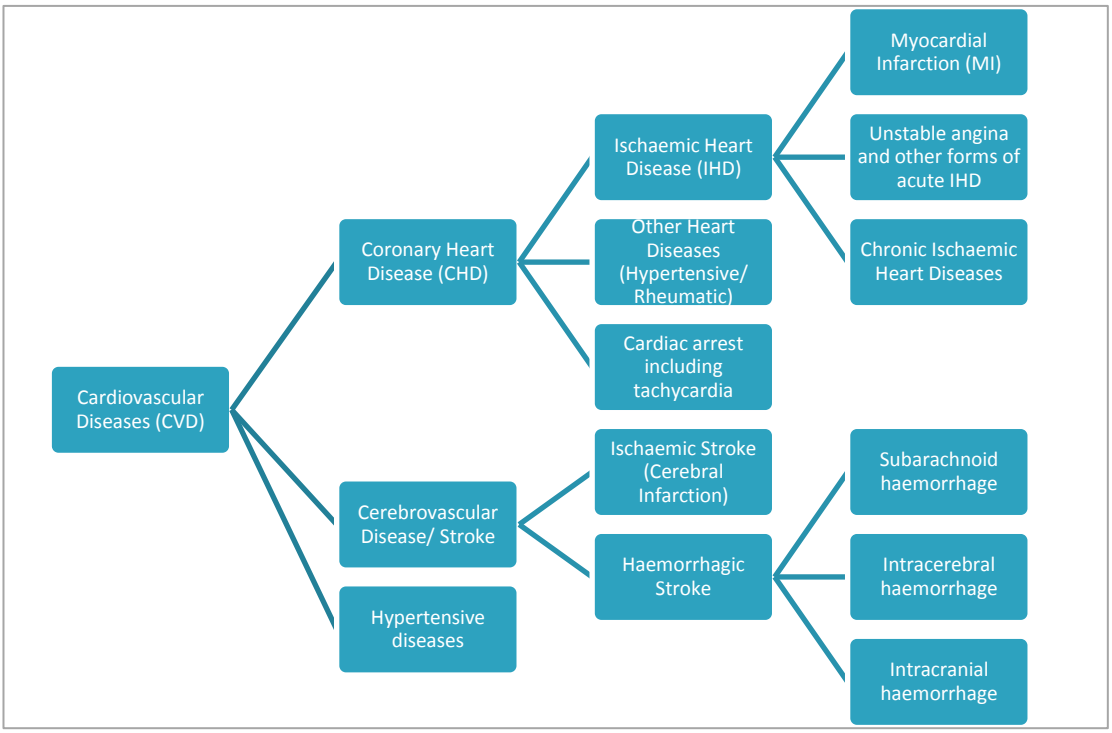


Figure 1.1 Overview of Cardiovascular Disease Sub-types

Coronary artery disease is almost always due to atheromatus narrowing and subsequent occlusion of vessels (Grech, 2003) and the same atherosclerotic plaques in cerebrovascular vessels are responsible for the majority of stroke events (Frizzell, 2005). Athoerosclerosis is a dynamic process that develops and worsens over several decades. The sequence of atherosclerotic plaque development is covered in detail by George and Lyon (George and Lyon, 2010) and is briefly summarised below:

1) Endothelial dysfunction, caused by many CVD risk factors, results in lipids and inflammatory cells being allowed into the artery wall.
2) Monocytes differentiate into macrophages and these then take in the excess lipids to become foam cell macrophages.
3) Engorged with lipids, the foam cells begin to die, resulting in the formation of a necrotic core within lesions. The release of cytoplasmic contents of the necrotic foam cells causes extracellular lipid accumulation and growth factors that cause inflammation.
4) Vascular smooth muscle cells proliferate and migrate to form a fibrous cap around lesions that protrude into the vessel lumen.
5) Plaques may rupture, triggering a cascade that produces a thrombus and this can partially or completely occlude blood flow (George and Lyon, 2010).

Endothelial damage, inflammation and excess lipids are the triggers for atherosclerosis, one of the main causes of CVD development and these factors are all influenced through modifiable lifestyle behaviour such as diet. Potential mechanisms for the effect of dietary fibre on these risk factors for CVD are described below (Section 1.3).

1.1.3 Women and cardiovascular disease

Heart disease was once considered as a ‘man’s disease’ but significant changes over the past few decades have debunked this myth (Mosca et al., 2011) and now there is recognition of the importance of CVD in women (Stranges and Guallar, 2012). Differences in CVD progression and age of onset are observed between the sexes (Novella et al., 2012) but CVD remains as much a serious concern for women as men. There are suggestions that risk of CHD in women accelerates during the menopause but recent work suggests that it may be the deceleration in rates of male CHD mortality and not an acceleration in females that explains the apparent imbalance between sexes in age of CVD onset (Lam and Little, 2012, Vaidya et al., 2011).

It has been suggested that this difference in risk is partly due to hormonal changes during the menopause, potentially the loss of vascular protective effects exerted by oestrogen (Novella et al., 2012). However, a study that employed data from the UK and US to model CHD risk identified that heart disease mortality increased in women of all ages, with no additional notable increase around menopausal age. The concern about CVD risk in women is therefore justified but focus should be placed on overall lifetime risk rather than increased risk surrounding the menopause (Vaidya et al., 2011). Additionally, suggestions that changing hormone levels throughout menopause are responsible for this increased risk are disputed in a recent narrative synthesis of the menopause and CHD (Barrett-Connor, 2013). This work indicates that age-related changes in weight, blood pressure, cholesterol and waist circumference, may determine hormonal changes and the age of menopause and not vice versa. A systematic review of observational studies also indicated risk of CVD was not greater in postmenopausal compared to premenopausal women after controlling for age and smoking, although there was a high degree of heterogeneity between the pooled studies (Atsma et al., 2006). Despite the lack of an apparent increase in heart disease mortality in women at menopausal ages, risk continued to increase exponentially as age increased (Vaidya et al., 2011) and while women appear to be at lower CVD risk than men, this is misleading as risk is deferred by 10 years rather than avoided (Perk et al., 2012).

Women have historically been underrepresented in randomised controlled trials (RCT) of lifestyle and pharmacological interventions for prevention of CVD but there has been a rebalance over recent years and women are now well represented in clinical studies (Stranges and Guallar, 2012). Clinical studies are however limited in that they only provide information on short-term prevention for CVD, while recent guidelines for the prevention of CVD in women place emphasis on overall lifetime risk and prevention of risk factors throughout the life course (Mosca et al., 2011). As a consequence, observational study data and extrapolations from short-term trials will need to be relied upon for preventative strategies for CVD (Stranges and Guallar, 2012).

1.1.4 Risk factors for cardiovascular disease

As discussed above, sex is a general risk factor for CVD and influences the age of symptom onset and potentially mediates disease development. Other non-modifiable risk factors include age and family history of premature CVD (Mosca et al., 2011). However, CVD risk is considered largely modifiable (Stampfer et al., 2000, Mosca et al., 2011), with much of the improvement

in CVD rates in the UK between 1981 and 2000 being attributed to reductions in smoking and improvements in other lifestyle behaviour such as the reduction in total cholesterol level of the UK population (Unal et al., 2004). The presence of non-modifiable and modifiable risk factors for stroke, CHD or overall CVD may be used to calculate individual risk for an event, over a specified time frame, such as applying the Framingham score or European CVD risk assessment model 'SCORE' (Systematic Coronary Risk Evaluation) (Perk et al., 2012, NIH, 2013). In the UK, the QRISK2 score is more commonly applied to estimate CVD risk (Hippisley-Cox et al., 2008).

Non-modifiable risk factors for stroke include, greater age, sex (being male) and family history of stroke or other genetic risk factors but the primary modifiable risk factor for stroke is the presence of hypertension (Goldstein et al., 2011). This, in addition to smoking, poor glycaemic control or diabetes, dyslipidaemia, poor diet, physical inactivity and other risk factors may be used to develop strategies for reducing risk of first stroke occurrence (Goldstein et al., 2011). The Interstroke study collated data from 22 countries worldwide and identified that 90% of stroke risk was associated with 10 risk factors; history of hypertension, smoking, waist to hip ratio, diet risk score, regular physical activity, diabetes mellitus, alcohol, psychosocial stress, depression, cardiac causes and ratio of apolipoprotein B:A1 (O'Donnell et al., 2010).

Risk factors for CHD are similar to those for stroke and include hypercholesterolemia, hypertension, obesity and type two diabetes (Lattimer and Haub, 2010). According to the Cardio and Vascular Coalition report that modelled the UK burden of CVD to 2020, the most important modifiable factors for determining total CVD risk are smoking, elevated cholesterol, hypertension, diabetes, obesity and deprivation, with other smaller risk factors making only a minimal contribution to overall risk (Capewell et al., 2008).

It is known that risk varies by geographical location and socio-economic status (SES) and it is therefore accepted that lifestyle or dietary factors must influence the variation in risk, aside from genetic influences or the effects of ageing (British Nutrition Foundation, 2005). Using data from the Nurses' Health Study, an estimated 82% of CHD events were attributed to lack of adherence to a low risk pattern for diet, physical activity and cigarette use (Stampfer et al., 2000). These poor lifestyle habits and increased risk may be mediated through socio-economic position and for example in the Framingham study, low socio-economic position in childhood and over the life course was associated with risk factors for CHD that the authors thought were potentially mediated through smoking, hypertension, diabetes and obesity (in women) (Loucks

et al., 2009). In UK cohort studies, lower socio-economic position is also associated with worse health outcomes (McFadden et al., 2008, Elovainio et al., 2011). However, in the UK Whitehall II Study the association was found to be bi-directional with socio-economic position in adult life determining metabolic health disparity but heath status at younger ages was associated with the degree of social mobility (Elovainio et al., 2011).

The existing disparity between CVD rates in South-East England (lower) compared to the rest of the UK (higher) indicates that large improvements in CHD mortality are still attainable within the UK (Capewell et al., 2008). Given the large influence of lifestyle factors on overall risk, the key prevention strategy for morbidity free survival is to target smoking, physical activity and promote healthy diets.

1.1.4.1 Diet as a risk factor

Diet is a central modifiable risk factor for the prevention and risk reduction of CVD. Early research in this area focused on the relationship of single nutrients to risk profiles such folic acid and vitamin E, which proved to be ineffective when tested in clinical trials (Bhupathiraju and Tucker, 2011). The focus shifted to examine whole food intake and in the past decade a paradigm shift has been seen in examining associations between dietary patterns and health and evidence now indicates that it is a complicated set of different nutrients that may interact to influence risk (Bhupathiraju and Tucker, 2011).

Fruits and vegetables have been consistently associated with lower CHD risk and although the mechanisms of action are not entirely clear they likely include antioxidant and anti-inflammatory properties (Bhupathiraju and Tucker, 2011). Higher intake of fatty fish, nuts and other plant based n-3 fatty acids such as those found in rapeseed oil and soy beans have also been associated with lower CHD risk (Van Horn et al., 2008, Bhupathiraju and Tucker, 2011). The cardio-protective mechanisms of long chain n-3 polyunsaturated fats thought to be involved include the displacement of undesirable fats (saturated and trans fats) in the diet or more direct influences on lowering triacylglycerol levels and inhibition of pro-inflammatory cytokines (de Roos et al., 2009). Moderate alcohol intake is also associated with lowering of CVD risk, potentially because of a favourable effect on thrombolytic and coagulation processes or by indirectly inhibiting removal of high density lipoprotein cholesterol (HDL-C) (Rimm et al., 1999).

Whole-grains have been associated with lower CHD risk for many decades (Trowell, 1972), regularly feature in dietary recommendations and are a feature of dietary patterns that are associated with lower CVD risk, such as the Mediterranean diet (Bach et al., 2006). The potentially protective components within whole-grains that are removed or reduced during processing include minerals such as magnesium and phytosterols, which are associated with reducing serum cholesterol levels (Slavin, 2003). The removal of potentially beneficial non-starch polysaccharides (NSP) during grain processing is also a concern and dietary fibre is discussed at length below. In order to direct efforts towards minimising the loss of grain constituents that are potentially physiologically important, the most protective components of whole-grains need to be identified (Slavin, 2003).

1.2 Dietary fibre

Dietary fibre is a broad term that encompasses a heterogeneous mix of plant components that are indigestible in the small intestine (Dreher, 2001). The concept of 'dietary fibre' is not so clear-cut with definitions, chemical analysis methods and recommended intakes differing greatly across the world (Lunn and Buttriss, 2007). There are disagreements as to which specific carbohydrates should be included in definitions and which analytical methods should be used to assess them (Buttriss and Stokes, 2008). The following sections cover each of these different issues but firstly, the separate components of plants, which may be classified as dietary fibre, are discussed.

1.2.1 Constituents of dietary fibre

Non-digestible oligosaccharides

Oligosaccharides are carbohydrates with chain lengths between 3-15 monomer units (Lunn and Buttriss, 2007, Lean, 2006), are not generally digested in the small intestine, but may be broken down in the colon by bacterial enzymes (Lean, 2006, Cummings and Mann, 2007, Buttriss and Stokes, 2008, Lunn and Buttriss, 2007). Oligosaccharides include raffinose, stachyose and verbascose and main food sources of non-digestible oligosaccharides include onions, chicory and Jerusalem artichokes (Buttriss and Stokes, 2008, Lunn and Buttriss, 2007). Oligosaccharides may also be chemically or industrially produced from enzymatic hydrolysis of polysaccharides or from mono- and disaccharides (Buttriss and Stokes, 2008).

Non-starch polysaccharides

NSPs are defined as carbohydrates with longer chain length than oligosaccharides although there is a degree of overlap. NSP comprises many separate components, all of which are principally found in the pant cell wall: cellulose, hemicelluloses, pectins, arabinoxylans, beta-glucan, glucomannans, plant gums, mucilages and hydrocolloids (Cummings and Mann, 2007). Plant cell walls are made from a rigid scaffolding of cellulose fibres embedded amongst a jelly-like matrix of water soluble gums such as pectin (Lean, 2006). NSP may be considered in two broad categories, those that are insoluble in water (cellulose and hemicelluloses) and water-soluble types (pectin, gums, mucilage and hemicelluloses) (Lyons-Wall, 2007).

Soluble fibre can be found at high levels in foods such as oats, fruits, vegetables and pulses (British Nutrition Foundation, 2009). In the small intestine, viscous forms of soluble fibre may slow the absorption of lipids and glucose (SACN, 2008) and in the large intestine the fibres may undergo significant fermentation, by bacterial action (Lean, 2006, Cummings and Mann, 2007).

Insoluble fibre components have a greater influence on bowel habits and also undergo partial fermentation in the colon (SACN, 2008). The digestion of soluble and some insoluble fibres, by bacteria, results in the production of short-chain fatty acids (SCFA) (Lean, 2006, Cummings and Mann, 2007). These SCFAs make their own nutritional contribution and it is estimated that NSP provides 2-3kcal/g when digested, although estimates vary by cooking method and bowel absorption (Lean, 2006). In the UK about half of NSP is provided by fruit and vegetables (Lean, 2006) and cereal grain foods contribute 37% of total NSP intake in British adults (Bates et al., 2009). Grains such as rice, wheat and maize provide mainly insoluble NSP, whilst oats, barley and rye also provide soluble NSP (Lean, 2006).

The division of fibre into soluble and insoluble is not always appropriate as the classification is extremely pH dependent (SACN, 2008). Some insoluble fibres are completely fermented whilst not all soluble fibres have effects on glucose and lipid absorption (SACN, 2008, Lunn and Buttriss, 2007). In addition, other food components, such as resistant starch, have similar physiological properties to some soluble or insoluble fibres but are not classified as either.

Resistant starch

Resistant starch is not available for normal digestion for a number of reasons, either it is contained within the food matrix or within starch granules or it is retrograde starch, produced

during food manufacture and preparation, which leaves starch crystals resistant to enzymatic digestion (Englyst et al., 2007, Lunn and Buttriss, 2007). Resistant starch is present in foods such as in unripe bananas, legumes, whole-grains and cooked then cooled potatoes and is unavailable for digestion in the small intestine and so passes to the colon, where it is partially digested in a similar way to soluble fibre (Lyons-Wall, 2007, Lunn and Buttriss, 2007). The exact quantification of resistant starch in food is difficult because storage and cooking methods for foods determine the levels present (Buttriss and Stokes, 2008).

Lignin

Lignin is not classified as a carbohydrate as it is a polymer of phenylpropane units and it is chemically linked with hemicelluloses in plant cell walls (Lunn and Buttriss, 2007). Lignin is considered as a component of dietary fibre, using some definitions (discussed below).

1.2.2 Dietary fibre chemical analysis methods

The chemical diversity of molecules classed as dietary fibre makes laboratory analysis challenging and hence, a number of different techniques have developed (Lyons-Wall, 2007). All of the approaches use a dried and defatted food sample but each method measures a different chemical fraction. The methods may be broadly categorised as enzymatic-chemical and enzymatic-gravimetric (Lunn and Buttriss, 2007).

Enzymatic-chemical methods include the approaches of Englyst & Cummings (1988) and Southgate (1969). The 'Englyst' method identifies NSP present in foods, whilst the 'Southgate' approach is similar and also estimates the lignin fraction of dietary fibre (Lunn and Buttriss, 2007). In the UK, the common technique used to determine 'fibre' in foods is the method developed by Englyst and Cummings, which can be used to distinguish between soluble and insoluble NSP. In this method, starch is initially removed with strong amylases then free sugars are measured after acidic hydrolysis for both the soluble and insoluble fractions. The sum of the two fractions produces the total NSP value. In this method, lignin is not detected because it is not a carbohydrate and resistant starch is also not captured with this method (Lyons-Wall, 2007).

Enzymatic-gravimetric methods of fibre estimation include those used by the Association of Official Analytical Chemists (AOAC). These methods attempt to estimate the fraction of food components that resist digestion in the gastrointestinal tract and the method therefore

measures a variety of components. Again, enzymes are used to mimic digestion then the remaining fraction is weighed (Lunn and Buttriss, 2007). This approach involves less analytical work and therefore is much more economical to use (Lyons-Wall, 2007). The AOAC method gives a value including soluble and insoluble NSP and lignin combined (Lyons-Wall, 2007).

The presence of resistant starch complicates analytical methods for determining dietary fibre but the AOAC and Southgate methods both include some resistant starch in the estimated fibre value (Lyons-Wall, 2007).

For many vegetables, fruits and many unprocessed cereals, the values generated by the Englyst and AOAC methods produce similar results but values for heat processed cereals are far higher when the AOAC method is used (Lean, 2006). The Southgate and AOAC methods produce notably higher values for foods that are good sources of resistant starch such as legumes, compared to the Englyst method (Lyons-Wall, 2007). On average, over different food groups, AOAC values are 1.33 times greater than NSP estimates (Lunn and Buttriss, 2007).

1.2.3 UK and international definitions of 'dietary fibre'

Although the term 'dietary fibre' is practical for public health messages, it is not necessarily useful for classifying carbohydrates based on their molecular composition but rather, is useful in classifying groups of molecules with similar physiological properties.

In 2008 the UK Scientific Advisory Committee on Nutrition (SACN) panel issued a statement on dietary fibre and considered that a material could be classified as dietary fibre if it was "resistant to digestion and absorption in the small intestine and has a demonstrable physiological effect potentially associated with health benefits in the body such as increasing stool bulk, decreasing intestinal transit time or decreasing post prandial glycaemia". The panel considered that evidence only of fermentation in the gut was not sufficient to be included within the definition without an associated physiological effect (SACN, 2008).

SACN commented that using this definition and available evidence, NSP and soluble fibre would be the only two components included in the definition, without the need for further evidence of physiological effect but for other components to be included further evidence is required (SACN, 2008).

In contrast to this, the European Food Safety Authority (EFSA) definition is more inclusive, with no necessity for an observed physiological effect, but rather the definition focuses on chemical composition. Dietary fibre is therefore defined as any non-digestible carbohydrate: cellulose, hemicelluloses, pectins, hydrocolloids (i.e. gums, mucilages, beta-glucans), resistant oligosaccharides (fructooligosaccharides, galactooligosaccharides, other resistant oligosaccharides) and resistant starch (including physically enclosed starch, some types of raw starch granules, retrograde amylase, chemically and/or physically modified starches) and lignin (where this is associated with dietary fibre polysaccharides) (EFSA, 2010).

For a time, fibre was commonly categorised as insoluble (resisted fermentation in the large bowel) or soluble fractions (does not resist fermentation), but molecular structures and therefore solubility exists on a scale and this simplistic categorisation has since been recognised as misleading (Buttriss and Stokes, 2008). Additionally, newer carbohydrate-derived components of interest, such as resistant starches and oligosaccharides, do not fit completely into either category (Buttriss and Stokes, 2008).

A comprehensive article reviewed definitions and guidelines for dietary fibre in many countries across the world and summarised that:

“a workable definition of dietary fibre should: clarify the constituent makeup of dietary fibre; recognise that a primary characteristic is resistance to digestion and absorption in the small intestine and fermentation in the large intestine; and demonstrate that fibre has physiological properties” (Lunn and Buttriss, 2007).

1.2.4 UK and international intakes of dietary fibre and recommendations

The lack of a universal definition for carbohydrates that resist digestion has lead to complications in both establishing and communicating consistent recommendations, health claims and food labels (Buttriss and Stokes, 2008). Both the recommended intake levels and the way that dietary fibre is defined varies greatly across different countries around the world (Lunn and Buttriss, 2007). An additional complication is that current estimation methods may underestimate fibre content of foods, such as using the definition advised by the EFSA because resistant oligosaccharides and inulin are not currently captured by assessment methods and must be assessed separately (Buttriss and Stokes, 2008).

In the United States (US), the average fibre (AOAC) intake is estimated at around 15g/day, significantly lower than the recommended intake level of 25g/day for women and 38g/day for men (approximately 14g/1000kcal/day) (USDA/HHS, 2010). An EFSA panel consensus was that the role of fibre in bowel function was the most suitable criterion for setting recommended intakes and based on available evidence considered 25g/day to be adequate for normal laxation in adults (EFSA, 2010).

In the UK, dietary recommendations appear lower due to fewer components of plant cells walls being classified as fibre, namely the exclusion of lignin and resistant starch as these are not detected with the Englyst method of fibre estimation. The recommended intake level for British adults is 18g/day (COMA, 1991). This value corresponds to recommendations of around 24g/day using the AOAC method (Buttriss and Stokes, 2008). However it is estimated that the average person in the UK doesn't meet the recommended intake level. In the most recent National Diet and Nutrition Survey (NDNS) mean fibre intakes were estimated well below recommended levels at 13.3-13.8g/day (Bates et al., 2012).

1.3 Potential mechanisms for fibre and CVD risk factors

As dietary fibre encompasses a range of non-digestible carbohydrates, many mechanisms for the protective action of dietary fibre have been proposed (Liu et al., 2002a). Dietary fibre intake has been associated with improvements in key modifiable risk factors for CVD such as overweight (Du et al., 2010), hypertension (Ludwig et al., 1999, Ascherio et al., 1992) and hypercholesterolaemia (Brown et al., 1999) and these key risk factors are considered, in turn, below.

1.3.1 Circulating lipid levels

Bacterial fermentation of resistant starch, soluble fibres and some insoluble types of fibre, in the large intestine, produces SCFAs (principally butyrate, proprionate and acetate) (Lunn and Buttriss, 2007). These SCFAs inhibit hepatic cholesterol synthesis, consequently lowering serum levels (Coultate, 2009, Lunn and Buttriss, 2007, British Nutrition Foundation, 2005).

In addition to the partial fermentation of insoluble fibre molecules discussed above, bile-acids present in the gastrointestinal tract also physically bind to insoluble fibre molecules. This binding, together with the presence of viscous soluble fibre gels in the gut, slows the rate of unbound bile acid reabsorption into the blood stream. Bile acids contain cholesterol and when

absorption is slowed, blood cholesterol is shunted into bile acid production, thus lowering circulating levels (James et al., 2003, Lunn and Buttriss, 2007).

Wheat bran has been shown to have little effect on plasma cholesterol levels but in contrast to this, oat bran seems to reduce total plasma and low-density lipoprotein cholesterol (LDL-C) levels, possibly via the increased gut viscosity seen with soluble fibre intake. This high viscosity may interfere with bile acid reabsorption, resulting in a negative sterol balance (Truswell, 2002, Van Horn et al., 2008).

Meta-analyses of trials using soluble fibre have found consistent results with respect to blood cholesterol levels. Brown and colleagues identified that greater soluble fibre intake was associated with reductions in both total and LDL-C, with no influence on HDL-C or triglyceride levels. The authors however concluded that soluble fibre intake can form only a part of any therapy to reduce blood cholesterol levels given the relatively small effect sizes observed in the studies (Brown et al., 1999). In a more recent meta-analysis of fibre derived from barley (high in soluble fibre), reductions were also observed in total cholesterol, LDL-C and triglycerides, but no changes in HDL-C were observed (Talati et al., 2009).

1.3.2 Overweight

There appears to be an inverse association between fibre intake and body weight or weight gain (Slavin, 2005) and it is well documented that increasing obesity is a major factor contributing towards CVD risk (Van Horn et al., 2008, Logue et al., 2011).

Soluble fibres form gel-like substances in the stomach and small intestine in the presence of water and these gels can slow the rate of gastric emptying, contributing to greater feelings of satiety and could ultimately contribute to lower weight gains (James et al., 2003, Lunn and Buttriss, 2007). Soluble fibre gels also moderate the absorption of nutrients in the small intestine and the slower rate at which glucose is received in the blood stream may also contribute to improved feelings of satiety (Lunn and Buttriss, 2007). In addition, dietary fibre may act as a physical barrier to normal enzymatic digestion of other macronutrients, resulting in lower energy absorption (Du et al., 2010). The fermentation of soluble fibre is also believed to influence hormones associated with inducing satiety, glucagon-like peptide and peptide YY (Johansson et al., 2013, Nilsson et al., 2013, Reimer et al., 2010).

Few epidemiological studies have examined the effects of different sources or types of fibre on weight gain but pooled data from five countries in the European Prospective Investigation into Cancer and Nutrition (EPIC) study, including the UK, indicates that greater total fibre intake and greater cereal fibre intake were associated with lower annual weight gains, for each 10g/day greater intake of total fibre, weight gain was -39g/year (95% confidence interval (CI) -71 to -7 g/year). Greater intake of total fibre, cereal fibre and fibre from fruit and vegetables were also associated with lower annual waist circumference gains but higher intake of fruit and vegetable fibre were not associated with lower weight gains (Du et al., 2010).

Findings relating fibre intake to appetite, weight changes and energy intake are relatively mixed, with a recent systematic review identifying that the effects of fibre on energy intake and weight were relatively small and distinct dose-response relationships were not observed from the included studies. More viscous types of fibre were more often associated with reduced appetite than non-viscous fibres but this review focussed primarily on fibre isolates rather than high fibre diets (Wanders et al., 2011).

1.3.3 Blood pressure

Greater intake of dietary fibre has been associated with lower blood pressure in some observational and intervention studies (Streppel et al., 2005, Ascherio et al., 1992), although not in one cohort study of women in the US (Ascherio et al., 1996). Significant reductions in diastolic blood pressure and a decrease, although not significant for systolic blood pressure were observed with greater fibre intake in two meta-analyses, with the effects being more pronounced in older and hypertensive individuals (Whelton et al., 2005, Streppel et al., 2005). Little is actually known about the potential mechanism for the effects of fibre on blood pressure and the observations may be attributed to concurrent increases in potassium and magnesium with greater fibre intake (Streppel et al., 2005) or reductions in bodyweight which were seen in many of the included studies.

1.3.4 Glucose metabolism and diabetes

In a recent review paper, Lattimer and Haub succinctly discuss the many potential mechanisms through which dietary fibre may act on components of metabolic health. Briefly, the higher glycaemic index (GI) of foods is related to higher blood glucose levels and over the long-term this could lead to pancreatic beta cell dysfunction and decreased insulin release. Additionally, cell tissues may become resistant to insulin with chronic hyperglycaemia (Lattimer and Haub,

2010). The consequence of displacing high GI foods for fibre-rich and lower GI foods may therefore be one potential mode of action for fibre-rich diets.

The potential action of soluble-type fibre gels on blood glucose levels, noted above, may be through the slowed nutrient absorption from the small intestine into the blood stream, thus attenuating post-prandial blood glucose levels (James et al., 2003, Lunn and Buttriss, 2007).

Insoluble fibre may act to moderate long term risk of diabetes by stimulating accelerated secretion of glucose-dependent insulinotropic polypeptide, which stimulates postprandial insulin release, or through its effect on appetite reduction and thereby lowering weight gain. Additionally, SCFAs produced during the fermentation of soluble and some insoluble fibres may reduce post-prandial glucose response (Lattimer and Haub, 2010).

1.3.5 Inflammation

As noted above, there are many potential mechanisms via which fibre intake can influence energy intake and body weight. Another biological effect of obesity, in addition to traditional associations with risk factors for CVD is the influence on inflammation (Logue et al., 2011). Inflammation mediates the well known but poorly understood links between obesity, cardiovascular pathology and common comorbidities such as hypertension, diabetes and dyslipidaemia (Berg and Scherer, 2005).

Inflammation is not simply a reactive response to atherosclerosis but is itself an important contributor to CVD risk. Adipose tissue acts as an endocrine organ and directly augments systemic inflammation by releasing proinflammatory cytokines. The inflammatory proteins that are secreted from adipocytes and adipose tissue work in a complex and reciprocal matrix and it appears that these circulating mediators of inflammation are directly involved in the mechanism of vascular damage and atheromatous changes that progress into CVD (Berg and Scherer, 2005).

1.4 Gaps in current epidemiological research

Many studies have examined the relationship between dietary fibre intake and CVD risk (refer to Chapter 2 where other literature is discussed in depth). However, few studies have explored key important elements in addressing this question. Firstly, as the physiological consequences of diets with high fibre content may depend on the types of fibre and the food source (Rimm

et al., 1996), it is crucial to address this and explore which types or sources of fibre may be more closely associated with potential benefits. In identifying potentially beneficial components, a more targeted approach to disease prevention could be taken, thus maximising any benefits. Secondly, many studies combine CVD cases when assessing risk associations and do not explore risk in disease sub-types. For example, many studies present only 'total stroke' cases and not haemorrhagic or ischaemic types separately but it has been noted that risk factors differ for these conditions (Andersen et al., 2009).

Exploring different fibre sources and types in relation to total CVD and sub-classifications of the disease could lead to tailored advice for disease prevention, potentially targeting low consumers of specific fibre types or those at greater risk of different types of CVD.

Many studies from the US, other parts of Europe, Australia and Japan provided evidence on the association between both total fibre and sources of fibre in relation to CVD (refer to Chapter 2) but diets and population characteristics within these countries may differ significantly from the UK. The UK Women's Cohort Study (UKWCS) therefore has the potential to be extremely useful for exploring associations between dietary fibre intake and CVD in the context of the British diet and population.

1.5 Aims

Given the potential benefits in CVD risk reduction linked to dietary fibre intake, indicated with evidence from both intervention and observational studies, the primary aim for this work is to explore this association using data collected as part of the UKWCS and thus contribute to the growing body of evidence on this topic. The specific project aims include:

1) Update existing systematic reviews: Identify literature from observational studies concerning dietary fibre intake and risk of CHD, stroke and CVD (Chapter 2).
2) Obtain mortality records, in addition to non-fatal events recorded within the Myocardial Infarction National Audit Project (MINAP) and Hospital Episode Statistics (HES) and link this CVD event data with lifestyle variables from the UKWCS (Chapter 4).
3) Compare and thus assess the completeness of event data from different sources using a capture-recapture approach (Chapter 4).
4) Explore the association between CVD risk and fibre intake estimated using food frequency questionnaires (FFQ) (Chapters 5 and 6) and food diaries (Chapter 7).
5) Examine associations between total fibre or key types and sources of fibre with CVD risk.

1.6 Summary

The human and economic impact of CVD is difficult to overstate and since most CVD risk is modifiable, prevention is possible and is paramount to the health of every nation (Mosca et al., 2007). Interplay between key risk factors for CVD such as hypertension, hypercholesterolaemia, poor glucose regulation and overweight make pinpointing specific mechanistic pathways for the action of fibre molecules on total CVD risk a challenge.

The collective evidence from studies examining risk factors for CVD suggests that there are many modes of action of different dietary fibre components. Examining soluble and insoluble fibre along with fibre from key food sources, where the relative ratio of soluble and insoluble types of fibre will differ, may elucidate specific associations with CHD or stroke and hint at potential mechanisms for the action of fibre on overall risk. Additionally, quantifying risk reduction in relation to specific levels of fibre intake, and with different sources or types of fibre, may allow for tailored recommendations for CVD prevention to be developed. At the very least, it is apparent that CVD risk is largely modifiable and maintaining optimum health over the life course may contribute to lower overall risk and lengthen disease-free survival time. Exploring this question in a sample of British women, where intakes in the sample are diverse, will contribute to the growing body of evidence in this important area.

Chapter 2 Systematic literature review and meta-analysis

2.1 Chapter overview

Observational study data relating to dietary fibre intake and primary CVD events have been identified through systematically reviewing literature published since 1990. This chapter details the methods used in both the main systematic review (noted in the Declaration and Acknowledgments on pages 3-6) plus update searches, to highlight my separate contribution to this work. The role of the review team was to collate evidence and present it to SACN, for their interpretation. Thus, aside from using the identified literature, any interpretation, discussion and conclusions drawn are my own. Findings are presented separately for total CVD, CHD and stroke, where sufficient data were identified.

Results from this chapter have subsequently been further extended with additional, up-to-date literature searches and have been published (Threapleton et al., 2013d, Threapleton et al., 2013e). One of these publications includes results for the UKWCS which are presented in Chapter 5 (Threapleton et al., 2013b) and these more recently published results are therefore not included in the work presented here as were not available at the time. Both full manuscripts were written largely by me, with the exception of the statistical methods sections which were written mainly by Dr Darren Greenwood. Other author and reviewer suggestions on style and content were also incorporated into the final manuscript submissions. Abstracts presenting work from this chapter were also submitted and accepted for presentation at the winter meeting of the Nutrition Society 2012 (Threapleton et al., 2012c, Threapleton et al., 2012d).

Total fibre intake was associated with CVD, CHD and stroke risk reduction and insoluble fibre, vegetable and cereal sources of fibre were also associated with CHD risk reduction. Unfortunately, for stroke or CVD outcomes and some subtypes of fibre presented in this chapter, too few studies were identified that reported data in a suitable format to permit meta-analysis. However, in the recent publication discussed above (Threapleton et al., 2013e) additional studies were included from a further updated literature search and thus more meta-analyses could be conducted.

2.2 Background

In recent years, a decline in total CVD and CHD incidence has been seen in many developed countries (AHA, 2011, Allender et al., 2008). Although rates of CVD have long since peaked for many developed countries and CVD mortality is declining (Unal et al., 2004), it remains a significant issue, accounting for a third (34.3%) of all fatalities in the US and almost half (48%) of all deaths in Europe (AHA, 2011, Allender et al., 2008).

A similar trend for declining stroke incidence is also reported in many of the worlds developed countries, a reduction largely attributed to improvements in hypertension management. However, the absolute number of strokes continues to increase with the expansion of the aging population in these counties (Mackay and Mensah, 2004). Stroke and other cerebrovascular diseases are the second most common cause of death worldwide, and in 2008 accounted for 6.2 million deaths (11% of fatalities) (WHO, 2008). Data from the US suggests that 78% of strokes are first attacks (Roger et al., 2012) with ischaemic stroke being ten times more common than haemorrhagic stroke in most western countries (Andersen et al., 2009). Moreover, stroke is the leading cause of disability in many developed countries and its primary prevention should, therefore, be a key public health priority (He et al., 2006).

The previous chapter includes detailed discussion of the many proposed mechanisms for the action of dietary fibre on various risk factors for CVD. Many experimental studies have examined the relationship between dietary fibre or fibre-rich foods and CVD risk factors such as hypertension, central obesity, insulin sensitivity and elevated plasma cholesterol (Ludwig et al., 1999, Brown et al., 1999, Truswell, 2002, Van Horn et al., 2008). A number of literature reviews published in the past decade have also explored the association between dietary fibre and CVD or CHD risk using observational study data (Liu et al., 2002a, Pereira et al., 2004, Mente et al., 2009, Hauner et al., 2012, Ye et al., 2012) but although many individual epidemiological studies have examined stroke risk in relation to dietary fibre intake (Ascherio et al., 1998, Bazzano et al., 2003, Eshak et al., 2010, Kaushik et al., 2009, Kokubo et al., 2011, Larsson et al., 2009, Oh et al., 2005, Wallstrom et al., 2012), when this work was undertaken there were no published meta-data relating to stroke occurrence.

Taking the existing evidence base from observational study reviews and meta-analyses together, there appears to be an inverse association between total dietary fibre and CVD risk. One meta-analysis of 9 publications identified a 7% risk reduction for every 10g/day increase in

fibre, RR 0.83 (95% CI: 0.78 to 0.89) (Liu et al., 2002a), although studies were not identified through a systematic review of literature. A pooling project, using raw study data from 11 cohorts also reported a risk reduction for total fibre 0.81 (95% CI: 0.73 to 0.91) and additionally for fruit fibre 0.84 (95% CI: 0.70 to 0.99), but not for cereal fibre 0.90 (95% CI: 0.77 to 1.07) or vegetable fibre 1.00 (95% CI: 0.88 to 1.13) (Pereira et al., 2004).

Of the publications reporting systematically reviewed literature in this area, two meta-analyses identified around 20% lower CHD risk in high compared to low fibre consumers, RR 0.78 (95% CI: 0.72 to 0.84) (Mente et al., 2009) and RR 0.81 (95% CI: 0.77 to 0.86) (Ye et al., 2012). A protective association of similar magnitude was also reported for high compared to low cereal fibre consumers 0.80 (95% CI: 0.73 to 0.88) (Ye et al., 2012) although no other sources of fibre or types (soluble or insoluble fibre) were examined in these publications.

A narrative synthesis of systematically reviewed literature used to inform evidence-based guidelines for the German Nutrition Society was also recently published (Hauner et al., 2012). Increased dietary fibre intake and greater whole-grain intake were judged as 'probably' associated with primary prevention of CHD and evidence of cereal fibre, soluble fibre and insoluble fibre was judged as 'possibly' being inversely related. For vegetable fibre, evidence was judged as 'possibly' indicative of there being no association (Hauner et al., 2012).

The aim of this work was to review literature published since 1990, in generally healthy populations, concerning dietary fibre intake and cardiovascular disease risk, to update reports published in the UK by the Committee on Medical Aspects of Food Policy, in the early 1990's (COMA, 1991, COMA, 1994). The aim was to systematically review the evidence base and combine study data in order to calculate dose-response estimates for total dietary fibre in addition to fibre from major food sources, thus improving upon previously published reviews that were either not systematic (Liu et al., 2002a, Pereira et al., 2004), calculated risk in high compared to low consumers (Mente et al., 2009, Ye et al., 2012) or did not explore key sources or types of dietary fibre (Liu et al., 2002a, Mente et al., 2009). A further aim was to report on potential sources of heterogeneity between studies to give insight into population characteristics or study design issues that may introduce bias and influence whether significant associations were observed.

Systematically reviewing literature using a strict protocol is important to ensure the quality of a review and to minimise bias in study identification and inclusion and reduce random errors

(Egger et al., 2008a, Bowers et al., 2006b). The term meta-analysis, used to describe the integration of analyses from multiple studies, was coined in the 1970's (Glass, 1976) and has become a widely accepted method for orderly summarising information from many studies, which may present opposing findings.

Strengths of combining data from RCTs include the greater statistical power achieved by adding many smaller studies together, whilst maintaining the validity of results, because of the use of randomised groups. However, issues with validity in observational studies are largely a reflection of confounding and bias (Willett, 1998b). Meta-analysis may therefore produce precise but spurious results because the issues of confounding are not addressed through the statistical combination of data (Egger et al., 2008b). The estimate of effect resulting from the combination of data should therefore not outweigh the careful examination of potential sources of heterogeneity between the results of observational studies (Egger et al., 2008b). Heterogeneity, or differences in observed findings, may occur because of different methodological features between studies such as dietary assessment method, outcome assessment and length of follow-up. Meta-analysis can provide great value in examining the potential reasons for heterogeneity through meta-regression and also allows for questions not posed in individual studies to be answered, such as establishing dose-response relationships (Willett, 1998b).

2.3 Methods

2.3.1 Literature searches

Main review: The following online databases were searched for prospective cohort studies and RCTs published in English language from 1st January 1990 to November 2009: The Cochrane Library, MEDLINE, MEDLINE in-process, Embase, CAB Abstracts, ISI Web of Science and BIOSIS. Search strategies for the Cochrane Library and Medline (and Medline in-process articles) were developed by an information specialist. I adapted the Medline search strategy for CAB full-text articles and the Web of Science databases whilst other colleagues adapted the strategy to Embase and BIOSIS.

Hand-searching of the following selected journals was carried out to supplement the electronic searches: Journal of Nutrition, Journal of the American Dietetic Association, American Journal of Clinical Nutrition, Diabetes Care, European Journal of Clinical Nutrition and British Journal of Nutrition. Additionally, bibliographies of relevant published systematic literature reviews and

meta-analyses were cross checked against the articles identified from electronic search databases.

Update: The majority of studies identified in the main review were found via Medline and Embase databases. These two resources were therefore used for the update search, in addition to the Medline in-process and non-indexed citations database. Databases were searched from 1st January 2009 to 31st May 2012. Search strategies used for Medline and Embase were based on those used during the large systematic review, but only search terms for fibre and cardiovascular disease were retained (Appendix I). Key terms used for the update search included, among others, 'fibre' 'fiber', 'cellulose', 'lignin', various other fibre sub-fractions and sub-types, 'cardiovascular diseases', 'coronary diseases', 'myocardial ischaemia', 'stroke' and 'transient ischaemic accident/incident'.

2.3.2 Screening articles

Main review: The electronically retrieved bibliographies were downloaded into Reference Manager software and automated removal of duplicate references was carried out. Titles and abstracts of unique references were scanned for non-relevancy by one of several review team members, using structured guidelines (Appendix II), to ensure consistency. A 10% sample of the articles deemed 'not-relevant' were cross-checked by a second team member to ensure an acceptable level of agreement (>99%). Where the title and abstract did not provide enough information to determine that the article was clearly not relevant to the review, it was marked as potentially relevant.

Full-text versions of all articles not excluded during screening were retrieved and reviewed independently by two members of the research team. Any disagreements were resolved by a third reviewer. A structured flow-chart plus detailed information sheets were used to determine eligibility for inclusion (Appendix III).

Prospective cohort studies with follow-up duration less than three years were excluded. Multiple reports from the same cohort study were reviewed for instances where papers presented the same results at multiple follow-up periods. Papers with the longest follow-up for identical results were preferentially included unless the paper with longer follow-up did not report sufficient information to allow inclusion in dose-response meta-analysis.

The focus of this review was prevention of cardio-metabolic diseases (not secondary prevention or reversal of CVD risk factors) and as such, studies were only included where participants were considered generally healthy or had an intermediate stage of illness at baseline. Studies where greater than 50% of participants were chronically ill, or where data on the 'healthy' participants were not presented separately, were excluded. Examples of ill populations include studies where >50% of participants had a history of CVD, diabetes, cancer, hypertension, hypercholesterolaemia or hyperinsulinaemia. However, studies including pre-hypertensive, glucose intolerant or obese participants were permitted. In addition to this, studies where the majority (>50%) of participants were taking medications for chronic illnesses (e.g. lipid-lowing or anti-hypertensive medication) were excluded.

Update: Duplicate articles were manually identified and removed from the reference database. Titles and abstracts of unique references were screened for relevancy using the same criteria as applied for the main review. Screening was undertaken independently by Dr Victoria Burley and I, with any disagreements being settled through discussion. Any articles identified as potentially relevant by either of us were obtained in full-text and were further screened, in duplicate, using the same formal inclusion criteria as applied for other studies in the main review.

2.3.3 Data Extraction

Data were extracted into a Microsoft Access database, with pre-defined fields that capture aspects of study design and quality as well as individual results (RRs and CIs), including exposure type and quantity, case numbers, definition of outcome and adjustments used within analyses. This method of data extraction was based on the approach used for the World Cancer Research Fund Second Expert Report (WCRF, 2007). Methodological quality of studies was not evaluated using a formal scoring approach but aspects of study quality, such as follow-up duration, case ascertainment and adjustment for various important confounders were extracted for investigation through meta-regression.

2.3.4 Statistical analysis

Main and update review: Dose-response trends for fibre exposures in relation to CVD outcomes were generated for each cohort (Greenland and Longnecker, 1992, Orsini et al., 2006) and these were then combined using random effects models. Random effects models incorporate an estimate of between-study variation into the combined effect calculation and

are therefore appropriate to use when heterogeneity between studies is likely to exist, as it is with observational studies (Deeks et al., 2008).

Summary estimates are only considered reliable when heterogeneity (I^2) does not exceed 75% and are only presented when studies had included appropriate adjustments because unadjusted results are liable to potentially substantial bias. Where results were presented by diagnostic category only e.g. fatal and non-fatal events presented separately, the method of Hamling and colleagues was used to first combine data in a random effects meta-analysis (2008). This approach allows for a more consistent comparison between different study results and thus potentially improves reliability of a meta-analysis (Hamling et al., 2008).

For comparability, fibre increments presented in the dose-response figures were chosen to be approximately one standard deviation (SD) of population intake level. Fibre intake SDs were obtained from multiple sources and reflect a realistic increase in daily intakes (Bates et al., 2009, Larsson et al., 2009, Streppel et al., 2008, Pietinen et al., 1996). The SDs for fibre estimated as NSP were multiplied by 1.33, a standard conversion factor (Lunn and Buttriss, 2007) in order to be comparable to fibre values estimated using the AOAC method.

Restricted cubic splines were used to model the nonlinear dose-response association between fibre exposures and CVD, CHD or stroke for each study (Orsini and Greenland, 2011), based on fixed knots at 10%, 50% and 90% through the distribution of intake, then combined using multivariate meta-analysis (White, 2009).

Heterogeneity between studies was tested using Cochran's Q statistic, alongside the proportion of total variation in study estimates that is due to heterogeneity (I^2) (Higgins and Thompson, 2002). The following methodological features were explored through pre-defined meta-regression: method used to assess fibre intake (AOAC/non-AOAC), whether results include non-fatal events, follow up length (<10years/≥10years), geographic location (US/European Union/Other) and also whether the results were adjusted for the following: age, alcohol, anthropometry, energy intake, physical activity and gender. All analyses were performed by Dr Darren Greenwood using Stata 12.1 (StataCorp, 2011).

2.4 Results

2.4.1 Included articles

Main review: In total 42,518 references were obtained from both electronic and hand-searching. After removal of duplicates, 23,165 unique references remained. On first screening 1,736 of these references were deemed to be potentially relevant and 21,429 were marked as not relevant for this review. Just over 10% of the references (2,214) marked as not relevant were screened independently by a second reviewer. Of this checking sample, 0.8% (17 articles) were identified as potentially relevant and these were re-marked as potentially relevant articles to include. The number identified in this check process was lower than the pre-specified cut-off of 1% and no further checking was carried out. At this stage, 16 additional unique references were identified during hand-searching and were included into the potentially relevant file, bringing the total of potentially relevant articles to 1,769.

Of the 17 papers identified during the quality check process, five were eventually included into the review. Four of these five papers would have been included in the review, had the checking process not identified them, as they were also identified during hand-searching of relevant journals and reference lists of relevant literature reviews. In summary, the screening and hand-searching processes seem acceptably thorough since just one article from the 10% (n=2,214 papers) check sample would have otherwise been missed from the full review.

In total 396 articles were included in the full carbohydrate and cardio-metabolic heath review and of these, dietary fibre intake in relation to cardiovascular events were reported in 17 publications from 14 cohorts.

Update searches: A total of 879 unique references were identified and from these, 19 were flagged as being potentially relevant to the review (Figure 2.1). After screening the 19 full-text articles, 8 were excluded; one was a cross-sectional study (Oba et al., 2010), one included diabetic participants only (He et al., 2010), three reported diet scores or foods rather than fibre as a single dietary component (Heroux et al., 2010, Van Horn et al., 2012, Hlebowicz et al., 2011), one paper published in 2009 had already been identified from the main review search (Kaushik et al., 2009) and two articles reported narrow and specific cardiovascular events (atrial fibrillation and venous thromboembolism) rather than CVD or CHD (Shen et al., 2011, Varraso et al., 2012).

From the update search, 11 articles from 8 cohorts were included. Five reported only total CVD (Akbaraly et al., 2011, Baer et al., 2011, Buyken et al., 2010, Park et al., 2011, Chuang et al., 2012), three reported only CHD (Bernstein et al., 2011, Crowe et al., 2012, Ward et al., 2012) and three reported incident CVD, CHD and stroke events (Eshak et al., 2010, Kokubo et al., 2011, Wallstrom et al., 2012).

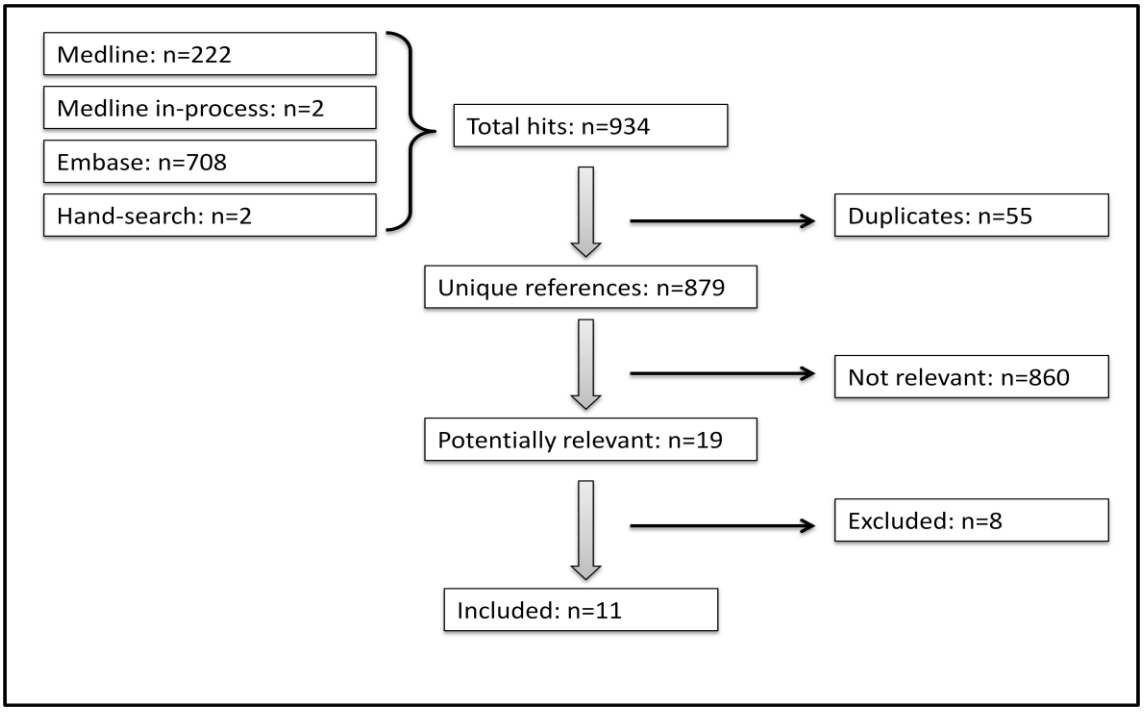


Figure 2.1 Flow chart for update search publication identification and inclusion

Searches combined: A total of 28 articles from the main plus update searches were identified from 1st January 1990 to 31st May 2012, reporting data from healthy cohorts concerning dietary fibre and incident cardiovascular events, with follow-up of at least 3 years (Figure 2.2).

Studies reported combinations of exposures and outcomes (Table 2.1). Four of the articles didn't report risk estimates but instead provided a description of baseline dietary fibre intakes in those that subsequently became a case or not. These four publications will not be discussed here further as results were only minimally adjusted or were not adjusted for any potential confounders and are likely subject to substantial bias (Drogan et al., 2007, Fehily et al., 1993, Knekt et al., 1994, Laaksonen et al., 2005).

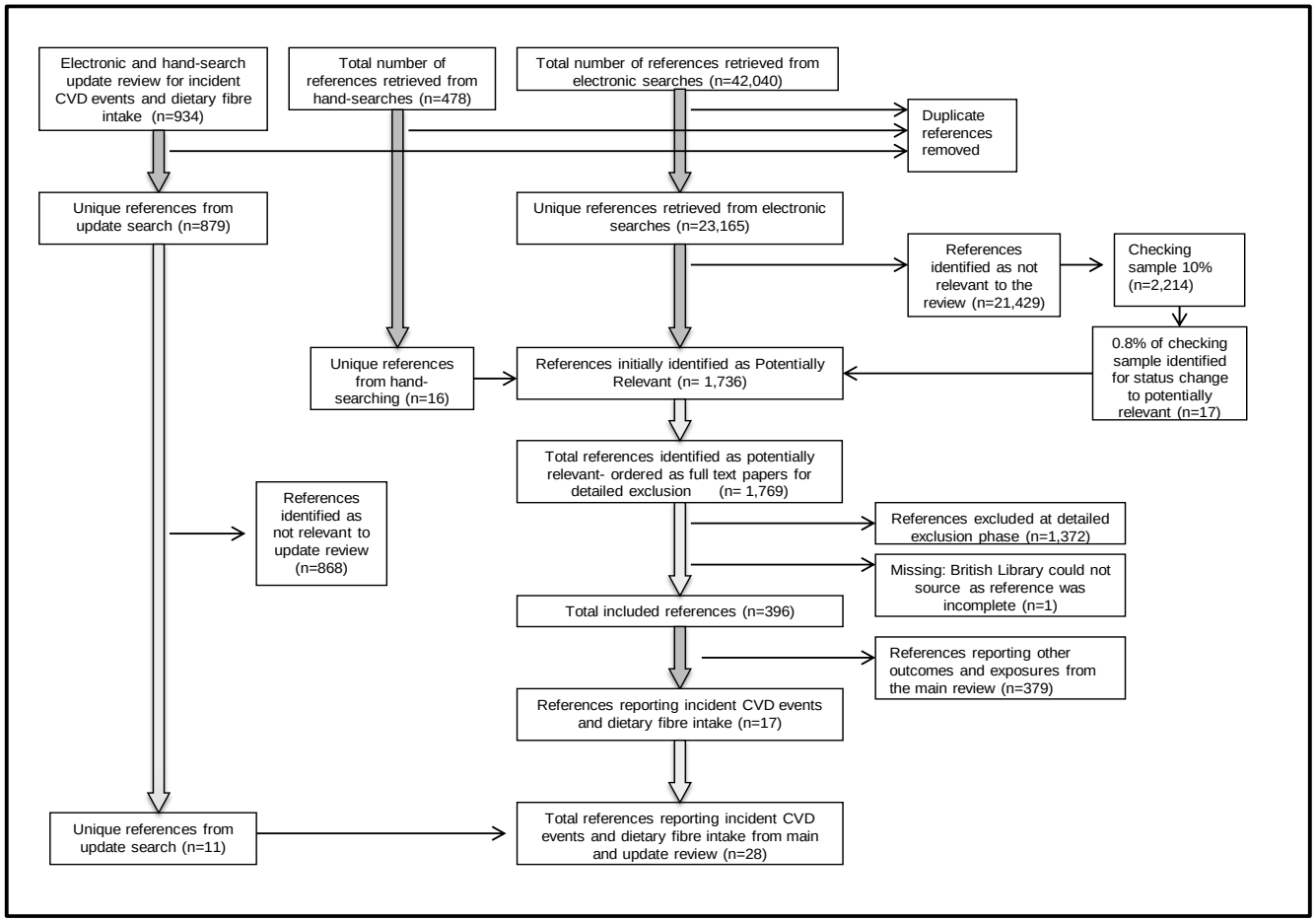


Figure 2.2 Main plus update search flow diagram for publication identification and inclusion for CHD, stroke and CVD outcomes combined

Table 2.1 Summary of outcomes and exposures reported in cohort studies identified during the systematic review and update searches

Reference/ Authors	Cohort name	Reported study outcomes			Dietary exposures											
		Total CVD	CHD	Stroke	Dietary fibre	Fibre density	Soluble fibre	Insoluble fibre	Fibre from specific food sources						Fibre fractions	
									Fruit	Veg	Cereal	Le	Oth	Po	Cell'	Lignin
*Akbaraly et al., 2011	Whitehall II	X F			X											
Appleby et al., 1999	Oxford Vegetarian Study		X F		X											
Ascherio et al., 1998	Health Professionals Follow-Up Study			X C	X											
* Baer et al., 2011	Nurses' Health Study	X F									X					
Bazzano et al., 2003	NHANES I	X F,C	X F,C	X F,C		X	X									
*Bernstein et al., 2011	Nurses' Health Study		X C								X					
*Buyken et al., 2010	Blue Mountains Eye Study	X F			X				X	X	X					
* Chuang et al., 2012	EPIC	X F			X				X	X	X					
* Crowe et al., 2012	EPIC-Heart		X F		X				X	X	X		X			
*Eshak et al., 2010	Japan Collaborative Cohort Study	X F	X F	X F	X		X	X	X	X	X					
Kaushik et al., 2009	Blue Mountains Eye Study		X F	X F							X					
* Kokubo et al., 2011	Finnish Mobile Clinic Health Surveys	X C	X C	X C	X		X	X								
Larsson et al., 2009	Alpha-tocopherol beta-carotene Study			X C	X		X	X	X	X	X					
Liu et al., 2002	The Women's Health Study	X C	X C		X		X	X	X	X	X					
Mozaffarian et al., 2003	Cardiovascular Health Study		X C		X				X	X	X					
Oh et al., 2005	Nurses' Health Study			X C	X				X	X	X					
* Park et al., 2011	NIH-AARP Diet and Health Study	X F			X				X	X	X	X				
Pietinen et al., 1996	Alpha-tocopherol beta-carotene Study		X F,C		X		X	X	X	X	X				X	X
Rimm et al., 1996	Health Professionals Follow-Up Study		X F,NF,C		X		X	X	X	X	X					
Streppel et al., 2008	Zutphen Elderly Study		X F		X				X	X	X	X		X		
Todd et al., 1999	Scottish Heart Health Study		X C		X											
* Wallstrom et al., 2012	Malmo Diet and Cancer Cohort	X C †	X C †	X C †		X										
* Ward et al., 2012	EPIC-Norfolk		X C		X				X	X	X					
Wolk et al., 1999	Nurses' Health Study		X F,NF,C		X	X			X	X	X					

Key: *Identified in update review; † Events were ischaemic only; C=fatal and non fatal events combined; Cell=cellulose; F=fatal events; Le= Legume; NF=non-fatal events; Oth= other sources of fibre (fibre after cereal, fruit and vegetable fibre deducted); Po=potato fibre; Veg=vegetable; X=reported in the paper

Cohort characteristics: Identified publications all included adult participants and were mainly from Europe (10) and the US (10), but were also from Japan (2) and Australia (2) (Table 2.2). Studies varied greatly in follow-up duration, the shortest being the Women's Health study which reported at 6 years from baseline (Liu et al., 2002a), and the longest follow-up was in the Zutphen Elderly Study that reported at 40 years (Streppel et al., 2008). The mean follow-up duration of studies was around 14 years.

Most cohorts included males and female except two reporting on just women, the Nurses' Health Study (Baer et al., 2011, Bernstein et al., 2011, Oh et al., 2005, Wolk et al., 1999) and the Women's Health study (Liu et al., 2002a) and three reporting only on men, the Health-Professional's Follow-Up Study (HPFS) (Ascherio et al., 1998, Rimm et al., 1996), the Zutphen Elderly Study (Streppel et al., 2008) and the Finnish Alpha-Tocopherol Beta-Carotene (ATBC) Study of male smokers (Larsson et al., 2009, Pietinen et al., 1996).

Included cohorts varied greatly in terms of participant numbers and cases for the different outcomes. The smallest included study was the Zutphen Elderly Study with just 1,373 participants included at baseline but because of the long follow-up period, 348 fatal CHD events were reported at 40 years (Streppel et al., 2008). The pooled pan-European data presented in the two publications from EPIC included the largest number of participants at baseline, 519,978 and 518,408 and reported on 2,381 CHD and 4,604 CVD mortality cases (Crowe et al., 2012, Chuang et al., 2012).

Dietary intake was assessed with FFQs in the majority of studies, the exceptions being a single 24-hour recall used in the National Health and Nutrition Examination Survey I (NHANES) (Bazzano et al., 2003) and diet history or food-diaries being used in two Finnish studies (Knekt et al., 1994, Laaksonen et al., 2005) and in the Zutphen Elderly Study (Streppel et al., 2008). Intakes were assessed from nutrient tables that derived fibre intakes using the AOAC method in most of the cohorts. The methods used in the NHANES I study and Framingham Heart study were not reported but as these were both conducted in the US, are likely to have used AOAC methods (Bazzano et al., 2003, Shen et al., 2011). The two Japanese cohort studies reportedly employed analysis methods similar to AOAC methods (Eshak et al., 2010, Kokubo et al., 2011). The four British studies, the Malmo Diet and Cancer Cohort and the Finnish ATBC study assessed fibre as NSP (Akbaraly et al., 2011, Appleby et al., 1999, Todd et al., 1999, Ward et al., 2012, Larsson et al., 2009, Pietinen et al., 1996, Wallstrom et al., 2012).

Table 2.2 Characteristics of cohort studies identified during the systematic review and update searches

Reference	Cohort name (Country)	Population characteristics/ notes	Sex	Age at baseline years	Initial cohort size: and case numbers	Dietary assessment	Fibre estimation method	Outcome definition by ICD codes †	Follow up duration (loss)
*Akbaraly et al., 2011	Whitehall II (England)	London-based civil servants	M/F	Mean 49	7319: 141 CVD deaths	Semi-quantitative 127-item FFQ.	Englyst	CVD: 9th 390.0-458.9/ 10th I00-I99	17.7years
Appleby et al., 1999	Oxford Vegetarian Study (UK)	Half of sample were vegetarian	M/F	16-79 (mean 46)	11140: 525 IHD deaths	Simple validated FFQ	Not reported, likely Southgate	IHD: 9th edition ICD codes used to identify IHD (no further details reported)	13.3 years
Ascherio et al., 1998	Health professionals follow-up study (US)	Male health professionals free of CHD at baseline	M	40-47	51529: 328 fatal and non-fatal stroke cases	Validated 131 item FFQ referring to diet over previous year.	AOAC	Stroke: all stroke included and sub-classifications identified with criteria from National Survey of Stroke (embolism/ thrombosis)	8 years
* Baer et al., 2011	Nurses' Health Study (US)	Health professionals free of CHD at baseline	F	30-55	121700: 1026 CVD deaths	Validated 116 item FFQ administered 3 times.	AOAC	CVD: 8th:390.0-458.9 and 795.0-795.9	18 years
Bazzano et al., 2003	NHANES I (US)	Nationally representative sample free of CVD	M/F	25-74 (mean 49)	14407: 928 stroke, 1843 CHD events.	One 24 hour recall including portion size estimates.	Assessment not reported, likely AOAC	CHD: 9th 410-414/ CVD: 9th 390-459/ Stroke: 9th 430-438	19 years (4% loss)
* Bernstein et al., 2011	Nurses' Health Study (US)	Health professionals free of CHD at baseline	F	30-55	72266: 2500 CHD cases	Validated 116 item FFQ administered multiple times.	AOAC	CHD defined using criteria from the WHO, symptoms plus either ECG changes or elevated cardiac enzyme levels	22 years
* Buyken et al., 2010	Blue Mountains Eye Study (Australia)	Older age population cohort	M/F	Median 65	3654: 260 fatal CVD cases	Validated 145 item FFQ.	AOAC	Total CVD classification not reported in publication	13 years
*Chuang et al., 2012	EPIC (Europe)	Samples from 23 centres across 10 countries in Europe	M/F	Mean 50.8	518408: 4604 fatal CVD cases	Various: FFQ, semi-quantitative FFQ and diet history.	AOAC and standardised values	Circulatory disease 10th I00-I99	12.7
* Crowe et al., 2012	EPIC-Heart (Europe)	Samples from 23 centres across 10 countries in Europe	M/F	Mean 54	519978: IHD 2381 deaths	Various: FFQ, semi-quantitative FFQ and diet history.	AOAC and standardised values	IHD: 9th 410-414/ 10th I20-I25	11.5 years
*Eshak et al., 2010	Japan Collaborative Cohort Study (Japan)	Sample from the general population	M/F	40-79	110792: 2080 CVD, 422 CHD and 983 stroke deaths	Validated 40-item FFQ.	Method similar to AOAC (Prosky et al., 1988)	CHD: 10th I20-I25. Other CVD 10th I30-I52, total CVD 10th I01-I99. Stroke 10th I60-I69	14.3years (4.2% loss to follow-up)

Reference	Cohort name (Country)	Population characteristics/ notes	Sex	Age at baseline years	Initial cohort size: and case numbers	Dietary assessment	Fibre estimation method	Outcome definition by ICD codes †	Follow up duration (loss)
Kaushik et al., 2009	Blue Mountains Eye Study (Australia)	Older age population cohort	M/F	Median 65	3654: 95 stroke deaths, CHD cases not reported	Validated 145 item FFQ.	AOAC	Stroke 9th 430-438.9/ 10th I60.0-I69.9. IHD not defined	13 years (29% loss to follow-up)
*Kokubo et al., 2011	Japan Public Health Centre-based cohort (Japan)	Representative sample from many regions	M/F	40-69	133323: 2553 stroke, 684 CHD cases	Validated 138-item FFQ.	Method similar to AOAC (Prosky et al., 1985)	CVD I00-I99. Stroke confirmed with National Survey Stroke criteria. CHD, symptoms plus ECG or cardiac enzyme abnormalities	10.4years
Larsson et al., 2009	The ATBC study (Finland)	Male smokers recruited into RCT	M	50-69	29133: 2381 fatal +non-fatal stroke cases	Validated 276 item FFQ referring to diet over previous year.	Englyst	Stroke, 8th 430-434 and 436, 9th 430-431, 433-434, 436. 10th I60, I61, I63-I64. Excluding 8th 431.01 and 431.91 and 9th 4330X, 4331X, 4339X, 4349X	13.6 years
Liu et al., 2002	The Women's Health Study (US)	Health professionals in RCT for aspirin and Vitamin E supplementation	F	Mean 54	39876: 570 CVD and 171 MI cases	Validated semi-quantitative 131-item FFQ.	AOAC	CVD: MI, stroke, PTCA, CABG. IHD defined using WHO criteria for symptoms plus ECG changes or cardiac enzyme changes	6 years
Mozaffarian et al., 2003	Cardiovascular Health Study (US)	Randomly selected older participants from communities	M/F	>65	5201: 811 fatal and non-fatal CHD cases	Validated 99 item FFQ referring to diet over previous year.	AOAC	CHD: ICD codes not reported	8.6 years
Oh et al., 2005	Nurses' Health Study (US)	Health professionals free of CHD at baseline	F	30-55	121700: 1020 cases of stroke	61 and 116 item FFQs. Diet assessed 4 times between 1980-94.	AOAC	Stroke classified according to (Walker et al., 1981) and excluding infection, trauma or malignancy	18 years
* Park et al., 2011	NIH-AARP Diet and Health Study (US)	Representative sample from many US states	M/F	50-71	388122: 7665 CVD deaths	124 item FFQ. Intake over prior 12 months.	AOAC	CVD 10th I00-I78	9 years
Pietinen et al., 1996	The ATBC Study (Finland)	Male smokers recruited into RCT	M	50-69	29133: 1399 fatal and non fatal CHD events	Validated 276 item FFQ referring to diet over previous year.	Englyst	CHD 9th 410-414	6.1 years

Reference	Cohort name (Country)	Population characteristics/ notes	Sex	Age at baseline years	Initial cohort size: and case numbers	Dietary assessment	Fibre estimation method	Outcome definition by ICD codes †	Follow up duration (loss)
Rimm et al., 1996	Health Professionals' Follow-up study (US)	Male health professionals free of CHD at baseline	M	40-75	51529: 740 CHD events	Validated 131 item FFQ referring to diet over previous year.	AOAC	MI defined using WHO criteria. IDC codes used to define CHD were not reported	6 years (6% loss)
Streppel et al., 2008	Zutphen Elderly Study (The Netherlands)	Random sample of men from industrial town in Netherlands	M	Mean 49	1373: 348 fatal CHD events	Diet history, several times. Intake over prior 6 to 12 months.	AOAC	CHD 9th 410-414	40 years (0.2% loss to follow-up)
Todd et al., 1999	Scottish Heart Health Study (Scotland)	Recruited via GP surgeries in Scotland	M/F	40-59	11629: 292 male and 97 female CHD cases	Validated 60-item semi-quantitative FFQ.	Englyst and Southgate	CHD 9th 410-414	9 years (0.1% loss to follow-up)
* Wallstrom et al., 2012	Malmo Diet and Cancer Cohort (Sweden)	Adults living around Malmo identified from national registries	M/F	58	28098: 1764 ischaemic CVD cases, 743 strokes	Interview based diet history method.	Non-starch polysaccharide	Ischaemic CVD 10th I20-I25, I63, I64/ 9th 410-414, 434, 436. Ischaemic CHD 10th I20-I25/ 9th 410-414. Ischaemic stroke 10th I63, I64/ 9th 434, 436	13 years
* Ward et al., 2012	EPIC-Norfolk (England)	Recruited via GP registers	M/F	40-79	25639: 2151 CHD cases	FFQ and 7-day diaries.	Englyst	CHD 9th 410-414/ 10th I20-I25	11 years
Wolk et al., 1999	Nurses' Health Study (US)	Health professionals free of CHD at baseline	F	30-55	121700: 591 CHD cases	Validated 116 item FFQ. Diet assessed at least 3 times.	AOAC	ICD codes used to define CHD were not detailed	10 years (20% loss)

Key: *Identified during update search; † International Classification of Disease Codes, 8th and 9th versions start with a number, 10th edition codes start with the letter 'I'.

Abbreviations: AOAC Association of Official Analytical Chemists; CABG coronary artery bypass graft; CHD coronary heart disease; CVD cardiovascular disease; ECG electro-cardio graph; EPIC European Prospective Investigation into Cancer and Nutrition; F female; FFQ food frequency questionnaire; GP general practitioner; IHD ischaemic heart disease; M male; MI myocardial infarction; PTCA percutaneous transluminal coronary angioplasty; RCT randomised controlled trial; US United States; WHO world health organisation.

2.4.2 Meta-analyses and comparison across similar studies

In total, ten publications reported total circulatory or CVD events (Akbaraly et al., 2011, Baer et al., 2011, Bazzano et al., 2003, Buyken et al., 2010, Chuang et al., 2012, Eshak et al., 2010, Kokubo et al., 2011, Liu et al., 2002a, Park et al., 2011, Wallstrom et al., 2012), sixteen reported coronary events (Appleby et al., 1999, Bazzano et al., 2003, Bernstein et al., 2011, Crowe et al., 2012, Eshak et al., 2010, Kaushik et al., 2009, Kokubo et al., 2011, Liu et al., 2002a, Mozaffarian et al., 2003, Pietinen et al., 1996, Rimm et al., 1996, Streppel et al., 2008, Todd et al., 1999, Wallstrom et al., 2012, Ward et al., 2012, Wolk et al., 1999), and eight cohort studies were indentified reporting stroke events (Ascherio et al., 1998, Bazzano et al., 2003, Eshak et al., 2010, Kaushik et al., 2009, Kokubo et al., 2011, Larsson et al., 2009, Oh et al., 2005, Wallstrom et al., 2012). Individual study results are detailed in Appendix IV.

Individual study and combined estimates are displayed on forest plots (e.g. Figure 2.3a). The black squares and horizontal lines display individual study risk estimates and 95% CIs around the estimate. The area of the black boxes reflects the contributing weight of each study to the summary estimate and loosely reflects study size. The weight contribution of each study is related to the inverse of the variance (standard error) (Juni et al., 2008) and as larger studies tend to have smaller variance, larger studies are represented by larger squares. The combined estimate is represented with a diamond shape, with the left and right extremes showing the 95% CIs.

The cubic spline figures (e.g. Figure 2.3b) display dose-response associations. Marks on the *x*-axis indicate category mean fibre intakes reported from each study, so it is possible to see where, throughout the intake range, evidence is greatest and that data become sparse at lower and higher levels. The line of best fit indicates a summary estimate for risk along the range of intakes with 95% CIs. CIs meet the line of best fit at a point of no uncertainty, where the RR=1. This is the reference intake that was set according to the mean fibre intake reported in included studies.

Since data are sparse at lower and higher intakes, it is important not to extrapolate risk associations outside normal or plausible intake ranges and not to over interpret non-linearity of lines where the spread of data points and CIs widen. For the reasons discussed above, spline graphs remain mainly informative for displaying general patterns rather than indicating risk at specific intake levels.

Total dietary fibre

Total fibre and CVD: Nine publications reported total CVD risk and total dietary fibre intake and all were included in the dose-response meta-analysis (Liu et al., 2002a, Bazzano et al., 2003, Buyken et al., 2010, Chuang et al., 2012, Eshak et al., 2010, Akbaraly et al., 2011, Kokubo et al., 2011, Park et al., 2011, Wallstrom et al., 2012). Different criteria were employed by each study to classify CVD cases. Five studies included any circulatory condition (equivalent to the International Classification of Disease (ICD) codes 10th Edition I00-I99) (Buyken et al., 2010, Chuang et al., 2012, Eshak et al., 2010, Akbaraly et al., 2011, Kokubo et al., 2011) and four studies used narrower definitions (Liu et al., 2002a, Buyken et al., 2010, Park et al., 2011, Wallstrom et al., 2012).

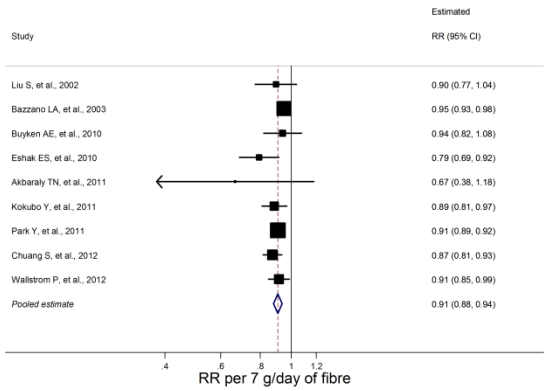
The summary RR per 7g/day increase in total fibre was 0.91 (95% CI: 0.88 to 0.94) with some evidence of heterogeneity between studies I^2=51% (95% CI: 0 to 77%) (Figure 2.3a). The dose-risk association for total fibre intake and CVD is displayed in Figure 2.3b where it appears that event risk steadily reduces with increasing total fibre intake.

Total fibre and CHD: Eleven of fourteen publications reporting coronary event risk and total fibre intake were included in the dose-response meta-analysis (Pietinen et al., 1996, Rimm et al., 1996, Wolk et al., 1999, Liu et al., 2002a, Bazzano et al., 2003, Mozaffarian et al., 2003, Streppel et al., 2008, Eshak et al., 2010, Kokubo et al., 2011, Crowe et al., 2012, Wallstrom et al., 2012). Results concerning fatal and non-fatal events from the Norfolk arm of the EPIC study (Ward et al., 2012) were not included as fatal cases were additionally included in another publication (Crowe et al., 2012). The Scottish Heart Health Study could not be included as only data for fibre density were reported and it was not possible to quantify intakes in each quartile (Todd et al., 1999). Results from the Oxford Vegetarian Study (Appleby et al., 1999) were also not included in the meta-analysis as participants from this cohort had been invited to participate in the Oxford branch of the EPIC study (Davey et al., 2003), which formed part of another included cohort (Chuang et al., 2012).

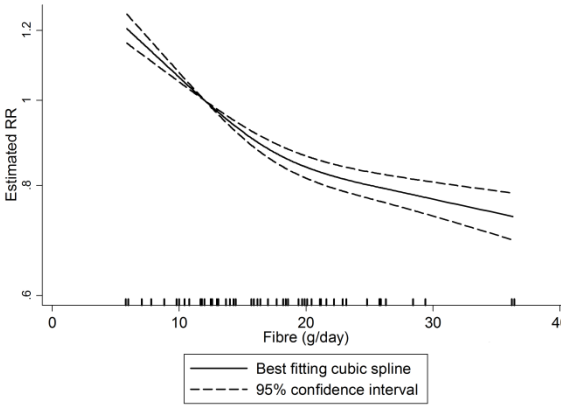
The combined risk estimate per 7g/day increase in fibre was 0.90 (95% CI: 0.87 to 0.94) and there was moderate evidence of heterogeneity between studies, I^2=38% (95% CI: 0 to 70%) Figure 2.3c. The dose-risk figure shows CHD risk steadily decreased with greater total fibre intake but CIs around the estimate widened towards the upper intake levels, where data are sparse, so interpretation of risk at these higher intakes should be undertaken cautiously (Figure 2.3d).

Total fibre and all stroke: Seven studies reported total dietary fibre intake in relation to stroke risk, all of which were included in the dose-response meta-analysis (Figure 2.3e) (Ascherio et al., 1998, Bazzano et al., 2003, Eshak et al., 2010, Kokubo et al., 2011, Larsson et al., 2009, Oh et al., 2005, Wallstrom et al., 2012). The combined RR per 7g/day increase was 0.93 (95% CI: 0.88 to 0.98) and there was some evidence of heterogeneity between studies I^2=59% (95% CI: 7 to 82%). Stroke risk appeared to steadily reduce with increasing total fibre intake (Figure 2.3f). Data points became especially sparse above 25g/day and so extrapolation of risk at higher intakes should be undertaken with caution.

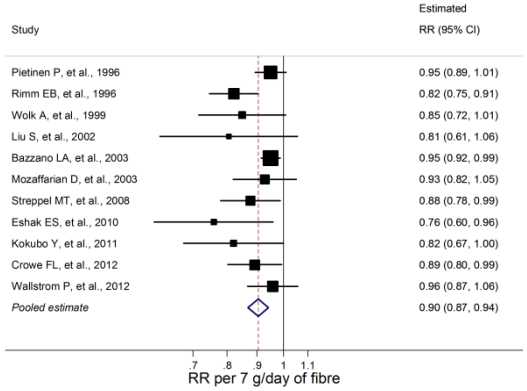
a. total fibre intake and total CVD



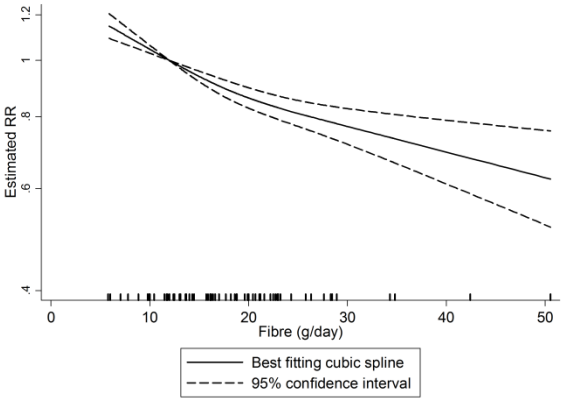
b. total fibre intake and total CVD



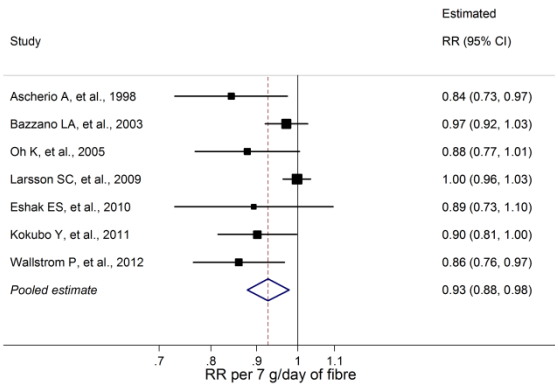
c. total fibre intake and CHD



d. total fibre intake and CHD



e. total fibre intake and stroke



f. total fibre intake and stroke

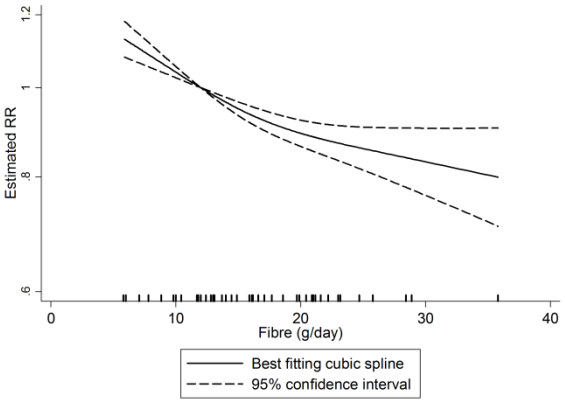


Figure 2.3 Forest plots and restricted cubic spline figures for total fibre intake and CVD, CHD and stroke

Total fibre and haemorrhagic or ischaemic stroke: Four cohorts reported results for total fibre intake and ischaemic stroke and three for haemorrhagic stroke (Table 2.3). Risk of intracerebral haemorrhagic stroke was significantly reduced with greater fibre intake in Japanese women (Kokubo et al., 2011), but this significance did not remain across the trend or for haemorrhagic stroke in the other two cohorts reporting this outcome (Larsson et al., 2009, Oh et al., 2005).

Case numbers for ischaemic stroke or cerebral infarction were greater than for haemorrhagic events and a significant risk reduction was observed with greater fibre intake, again in Japanese women (Kokubo et al., 2011), and also Swedish men (Wallstrom et al., 2012) but not in either the Nurses' Health Study or the ATBC study (Oh et al., 2005, Larsson et al., 2009).

Table 2.3 Total fibre intake and stroke risk by stroke subtype

Cohort	Fibre intake/ comparison	Outcomes measured	Cases	Relative Risk (95% Confidence Interval)	p-trend
Japan public health centre-based cohort (Kokubo et al., 2011)	M: 19.9 vs.6.0 F:21.6 vs. 7.8 g/day	Cerebral infarction	910 518	Male:0.94 (0.66, 1.34) Female: 0.73 (0.55, 0.97)	0.540 0.029
		Subarachnoid haemorrhage	133 226	Male:1.02 (0.45, 2.54) Female: 0.72 (0.37, 1.43)	0.672 0.419
		Intracerebral haemorrhage	456 310	Male:1.08 (0.66, 1.78) Female: 0.53 (0.28, 0.97)	0.588 0.100
Alpha-Tocopherol Beta- Carotene Study (Larsson et al., 2009)	35.8 vs. 16.1 g/day	Cerebral infarction	2702	Male: 1.01 (0.85, 1.19)	0.83
		Subarachnoid haemorrhage	196	Male: 0.86 (0.47, 1.59)	0.49
		Intracerebral haemorrhage	383	Male: 0.97 (0.61, 1.54)	0.63
Nurses' Health Study (Oh et al., 2005)	21 vs.10g/day	Ischaemic stroke	515	Female: 0.78 (0.56, 1.09)	0.09
		Haemorrhagic stroke	279	Female: 0.84 (0.54, 1.30)	0.34
Malmo diet and cancer cohort (Wallstrom et al., 2012)	M:11.4 vs. 5.8 F:12.9 vs. 6.5 g/1000kcal	Ischaemic stroke	397 346	Male: 0.69 (0.49, 0.96) Female: 0.73 (0.52, 1.04)	0.05 0.18

Key: M=male, F=female

Insoluble fibre

Insoluble fibre and CVD: Three studies reported water-insoluble fibre and risk of total CVD (Eshak et al., 2010, Kokubo et al., 2011, Liu et al., 2002a). Meta-analysis of results was not possible as one study did not report details of insoluble fibre intake (Kokubo et al., 2011). A protective association was observed in both Japanese cohorts (Eshak et al., 2010, Kokubo et al., 2011) but not in the American Women's Health Study (Liu et al., 2002a).

Insoluble fibre and CHD: Five publications reported incident CHD risk and insoluble fibre intake and four were included in the meta-analysis (Pietinen et al., 1996, Rimm et al., 1996, Liu

et al., 2002a, Eshak et al., 2010). Results from the Japan Public Health Centre-Based Cohort were not included in the meta-analysis because no estimate of insoluble fibre intake was presented (Kokubo et al., 2011).

The summary estimate per 7g/day increase in insoluble type fibre was 0.78 (95% CI: 0.62 to 0.97) and evidence of heterogeneity was reasonably high, I^2=79% (95% CI: 44 to 92%) (Figure 2.4a). From the spline graph (Figure 2.4b) it is possible to see CHD risk gradually decrease with increasing intakes of insoluble fibre. It is important to note that data across the range of intakes are sparse and are provided only by four studies.

Insoluble fibre and stroke: Three cohorts reported stroke risk and insoluble fibre (Eshak et al., 2010, Kokubo et al., 2011, Larsson et al., 2009) but a meta-analysis could not be conducted as one paper did not provide an estimate for insoluble fibre intake in the sample (Kokubo et al., 2011). Significant risk reduction was observed for total stroke, cerebral infarction and intracerebral haemorrhage in Japanese women in one study (results for men not presented in paper) (Kokubo et al., 2011) but this was not observed for total stroke risk in either men or women in another Japanese cohort (Eshak et al., 2010). Non-significant results were also seen in the participants of the Finnish ATBC study for ischaemic and haemorrhagic stroke (Larsson et al., 2009).

Soluble fibre

Soluble fibre and CVD: Two Japanese and two American studies reported risk estimates for water-soluble fibre and total CVD. One study from each country observed protective associations with greater soluble fibre intake (Bazzano et al., 2003, Kokubo et al., 2011) but the other studies did not (Eshak et al., 2010, Liu et al., 2002a).

Soluble fibre and CHD: Six cohorts reported incident CHD risk and soluble fibre intake, five of these were included in the meta-analysis (Pietinen et al., 1996, Rimm et al., 1996, Liu et al., 2002a, Bazzano et al., 2003, Eshak et al., 2010). Results from the Japan Public Health Centre-Based Cohort were not included in the meta-analysis because no estimate of soluble fibre intake was presented (Kokubo et al., 2011).

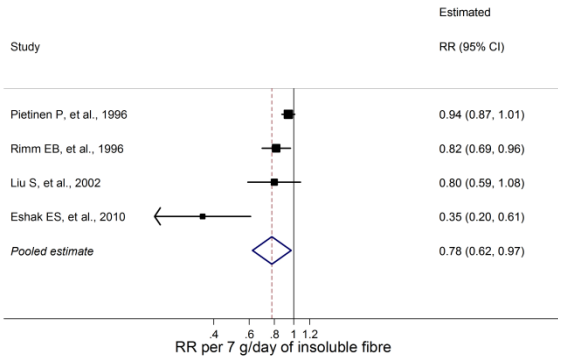
The combined risk estimate per 4g/day increase in soluble type fibre was 0.88 (95% CI: 0.75 to 1.04) and evidence of heterogeneity was moderate, I^2=57% (95% CI: 0 to 84%) (Figure 2.4c).

Figure 2.4d illustrates a trend for decreasing CHD risk with increasing soluble fibre intake but CIs around the estimate remain wide across intake levels as data are thinly spread.

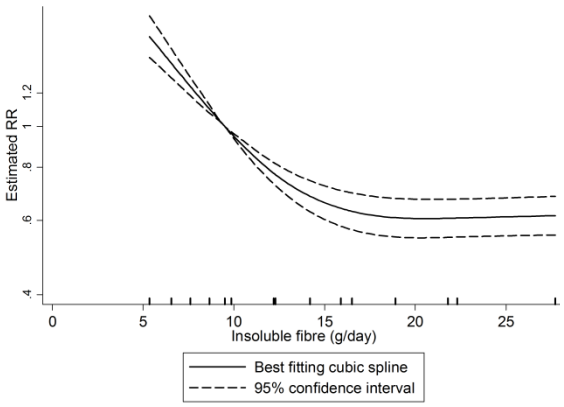
Soluble fibre and stroke: Four studies presented stroke risk in relation to soluble fibre intake (Bazzano et al., 2003, Eshak et al., 2010, Kokubo et al., 2011, Larsson et al., 2009) and all but one, which did not present an estimate of soluble fibre intake (Kokubo et al., 2011), were included in the meta-analysis (Figure 2.4e).

For each 4g/day increase in soluble fibre, risk was reduced by 6%: RR 0.94 (95% CI: 0.88 to 1.01). Evidence of heterogeneity between studies was relatively low, I^2=21% (95% CI: 0 to 92%) but since the summary estimate was based on only three studies, it should be interpreted with care. The study that could not be included did not observe significant risk reduction in total stroke or sub-types of stroke with soluble fibre intake and only reported results in women (Kokubo et al., 2011).

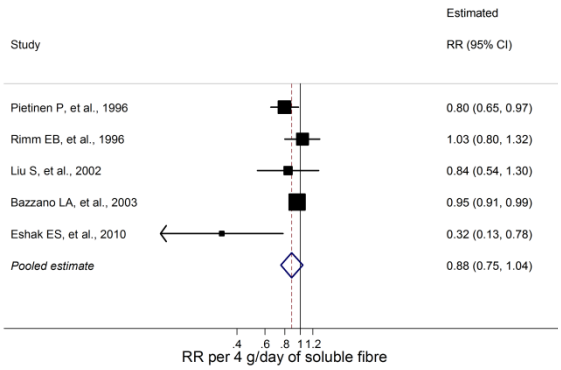
a. Insoluble fibre intake and CHD



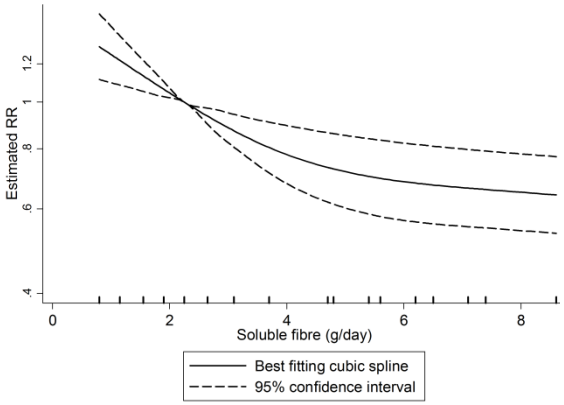
b. Insoluble fibre intake and CHD



c. Soluble fibre intake and CHD



d. Soluble fibre intake and CHD



e. Soluble fibre intake and stroke

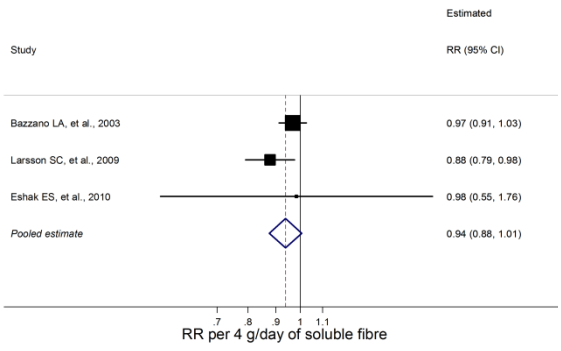


Figure 2.4 Forest plots and restricted cubic spline figures for insoluble and soluble fibre

Cereal fibre

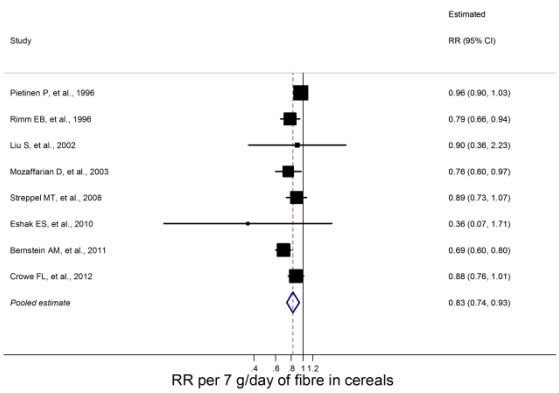
Cereal fibre and CVD: Three American studies, the pooled EPIC study and one Australian cohort reported total CVD risk and cereal fibre intake. Greater intake was significantly associated with risk reduction in three of the studies (Baer et al., 2011, Park et al., 2011, Chuang et al., 2012) and not in the others (Buyken et al., 2010, Liu et al., 2002a).

Cereal fibre and CHD: CHD risk in relation to cereal fibre intake was reported in 11 publications. Eight were included in the meta-analysis (Pietinen et al., 1996, Rimm et al., 1996, Liu et al., 2002a, Mozaffarian et al., 2003, Streppel et al., 2008, Eshak et al., 2010, Bernstein et al., 2011, Crowe et al., 2012). Results from the EPIC Norfolk study (Ward et al., 2012) were not included as this cohort was included in the pooled EPIC estimate (Crowe et al., 2012). Two results from the Nurses' Health study were identified and the results from Bernstein *et al.*(Bernstein et al., 2011) were included over Wolk *et al.* (Wolk et al., 1999) because of longer follow-up. Results from the Australian Blue Mountain Eye Study were not included because data presented were insufficient to derive a dose-response trend (Kaushik et al., 2009).

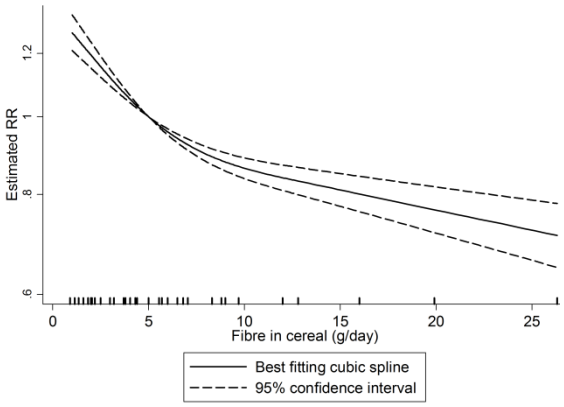
The combined estimate per 7g/day increase in fibre from cereal sources was 0.83 (95% CI: 0.74 to 0.93) and evidence of heterogeneity between studies was high, I^2=68% (95% CI: 33 to 85%) (Figure 2.5a). The dose-response curve (Figure 2.5b) illustrates that CHD risk continues to fall with increasing intakes of fibre from cereals, although data are concentrated around lower intake levels so less weight should be placed on risk estimates at higher intakes.

Cereal fibre and stroke: Three cohorts reported stroke risk and cereal fibre intake (Kaushik et al., 2009, Larsson et al., 2009, Oh et al., 2005). When combined, heterogeneity between studies was very high I^2=90% (95% CI: 73 to 96%) and a summary estimate is therefore not presented since this would be unreliable (Figure 2.5c). The Nurses' Health Study reported significant risk reduction for total and haemorrhagic, but not ischaemic stroke (Oh et al., 2005) and the Australian cohort also reported a significant risk reduction for total stroke with greater cereal fibre intake (Kaushik et al., 2009). No significant association was seen in the ATBC study (Larsson et al., 2009).

a. Cereal fibre intake and CHD



b. Cereal fibre intake and CHD



c. Cereal fibre intake and stroke

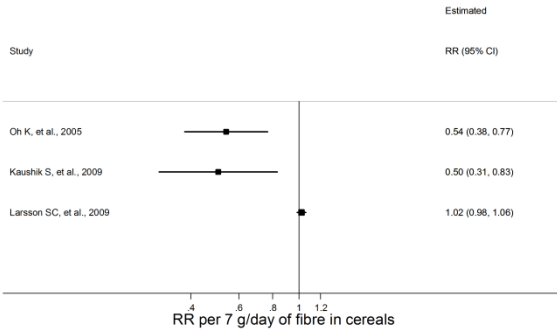


Figure 2.5 Forest plots and restricted cubic spline figures for cereal fibre

Vegetable fibre

Vegetable fibre and CVD: No significantly protective association was observed in three of the four cohorts reporting total CVD risk and vegetable fibre intake (Buyken et al., 2010, Liu et al., 2002a, Park et al., 2011). However, an inverse association for vegetable fibre and circulatory death disease was reported in the EPIC heart study, which included data pooled from 10 European countries (Chuang et al., 2012).

Vegetable fibre and CHD: Eight of nine publications that reported vegetable fibre intake and CHD risk were included in the meta-analysis (Pietinen et al., 1996, Rimm et al., 1996, Wolk et al., 1999, Liu et al., 2002a, Mozaffarian et al., 2003, Streppel et al., 2008, Eshak et al., 2010, Crowe et al., 2012), again the result from EPIC Norfolk was not included here (Ward et al., 2012).

The summary estimate per 4g/day increase in fibre from vegetable sources was 0.94 (95% CI: 0.89 to 1.00) and there was no evidence of heterogeneity between studies, I^2=0% (95% CI: 0 to 41%) (Figure 2.6a). CHD risk decreases with increasing intakes of vegetable fibre up to intakes around 6g/day, where the risk reduction appears to levels out, but again, data become sparse at upper intakes (Figure 2.6b).

Vegetable fibre and stroke: Stroke risk was reported in association with vegetable fibre in the ATBC study and also the American Nurses' Health Study. No significant associations were reported from the Nurses' Health Study for either total, haemorrhagic or ischaemic stroke (Oh et al., 2005). In the ATBC study, vegetable fibre only was associated with ischaemic stroke risk reduction and not haemorrhagic stroke (Larsson et al., 2009).

Fruit fibre

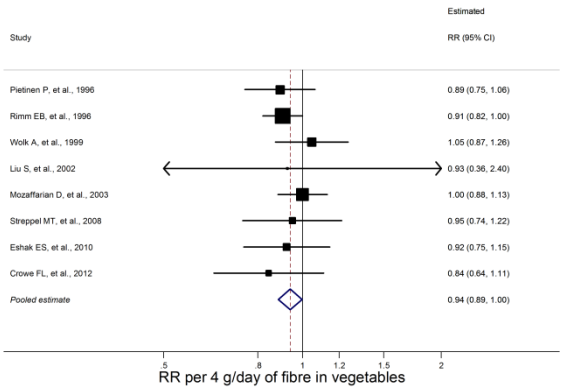
Fruit fibre and CVD: Four studies reported total CVD risk and fruit fibre intake. A significant CVD risk reduction was observed only in the male participants of the Australian Blue Mountain Eye Study (Buyken et al., 2010). No significant association was reported in the Women's Health Study, the NIH-AARP diet and health study or the EPIC heart study (Liu et al., 2002a, Park et al., 2011, Chuang et al., 2012).

Fruit fibre and CHD: Eight of nine publications that reported fibre intake from fruit and CHD risk were included in the meta-analysis (Pietinen et al., 1996, Rimm et al., 1996, Wolk et al., 1999, Liu et al., 2002a, Mozaffarian et al., 2003, Streppel et al., 2008, Eshak et al., 2010, Crowe et al., 2012). The results reported in the EPIC Norfolk paper were again not included as discussed earlier (Ward et al., 2012).

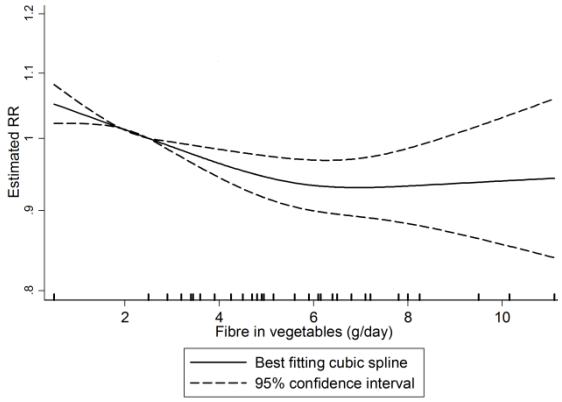
The combined risk estimate per 4g/day increase in fibre from fruit was 0.91 (95% CI: 0.82 to 1.02) and evidence of heterogeneity between studies was high, I^2=67% (95% CI: 30 to 84%) (Figure 2.6c). Similar to the dose-risk curve for vegetable fibre (Figure 2.6b), with fruit fibre (Figure 2.6d) there is some evidence that risk reduction continues with intakes of up to around 4 or 5g/day and there is some evidence of a possible threshold effect where the line flattens over higher intakes. As with vegetable fibre, data are sparse at upper intake levels.

Fruit fibre and stroke: Stroke risk was reported in association with fruit fibre in the ATBC study and also the Nurses' Health Study but no significant associations were reported in either (Oh et al., 2005, Larsson et al., 2009).

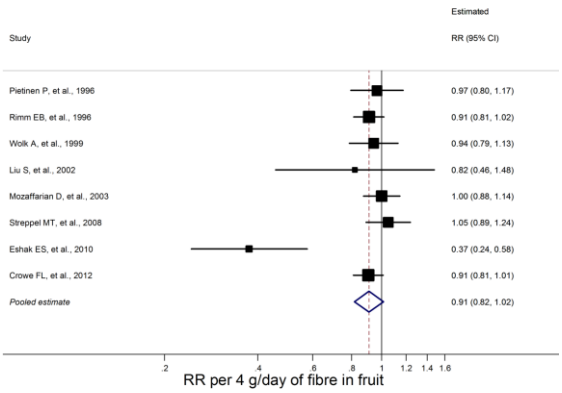
a. Vegetable fibre intake and CHD



b. Vegetable fibre intake and CHD



c. Fruit fibre intake and CHD



d. Fruit fibre intake and CHD

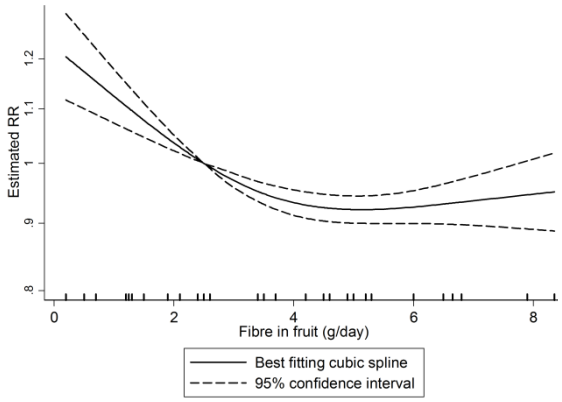


Figure 2.6 Forest plots and restricted cubic spline figures for fruit and vegetable fibre

Other sources of fibre

Too few studies reported results for fibre fractions or from other sources to permit meta-analysis but findings from the Finnish cohort study of male smokers suggest a possible protective association for cellulose and lignin intake for fatal CHD risk; however this association was not evident when non-fatal myocardial infarction events were combined with the fatal CHD cases (Pietinen et al., 1996).

The EPIC-Heart study did not see a protective association for 'other fibre' (non cereal, vegetable or fruit-derived) (Crowe et al., 2012), nor did the Zutphen Elderly Study which examined both long-term and recent legume and potato fibre intake (Streppel et al., 2008). An American study however did see a protective association for legume fibre in women but not men for fatal CVD risk (Park et al., 2011).

Meta-regression

Meta-regression was conducted to explore possible heterogeneity created by differing study characteristics or through adjustment for different confounding variables (Table 2.4). These results should be considered more exploratory than confirmatory because of the smaller numbers of studies combined to give risk estimates and also because of the increased potential for finding false positives, by chance, through multiple testing.

Gender:

There was no evidence of significant heterogeneity between subgroups of studies when results were combined for those reporting in males, females or mixed-gender ($p>0.05$). There was however, suggestion that fibre was protectively associated with all three outcomes in males and fibre also appeared to be protectively associated with CHD risk reduction in the two studies reporting results for women RR 0.84 (95% CI: 0.72 to 0.97).

Method used to assess fibre intake:

For total CVD and stroke outcomes, risk estimates were similar for studies estimating fibre as NSP and those estimating fibre as NSP plus resistant starch and lignin (AOAC method). For CHD, the protective association for fibre appeared stronger when calculated using the AOAC method, RR 0.90 (95% CI: 0.84 to 0.96) although the combined estimate for the two studies estimating fibre as NSP was also indicative of a protective association, with CIs just stretching to the line of no effect, RR 0.95 (95% CI: 0.90 to 1.00).

Fatal or total events:

Reporting fatal events only or incidence data marginally influenced risk estimates for total CVD and CHD but the combined study estimates remained indicative of a significant protective association. Only one study reported fatal stroke risk and greater fibre intake did not appear to be significantly associated with risk, RR 0.89 (95% CI: 0.73 to 1.10). This observation was in contrast to the combined estimate for the studies reporting stroke incidence data, RR 0.93 (95% CI: 0.88 to 0.98).

Length of follow-up:

Studies were divided based on follow-up duration being shorter or longer than 10 years. There was no marked difference in results using this criterion and there was no evidence of significant heterogeneity between studies when grouped in this way.

Geographic location:

For CVD and CHD risk, significant protective associations were reported for greater fibre intake in studies conducted in the US, Europe or other parts of the world. For stroke risk, the combined estimate for the two European studies indicated very high heterogeneity and therefore an unreliable not-significant estimate, RR 0.94 (95% CI: 0.81 to 1.08) I^2=82%. Stroke risk however was significantly lower with greater fibre intake in the studies conducted both in the US, RR 0.91 (95% CI: 0.83 to 1.00) or other parts of the world RR 0.90 (95% CI: 0.82 to 0.99).

Adjustment for confounding factors:

Almost all studies included adjustments for age, anthropometry, gender (where appropriate), smoking, physical activity and energy intake so it is not possible to explore study results based on adjustment for these factors.

Two studies reporting CVD events did not adjust for alcohol intake and the combined result for these two differed from the main result for CVD and was not significant because of wide CIs, RR 0.89 (95% CI: 0.69 to 1.14). Removing these two study results did not change the overall summary estimate for CVD events.

The majority of studies did not include adjustment for family history of CVD. One study reporting CVD events and that had included adjustment for family history, reported a non-significant inverse association but removing this result did not impact the overall summary estimate. Three studies reporting CHD included adjustment for family history of CVD and the summary estimate indicated a stronger inverse association RR 0.83 (95% CI: 0.76 to 0.90) compared to those studies not including this adjustment RR 0.93 (95% CI: 0.90 to 0.96), although both estimates were statistically significant and the within subgroup heterogeneity was low for both combined values. There was evidence of heterogeneity (p=0.02) between the subgroup combined estimates for CHD risk.

Two studies presenting stroke outcome data had included adjustment for parental myocardial infarction. The combined risk estimate for these indicated a stronger inverse association RR 0.86 (95% CI: 0.78 to 0.95) than those studies not adjusting for family history, where the risk estimate was weaker and CIs reached the line of no effect RR 0.95 (95% CI: 0.90 to 1.00).

Table 2.4 Study subgroup combined risk estimates for total fibre intake and CVD, CHD and stroke

Subgroup of studies		CVD					CHD					Stroke				
	Subgroup	RR (95% CI)	I^2	n	P_{het}*	P_{het}†	RR (95% CI)	I^2	n	P_{het}*	P_{het}†	RR (95% CI)	I^2	n	P_{het}*	P_{het}†
subjects' gender	Male	0.91 (0.89, 0.92)	0%	5	0.8	0.2	0.90 (0.84, 0.96)	53%	5	0.07	0.6	0.91 (0.83, 1.00)	75%	4	0.07	0.8
	Mixed	0.86 (0.73, 1.02)	73%	3	0.03		0.92 (0.87, 0.98)	33%	4	0.2		0.97 (0.92, 1.02)	0%	2	0.4	
	Female	0.90 (0.77, 1.04)		1			0.84 (0.72, 0.97)	0%	2	0.7		0.88 (0.77, 1.01)		1		
gender in same study	Male	0.91 (0.89, 0.92)	0%	5	0.8	0.9	0.90 (0.84, 0.96)	53%	5	0.07	0.3	0.91 (0.83, 1.00)	75%	4	0.07	0.6
	Female	0.90 (0.77, 1.04)		1			0.84 (0.72, 0.97)	0%	2	0.7		0.88 (0.77, 1.01)		1		
method used to	AOAC	0.91 (0.88, 0.94)	60%	7	0.02	1	0.88 (0.83, 0.93)	45%	9	0.07	0.2	0.92 (0.87, 0.97)	26%	5	0.2	0.5
assess fibre	not AOAC	0.89 (0.75, 1.06)	14%	2	0.3		0.95 (0.90, 1.00)	0%	2	0.09		0.94 (0.81, 1.08)	82%	2	0.2	
includes non-fatal	No	0.89 (0.82, 0.95)	36%	4	0.2	0.6	0.85 (0.75, 0.96)	71%	2	0.3	0.3	0.89 (0.73, 1.10)		1		0.8
events (incidence)	Yes	0.91 (0.88, 0.95)	46%	5	0.1		0.91 (0.88, 0.95)	36%	9	0.1		0.93 (0.88, 0.98)	65%	6	0.01	
length of follow-up	<10 years	0.91 (0.89, 0.92)	0%	2	0.9	0.2	0.89 (0.82, 0.97)	57%	4	0.07	0.9	0.84 (0.73, 0.97)		1		0.4
	≥10 years	0.90 (0.86, 0.95)	55%	7	0.04		0.91 (0.87, 0.96)	30%	7	0.2		0.94 (0.89, 0.99)	56%	6	0.05	
geographic location	US	0.93 (0.89, 0.96)	75%	3	0.02	0.7	0.89 (0.82, 0.96)	58%	5	0.05	0.4	0.91 (0.83, 1.00)	56%	3	0.1	0.6
	EU	0.89 (0.85, 0.94)	40%	3	0.4		0.93 (0.89, 0.97)	37%	4	0.5		0.94 (0.81, 1.08)	82%	2	0.2	
	Other	0.88 (0.80, 0.95)	32%	3	0.2		0.80 (0.68, 0.92)	0%	2	0.6		0.90 (0.82, 0.99)	0%	2	0.9	
adjusted for age	Yes	0.91 (0.88, 0.94)	51%	9	0.04		0.90 (0.87, 0.94)	38%	11	0.1		0.93 (0.88, 0.98)	59%	7	0.06	
	No															
adjusted for alcohol	Yes	0.91 (0.88, 0.94)	60%	7	0.02	0.9	0.90 (0.87, 0.94)	38%	11	0.1		0.93 (0.88, 0.98)	59%	7	0.06	
	No	0.89 (0.69, 1.14)	24%	2	0.2											
adjusted for	Yes	0.91 (0.88, 0.94)	51%	9	0.04		0.90 (0.86, 0.94)	44%	10	0.06	0.7	0.93 (0.88, 0.98)	59%	7	0.06	
anthropometry	No						0.93 (0.82, 1.05)		1							
adjusted for energy	Yes	0.91 (0.88, 0.94)	51%	9	0.04		0.90 (0.87, 0.94)	38%	11	0.1		0.93 (0.88, 0.98)	59%	7	0.06	
intake	No															
adjusted for family	Yes	0.90 (0.77, 1.04)		1		0.9	0.83 (0.76, 0.90)	0%	3	0.9	0.02	0.86 (0.78, 0.95) ‡	0%	2	0.7	0.2
history	No	0.91 (0.88, 0.94)	58%	8	0.02		0.93 (0.90, 0.96)	10%	8	0.4		0.95 (0.90, 1.00)	56%	5	0.06	
adjusted for physical	Yes	0.91 (0.88, 0.94)	58%	8	0.02	0.6	0.91 (0.87, 0.95)	41%	10	0.08	0.7	0.93 (0.88, 0.98)	59%	7	0.06	
activity	No	0.94 (0.82, 1.08)		1			0.88 (0.78, 0.99)		1							
adjusted for gender	Yes	0.91 (0.88, 0.94)	51%	9	0.04		0.90 (0.87, 0.94)	38%	11	0.1		0.93 (0.88, 0.98)	59%	7	0.06	
	No															
adjusted for smoking	Yes	0.91 (0.88, 0.94)	51%	9	0.04		0.90 (0.87, 0.94)	38%	11	0.1		0.93 (0.88, 0.98)	59%	7	0.06	
	No															

P_{het}* Heterogeneity within each subgroup; P_{het}† Heterogeneity between each subgroup; ‡Adjustment for parental history of myocardial infarction in both cases
Abbreviations: AOAC Association of Official Analytical Chemists; CI confidence intervals; EU European Union; n Number of studies; RR relative risk; US united states

2.5 Discussion

2.5.1 Result summary

A significant risk reduction of around 10% was seen for both CVD and CHD and a reduction of 7% for stroke risk was identified with every additional 7g/day of total fibre consumed (Figure 2.3). Findings are aligned with current recommendations to increase fibre intake and demonstrate a clinically significant risk reduction associated with an achievable increase in daily fibre intake. As studies included in meta-analyses mainly calculated fibre using the AOAC method, this increment relates to AOAC fibre. To place this in context, an additional 7g of fibre (AOAC) is contained within approximately one portion (70g) of wholemeal pasta plus two servings of fruit or vegetables (Lunn and Buttriss, 2007).

For stroke outcomes, studies focusing on cereal, fruit or vegetable sources of fibre were too few or too heterogeneous to draw sound conclusions. The summary estimate for soluble fibre intake and stroke risk indicates an inverse association but statistical significance was not quite reached (Figure 2.3e). CHD risk was significantly and inversely associated with insoluble fibre (Figure 2.4a), vegetable fibre (Figure 2.6a) and cereal fibre intake (Figure 2.5a) but not with soluble fibre (Figure 2.4b) or fibre from fruit (Figure 2.6c), despite risk estimates being in the same direction.

The findings here relate only to fibre consumed within, rather than extracted from, foods and any public health messages must therefore reflect this. It is not clear from these observations whether fibre consumed as an extract from certain foods may be beneficial.

Meta-regression highlighted possible differences with studies using AOAC methods or not, to assess fibre intake (Table 2.4). However, the small numbers of studies using non-AOAC methods, the notable geographical differences between these studies and the likely differences between the main sources of fibre in different cohort populations somewhat limits the ability to draw conclusions.

Grouping cohort results by whether family history of disease was used as an adjustment resulted in significant heterogeneity being seen between these studies, indicating that this factor influenced final risk estimates. Additionally, grouping studies based on this greatly reduced the within subgroup heterogeneity to 0% and 10%, indicating that this factor explains a degree of the heterogeneity observed in the main summary estimate for CHD risk.

2.5.2 Findings in context, other published reviews

Although no previous reviews were identified that examined fibre and stroke, one review on whole-grains found a similar lack of published data relating to stroke risk (Flight and Clifton, 2006). The review presented mixed findings in the few studies identified, but concluded there was a strong suggestion of a protective effect of whole-grain on stroke risk (Flight and Clifton, 2006). My findings are aligned with the observation for whole-grain diets, but whole-grains contain many other potentially protective components aside from having a high fibre content (Slavin, 2004). Other protective components of whole-grains include plant stannols and sterols, found in oilseeds, grains, nuts and legumes, which are associated with reducing both biliary and dietary cholesterol absorption and also unsaturated fatty-acids, found in whole-grain wheat and oats which additionally contribute towards lowering cholesterol levels (Slavin, 2003).

The inverse associations for both CVD and CHD with total fibre intake are consistent with those of previous reviews (Liu et al., 2002a, Pereira et al., 2004, Mente et al., 2009, Hauner et al., 2012, Ye et al., 2012), as discussed earlier, in the Background section of this chapter.

Less consistent associations are apparent when considering previous findings for soluble and insoluble fibre or cereal, fruit and vegetable fibre with results from the present meta-analyses. The German dietary guidelines published in 2012 were based on a review of literature from over 25 years and concluded ‘possible’ evidence of an inverse association for cereal, fruit, soluble and insoluble fibre with CHD risk and ‘possible’ evidence of no association for vegetable fibre (Hauner et al., 2012). The pooled data presented by Pereira et al., found an inverse association for fruit but not vegetable or cereal fibre (Pereira et al., 2004) and the recent systematic review by Ye et al., suggests a protective association for cereal fibre but vegetable or fruit sources of fibre are not discussed in this publication as the focus was whole-grain in the diet (Ye et al., 2012).

The findings from this meta-analysis of systematically sought literature identify possible inverse associations for insoluble, cereal and vegetable fibre but not for soluble or fruit fibre. The design of this work improves somewhat upon previously published reviews concerning sources of fibre, which were not systematic in identifying literature (Pereira et al., 2004), did not statistically combine data (Hauner et al., 2012) or selectively reported outcomes (Ye et al., 2012).

Just two previously conducted systematic reviews report exploration of between study heterogeneity. No associations were observed when considering the different mean fibre intakes in included cohorts in one review (Mente et al., 2009) and no significant change in magnitude or direction of estimates was reported in another review that examined studies based on sex, study quality, health status, study duration, dietary intake and outcome measurement, but combined study subgroup data were not reported (Ye et al., 2012).

Published results presented in Chapter 5, relating CVD mortality risk to dietary fibre intake in the UKWCS (Threapleton et al., 2013b) have been included in updated versions of CVD and CHD meta-analyses presented in this chapter (Threapleton et al., 2013e). Extending the literature search period and including additional studies in this publication has made possible some additional meta-analyses, especially for total CVD risk. Newer risk estimates generated for CHD with the various exposures changed only minimally and there was no shift in result significance. For CVD risk and total fibre intake there was no change in the risk estimate. New meta-analyses for CVD with increased intake of insoluble, fruit, vegetable and cereal fibre were significantly associated with risk reduction. However, the combined estimate for soluble fibre and CVD risk was non-significant, with CIs spanning the line of no effect despite the estimate being in the direction of indicating a protective association (Threapleton et al., 2013e).

2.5.3 Limitations

Limitations for this systematic review and meta-analysis include the problem of residual confounding which is an issue with observational work and therefore remains a limitation when data are statistically combined. Greater intake of dietary fibre is associated with other healthy behaviour such as greater physical activity and lower smoking rates (Eshak et al., 2010, Kokubo et al., 2011), both of which may independently influence CVD risk. It is difficult to estimate the extent to which other behaviours are accurately controlled for when used as adjustments in models and therefore we cannot ascribe causality to the associations from observational without additional RCT evidence. However, given the lengthy pathogenesis of the disease, trials of adequate duration and compliance would be costly and virtually impossible to run. Yet, most of the observational studies here did include important confounders such as age, sex, education/class and smoking status in their analyses but not all adjusted for other potentially important confounders such as physical activity or other dietary factors. Exploration of adjustment for factors such as Body Mass Index (BMI) or alcohol intake, through meta-regression, did not reveal that adjustment for such confounders sufficiently

explained the strength or direction of associations observed, though, this was mainly because the majority of studies had included these as adjustments so comparisons were not possible.

Furthermore, Kaushik *et al*, Wolk *et al* and Mozaffarian *et al* point out that although residual confounding may be a limitation for prospective cohorts, it is unlikely in their cases because different observations were seen for cereal, fruit and vegetable fibre, despite intakes being highly correlated (Wolk et al., 1999, Mozaffarian et al., 2003, Kaushik et al., 2009). In light of the differing associations observed, they argue that residual confounding is not likely to be an issue of concern since all of the fibre exposures are related to healthier lifestyles (Wolk et al., 1999, Mozaffarian et al., 2003, Kaushik et al., 2009). Additionally, in the case of the Nurses' Health Study, cereal fibre but not fibre from fruit or vegetables or total fibre was protective and this association was not explained by higher intakes of associated micronutrients (vitamin E, folate, vitamin B6, magnesium) or by vegetable or fruit intake (Wolk et al., 1999). Similarly, in the Cardiovascular Health Study, risk estimates were minimally modified by adjustment for social or lifestyle factors (Mozaffarian et al., 2003).

Another possible limitation is that the majority of dietary data were collected using FFQs which may adequately characterise dietary patterns but could be limited in terms of describing individual nutrient intakes. One study explored risk of CHD with diet assessed both using FFQs and 7-day food diaries and a protective association was seen with fibre assessed from food diaries but risk was attenuated with fibre estimated by FFQs (Ward et al., 2012). The authors suggest that FFQs may not capture sufficient heterogeneity within a single population but are appropriate in pooled analyses where a wider range of intakes are collated.

Studies assessing fibre intakes using different methods (AOAC or not) were combined. Although direct comparisons may not be appropriate between studies using different intake estimations, the direction of effect and to some extent the magnitude of the associations may be similar and combining results may therefore be informative when summarising data from multiple studies. For example, in the HPFS, the protective association observed between total MI and total fibre intake calculated with the Southgate method (RR per 10g/day increase 0.76 (95% CI: 0.65 to 0.88)) and the Englyst method (RR 0.74 (95% CI: 0.61 to 0.89)) were not substantially different from AOAC calculation (RR 0.81 (95% CI: 0.70 to 0.93)) (Rimm et al., 1996).

Inconsistent results between cohorts may also result from different dietary assessment methods. For example, the Japanese Collaborative Cohort Study only included 40 food items on the FFQ and this limited list may result in difficulty when differentiating between consumers of high or low levels of both soluble and insoluble dietary fibre. Although the ATBC used a comprehensive 276-item FFQ to assess diet, the inclusion of only smokers in this study means findings must be interpreted with caution as results may not apply to wider populations since smoking may modify the effect of fibre on CVD risk.

A further limitation in meta-analysing data is the problem of publication bias, where non-significant results may be largely unreported, leading to higher chance of false positive findings. Publication bias can arise from a number of sources such as a tendency for authors to write up and submit positive findings to journals, a tendency for journals to favour acceptance of studies showing positive findings and studies with positive findings being more likely to be published in English than studies reporting no evidence of associations (Bowers et al., 2006b).

2.5.4 Strengths

A major strength of this work is the inclusion of studies from multiple online database searches, covering published literature from over two decades. The prospective nature of the included studies also avoids issues caused by recall bias. The included cohorts additionally reported on large numbers of participants, had long follow-up durations and therefore many case observations.

The quality of the meta-analyses was strengthened by generating dose-response curves rather than comparisons of high and low fibre consumers. Combining studies that examined dietary sources of fibre is an additional strength of this work as the physiological consequences of a high-fibre diet may depend on the type of fibre and the food source (Rimm et al., 1996).

2.6 Summary

Diets high in fibre, specifically from cereal or vegetable sources and rich in insoluble-type fibre should be recommended for prevention of cardiovascular diseases. These data provide evidence relating to whole-food consumption and therefore do not support consumption of foods specifically enriched in cereal or vegetable-derived fibre, however RCT data may provide insight into whole food intake compared to fibre derived from these foods on CVD risk factors.

The best available research evidence should inform public policy to ensure recommendations are rooted in knowledge (CRD, 2008). In due course, the UK SACN will issue their own interpretation of findings from this systematic review and will potentially make new recommendations or strengthen existing recommendations for England, on the basis of the work carried out here.

Exploring food sources of fibre brings understanding of how different types of fibre or, foods containing different ratios of fibre molecules, may influence CVD risk and enables tailoring of nutritional recommendations for at-risk populations. In Chapters 5, 6 and 7, I build on findings from my systematic review to explore whether fibre intake for a relatively health-conscious sample of women (the UKWCS) remains protectively associated with CHD, stroke and total CVD risk, as observed in these meta-analyses. The findings reported in these coming chapters will additionally contribute to the small group of existing studies that report stroke risk in association with total fibre and major food sources of dietary fibre.

Chapter 3 Methods, the UK Women's Cohort Study and dietary fibre intake

3.1 Chapter overview

Briefly, this chapter gives a detailed background to the UK Women's Cohort Study. Sampling methods and details of dietary and lifestyle data collection are presented in addition to simple descriptive statistics. These dietary and lifestyle descriptive statistics are relevant to later chapters, where dietary fibre intake is assessed in relation to CVD mortality (Chapter 5) and incident CVD events (Chapters 6 and 7). This chapter does not provide details of cardiovascular outcome data, which is covered in depth the next chapter (Chapter 4), or details of statistical methods as these are presented within the methods sections of each relevant chapter (refer to method section of Chapters 5, 6 and 7).

3.2 Method

3.2.1 Study population

Participants were primarily drawn from respondents to the World Cancer Research Fund's (WCRF) direct mailing survey in the early 1990's. The mailing respondents numbered approximately 500,000 (85% female) and had indicated in the survey whether they were vegetarian or meat-eaters. This survey had identified about 16,000 vegetarians and non-red meat eaters that were between the study inclusion ages (35-69 years) and had also indicated that they were willing to participate in further research.

These 16,000 vegetarian and non-red meat eating women formed the basis for the cohort and recruitment was further boosted to include meat-eaters from the same list. Additional participants were recruited into the cohort from respondents to the baseline questionnaires who were asked to identify friends and relatives of a similar age who were meat-eaters or vegetarian (Cade et al., 2004a). Figure 3.1 details recruitment of participants and the proportion of respondents to each of the two contact phases.

Baseline data were collected between 1995 and 1998 and 35,692 of 61,000 (58%) women responded to the questionnaires. Participants completed a self-administered FFQs and also provided further dietary, lifestyle and health information at baseline.

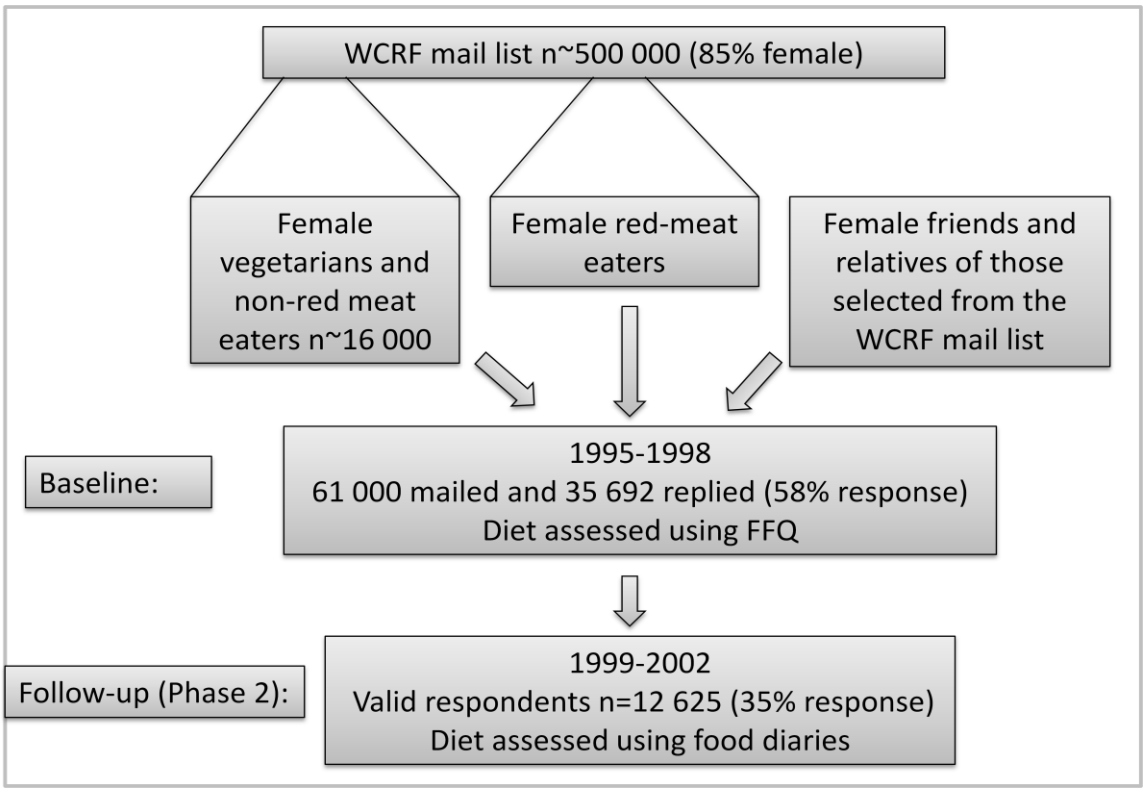


Figure 3.1 Recruitment of UKWCS participants

Around five years after baseline data were collected, the participants were asked to complete a four day food and one day exercise diary and again to provide updated health and lifestyle information. Over 14,000 women responded to this contact and of these, 12,625 returned completed food diaries (35% of baseline participants). All women that returned questionnaires and diaries were eligible for inclusion in follow-up of the cohort study.

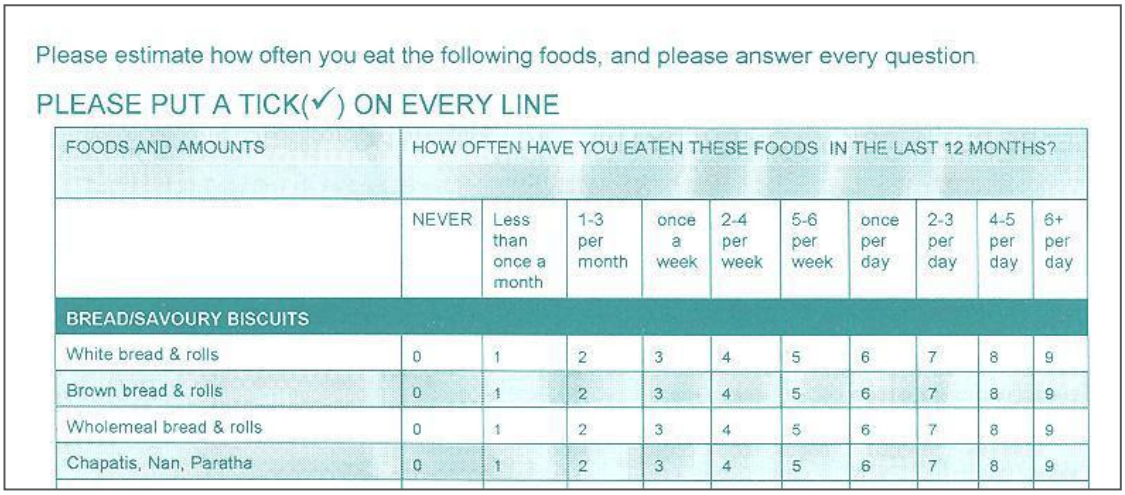
Women that provided accurate General Practitioner (GP) information or National Health Service (NHS) identification numbers at baseline (>98% of the full cohort) were successfully flagged through the Office of National Statistics (ONS), now NHS Information Centre (NHSIC) for health outcome episodes or death registration. Details of outcome event data are presented in Chapter 4.

The UKWCS population is not representative of British women, nor was it designed to be. The inclusion of high proportions of vegetarians and non-red meat eaters was intended to optimise power for exploration of foods such as fruit, vegetables and other related nutrients to disease. The motivation being that diet within a single, more representative, population may be too homogeneous to be able to detect effects of different dietary intakes (Cade et al., 2004a).

Participants are generally white (99%), middle class (63% professional or managerial), well educated (27% had degree), and married with children (86%) (Cade et al., 2004a).

3.2.2 Dietary assessment: Food Frequency Questionnaire

The FFQ sent to participants at study baseline included a list of 217 commonly consumed food items. Participants were asked to indicate their intake of each item over the previous 12 months by ticking an appropriate consumption category (from 10 choices) such as 'once per day' or '2-4 times per week' (Figure 3.2).



Please estimate how often you eat the following foods, and please answer every question

PLEASE PUT A TICK(✓) ON EVERY LINE

FOODS AND AMOUNTS	HOW OFTEN HAVE YOU EATEN THESE FOODS IN THE LAST 12 MONTHS?									
	NEVER	Less than once a month	1-3 per month	once a week	2-4 per week	5-6 per week	once per day	2-3 per day	4-5 per day	6+ per day
BREAD/SAVOURY BISCUITS										
White bread & rolls	0	1	2	3	4	5	6	7	8	9
Brown bread & rolls	0	1	2	3	4	5	6	7	8	9
Wholemeal bread & rolls	0	1	2	3	4	5	6	7	8	9
Chapatis, Nan, Paratha	0	1	2	3	4	5	6	7	8	9

Figure 3.2 Example section of baseline FFQ form

The FFQ was developed from that used in the Oxford arm of the EPIC study (Riboli and Kaaks, 1997) with modifications based on a pilot study, which was undertaken in a sample of vegetarian women (Cade et al., 2004a). Additional vegetable-based composite dishes were added to the FFQ based on food diaries completed during this pilot study and information relating to portion size estimates used for FFQ nutrient calculations were also contributed by this pilot. Nutrient intakes were calculated by multiplying each food item by the consumption frequency and estimated portion size. Portion sizes were an average calculated from three sources where the information was available (Calvert et al., 1997); 1) the pilot study food diaries; 2) food portion sizes for women from the NDNS (NDNS, 1994); 3) other published values (Crawley, 1993).

FFQ nutrient values were originally generated using data from McCance & Widdowson's *The Composition of Foods* (5th edition) (Holland et al., 1991). A mean value was created from multiple food items from *The Composition of Foods* for each FFQ row, to take account of type or preparation and cooking methods of different foods and dishes. In the FFQ, a number of

fruit items were under the heading 'seasonal' and participants were asked to mark the consumption of these foods when they were seasonally available. The number of months these items were available was taken into account when nutrient intakes were calculated.

Daily NSP values were generated for the total diet and also from key food sources (detailed below). Fibre intake was also calculated using the AOAC method. British nutrient tables do not include AOAC values for all foods and so to estimate AOAC fibre intake from the FFQ, AOAC estimates were sought from a number of sources. The following order of preference for sources of AOAC data was applied, and values in brackets indicate the proportion of data ultimately identified from each source: British reference values (7%) (Holland et al., 1991); a review article (12%) (Lunn and Buttriss, 2007); European databases (11%) (EuroFIR, 2010); United States Department of Agriculture databank (46%) (USDA, 2010); food packaging labels (18%); in-house recipe calculation (6%). A similar approach was taken to estimate soluble or insoluble dietary fibre intake from the FFQ where these values were missing in *The Composition of Foods*.

AOAC fibre, soluble and insoluble fibre values were estimated for each FFQ item by another doctoral student, Maryam Aldwairji, and values were manually input into the existing Microsoft Access nutrient database for the cohort. For accuracy, I cross checked each food item and new values to ensure all nutrients and food items were correctly matched to those used for FFQ items.

FFQ calculations for fibre from food sources

The original FFQ data for participants was converted into a Microsoft Access format and this has permitted the derivation of fibre from different food sources. Note that throughout this work, the terms, for example, 'fruit fibre', 'fibre from fruit' and 'fibre in fruit' are used interchangeably to identify fibre which is contributed to the diet by fruit. Since this work focuses on consumption of whole foods rather than extracted constituents, the terms and thus any findings relate only to fibre which is consumed within the whole food.

I grouped FFQ items to generate dietary fibre (NSP) from the following specific food groups (see Appendix V for details of which FFQ items were combined to form each food group):

- Total cereal foods
- Breakfast cereals

- Fruit (excluding juice)
- Vegetables (excluding potatoes)
- Legumes
- Nuts and seeds

Correlation of baseline FFQ fibre intakes, by type

Fibre intake assessed as either NSP or using the AOAC method were highly correlated 0.99 (95% CI: 0.99 to 0.99) (Figure 3.3). Fibre density of the diet (g/1000kcal/day) was generated for both NSP and AOAC fibre, using the total energy intake that was estimated from the FFQ. Correlation between the density values for both methods of fibre estimation was also high at 0.97 (Table 3.1).

Soluble and insoluble fibre intakes were highly correlated with each other (0.91). Fibre from the various food sources were generally not so highly correlated, except fibre from breakfast cereals and total cereal intake (0.71), which is to be expected since breakfast cereals are included within the total cereal category. For example, fruit fibre and fibre from cereal foods were not highly correlated, 0.19 (95% CI: 0.18 to 0.20), as can be seen in Figure 3.4.

Table 3.1 Correlation between fibre types and sources of fibre assessed using FFQ

	NSP	AOAC	NSP Density	AOAC Density	Soluble fibre	Insoluble fibre	Cereal fibre	Breakfast cereal fibre	Vegetable fibre	Fruit fibre	Legume fibre	Nuts/Seed fibre
NSP	1											
AOAC	0.99	1										
NSP Density	0.61	0.56	1									
AOAC Density	0.61	0.60	0.97	1								
Soluble fibre	0.95	0.96	0.51	0.56	1							
Insoluble fibre	0.98	0.98	0.62	0.64	0.91	1						
Cereal fibre	0.72	0.68	0.45	0.41	0.51	0.78	1					
Breakfast cereal fibre	0.46	0.44	0.37	0.36	0.34	0.55	0.71	1				
Vegetable fibre	0.64	0.62	0.48	0.47	0.69	0.58	0.17	0.07	1			
Fruit fibre	0.62	0.66	0.46	0.54	0.68	0.60	0.19	0.11	0.37	1		
Legume fibre	0.44	0.45	0.27	0.29	0.48	0.40	0.17	0.05	0.23	0.14	1	
Nuts/Seed fibre	0.25	0.26	0.08	0.09	0.24	0.24	0.06	0.01	0.13	0.16	0.11	1

Based on 35,262 observations, after excluding energy intake values outside the range 500-6000 kcal/day.

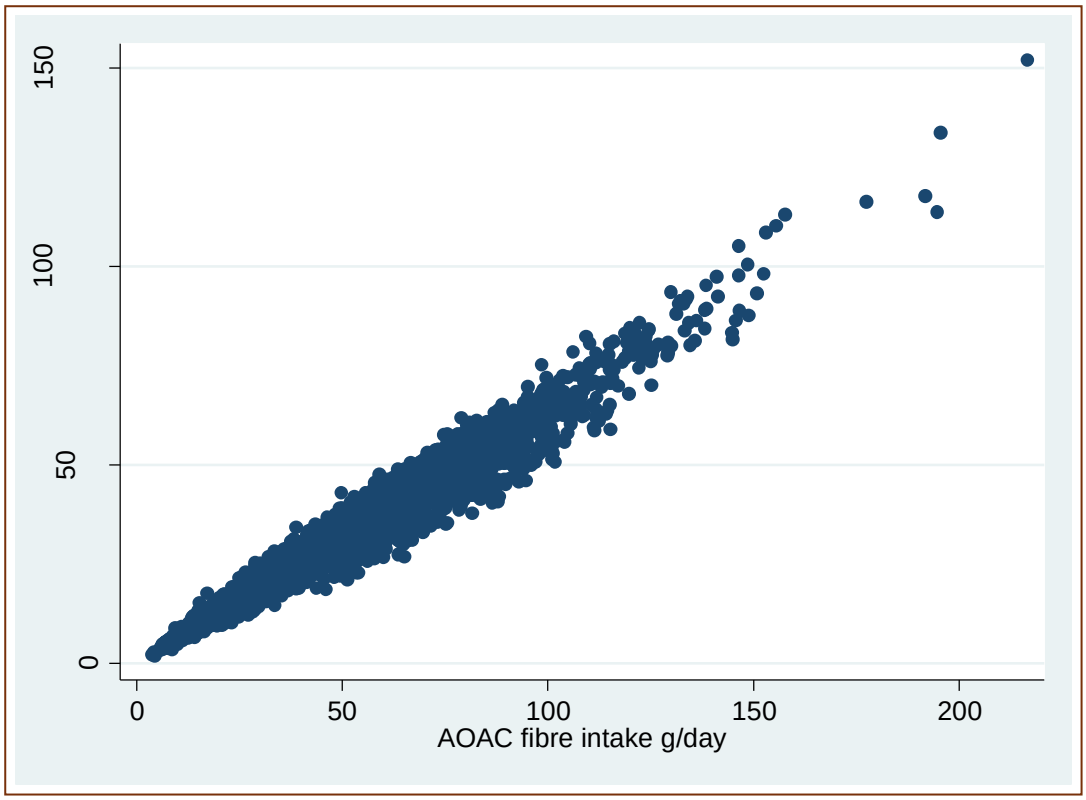


Figure 3.3 Correlation between NSP and AOAC fibre, estimated from FFQ

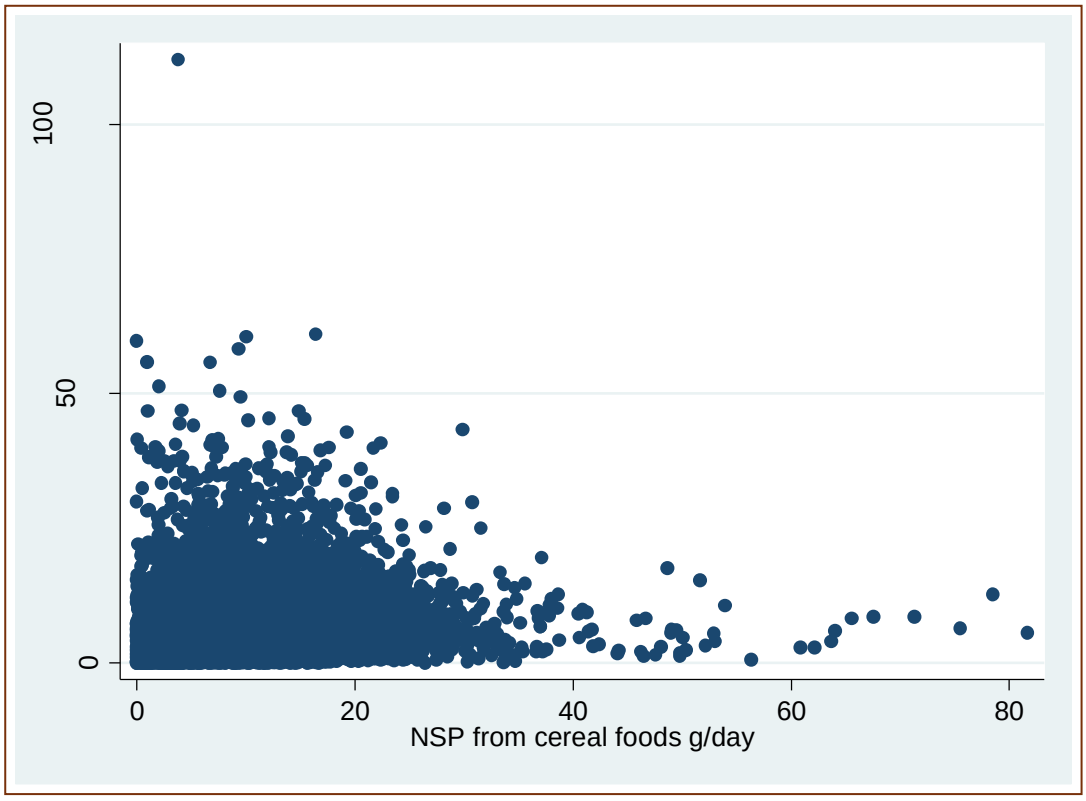


Figure 3.4 Correlation between NSP intake from fruit and cereal foods, estimated from FFQ

Characteristics of low and high fibre consumers as estimated from FFQ

I divided the sample into 5 equal-sized groups, based on NSP intake and NSP density and descriptive characteristics are presented in Table 3.2 and Table 3.3, respectively. Age increases and BMI decreases in each increasing NSP or NSP density category. The lowest NSP consumers (14.1g/day) had a median age of 50.5 years (interquartile range (IQR) 14.5) compared to the highest NSP consumers (38.6g/day) where the median age was 51.3 years (IQR 14.9).

The proportions of participants reporting history of hypertension or angina at baseline were similar across the groups, with 16% reporting history of hypertension and 2% reporting history of angina in the total sample.

The proportion of smokers, women with lower NS-SEC rankings and lower educational achievement were greatest in the lowest NSP or NSP density categories compared to the other four categories, among which the differences were not so apparent.

Physical activity levels were greater in women consuming highest levels of NSP, 17.0 vs. 12.2 Metabolic Equivalent of Task (MET) in the lowest NSP intake category. The large apparent difference in activity level is reduced once energy intake is taken into consideration, when looking at NSP density quintiles (Table 3.3). However greater activity levels are still reported with the highest NSP density intake.

Across the quintiles of NSP intake, greater intakes of carbohydrate, fat and protein were observed, likely due to the greater energy intakes in the higher NSP consumers (Table 3.2). When energy intake is accounted for, it becomes clear that those with the higher NSP density reportedly consume fewer calories than the lower NSP density groups (Table 3.3). In the higher NSP density groups, where total energy intake is lower, it appears that a greater proportion of the diet is made up of carbohydrate rather than fat, as compared with lower NSP density groups. The greater proportion of vegetarian participants in the higher NSP density group (35%) compared to the lowest NSP density group (7%) may explain some of the dietary macronutrient differences across groups.

The proportional contribution of fibre from different food groups to the total NSP intake, in the full sample of participants and NSP intake quintiles, is shown in Figure 3.5. Other sources of fibre, aside from cereals, fruit, vegetables, legumes and nuts/seeds are likely to include fibre from potatoes and also from mixed dishes, where it was not possible to determine the

contribution of separate foods groups. A notably larger proportion of NSP is provided by 'other' fibre in the lowest NSP quintile (35%), compared to the highest NSP quintile (17%). The proportion of NSP from vegetables and legumes is comparable across the groups at around 21% and 5%, respectively. NSP from fruit and cereal sources contributed less towards the total NSP intake in the lowest intake group, compared to the categories of higher NSP intake.

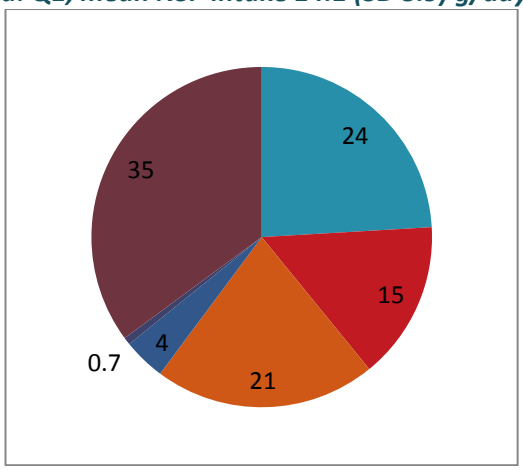
Table 3.2 Characteristics of all participants, groups based on total NSP intake (FFQ)

		Q1	Q2	Q3	Q4	Q5	All
N		6925	6929	6950	6944	6821	34569
NSP, g/day		14.1 (3.9)	19.5 (2.3)	24.0 (2.4)	29.3 (3.1)	38.6 (8.5)	23.9 (12.5)
NSP density, g/1000kcal/day		8.2 (2.9)	9.9 (3.0)	11.0 (3.2)	12.1 (3.4)	13.8 (3.8)	11.0 (4.2)
AOAC fibre, g/day		21.9 (6.1)	30.1 (3.9)	36.8 (4.3)	44.7 (5.4)	59.1(13.7)	36.8 (18.9)
AOAC density, g/1000kcal/day		12.8 (4.2)	15.4 (4.4)	16.9 (4.7)	18.6 (4.9)	21.0 (5.7)	16.9 (6.2)
Fibre types or sources, g/day	Soluble	6.5 (1.9)	8.8 (1.6)	10.5 (1.9)	12.5 (2.2)	16.3 (4.4)	10.4 (5.1)
	Insoluble	8.4 (2.7)	12.2 (2.0)	15.4 (2.1)	19.1 (2.6)	25.7 (6.1)	15.4 (8.9)
	Cereal	3.4 (2.6)	5.8 (3.70	8.0 (4.5)	10.5 (5.6)	14.3 (7.6)	7.6 (7.1)
	Breakfast cereal	0.4 (1.4)	1.4 (3.0)	2.1 (3.4)	3.2 (5.0)	4.0 (5.9)	1.8 (3.7)
	Fruit	2.1 (2.1)	3.4 (2.5)	4.2 (2.9)	5.3 (3.60	7.6 (5.9)	4.2 (3.9)
	Vegetables	3.0 (2.0)	4.2 (2.4)	5.1 (2.8)	6.0 (3.4)	8.0 (4.9)	4.9 (3.7)
	Legumes	0.6 (0.9)	0.9 (0.9)	1.1 (1.1)	1.3 (1.5)	1.9 (2.6)	1.1 (1.3)
	Nuts/seeds	0.07(0.17)	0.08(0.20)	0.08(0.24)	0.15(0.34)	0.24(0.62)	0.08 (0.29)
Age, years		50.5(14.5)	50.8(14.8)	51.2(14.8)	51.1(14.9)	51.3(14.9)	51.0 (14.8)
BMI, kg/m^2		24.0 (5.2)	23.8 (4.7)	23.7 (4.7)	23.4 (4.6)	23.2 (4.4)	23.7 (4.8)
Hypertension at baseline (%)	Yes	1165 (17)	1159 (17)	1093 (16)	1122 (16)	1113 (16)	5652 (16)
	No	5760 (83)	5770 (83)	5857(84)	5822 (84)	5708 (84)	28917 (84)
Angina at baseline (%)	Yes	136 (2)	119 (2)	132 (2)	141 (2)	176 (3)	695 (2)
	No	6789 (98)	6810 (98)	6827 (98)	6803 (98)	6645 (97)	33874 (98)
Smoking status (%)	Never smoker	3598 (53)	3931 (58)	4100 (60)	4073 (60)	4070 (61)	19772 (59)
	Current smoker	1194 (17)	779 (11)	623 (9)	538 (8)	501 (8)	3635 (11)
	Former smoker	1964 (29)	2068 (31)	2063 (30)	2137 (32)	2059 (31)	10291 (31)
Diet group (%)	Meat-eaters	5426 (78)	4932 (71)	4634 (67)	4011 (58)	3323 (49)	22326 (65)
	Poultry-eaters	147 (2)	163 (2)	191 (3)	215 (3)	267 (4)	983 (3)
	Fish-eaters	467 (7)	694 (10)	857 (12)	1056 (15)	1308 (19)	4382 (13)
	Vegetarian	885 (13)	1140 (16)	1268 (18)	1662 (24)	1923 (28)	6878 (20)
Socio-economic status NS-SEC (%)	Professional/ managerial	4092 (61)	4189 (62)	4297 (63)	4371 (64)	4414 (66)	21363 (63)
	Intermediate	1984 (29)	1933 (28)	1903 (28)	1824 (27)	1693 (25)	9337 (28)
	Routine and manual	679 (10)	678 (10)	605 (9)	594 (9)	570 (9)	3126 (9)
Highest educational achieve-ment (%)	No formal record	1244 (20)	1044 (17)	997 (16)	901 (15)	1001 (16)	5187 (17)
	O-level	2075 (34)	1945 (32)	1968 (32)	1857 (30)	1809 (30)	9654 (31)
	A-level	1371 (22)	1511 (24)	1567 (25)	1630 (26)	1528 (25)	7606 (25)
	Degree	1467 (24)	1672 (27)	1712 (27)	1815 (29)	1746 (29)	8412 (27)
Menopause status (%)	Post	2436 (37)	2418 (36)	2524 (38)	2494 (38)	2538 (39)	12410 (37)
	Pre	2654 (40)	2771 (41)	2669 (40)	2691 (40)	2562 (39)	13347 (40)
	NA †	1553 (23)	1509 (23)	1500 (22)	1463 (22)	1428 (22)	7453 (22)
Ethanol, g/day		5.4 (12.9)	5.5 (12.2)	5.6 (11.4)	5.3 (11.0)	4.7 (10.5)	5.3 (11.5)
Physical activity, MET-hrs/wk		12.2(12.3)	13.7(12.3)	14.4(12.6)	15.2(12.8)	17.0(14.7)	14.5 (13.2)
Energy intake, kcal/day		1636(578)	1949(581)	2177(631)	2412(679)	2882(878)	2189 (864)
Carbohydrate intake, g/day		211 (76)	260 (73)	295 (79)	335 (83)	416 (116)	299 (127)
Protein intake, g/day		68 (26)	79 (26)	87 (29)	94 (30)	110 (36)	86 (34)
Fat intake, g/day		64 (31)	74 (33)	80 (35)	87 (39)	99 (46)	80 (39)
Saturated fat intake, g/day		23 (14)	26 (15)	27 (15)	28 (16)	31 (18)	27 (16)

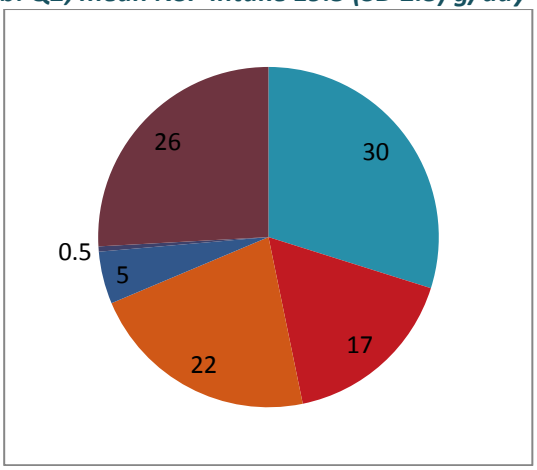
Values are median (interquartile range) or frequency (percent)

† Pregnant, taking the contraceptive pill or hormone replacement therapy

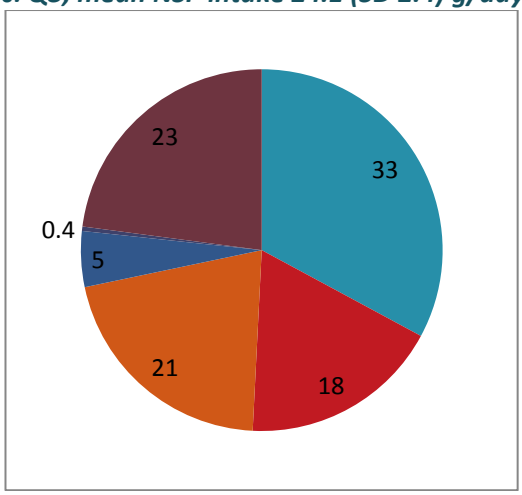
a: Q1, mean NSP intake 14.1 (SD 3.9) g/day



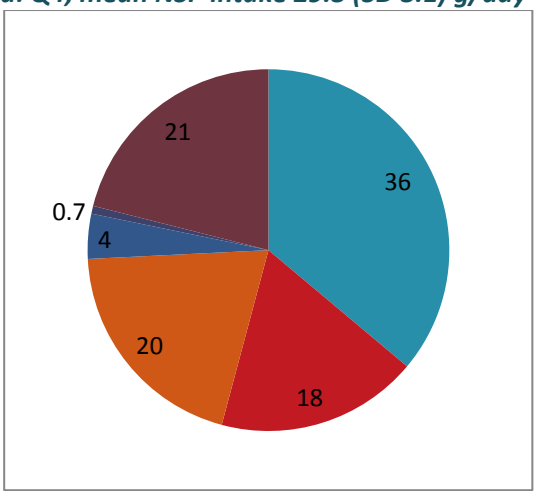
b: Q2, mean NSP intake 19.5 (SD 2.3) g/day



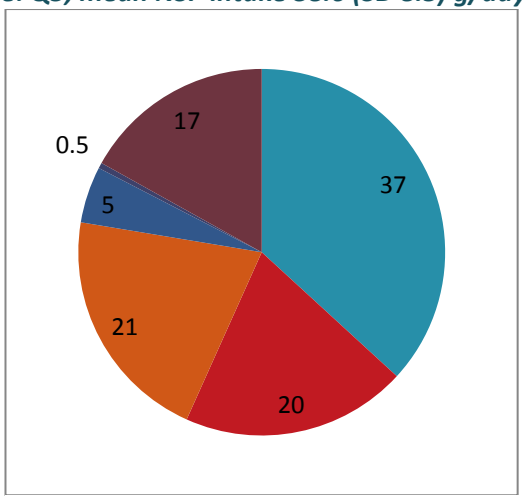
c: Q3, mean NSP intake 24.1 (SD 2.4) g/day



d: Q4, mean NSP intake 29.3 (SD 3.1) g/day



e: Q5, mean NSP intake 38.6 (SD 8.5) g/day



f: All participants, mean NSP intake 23.9 (SD 12.5) g/d

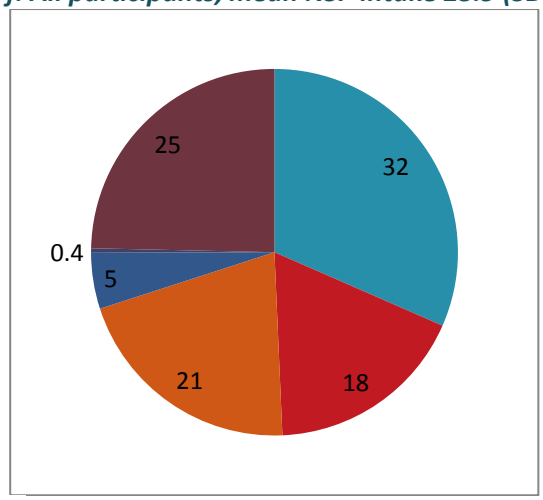


Figure 3.5 Percent contributions of food groups calculated from FFQ to total NSP intake in all participants and NSP intake quintiles (as presented in Table 3.2)

Table 3.3 Characteristics of all participants, groups based on NSP density of the diet (FFQ)

		Q1	Q2	Q3	Q4	Q5	All
N		6923	6924	6913	6918	6891	34569
NSP, g/day		16.4 (7.5)	21.0 (8.4)	24.4 (9.3)	27.7(11.0)	32.8(13.5)	23.9(12.5)
NSP density, g/1000kcal/day		7.4 (1.5)	9.4 (0.8)	11.0 (0.8)	12.7 (1.0)	15.4 (2.3)	11.0 (4.2)
AOAC fibre, g/day		26.0(12.3)	32.6(13.3)	37.4(14.9)	42.0(16.7)	49.2(20.9)	36.8(18.9)
AOAC density, g/1000kcal/day		11.7 (2.2)	14.6 (1.5)	16.9 (1.6)	19.4 (1.8)	23.4 (3.7)	16.9 (6.2)
Fibre types or sources, g/day	Soluble	7.8 (3.6)	9.5 (3.8)	10.5 (4.2)	11.6 (4.7)	13.4 (5.9)	10.4 (5.1)
	Insoluble	9.8 (5.1)	13.1 (5.8)	15.7 (6.6)	18.1 (7.5)	21.7 (9.4)	15.4 (8.9)
	Cereal	4.4 (3.5)	6.4 (5.0)	8.1 (6.2)	9.7 (6.9)	11.5 (7.9)	7.6 (7.1)
	Breakfast cereal	0.6 (1.6)	1.6 (3.0)	2.1 (3.5)	3.0 (4.8)	3.6 (6.7)	1.8 (3.7)
	Fruit	2.4 (2.4)	3.6 (2.8)	4.3 (3.3)	5.2 (4.0)	6.6 (5.7)	4.2 (3.9)
	Vegetables	3.4 (2.3)	4.4 (2.9)	5.1 (3.3)	5.7 (3.7)	7.1 (4.9)	4.9 (3.7)
	Legumes	0.8 (0.9)	1.0 (1.0)	1.1 (1.2)	1.2 (1.5)	1.5 (2.5)	1.1 (1.3)
	Nuts/seeds	0.07(0.18)	0.08(0.24)	0.11(0.29)	0.13(0.35)	0.12(0.36)	0.08(0.29)
Age, years		50.6(15.2)	50.7(15.1)	50.7(14.4)	51.3(14.6)	51.7(14.3)	51.0(14.8)
BMI, kg/m^2		24.1 (5.2)	23.8 (4.7)	23.6 (4.7)	23.5 (4.5)	23.2 (4.3)	23.7 (4.8)
Hypertension at baseline (%)	Yes	1191 (17)	1139 (16)	1097 (16)	1154 (17)	1071 (16)	5652 (16)
	No	5732 (83)	5785 (84)	5816 (84)	5764 (83)	5820 (84)	28917(84)
Angina at baseline (%)	Yes	118 (2)	141 (2)	142(2)	126 (2)	168 (2)	695 (20)
	No	6805 (98)	6783 (98)	6771 (98)	6792 (98)	6723 (98)	33874(98)
Smoking status (%)	Never smoker	3761 (56)	4048 (60)	3992 (59)	4024 (60)	3947 (59)	19772(59)
	Current smoker	1185 (18)	740 (11)	672 (10)	544 (8)	494 (7)	3635 (11)
	Former smoker	1805 (27)	1969 (29)	2095 (31)	2175 (32)	2247 (34)	10291(31)
Diet group (%)	Meat-eaters	6075 (88)	5181 (75)	4527 (65)	3794 (55)	2749 (40)	22326(65)
	Poultry-eaters	90 (1)	164 (2)	183 (3)	224 (3)	322 (5)	983 (3)
	Fish-eaters	289 (4)	642 (9)	886 (13)	1160 (17)	1405 (20)	4382 (13)
	Vegetarian	469 (7)	937 (14)	1317 (19)	1740 (25)	2415 (35)	6878 (20)
Socio-economic status NS-SEC (%)	Professional/ managerial	3849 (57)	4195 (62)	4415 (65)	4492 (66)	4412 (65)	21363(63)
	Intermediate	2109 (31)	1953 (29)	1752 (26)	1744 (26)	1779 (26)	9337 (28)
	Routine and manual	809 (12)	642 (9)	598 (9)	526 (8)	551 (8)	3126 (9)
Highest educational achieve-ment (%)	No formal record	1238 (20)	1024 (17)	942 (15)	912 (15)	1071 (17)	5187 (17)
	O-level	2143 (35)	1913 (31)	1850 (30)	1889 (30)	1859 (30)	9654 (31)
	A-level	1387 (23)	1540 (25)	1557 (25)	1583 (26)	1539 (25)	7606 (25)
	Degree	1371 (22)	1722 (28)	1841 (30)	1817 (29)	1661 (27)	8412 (27)
Menopause status (%)	Post	2462 (37)	2427 (36)	2383 (36)	2491 (37)	2647 (40)	12410(37)
	Pre	2681 (40)	2742 (41)	2804 (42)	2636 (40)	2484 (38)	13347(40)
	NA †	1497 (23)	1488 (22)	1485 (22)	1534 (23)	1449 (22)	7453 (22)
Ethanol, g/day		5.0 (11.6)	5.9 (12.0)	6.0 (11.8)	5.4 (11.3)	4.0 (10.9)	5.3 (11.5)
Physical activity, MET-hrs/wk		13.8(13.7)	14.4(13.1)	14.5(13.4)	14.5(13.1)	15.0(13.0)	14.5(13.2)
Energy intake, kcal/day		2275(951)	2227(867)	2209(848)	2162(847)	2077(817)	2189(864)
Carbohydrate intake, g/day		288 (131)	293 (121)	300 (122)	303 (126)	311 (131)	299 (127)
Protein intake, g/day		89 (35)	88 (34)	87 (35)	85 (33)	82 (34)	86 (34)
Fat intake, g/day		93 (45)	86 (39)	82 (37)	76 (35)	65 (32)	80 (39)
Saturated fat intake, g/day		35 (19)	31 (16)	28 (14)	24 (12)	19 (11)	27 (16)

Values are median (interquartile range) or frequency (percent)

† Pregnant, taking the contraceptive pill or hormone replacement therapy

Exploring linear dose-response associations with CVD for each fibre exposure

Incremental units were derived for each of the various fibre intake exposures to be used for modelling linear dose-response associations with CVD risk (full statistical methods are detailed in Chapters 5, 6 and 7). The exposure increments to be used were generated to reflect intakes reported within the cohort sample (Table 3.4). For each exposure, the sample was divided into five approximately equal groups based on intake level and median intakes were calculated within each group. The mean difference between these median values was then used in linear dose-response models to provide a realistic increment for that exposure. The mean difference between the fifths was rounded to the nearest gram, where practical, for ease of interpretation.

Taking total NSP intake as an example, the mean difference between each median value, when divided into 5 categories was 6.05g/day. Linear dose-response models using NSP estimates from the FFQ will therefore use an increment of 6g/day rather than simply using 1g/day. This will reflect a realistic increase in NSP intake within the sample and mean results may be more easily interpreted.

Table 3.4 Median intakes and mean difference between categories for different dietary exposures estimated from the FFQ

Exposure		Median intake in each fifth (Interquartile range) Q1	Q2	Q3	Q4	Q5	Mean difference between Qs	Continuous increment for model
NSP, g/day		14.1 (3.9)	19.4 (2.3)	23.8 (2.3)	29.1 (3.1)	38.3 (8.6)	6.05	6
AOAC, g/day		21.0 (5.9)	30.0 (3.4)	36.8 (3.5)	44.8 (4.8)	63.0 (13.5)	10.5	11
NSP density, g/1000kcal/day		7.4 (1.5)	9.4 (0.8)	11.0 (0.8)	12.7 (1.0)	15.4 (2.3)	2.0	2
AOAC density, g/1000kcal/day		11.3 (2.1)	14.6 (1.2)	16.9 (1.1)	19.4 (1.4)	24.3 (3.6)	3.25	3
Soluble fibre, g/day		6.4 (1.6)	8.6 (0.9)	10.4 (0.9)	12.5 (1.2)	16.4 (3.8)	2.5	3
Insoluble fibre, g/day		8.4 (2.6)	12.4 (1.6)	15.3 (1.6)	19.1 (2.2)	25.6 (6.0)	4.3	4
Fibre from food sources, g/day	Cereals	2.8 (1.4)	5.1 (1.1)	7.6 (1.4)	10.7 (1.8)	15.7 (4.5)	3.23	3
	Breakfast cereals	0.05 (0.14)	0.5 (0.4)	1.8 (0.7)	3.5 (0.7)	7.6 (2.6)	1.89	2
	Fruit	1.4 (0.9)	2.9 (0.7)	4.2 (0.7)	5.8 (1.1)	9.5 (4.1)	2.03	2
	Vegetables	2.3 (0.9)	3.7 (0.6)	4.9 (0.7)	6.6 (1.0)	9.5 (3.0)	1.8	2
	Legumes	0.2 (0.2)	0.65 (0.20)	1.11 (0.18)	1.66 (0.39)	3.6 (1.4)	0.85	1
	Nuts/seeds	0 (0.01)	0.06 (0.01)	0.08 (0.05)	0.27 (0.13)	0.85 (0.91)	0.21	0.2

3.2.3 Dietary assessment: Food Diary

Four-day weighed food diaries were collected from 12,625 respondents to the follow-up phase (1999-2004). Participants were asked to list all food and drinks and estimate the weight of items or weigh foods, where possible. Diary nutrient intakes were calculated using an in-house package which was developed at the University of Leeds, Nutritional Epidemiology Group. This package 'DANTE' contains standard nutrient values from McCance & Widdowson's *The Composition of Foods* (5th edition) (Holland et al., 1991) and also supplemental data from manufacturers and recipe information. The software allows coders to search for foods and provides information on standard servings or portions sizes, so this can be selected when it is absent from the diary. The package also allows coders to examine nutrient information for all foods so that closely matching items may be substituted if any foods recorded in the diary do not match existing items in the software. All recipe information provided by participants was carefully used to calculate the exact serving proportion and coders took care to select the appropriate cooking methods for foods, where relevant.

Because coding diaries is very labour intensive, only a fraction of the 12,625 available diaries could be coded within the scope of this project. Diary selection and statistical methods relating to analysis of food diary data are presented in depth within Chapter 7, which focuses entirely on diet assessed using food diaries and CVD risk in the UKWCS.

Characteristics of responders to follow-up data collection

Descriptive characteristics, collected at baseline, for responders and non-responders to the second phase of contact are presented in Table 3.5. Responders are classified as those providing both valid lifestyle information and a completed four day food diary (n= 12,625).

Characteristics of responders to the phase 2 contact appear largely similar in comparison to non-responders. Responders were marginally older (0.4 years older), had slightly lower BMI (0.5 units), higher energy intake (46 kcal/day) and higher physical activity (0.6 MET-hours/week). A greater proportion of the non-responders were current smokers, meat-eaters and had lower educational achievement however these differences in proportions between responders and non-responders were not large.

Table 3.5 Baseline characteristics of participants that provided complete dietary and lifestyle information at phase 2, follow-up.

		Baseline participant not responding at phase 2	Baseline participant responding at phase 2	All baseline participants†
N		23121	12625	35746
Age, years		50.9 (15.2)	51.3 (14.0)	51.0 (14.8)
BMI, kg/m^2		23.8 (4.9)	23.3 (4.4)	23.7 (4.8)
Energy intake, kcal/day		2174 (879)	2220 (848)	2191 (870)
Physical activity, MET-hours/week		14.2 (13.5)	14.8 (12.6)	14.4 (13.2)
NSP intake g/day		23.4 (12.4)	25.1 (12.7)	24.0 (12.5)
NSP density g/1000kcal/day		10.8 (4.3)	11.3 (4.2)	11.0 (4.2)
Smoking status (%)	Never smoked	12232 (57)	7558 (63)	19790 (59)
	Current smoker	2676 (12)	964 (8)	3640 (11)
	Ex-smoker	6725 (31)	3571 (30)	10296 (31)
Diet group (%)	Meat-eaters	15255 (67)	7564 (60)	22819 (65)
	Poultry-eaters	645 (3)	364 (3)	1009 (3)
	Fish-eaters	2774 (12)	1718 (14)	4492 (13)
	Vegetarian	4159 (18)	2894 (23)	7053 (20)
Socio-economic status NS-SEC (%)	Professional/ managerial	13710 (62)	8142 (66)	21852 (63)
	Intermediate	6314 (28)	3201 (26)	9515 (28)
	Routine and manual	2210 (10)	993 (8)	3203 (9)
Highest educational achievement (%)	No formal record	3735 (19)	1455 (13)	5190 (17)
	O-level	6298 (32)	3363 (30)	9661 (31)
	A-level	4690 (24)	2921 (26)	7611 (25)
	Degree	5012 (25)	3409 (31)	8421 (27)
Menopause status (%)	Post	7931 (37)	4496 (38)	12427 (38)
	Pre	8551 (40)	4803 (40)	13354 (40)
	Not applicable‡	4780 (22)	2677 (22)	7457 (22)
History of hypertension (%)	Yes	3839 (18)	1916 (16)	5755 (17)
	No	17754 (82)	10003 (84)	27577 (83)
History of angina (%)	Yes	521 (2)	197 (2)	718 (2)
	No	20383 (98)	11489 (98)	31872 (98)

Values are median (IQR) or frequency (percent)
† No exclusions were used and this data represents the full sample of baseline participants
‡Pregnant, taking the contraceptive pill or hormone replacement therapy

Diary data quality and cleaning

Coders were instructed to enter food items into DANTE in the form eaten (e.g. cooked rice) but often participants report weight of foods before cooking. Despite the DANTE data entry protocol stating that foods must be entered 'as eaten', it is relatively common for coders to mistakenly enter this information incorrectly. This is especially an issue for cereals where the food mass can multiply four or five-fold during cooking but also concerns other foods such as meat where mass is lost during cooking. To help reduce the errors in coding, I designed an

Excel spreadsheet with conversion factors programmed in, so coders could easily enter the raw weight for foods and this would convert to cooked weights. Cooking conversion multiplication factors were sourced from The 5th Edition of *The Composition of Foods* (Holland et al., 1991).

I was responsible for coordinating and training coders on how to use the DANTE package and dealt with issues relating to any food items which were difficult to code. I also implemented a program of cross-checking between coders to ensure a high standard of data entry and make sure corrections were completed for any errors in food coding.

Food diary calculations for fibre from food sources

It was possible to generate values for NSP from various food sources using food diary data. The following food sub categories, as defined in The 5th Edition of *The Composition of Foods* (Holland et al., 1991), were combined to produce fibre values for key food sources:

1) Cereal fibre
    - Biscuits
    - Bread
    - Breakfast cereals
    - Buns and pastries
    - Cakes
    - Flours, grains and starches
    - Pasta
    - Pastry
    - Puddings
    - Rice
    - Rolls
    - Savouries
2) Breakfast cereal fibre
    - Breakfast cereals
3) Fruit fibre (excluding fruit juice)
    - Fruit, general
4) Vegetable fibre (excluding potatoes and potato products)
    - Vegetables, dried
    - Vegetables, general
    - Vegetable dishes
    - Peas
5) Legume fibre
    - Beans and lentils
    - Peas
6) Fibre from nuts and seeds
    - Nuts and seeds

Correlation of food diary fibre intakes

As expected, total NSP intake and NSP density correlate highly 0.74 (95% CI: 0.72 to 0.76) (Table 3.6, Figure 3.6). However, correlation between the various sources of fibre assessed using food diaries was relatively low and ranged from 0.28 for legume and vegetable fibre to almost 0 for other comparisons. The highest correlation for legume and vegetables sources of fibre likely exists as peas are counted within both categories. The correlation between fruit and vegetable fibre as assessed using food diaries was relatively low at 0.19 (95% CI: 0.72 to 0.24) (Figure 3.7).

Table 3.6 Correlation between total and sources of fibre assessed using food diaries

	NSP	NSP Density	Cereal fibre	Breakfast cereal fibre	Fruit fibre	Vegetable fibre	Legume fibre	Nut/Seed fibre
NSP	1							
NSP density	0.74	1						
Cereal fibre	0.67	0.46	1					
Breakfast cereal fibre	0.45	0.38	0.72	1				
Fruit fibre	0.60	0.50	0.14	0.11	1			
Vegetable fibre	0.53	0.48	0.04	-0.01	0.20	1		
Legume fibre	0.31	0.27	-0.02	-0.03	0.08	0.28	1	
Nut/ seed fibre	0.30	0.15	0.04	-0.04	0.21	0.14	0.04	1

Based on 1,878 observations, after excluding diaries with fewer than 3 full days.

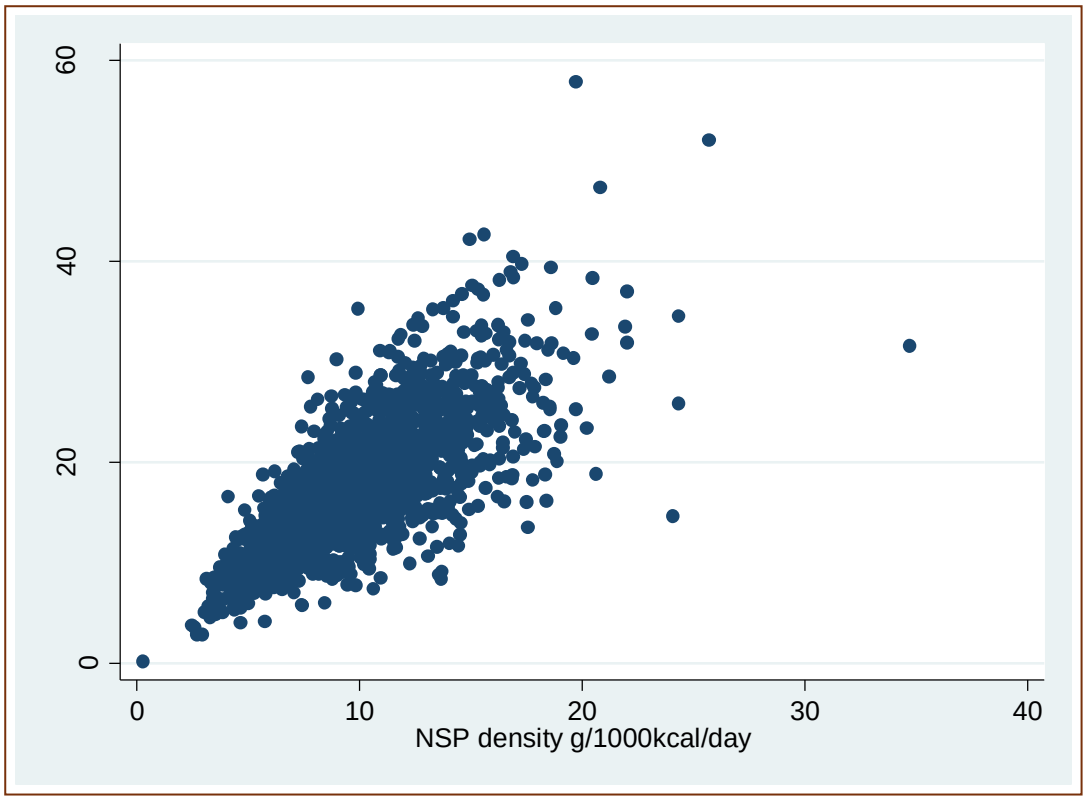


Figure 3.6 Correlation between NSP intake and NSP density estimated from food diaries

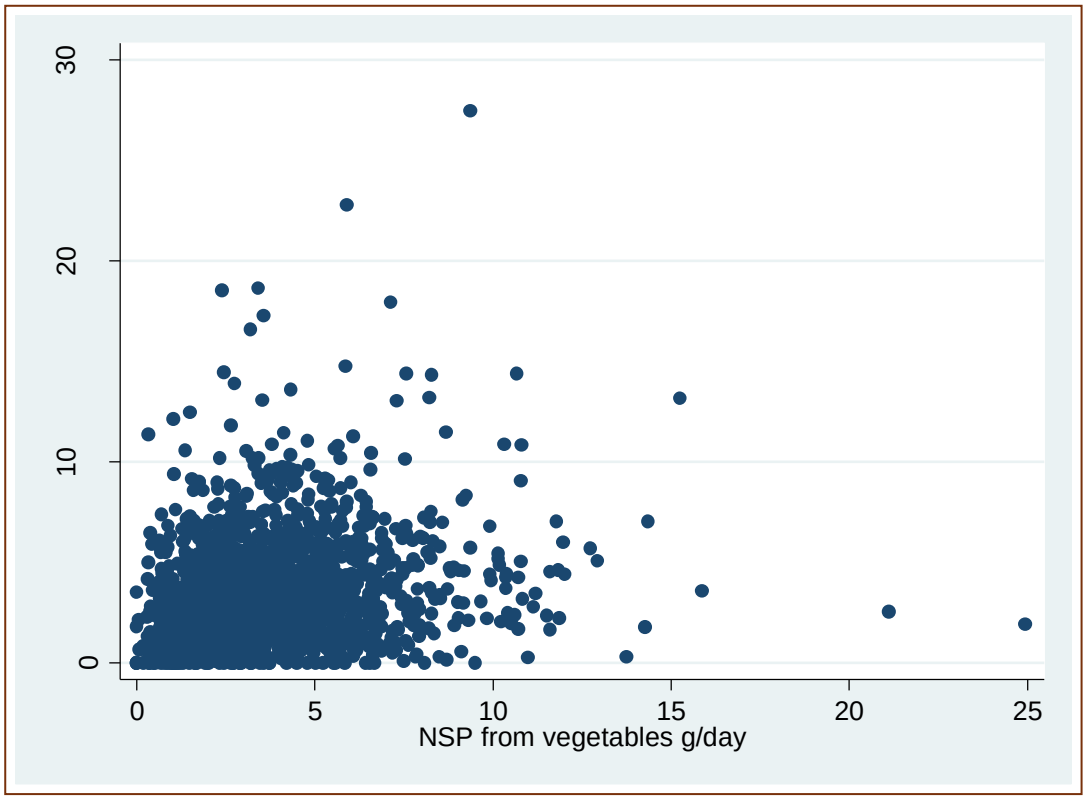


Figure 3.7 Correlation between fruit and vegetable fibre estimated from food diaries

Characteristics of low and high fibre consumers as estimated from food diaries

Of the 12,625 available food diaries, nutrients for 1,844 have so far been processed. Some descriptive statistics using the 1,844 diaries are presented below in addition to descriptive statistics for the 451 diaries which were selected and used in Chapter 7 (refer to Chapter 7 method selection for diary selection procedure and statistical methods used to analyse data).

Characteristics were explored in women who met the UK dietary recommendation to consume a minimum 18 g/day of NSP and those who did not (Table 3.7). The median NSP intake in the total sample of case and non-case diaries (n=451) was just under the recommended level, at 16.9g/day (IQR 8.3). Of these 451 women, 43% reached the 18 g/day recommendation, having average fibre intakes of 21.6g/day (IQR 5.6) compared to the 57% that did not reach the recommended intake, where the median intake was 13.3g/day (IQR 4.8).

Median BMI was marginally greater in the lower consumers 23.8 Kg/m^2 (IQR 4.4) compared to 23.4 Kg/m^2 (IQR 3.9) in high NSP consumers. Markedly lower energy intake was reported in lower fibre consumers 1679 kcal/day (IQR 557) compared to higher consumers 1904 kcal/day (IQR 456), though this may simply reflect under-reporting. Differing activity levels may also explain the BMI and energy intake differences as higher NSP consumers reported greater levels of vigorous activity and lower levels of moderate activity.

Age was also only marginally different in the two groups with the lower fibre intake group being slightly older (53.8 years) than the high fibre group (53.2 years). As expected, a greater proportion of the high fibre group were vegetarian.

Education and socio-economic profiles and saturated fatty acid intake were not widely different between the two groups and the relative proportion of women in each group with history of hypertension was similar.

The relative proportion of NSP intake from various food groups to the total intake is presented in Figure 3.8. Cereal foods make the largest contribution to NSP intake (35%), followed by other foods, which includes potatoes and mixed dishes (22%), vegetables (20%) and fruit (17%). These proportion contributions are remarkably similar to diet assessed using FFQs, where the relative contribution of each food group was: cereals 32%, fruit 18%, vegetables 21%, legumes 5%, nuts/seeds 0.4% and other sources 25% (Figure 3.5).

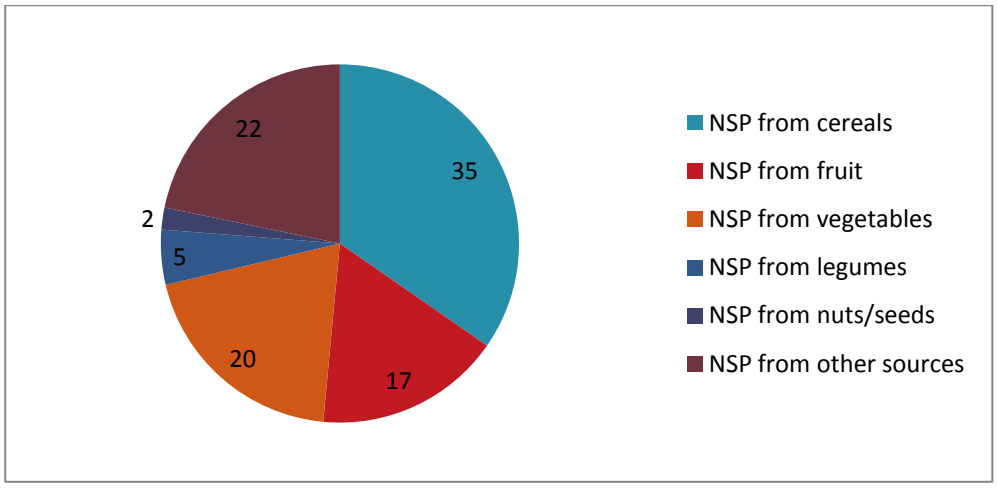


Figure 3.8 Percent contributions of food groups calculated from food diary to total NSP intake (as presented in Table 3.7)

When the sample was divided into categories based on NSP intake (Table 3.8) or NSP density (Table 3.9) patterns of characteristics across increasing intake groups were apparent. Energy intake increased with each NSP intake category from a median of 1556kcal/day (IQR 490) in Q1 to 2050kcal/day (IQR 538) in Q5. With the greater energy intake, ever greater carbohydrate, protein and fat intakes were also reported, although the level of saturated fat intake across the groups remained relatively constant around the median sample value of 22.3g/day.

When energy intake was accounted for (Table 3.9), these differences in macronutrient intakes were less apparent with protein and carbohydrate intakes being similar, although intake of total fat and saturated fat were lower in the greater NSP density groups. Energy intakes were similar across the NSP density categories except the highest group where the median energy intake was 1531 (IQR 656) compared to the sample median 1778 (IQR 563).

In contrast to the categories which were made for NSP assessed from FFQs, there was little difference in median age or BMI in the NSP or NSP density categories calculated from food diaries.

Fewer women reported high activity levels, high educational achievement and a higher proportion of women were classed in the lower NS-SEC category in women with lower NSP intakes as compared to women with greater NSP intakes.

Table 3.7 Characteristics of all case and sub-cohort participants who meet the UK dietary NSP recommended intake level, as assessed using food diaries

		NSP intake under 18g/day	NSP intake ≥ 18g/day	All diaries
N		258	193	451
Food and nutrient intake from phase 2 diary, g/day	NSP	13.3 (4.8)	21.6 (5.6)	16.9 (8.3)
	NSP density g/1000kcal/day	7.9 (3.1)	11.7 (3.7)	9.5 (4.5)
	Ethanol	5.7 (15.8)	4.2 (13.5)	4.8 (14.9)
	Protein	65.3 (21.2)	72.8 (23.5)	69.1 (21.6)
	Total fat	61.1 (33.3)	68.6 (30.8)	63.9 (33.1)
	Carbohydrate	201.4 (70.1)	249.6 (65.3)	223.1 (79.4)
	Saturated fat	22.5 (14.6)	22.0 (12.2)	22.3 (13.6)
	Cereal NSP	4.7 (3.6)	8.7 (5.7)	5.9 (5.1)
	Breakfast cereal NSP	0.5 (1.9)	2.3 (5.1)	1 (3.3)
	Fruit NSP	2.2 (2.2)	4.2 (2.5)	2.9 (2.8)
	Vegetable NSP	2.7 (2.3)	4.2 (2.9)	3.3 (2.7)
	Legume NSP	1.6 (0.4)	1.3 (1.9)	0.8 (1.9)
	Nut/Seed NSP [mean (SD)]	0.5 (1.4)	0.4 (1.0)	0.3 (0.7)
Age at phase 2, years		53.8 (15.9)	53.2 (15.8)	53.5 (16.0)
BMI at phase 2, kg/m^2		23.8 (4.4)	23.4 (3.9)	23.5 (4.4)
Energy intake at phase 2 from diary, kcal/day		1679 (557)	1904 (456)	1778 (563)
Smoking status at phase 2 (%)	Not a current smoker	229 (89)	182 (94)	411 (91)
	Current smoker	29 (11)	11 (6)	40 (9)
Diet group at baseline (%)	Meat-eaters	171 (66)	95 (49)	266 (59)
	Poultry-eaters	8 (3)	6 (3)	14 (3)
	Fish-eaters	35 (14)	28 (15)	63 (14)
	Vegetarian	44 (17)	64 (33)	108 (24)
Socio-economic status NS-SEC at baseline (%)	Professional/ managerial	160 (64)	136 (71)	296 (67)
	Intermediate	72 (29)	50 (26)	122 (28)
	Routine and manual	19 (8)	6 (3)	25 (6)
Highest educational achievement at baseline (%)	No formal record	40 (17)	22 (12)	62 (15)
	O-level	60 (26)	45 (25)	105 (25)
	A-level	65 (28)	46 (25)	111 (27)
	Degree	70 (30)	69 (38)	139 (33)
Menopause status p2 (%)	Post	194 (75)	131 (68)	325 (72)
	Pre	66 (25)	62 (32)	128 (28)
History of hypertension at phase 2 (%)	Yes	61 (25)	42 (23)	103 (24)
	No	181 (75)	139 (77)	320 (75)
Physical activity at phase 2 (%)	No weekly activity	27 (11)	9 (5)	36 (9)
	Light/moderate activity	122 (51)	86 (47)	208 (49)
	Vigorous activity 1-2 times/week	59 (25)	46 (25)	105 (25)
	Vigorous activity ≥3 times/week	31 (13)	41 (23)	72 (17)

Values are median (IQR) or frequency (percent) unless otherwise stated

Table 3.8 Characteristics of all case and sub-cohort participants, grouped based on total NSP intake, as assessed using food diaries

		Q1	Q2	Q3	Q4	Q5	All diaries
N		90	90	90	90	91	451
NSP at phase 2 from diary, g/day		9.8 (2.4)	13.5 (1.7)	16.9 (1.7)	20.1 (1.8)	25.5 (5.3)	16.9 (8.3)
NSP density at phase 2 from diary, g/1000kcal/day		5.9 (2.0)	8.4 (2.6)	9.0 (2.8)	11.0 (2.5)	12.5 (3.8)	9.5 (4.5)
Frequency of incident MI/ acute coronary syndrome cases*		12/ 18	8/ 17	12/ 20	10/ 18	11/15	53/ 88
Frequency of fatal IHD / stroke cases*		6/ 12	2/ 8	5/ 10	4/ 8	8/ 4	25/ 42
Frequency of random sub-cohort diaries		59	64	60	64	67	314
Age at phase 2, years		53.7 (17.9)	53.4 (13.7)	54.6 (17.4)	51.8 (15.2)	54.8 (17.5)	53.5 (16.0)
BMI at phase 2, kg/m^2		24.1 (4.1)	23.0 (3.9)	24.1 (5.0)	23.5 (4.3)	23.0 (3.6)	23.5 (4.4)
Current smoker at phase 2 (%)		14 (16)	12 (13)	3 (3)	5 (6)	6 (7)	40 (9)
Diet group at baseline (%)	Meat-eaters	70 (78)	58 (64)	54 (60)	49 (54)	49 (54)	280 (62)
	Fish-eaters	10 (11)	13 (14)	15 (17)	10 (11)	15 (16)	63 (14)
	Vegetarian	10 (11)	19 (21)	21 (23)	31 (34)	27 (30)	108 (24)
Socio-economic status NS-SEC at baseline (%)	Professional/ managerial	48 (55)	60 (69)	62 (70)	65 (72)	61 (68)	296 (67)
	Intermediate	34 (39)	23 (26)	17 (19)	22 (24)	26 (29)	122 (28)
	Routine and manual	5 (6)	4 (5)	10 (11)	3 (3)	3 (3)	25 (6)
Highest educational achievement at baseline (%)	No formal record	19 (22)	13 (16)	9 (12)	7 (9)	14 (16)	62 (15)
	O-level	30 (35)	15 (18)	18 (23)	24 (30)	18 (20)	105 (25)
	A-level	19 (22)	24 (29)	24 (31)	20 (25)	24 (27)	111 (27)
	Degree	18 (21)	31 (38)	27 (35)	30 (37)	33 (37)	139 (33)
Menopause status phase 2 (%)	Post	64 (71)	73 (80)	63 (70)	56 (62)	69 (76)	325 (72)
	Pre	26 (29)	18 (20)	27 (30)	35 (38)	22 (24)	128 (28)
History of hypertension at phase 2 (%)	Yes	26 (30)	19 (23)	20 (24)	20 (24)	18 (21)	103 (24)
	No	60 (70)	65 (77)	64 (76)	62 (76)	69 (79)	320 (75)
Physical activity at phase 2 (%)	No weekly activity	13 (15)	6 (7)	9 (11)	3 (3)	5 (6)	36 (9)
	Light/moderate activity	50 (60)	41 (50)	33 (40)	43 (49)	41 (49)	208 (49)
	Vigorous activity 1-2 /week	14 (17)	22 (27)	27 (33)	21 (24)	21 (25)	105 (25)
	Vigorous 3 or more/week	7 (8)	13 (16)	14 (17)	21 (24)	17 (20)	72 (17)
Ethanol at phase 2 from diary, g/day		7.4 (16.7)	7.3 (17.3)	4.4 (10.9)	5.0 (13.5)	2.8 (14.0)	4.8 (14.9)
Energy intake at phase 2 from diary, kcal/day		1556 (490)	1645 (544)	1817 (517)	1844 (381)	2050 (538)	1778 (563)
Protein intake from food diary, g/day		64.2 (23.0)	62.0 (21.3)	69.8 (21.0)	69.9 (20.3)	76.3 (22.8)	69.1 (21.6)
Total fat intake from food diary, g/day		59.1 (31.3)	58.2 (30.4)	66.7 (35.9)	62.8 (28.4)	72.9 (27.1)	63.9 (33.1)
Carbohydrate intake from food diary, g/day		179.1 (62.0)	197.4 (63.8)	231.1 (58.5)	235.2 (56.1)	268.6 (94.3)	223.1 (79.4)
Saturated fat intake at phase 2 from diary, g/day		22.4 (16.1)	22.4 (13.7)	22.8 (13.6)	21.4 (12.8)	22.1 (13.6)	22.3 (13.6)

*Case definitions are detailed in Chapter 4; Values are median (IQR) or frequency (percent), unless otherwise stated

Table 3.9 Characteristics of all case and sub-cohort participants, grouped based on NSP density of the diet, as assessed using food diaries

		Q1	Q2	Q3	Q4	Q5	All diaries
N		90	90	90	90	91	451
NSP density at phase 2 from diary, g/1000kcal/day		5.8 (1.2)	7.8 (1.0)	9.5 (0.9)	11.2 (1.0)	14.4 (3.1)	9.5 (4.5)
NSP at phase 2 from diary, g/day		10.3 (3.9)	15.1 (5.1)	16.3 (4.7)	20.7 (4.3)	22.3 (7.5)	16.9 (8.3)
Frequency of incident MI/ acute coronary syndrome cases*		10/ 16	11/ 20	8/ 14	9/ 17	15/ 21	53/ 88
Frequency of fatal IHD/ stroke cases*		5/ 10	4/ 11	2/ 9	5/ 5	9/ 7	25/ 42
Frequency of random sub-cohort diaries		64	57	68	67	58	314
Age at phase 2, years		53.7 (13.7)	53.6 (14.7)	52.7 (16.8)	52.2 (15.4)	53.5 (16.0)	53.5 (16.0)
BMI at phase 2, kg/m^2		24.1 (4.0)	24.1 (4.5)	23.6 (4.3)	23.1 (4.4)	23.2 (4.4)	23.5 (4.4)
Current smoker at phase 2 (%)		14 (16)	10 (11)	5 (6)	5 (6)	6 (7)	40 (9)
Diet group at baseline (%)	Meat-eaters	74 (82)	62 (69)	55 (61)	45 (50)	44 (49)	280 (62)
	Fish-eaters	8 (9)	11 (120	16 (18)	17 (19)	11 (12)	63 (14)
	Vegetarian	8 (9)	17 (19)	19 (21)	28 (31)	36 (40)	108 (24)
Socio-economic status NS-SEC at baseline (%)	Professional/ managerial	46 (53)	67 (74)	61 (70)	62 (70)	60 (67)	296 (67)
	Intermediate	34 (39)	19 (21)	21 (24)	21 (24)	25 (28)	122 (28)
	Routine and manual	7 (8)	4 (4)	3 (3)	6 (7)	5 (6)	25 (6)
Highest educational achievement at baseline (%)	No formal record	19 (23)	7 (9)	9 (11)	11 (13)	16 (19)	62 (15)
	O-level	26 (31)	21 (26)	18 (21)	15 (18)	25 (29)	105 (25)
	A-level	19 (23)	20 (25)	28 (33)	20 (24)	24 (28)	111 (27)
	Degree	19 (23)	32 (40)	29 (35)	38 (45)	21 (24)	139 (33)
Menopause status phase 2 (%)	Post	65 (72)	70 (77)	62 (69)	57 (63)	71 (78)	325 (72)
	Pre	25 (28)	21 (23)	28 (31)	34 (37)	20 (22)	128 (28)
History of hypertension at phase 2 (%)	Yes	23 (27)	23 (27)	15 (18)	22 (26)	20 (24)	103 (24)
	No	62 (73)	61 (73)	69 (82)	64 (74)	64 (76)	320 (75)
Physical activity at phase 2 (%)	No weekly activity	10 (12)	9 (11)	8 (9)	4 (5)	5 (6)	36 (9)
	Light/moderate activity	49 (60)	39 (47)	39 (46)	39 (46)	42 (49)	208 (49)
	Vigorous activity 1-2 /wk	16 (20)	22 (27)	30 (35)	21 (25)	16 (19)	105 (25)
	Vigorous 3 or more/week	7 (9)	13 (16)	8 (9)	21 (25)	23 (27)	72 (17)
Ethanol at phase 2 from diary, g/day		8.2 (20.3)	7.4 (16.0)	6.2 (15.7)	5.5 (15.3)	0 (4.6)	4.8 (14.9)
Energy intake at phase 2 from diary, kcal/day		1841 (501)	1927 (627)	1748 (488)	1856 (396)	1531 (656)	1778 (563)
Protein intake from food diary, g/day		71.7 (18.0)	69.7 (26.5)	67.8 (19.2)	70.0 (21.8)	63.5 (20.1)	69.1 (21.6)
Total fat intake from food diary, g/day		73.9 (32.2)	73.9 (33.0)	62.9 (26.2)	61.5 (28.9)	47.1 (30.1)	63.9 (33.1)
Carbohydrate intake from food diary, g/day		211.3 (80.1)	236.1 (82.6)	222.2 (65.0)	236.0 (55.1)	209.3 (76.3)	223.1 (79.4)
Saturated fat intake at phase 2 from diary, g/day		28.0 (14.0)	24.9 (13.9)	23.5 (9.7)	21.0 (10.0)	16.4 (9.6)	22.3 (13.6)

*Case definitions are detailed in Chapter 4; Values are median (IQR) or frequency (percent), unless otherwise stated

Exploring linear dose-response associations with CVD, for each fibre exposure

Incremental units for exploring linear, dose-response, associations between fibre exposures and CVD were generated using the same approach as for fibre estimated from the FFQ (Table 3.4). In order to reflect the trajectory of intakes reported within the cohort sample, the mean differences between quintile median intakes were calculated. These calculations were carried out using all case and non-case diaries (n=451) with the only exclusion being of diaries where fewer than 3 full days had been completed (Table 3.10).

Table 3.10 Median intakes and mean difference between categories for different dietary exposures estimated from food diaries

Exposure		Median intake in each fifth (Interquartile range) Q1	Q2	Q3	Q4	Q5	Mean difference between Qs	Continuous increment for model
NSP, g/day		9.8 (2.4)	13.5 (1.7)	16.9 (1.7)	20.1 (1.8)	25.5 (5.3)	3.9	4
NSP density, g/1000kcal/day		5.8 (1.2)	7.8 (1.0)	9.5 (0.9)	11.2 (1.0)	14.4 (3.1)	2.2	2
Fibre from food sources, g/day	Cereal	2.4 (1.3)	4.2 (0.9)	5.9 (1.20	8.0 (1.2)	12.1 (5.3)	2.9	3
	Breakfast cereals†	/	0 (1.1)	1.0 (0.5)	2.7 (1.2)	6.2 (4.7)	2.1	2
	Fruit	0.7 (0.8)	1.9 (0.4)	2.8 (0.6)	4.1 (0.6)	6.0 (2.0)	1.3	1
	Vegetables	1.2 (0.8)	2.4 (0.6)	3.3 (0.4)	4.4 (0.6)	6.6 (2.7)	1.4	1
	Legumes†	/	0 (0.2)	0.8 (0.3)	1.7 (0.5)	3.0 (1.6)	1.0	1
	Nuts/seeds †	/	/	/	0 (0)	0.9 (0.9)	0.9	1

† Categories 1, 2 and 3 include a high proportion of non-consumers so means are derived using comparison between fewer categories, for consumers of the specific source of fibre.

3.3 FFQ values compared to food diaries

To approximately estimate the degree of difference between the average energy and fibre intakes as calculated by the two dietary assessment methods, intakes were compared using all available FFQ data plus all coded diaries to date (n=1,844). Participants with extreme calorie intakes (<500kcal/day or >6000kcal/day), as estimated by the FFQ, were excluded as were food diaries that had not been completed for at least 3 days.

There were 35,260 participants with valid FFQ data and for the whole cohort to date and 1,844 food diaries have so far been processed.

Using just those women with available data from both FFQs and food diaries, the mean NSP intake, as estimated by food diaries, was on average 8.9g/day (SD 9.9) lower than assessed by FFQs (Table 3.11). The mean NSP intake from FFQs was 25.5g/day (SD 10.3) and was 17.4g/day (SD 6.3) from food diaries. Energy intakes were 530kcal/day (SD 739) higher with FFQs compared to mean energy intake estimated from food diaries. Although both fibre and energy values were far greater when estimated using the FFQ, the fibre density values were closer when methods were compared, there was a difference of 1.8 (SD 3.2) g/1000kcal/day. The mean fibre intake calculated as AOAC was 13.7g/day (SD 6.0) higher than mean NSP intake.

Table 3.11 Energy and fibre intakes estimated from FFQs and food diaries

		N	Mean (SD)	Min	Max
FFQ	Energy intake from FFQ, kcal/day	35260	2342 (713)	510	5997
	NSP intake from FFQ, g/day	35260	25.5 (10.3)	1.8	151.9
	NSP density from FFQ, g/1000kcal/day	35260	11.3 (3.2)	1.9	32.8
	AOAC intake from FFQ, g/day	35260	39.2 (15.9)	3.7	216.6
	AOAC density from FFQ, g/1000kcal/day	35260	17.3 (4.7)	3.4	56.8
	AOAC intake minus NSP intake, g/day	35260	13.7 (6.0)	-0.4	81.1
	AOAC density minus NSP density, g/1000kcal/day	35260	6.1 (1.8)	-0.2	28.5
Diary	Energy intake, kcal/day	1844	1811 (422)	607	4043
	NSP intake, g/day	1844	17.4 (6.3)	0.2	57.9
	NSP density, g/1000kcal/day	1844	9.8 (3.4)	0.3	34.7
FFQ-Diary comparison	NSP, FFQ minus Diary, g/day	1844	8.9 (9.9)	-22.4	62.9
	NSP density, FFQ minus Diary, g/1000kcal/day	1844	1.8 (3.2)	-18.9	17.5
	Energy intake, FFQ minus Diary, kcal/day	1844	530 (739)	-1640	4306

3.4 Discussion

3.4.1 How FFQs compare to diary methods

The two different dietary assessment methods employed in the UKWCS have distinct strengths and weaknesses. The use of a FFQ has allowed dietary habits to be examined in a large sample of women. The tool captures an impression of usual intakes over the previous 12 months and so gives an estimation of long-term diet. The diary, by contrast, captures very detailed dietary intake and represents more of a 'snap-shot' within the normal variation in dietary intakes. Despite food diaries giving more precise estimates of intakes, they do not capture long-term intake unless multiple diaries are completed throughout the year. Additionally, because of limited resources it has not been possible to generate nutrient values for the whole cohort and so only a sub-sample of diaries could be processed.

Energy and NSP intakes estimated here using both methods indicate that FFQs tend to over-estimate intakes, assuming that food diaries are more precise and not vice versa. Greater energy and fibre intakes were also observed from FFQs, compared to food diaries, in other British cohort studies (Bingham et al., 1997, Brunner et al., 2001). Dietary assessment validation for the UK arm of the EPIC study and the Whitehall II study identified good correlation between fibre intakes estimated from FFQs and 16-day weighed records (0.57) (Bingham et al., 1997) or between energy adjusted intakes from FFQs and 7-day diaries (0.60) (Brunner et al., 2001). Agreement between quartile classification for NSP or Southgate fibre intake assessed with FFQs and 16-day or 7-day diaries, respectively, were also assessed and reported to be around 40% in both studies (Bingham et al., 1997, Brunner et al., 2001).

Agreement between FFQs and food diaries was not explored in depth here as the two methods were used around 5 years apart. Any differences observed between the two methods may well be due to actual changes in diet over this time. Additionally, a small validation study was carried out in a sub-group of UKWCS participants a few years after baseline dietary assessment (Spence et al., 2002). FFQs and food diaries were completed by 303 participants and the correlation for key macronutrients assessed with the two methods was not high at around 0.35.

3.4.2 Fibre intakes in UKWCS compared to other study populations

The mean NSP intake, as assessed by food diaries (n=451) was 16.9 g/day (SD 8.3) and of those women meeting the UK dietary recommended intake, the mean was 21.6 g/day (SD 5.6) (Table

3.7). Even in those participants that did not meet the recommendation, the mean intake was 13.3 (SD 4.8) g/day, which is still greater than the mean intake in UK women estimated in 2001, around the same time as diaries were completed for UKWCS participants. In the NDNS survey 2001 the mean intake in women was 12.6g/day and was slightly higher in the older sample of participants, 50-64 years where the mean NSP intake was 14.0g/day (Henderson et al., 2003).

Average NSP intake in the UKWCS, assessed by FFQ, was around 24g/day, much higher than the 18g/day NSP intake found in another UK cohort, where diet was also assessed by FFQ (Ward et al., 2012). After accounting for the higher daily energy intake in the UKWCS (2342kcal) compared to EPIC Norfolk (2103kcal), the NSP density of the UKWCS, at 11.3g/1000kcal/day was still greater than EPIC Norfolk, where I estimate NSP density as 8.8g/1000kcal/day (Ward et al., 2012). A Finnish study of male smokers reported mean NSP intake closer to the UKWCS intake, at around 25g/day (Larsson et al., 2009). Mean AOAC fibre intake in the Nurses' Health Study however was drastically lower than the 36g/day AOAC fibre intake observed here at around 15g/day (Oh et al., 2005). Focusing on NSP estimated from food diaries, the EPIC Norfolk study again reported lower intakes at 15.4g/day (SD 5.5) [NSP density 7.6g/1000kcal/day] (Ward et al., 2012) compared to the UKWCS where mean intake was 17.4g/day (SD 6.3) [NSP density 9.8g/1000kcal/day].

The high fibre intakes in this cohort, by comparison to representative study populations from the UK or US, could simply reflect the healthy characteristics of participants here and the large proportion of vegetarians in the UKWCS. Average fibre intake estimated with FFQs in the Oxford arm of the EPIC study, which also recruited a large proportion of non-meat eaters, was closer to that seen for the UKWCS and in women, mean NSP was 20.4 (SD 7.7) g/day (Davey et al., 2003). In addition to the greater proportion of non meat-eaters, it is possible that higher NSP intake may result from the large number of individual fruit and vegetable items and the inclusion of additional composite vegetable dishes listed on the FFQ, causing participants to over-estimate their intake of these foods, leading to inflated fibre values. The issue of exaggerated vegetable intake using FFQ methods, compared to other methods of dietary assessment, has also been noted in other British studies (Bingham et al., 1997, Brunner et al., 2001). In the UK arm of the EPIC study, authors ascribe higher nutrient values from the FFQ, as compared with weighed diaries completed at four time points throughout the year, partly to the 120g/day greater vegetable intake calculated from FFQs (Bingham et al., 1997). In the Whitehall II study, reported intake of vegetable foods also appeared to be over-estimated

compared to food diaries when biomarkers such as beta-carotene were assessed. The authors similarly suggest this may occur because of the large number of vegetables items on their FFQ (Brunner et al., 2001).

Differences observed between study populations may also result not only from actual differences but also the method of coding diaries and sources of nutrient information. Data from the UKWCS was pooled with six other cohort studies in the UK to investigate diet and associations with cancer. All other cohorts assessed diet using 'DINER' software but not all diaries for the UKWCS were assessed with DINER and some were assessed using DANTE. In an investigation into fibre and colorectal cancer risk, Dahm and colleagues performed a comparative analysis of the DINER and DANTE methods for 100 randomly selected UKWCS diaries and found that the geometric mean difference in energy and carbohydrate intake between the methods was 2% (95% CI: 0 to 5%). The geometric mean difference for fibre assessed by the two methods was 8% (95% CI: 4 to 12%) which was estimated to be equivalent to an arithmetic mean difference of 1.3g/day (Dahm et al., 2010).

3.4.3 Limitations and strengths of dietary assessment methods

The effects of systematic bias from FFQs are debated, with some believing that bias does not generally hinder the ability to identify important associations in epidemiologic research and is lessened through categorising participants into intake fifths (Willett and Stampfer, 1998). However, others argue that the substantial measurement error can profoundly influence the interpretation of epidemiologic studies and the attenuation could be so severe as to preclude useful results (Freedman et al., 2011, Kipnis et al., 2003). Solutions include adjusting for energy intake (Freedman et al., 2011) and other confounders. However measurement error in assessing confounders is also an issue of concern and this bias can be large and work to inflate or attenuate associations (Greenland, 1980, Willett, 2013a). Findings from validation studies using biomarkers suggest that measurement error in dietary exposures will often result in attenuated estimates (Kipnis et al., 2003, Freedman et al., 2011). Thus, moderate diet-disease associations in the order of 2.0 for risk would appear close to 1.3 using FFQ-based energy adjusted values (Kipnis et al., 2003). However, the impact of measurement error is less severe after adjustment for energy intake (Kipnis et al., 2003, Freedman et al., 2011).

Energy intake can be accounted for by both dividing nutrients by energy or adjusting for total energy intake in models, along with other potential confounders. When energy intake is unrelated to disease outcomes, dividing nutrients by energy intake can be beneficial in

reducing the variation in nutrient intake that is due to differences in body size or net activity. However, if the nutrient correlates with energy intake, as with fibre, dividing by energy intake creates a variable that is highly related to the factor we wish to account for i.e. energy intake. The correlation between NSP and energy intake assessed using the FFQ was 0.68 and for the diary was 0.43. The issue of nutrient density variables being highly related to energy intake may be addressed with additional adjustment for energy intake in density analyses (Willett, 2013b) and so will be applied here in analyses using fibre density values. Models and covariate adjustments are described in detail in the method section of Chapter 5.

Other limitations of assessing diet at one time point using a food diary is the day-to-day or week-to-week variation and the changes in diet during the week or weekend. Although individual diets may be influenced by day of the week, diaries were issued with staggered instructions for the start day so that this issue would be negated somewhat in the UKWCS dataset. Additionally, fibre intake is unlike certain micronutrients which may be very heavily influenced by seasonally available foods, although intake of fibre from seasonally available fruit or vegetables may vary. Since fibre intake is not heavily concentrated in just a few, sporadically consumed foods, but is rather a feature of many foods in the diet, using a food diary to assess intake of fibre should give a reasonable impression of usual intake (Willett, 2013c). However, Bingham has estimated that 10 days of diary records are necessary to give robust estimates for dietary fibre intake at the individual, rather than group, level in order to be ±10% of the average intake (Bingham, 1987).

To reduce error in dietary estimation that is attributable to daily variation, a greater number of recording days would be needed (Willett, 2013c), ideally spaced throughout the year to reduce error from seasonal variation. The restricted number of diary days available for the UKWCS is a limitation and the introduction of electronic dietary assessment in newer studies is a huge leap forward for speeding dietary assessment and reducing errors which may be introduced in coding foods.

While diaries capture detailed intakes, a drawback is the risk of reactivity for participants (Baranowski, 2013) as the high burden may cause participants to simplify their intake to make recording of their diet easier or they may omit foods or recipes which are difficult to record. Additionally self-monitoring may increase the chance of selecting or reporting more socially acceptable foods (Baranowski, 2013).

A strength of the food diary approach is that it is totally open ended, allowing participants to record as much or as little of any foods they consumed. The FFQ, by contrast is limited in terms of the items described and the fact that standard, rather than individual portion sizes, must be ascribed to the various food items. Diaries also have the benefit of not relying so heavily on memory as FFQs since foods can be recorded as they are eaten or in the near past (Baranowski, 2013).

A further source of error with diaries may be the lack of participant or coder motivation to accurately record or code diet (Baranowski, 2013). Little may be done to address lack of motivation in participants, although the recruitment of self-selecting participants means they may be more highly motivated that the general population. However, implementing cross-checking procedures between coders will go some way towards limiting any bias caused through lack of motivation from coders, such as the checking protocol applied for coding diaries here.

3.5 Summary

This chapter outlines the design and data collection of the UKWCS. Details of the two dietary assessment methods are described, along with details for how fibre values were derived for the total diet and from key food sources of fibre. Correlation between the different sources of fibre are explored within each method and study population characteristics are presented for low and high fibre consumers, as assessed using both methods.

Calculations for fibre and energy from the two methods are briefly compared and the high fibre intakes observed in the UKWCS are compared with fibre intakes estimated from other cohort study populations. Results from validation studies of other British cohorts are additionally discussed here as well as both strengths and limitations of the FFQ and food diary approaches to estimate usual intake.

This chapter has presented a background to the UKWCS and dietary fibre intakes. The next chapter (Chapter 4) focuses on cardiovascular outcome data; the different event sources used and how data were processed. Detailed methods describing the completeness of event reporting, or capture, within these different sources are also presented in the next chapter. Subsequent chapters (5, 6 and 7) then utilise the dietary data presented in this chapter along with cardiovascular event data presented in Chapter 4 to examine CVD risk in association with dietary fibre intake.

Chapter 4 Sources of cardiovascular event data and case ascertainment rates

4.1 Chapter overview

The three different sources of cardiovascular event data obtained for participants of the UKWCS are described. Mortality data were obtained in addition to CVD cases from Hospital Episode Statistics (HES) and acute coronary events from the Myocardial Ischaemia National Audit Project (MINAP). Cases from each dataset were identified and this preparation, in terms of the definitions applied for both CHD and stroke cases within each source, is described. The ethical approval process and data security arrangements are also detailed in this chapter.

In this chapter the potential number of missing cases, i.e. CVD cases not identified in any of the datasets, and therefore the degree of complete case capture within the three sources has been estimated using capture-recapture analysis.

4.2 Background

Over recent years there have been initiatives to expand the use of electronic health records for research in the UK, such as the Strategic Framework for Health Informatics in Support of Research, whose aims include maximising the potential for research through linking patient databases in the NHS (MRC, 2010). Unlike recording of cancer incidence, no single or complete register exists for CVD outcomes in the UK and so different resources must be combined in order to estimate incidence with as much accuracy as possible.

Combining different sources of outcome data for participants of the UKWCS is challenging because the available data cover different geographical areas and sources become 'complete' over different time frames. At study inception, no clinical register for CVD-related outcomes was available. CVD cases may therefore only be obtained from death certificates until the late 1990's when hospital records become available. The quality of routinely collected hospital data has been improving (Burns et al., 2012) and as a result, these data are becoming increasingly useful for research. Additionally, in the past decade a clinical registry for acute coronary admissions in English and Welsh hospitals has been established and now reports promising levels of data completeness (Herrett et al., 2010).

Although no case list can be proven to identify all cases in a given population, through combining lists it is possible to estimate the number of missed cases. Capture-recapture methods, which have their origins in ecology, can be applied to human populations. Originally, wildlife population sizes were estimated through capturing a sample, marking, releasing and later recapturing another sample of the same species. Using the numbers of recaptures and the number in each sample, it was possible to estimate the number not caught in either sample and thus the total population size (IWGDMF, 1995a). In human disease, these methods began to catch on in the late 1980s and by the 1990's epidemiologists were applying log-linear modelling methods to capture-recapture analyses (IWGDMF, 1995b). The application of these methods in epidemiology presents some difficulties because assumptions which should be met include that the population size is closed and lists are independent, rarely the case in epidemiological study settings. In epidemiology and public health, capture-recapture methodology can be applied to estimate the extent of incomplete case ascertainment using lists from different sources with overlapping cases (Hook and Regal, 1995). Log-linear analysis essentially compares the observed frequency of cases with the frequency that is expected to occur by chance (Cramer, 2003). The approach is deemed to be the method of choice for assessing completeness of data from multiple sources (Hook and Regal, 1995) and was therefore applied to the observed case frequency data for the UKWCS (details below in Section 4.3.5).

4.2.1 Ethical considerations and approval for access to data

At inception of the UKWCS in 1993, in the absence of a more centralised system, individual ethical approval was sought and obtained from 174 local ethics committees within the UK. Approval was granted from each local authority for the study to follow participants for cases of cancer and other diseases. At this time individual consent forms were not required by the ethics committees, therefore those women who returned questionnaires with a completed back page were considered to have provided consent for participation. The back page of the questionnaire informed participants that the purpose of the study was to examine "the occurrence of certain diseases such as cancer which are registered by the National Health Service" and participants were asked to provide their NHS number and GP address in order for their medical records to be accessed.

Obtaining the appropriate approvals for cardiovascular event data linkage with the UKWCS was not straightforward and the various processes are detailed below and shown in Figure 4.1:

1) In January 2011 I applied to the Data Access Advisory Group (DAAG) for access to HES data relating to cohort participants.
2) In May 2011 I received a letter from the DAAG informing me that the application was not approved and the group indicated that the return of a questionnaire at study baseline did not constitute appropriate consent for a project of this type.
3) Section 251 approval was therefore sought from the Ethics and Confidentiality Committee (ECC) since it was not possible to obtain additional consent, more aligned with current standards, directly from participants. 'Section 251 of the NHS Act 2006 allows the common law duty of confidentiality to be set aside in specific circumstances where anonymised information is not sufficient and where patient consent is not practicable' (NIGB, 2011). The application for approval was made in June 2011.
4) The ECC granted Section 251 approval for access to HES data for cohort participants in August 2011. This approval was subject to two specific conditions:
    - Data security arrangements had to be in place and confirmed
    - Confirmation was required to show that the original ethics approvals for the cohort covered the linkage with HES data or a new favourable research ethics committee (REC) opinion was needed for access to HES data.
5) After consultation with the National Research Ethics Committee (NRES), it was decided that a REC local to Leeds should be contacted regarding a Notice of Substantial Amendment (NoSA) for the new data linkage as no singular REC had been appointed for the cohort since the centralised system for ethical approval was introduced nationally.
6) I submitted a NoSA to Leeds East REC in November 2011. The NoSA related to access to and linkage with cardiovascular event data since this had not previously been accessed for or linked to the cohort participants. Leeds East REC considered the NoSA in December 2011 and approval for the linkage was granted (approval letter in Appendix VI).
7) The ECC were informed of the favourable ethics outcome and data security arrangements and full section 251 approval was granted (approval letter in Appendix VII).
8) Confirmation of this approval was provided to the DAAG and linked HES data were provided for cohort participants.

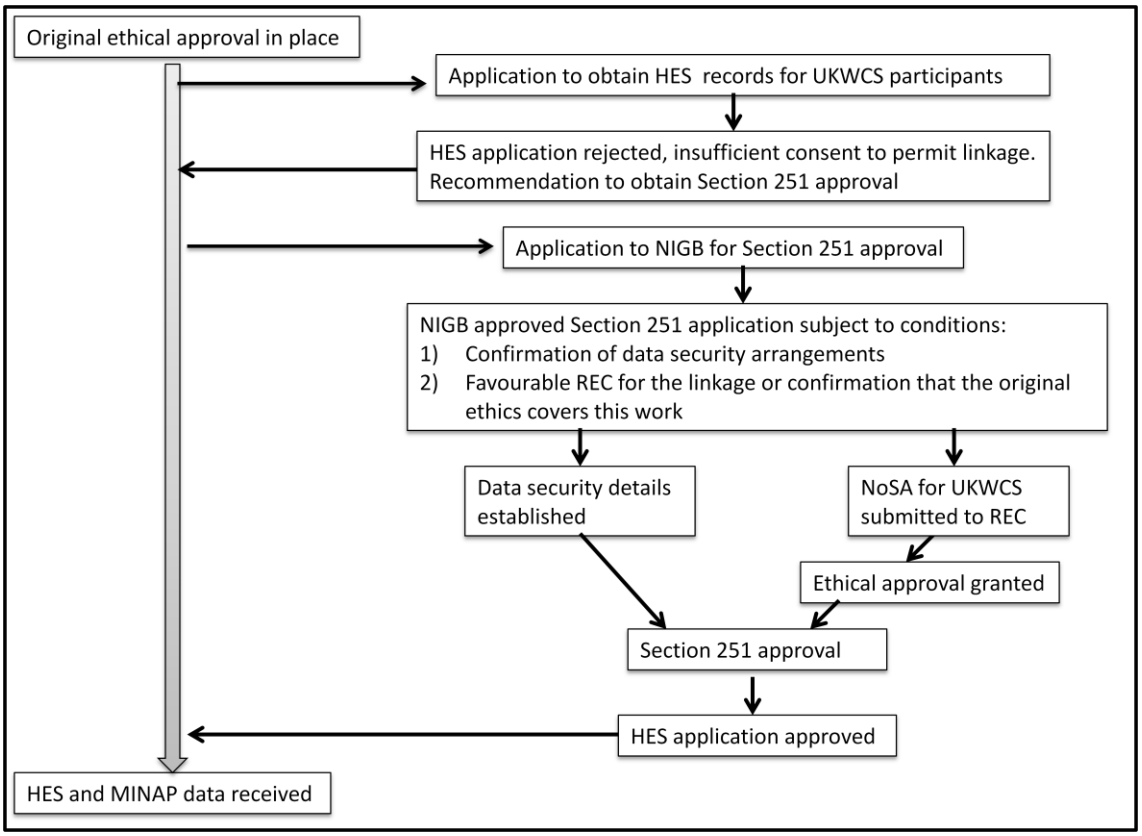


Figure 4.1 Process of obtaining ethical approval for access to and linkage with cardiovascular event data

4.2.2 Data security and anonymity

Data files including HES or MINAP outcome data have been either stored on an encrypted laptop or a secure drive, where access to the file was restricted and files were protected through secure firewalls.

Third party record linkage to UKWCS participant identifiers was carried out for both MINAP and HES to ensure anonymity for cohort participants. Identifiers (unique identification number, name, date of birth, NHS number) were provided to the trusted third parties (NHS information centre/MINAP clinical director) so that cohort participants could be identified from within these national registries. Relevant event records for participants were then returned with only a unique identification number so that records could be matched up with lifestyle information for cohort participants.

4.3 Methods

4.3.1 Overview of the three sources of outcome data

Three main sources of cardiovascular disease case data were available for the UKWCS and these span different timeframes (Figure 4.2). Mortality records for participants are available from study baseline to present (from the NHS central register/ NHS information centre), in-patient HES records are available for all English hospitals since 1998 and the MINAP clinical registry data became complete in 2003 and records spanning up to 2011 have been obtained.

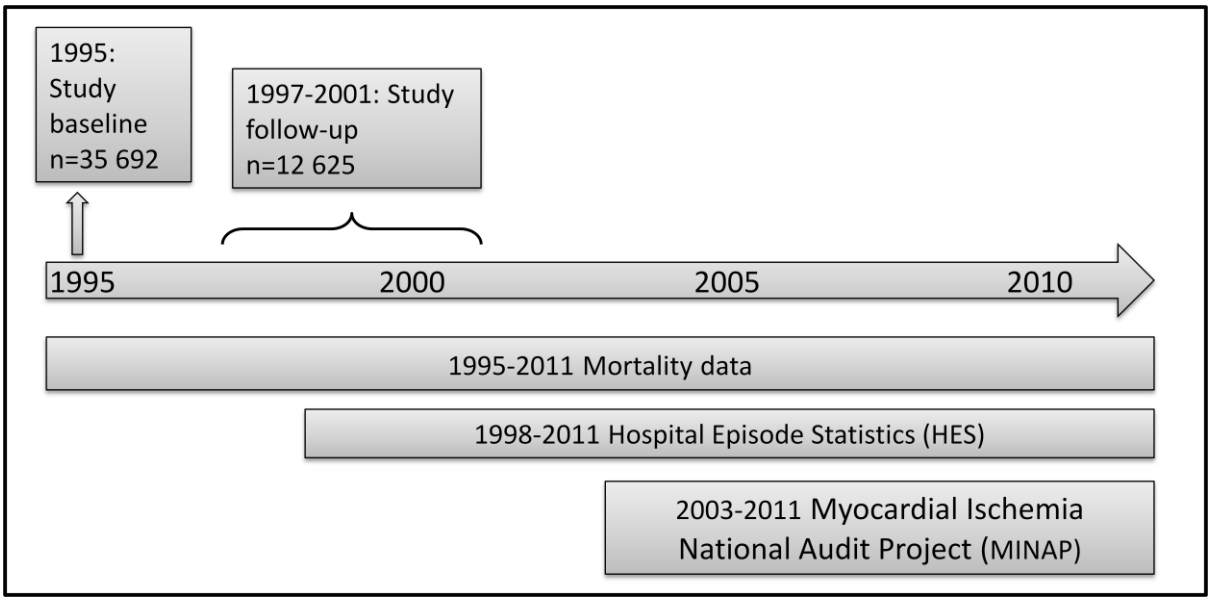


Figure 4.2 Time periods covered by different case datasets

Cohort participants were primarily from England but also from Scotland and Wales. MINAP covers only English and Welsh hospitals and HES data were only obtained from English hospitals as both Scottish and Welsh data are separately generated and stored. In analyses where only mortality data are used, all participants are retained in the sample but where incidence data, from HES or MINAP are used, the sample has been restricted to just those women whose address at baseline was within England. This ensures that datasets are comparable and as complete as possible for examining incidence in the largest regional group within the cohort.

In each of the following three sections, the different datasets are discussed including preparation of the data and for HES and MINAP, the number of cases whose address was listed in Scotland or Wales is presented.

4.3.2 Mortality data

Source summary

Mortality information has been received since baseline for all participants of the cohort who provided sufficient information (name, date of birth, NHS number) to allow record linkage through the then ONS, now NHSIC. Ninety eight percent of baseline participants were successfully traced to allow record linkage.

Records have been received regularly since study baseline and from these both fatal CHD and stroke cases have been identified.

Case definitions

Cases were classified using the 'original underlying' field from death certificates and using ICD 9th edition and 10th edition codes. Fatal cerebrovascular events were identified with ICD9 codes 430 to 438 or ICD10 codes I60 to I698 and fatal CHD events with ICD9 codes 410 to 4149 and ICD10 codes I20 to I259. For analyses using mortality data, total CVD cases were classed as those where the cause of death was either ascribed to cerebrovascular or heart disease. Note that fatal CHD cases identified here may all be classified as ischaemic heart disease (IHD) cases as no participants were identified with main cause of death classified as I46 or I47 (Cardiac arrest or Paroxysmal tachycardia).

Data summary

For analyses using only mortality outcome data, which are presented in the next chapter (Chapter 5), records span from study baseline to February 2011 and include English, Welsh and Scottish residents. Cases in later chapters, which examine total (fatal plus non-fatal) CVD risk and include cases from MINAP and HES (Chapters 6 and 7), extend up to June 2011 as additional mortality data were available. Table 4.1 details the IHD, stroke and CVD case numbers used in the different stages of work.

Table 4.1 Fatal IHD, stroke and CVD case numbers used in examining associations between dietary fibre and CVD mortality and CVD incidence

	IHD mortality case frequency	Stroke mortality case frequency	CVD mortality case frequency
Chapter 5: mortality cases span to 28th Feb 2011 for English, Welsh and Scottish residents	208	175	383
Chapter 6/7: mortality cases used in incident CVD analyses span to 30th June 2011 for English residents only	196	158	354

Note: case numbers presented here are before any exclusions have been applied to the sample

4.3.3 Myocardial Ischaemia National Audit Project

Source summary

MINAP is a clinical database specifically designed to record acute coronary syndromes (ACS) within English and Welsh hospitals (Herrett et al., 2010). ACS data are collected prospectively at each hospital, electronically encrypted and transferred on-line to a central database. By means of many data fields, each patient entry gives details of the patient journey, including the method and timing of admission, inpatient investigations, results and treatment, final diagnosis and (if applicable) date of death (from linkage to NHSIC). Data entry is subject to routine on-line error checking.

The registry was established in 1998, data collection began in 2000 and by mid-2002 all acute hospitals in England and Wales were participating (Herrett et al., 2010). Records for UKWCS participants were therefore obtained from 2003 onwards, when the register coverage was nationwide. All ACS data relating to participants of the UKWCS were identified in July 2011 through a third-party record linkage service. Event data therefore span from 1st January 2003 to 30th June 2011.

Case definitions

Within the MINAP dataset there is a 'final diagnosis' field which is completed for all admissions. Final diagnosis is defined for each participant using standard definitions (MINAP, 2010), which are detailed Appendix VIII. Each of the seven classifications used are listed in Table 4.2 below and were used to determine record inclusion with the help of Dr Chris Gale, a senior lecturer in Cardiovascular Health Sciences and Honorary Consultant Cardiologist, who

indicated which of the diagnosis types should be included in the definitions for MI and ACS (see Table 4.2).

Table 4.2 Final diagnosis frequency and outcome classification in all MINAP records and for English participants only

Final diagnosis classifications used in MINAP dataset	Total records received	Number of first diagnosis records in English residents*	Diagnosis type included in MI classification?	Diagnosis type included in ACS classification?
ST segment elevation MI	67	58	X	X
Threatened infarction	1	1	-	X
ACS troponin +ve	138	114	-	X
ACS troponin –ve	29	23	-	X
Chest pain, cause uncertain	6	5	-	-
Other diagnosis	18	16	-	-
ACS troponin not stated	2	2	-	X
Blank field	8	7	-	-
Total records	269	226	58	198

X= final diagnosis category included in classification of MI or ACS

*Subsequent events removed for each participant

Data cleaning and summary

In total, 269 event records were received, matching to 236 UKWCS participants (after removing any subsequent events for each participant, where multiple records existed). Records for 8 Welsh participants (12 event records) were then removed. The earliest event record for each participant was identified through manual screening and if a second event was listed, the earliest was retained, providing that the 'final diagnosis' for the earliest record was either MI or ACS. In all cases where more than one record existed, the earliest recorded event was either MI or ACS.

Case frequencies for each of the diagnosis types before and after removal of records for Welsh participants and any subsequent events are listed in Table 4.2. In total 58 MI and 198 ACS records were identified for English participants within the MINAP dataset.

Admission dates for MINAP records were examined to check whether a lag-time in event recording existed for the more recent events i.e. to check that recent event reporting was not incomplete because of recording lag. In general, consistent case frequencies each month followed by a reduction in event cases in the most recent month would indicate a lag-time in

event recording. However, when examining a histogram of MINAP record frequency reporting (Figure 4.3) there was no evidence of a drop-off in event reporting because the number of cases each month was small and there was a large degree of variation in case numbers between months. All MINAP events were therefore included in analyses. The latest event date in MINAP was chosen as the censor date for all outcome data sources.

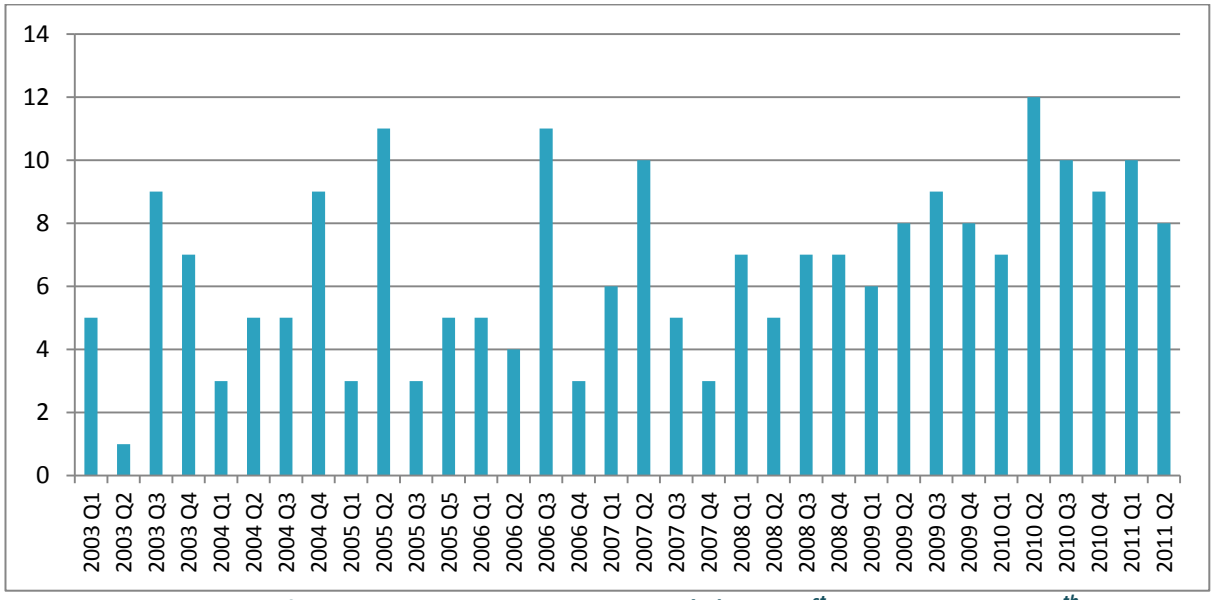


Figure 4.3 MINAP case frequency reported each quarter (Q) since 1st January 2003 to 30th June 2011, for English residents

4.3.4 Hospital Episode Statistics

Source summary

HES records have been routinely collected by all English hospitals since around 1998. These data primarily exist for economic purposes and to track service use but diseases or diagnoses are also recorded for each admission. The inpatient data for CHD and stroke obtained for the UKWCS span from 1998 to 30th December 2011.

Case definitions

All records for UKWCS participants with any of the CVD ICD10 codes listed in Table 4.3 were identified within the main HES database and were extracted by a trusted third party for data linkage with the cohort.

Table 4.3 ICD10 codes used to identify relevant HES records

Disease category	ICD 10 Code	General description (sub-category descriptions not included here)
Hypertensive diseases	I10	Essential (primary) hypertension
	I11	Hypertensive heart disease
	I15	Secondary hypertension
Ischaemic heart disease	I20	Angina pectoris
	I21	Acute myocardial infarction
	I22	Subsequent myocardial infarction
	I23	Certain current complications following acute myocardial infarction
	I24	Other ischaemic heart diseases
	I25	Chronic ischaemic heart disease
Other forms of heart disease	I46	Cardiac arrest
	I47	Paroxysmal tachycardia
	I48	Atrial fibrillation and flutter
	I49	Other cardiac arrhythmias
	I50	Heart failure
Cerebrovascular diseases	I60	Subarachnoid haemorrhage
	I61	Intracerebral haemorrhage
	I62	Other non-traumatic intracranial haemorrhage
	I63	Cerebral infarction
	I64	Stroke not specified as haemorrhage or infarction

Four CHD event types were identified from HES records using the relevant ICD codes (Table 4.4) in consultation with Dr Chris Gale (Cardiologist). Separating events by sub-type allows for exploration of fibre and risk associations, where sufficient cases exist in each sub-group:

- Total CHD: All CHD-related events
- Myocardial Infarction (MI)
- ACS: includes all acute coronary events
- Chronic: includes all chronic and not acute events
- Other: includes other cardiac events suggestive of heart disease. These are mostly due to chronic heart disease but these diseases can also have other, non heart disease related causes.

Table 4.4 Coronary heart disease classification codes and outcome grouping

ICD10 Code	Code description	Total CHD	MI	ACS	Chronic	Other
I20	Angina pectoris					
I20.0	Unstable angina	X		X		
I20.1	Angina pectoris with documented spasm	X			X	
I20.8	Other forms of angina pectoris	X			X	
I20.9	Angina pectoris, unspecified	X			X	
I21	Acute myocardial infarction					
I21.0	Acute transmural myocardial infarction of anterior wall	X	X	X		
I21.1	Acute transmural myocardial infarction of inferior wall	X	X	X		
I21.2	Acute transmural myocardial infarction of other sites	X	X	X		
I21.3	Acute transmural myocardial infarction of unspecified site	X	X	X		
I21.4	Acute subendocardial myocardial infarction	X	X	X		
I21.9	Acute myocardial infarction, unspecified	X	X	X		
I22	Subsequent myocardial infarction					
I22.0	Subsequent myocardial infarction of anterior wall	X	X	X		
I22.1	Subsequent myocardial infarction of inferior wall	X	X	X		
I22.8	Subsequent myocardial infarction of other sites	X	X	X		
I22.9	Subsequent myocardial infarction of unspecified site	X	X	X		
I24	Other acute ischaemic heart diseases					
I24.8	Other forms of acute ischaemic heart disease	X		X		
I24.9	Acute ischaemic heart disease, unspecified	X		X		
I25	Chronic ischaemic heart disease					
I25.0	Atherosclerotic cardiovascular disease, so described	X			X	
I25.1	Atherosclerotic heart disease	X			X	
I25.2	Old myocardial infarction	X			X	
I25.3	Aneurysm of heart	X			X	
I25.5	Ischaemic cardiomyopathy	X			X	
I25.6	Silent myocardial ischaemia	X			X	
I25.8	Other forms of chronic ischaemic heart disease	X			X	
I25.9	Chronic ischaemic heart disease, unspecified	X			X	
I46	Cardiac arrest					
I46.0	Cardiac arrest with successful resuscitation	X				X
I46.1	Sudden cardiac death, so described	X				X
I46.9	Cardiac arrest, unspecified	X				X
I47.0	Paroxysmal tachycardia	X				X
I47.2	Ventricular tachycardia	X				X
I49.0	Other cardiac arrhythmias	X				X

Key: MI Myocardial Infarction; ACS Acute Coronary Syndrome

Stroke data were similarly grouped into outcome classifications to allow exploration of the associations between fibre and different types of stroke (Table 4.5):

- Total stroke: Any ICD code spanning I60-I64X
- Subarachnoid haemorrhage (SH)
- Intracerebral haemorrhage (IBH)
- Intracranial haemorrhage (ICH)
- Cerebral infarction (CIF)
- Other: diagnosis of stroke but no distinction was made between infarction or haemorrhage

Table 4.5 Stroke classification codes and outcome grouping

ICD10 Code	Code description	Total stroke	SH	IBH	ICH	CIF	Other
I60	Subarachnoid haemorrhage (SH)						
I600	SH from carotid siphon and bifurcation	X	X				
I601	SH from middle cerebral artery	X	X				
I602	SH from anterior communicating artery	X	X				
I603	SH from posterior communicating artery	X	X				
I604	SH from basilar artery	X	X				
I605	SH from vertebral artery	X	X				
I606	SH from other intracranial arteries	X	X				
I607	SH from intracranial artery, unspecified	X	X				
I608	Other SH	X	X				
I609	SH, unspecified	X	X				
I61	Intracerebral haemorrhage (IBH)						
I610	IBH in hemisphere, subcortical	X		X			
I611	IBH in hemisphere, cortical	X		X			
I612	IBH in hemisphere, unspecified	X		X			
I613	IBH in brain stem	X		X			
I614	IBH in cerebellum	X		X			
I615	IBH, intraventricular	X		X			
I616	IBH, multiple localised	X		X			
I618	Other IBH	X		X			
I619	IBH, unspecified	X		X			
I62	Other nontraumatic intracranial haemorrhage (ICH)						
I620	Subdural haemorrhage (acute) (nontraumatic)	X			X		
I621	Nontraumatic extradural haemorrhage	X			X		
I629	ICH (nontraumatic), unspecified	X			X		
I63	Cerebral Infarction (CI)						
I630	CI due to thrombosis of precerebral arteries	X				X	
I631	CI due to embolism of precerebral arteries	X				X	
I632	CI due to unspecified occlusion or stenosis of precerebral arteries	X				X	
I633	CI due to thrombosis of cerebral arteries	X				X	
I634	CI due to embolism of cerebral arteries	X				X	
I635	CI due to unspecified occlusion or stenosis of cerebral arteries	X				X	
I636	CI due to cerebral venous thrombosis, nonpyogenic	X				X	
I638	Other CI	X				X	
I639	CI, unspecified	X				X	
I64X	Stroke, not specified as haemorrhage or infarction	X					X

Data cleaning and summary

Within the HES database, for UKWCS participants, 25,787 record rows were identified with the disease codes listed in Table 4.3. Each episode in HES or each admission into hospital is presented on multiple rows within the dataset. ICD10 codes are also presented in a 'main diagnosis' field and in multiple 'other diagnosis' fields and so taking only the main field for case identification results in the appearance of fewer cases compared to assessing ICD10 codes listed in any of the 'other diagnosis' fields.

Removal of non-English residents

As HES data were only obtained for English hospitals and MINAP covers only England and Wales, all participants whose address at baseline was listed outside of England were removed. The numbers and proportion of participants not matching to any CVD-related HES records within each region of the UK was calculated (Table 4.6). Received HES records were compared against the 36,126 UKWCS participants who were listed as successfully traceable via the NHSIC.

Firstly, missing geographical data (government office region) for participants was filled in where this was not already assigned in the dataset. In the existing UKWCS dataset 1,149 women had no government office region assigned. Region assignment was based on address information and so missing data were likely because of incorrect postcode or format of address. Addresses for these women were visually scanned and participants were either grouped with existing Scottish, Welsh and Northern Ireland participants or grouped as 'unassigned English region'.

In total, 7,841 women of 36,126 (21.7%) had not been identified in any CVD-related HES record since 1995, when both the main or other HES diagnosis fields were examined for the relevant ICD10 codes listed in Table 4.3.

For each of the English regions, the proportion of participants not appearing in HES was much lower than the average unmatched proportion of 21.7% (discussed above), at around 10% (8.4 to 11.9%). Despite HES records relating to only English hospitals, a small proportion of women whose address at baseline was listed as in Northern Ireland, Wales or Scotland appear in the dataset. The relative proportion of Welsh and Northern Ireland residents being unmatched in HES was similar, 76.2% and 67.9%, respectively and almost all of the Scottish residents were not matched in HES (96.1%). Of those women without an assigned region but not resident in Scotland, Wales or NI (n=961), 12.8% (n=123) were not matched to any HES record.

Hospital records for Scotland, Wales and Northern Ireland are not centrally held with English data and so fewer HES records would be expected for these participants. Examining record matching in this way confirmed that records were much less complete for non English residents and analyses using incidence data should therefore include only English residents.

Table 4.6 Proportion of participants with no CVD-related HES records, by region

Government office region (based on reported postcode or address at baseline)	Participant frequency in each region	Proportion of total sample in each region (%)	Frequency of unmatched* participants	Proportion of each region unmatched* to HES (%)	Proportion of total 7841 unmatched* participants from each region (%)
North East	1 102	3.05	93	8.4	1.19
North West	3 319	9.19	334	10.1	4.26
Yorkshire and Humber	2 702	7.48	226	8.4	2.88
East Midlands	2 028	5.61	183	9.0	2.33
West Midlands	2 739	7.58	327	11.9	4.17
East of England	2 942	8.14	327	11.1	4.17
Greater London	4 139	11.46	447	10.8	5.70
South East	6 939	19.21	812	11.7	10.36
South West	4 208	11.65	383	9.1	4.88
Unassigned English region	961	2.66	123	12.8	1.57
Wales	1 291	3.57	984	76.2	12.55
Northern Ireland	28	0.08	19	67.9	0.24
Scotland	3 728	10.32	3 583	96.1	45.70
	36 126	100	7 841		100

*Unmatched refers to participants that were not identified from the English HES record database for any CVD related record.

Identifying earliest event dates within HES

Within HES, multiple rows of data exist for each participant. These detail stages of treatment through an in-patient stay and also possible multiple inpatient experiences. The data were collapsed into a single row in order to link with cohort variables. New variables were generated for each disease diagnosis (detailed in Tables 4.4 and 4.5) and a distinction was also made as to whether the relevant diagnosis codes had been recorded in the main or other diagnosis fields. This distinction results in narrow and broad criteria for identifying cases from HES.

Each of the 20 new variables was separately condensed, using participant ID, to identify the earliest occurrence of each type of diagnosis, by participant. This allowed multiple diagnosis records, occurring at different dates, to exist on a single row for each participant (see example dataset in Table 4.7).

Table 4.7 Example dataset displaying multiple diagnoses per participant

ID	ACS case from main diagnosis	ACS main event date	ACS case from any diagnosis	ACS any event date	Stroke case from main diagnosis	Stroke main event date	Stroke case from any diagnosis	Stroke any event date
55550	no	-	yes	05/05/05	yes	01/01/01	yes	01/01/01
55551	no	-	no	-	no	-	yes	03/03/03
55552	yes	06/06/06	yes	06/06/03	no	-	no	-

ACS = acute coronary syndrome

Cases identified from HES, by diagnosis type

The numbers of data rows and individual cases for each diagnosis type using both the main diagnosis field and all diagnosis fields are displayed in Table 4.8. In total, 1178 CHD events were identified using the main and 1937 using all diagnosis fields. For stroke, 494 cases were observed using the main and 546 using any diagnosis field.

Table 4.8 Row and case frequency since study baseline of different CVD outcomes using the primary and multiple diagnosis fields from HES

	Identified using HES primary diagnosis field		Identified using all HES diagnosis fields	
	Data row frequency*	Case frequency[#]	Data row frequency*	Case frequency[#]
Total CHD	2675	1178	8143	1937
Myocardial Infarction	535	312	698	374
Acute coronary syndrome	1076	556	1446	666
Chronic cardiac events	1512	822	7180	1696
Other cardiac events	87	51	219	150
Total stroke	1007	494	1126	546
Subarachnoid haemorrhage	116	58	126	61
Intracerebral haemorrhage	144	73	170	89
Intracranial haemorrhage	56	30	78	41
Cerebral Infarction	492	258	554	283
Other cerebrovascular events	199	124	227	144

*Multiple rows/events per participant

[#] Subsequent events per participant removed.

Note numbers for total stroke and total CHD are not equal to the sum of each sub-type because first event type was considered where multiple events exist for participants.

Incident cases after study follow-up

To identify the number of incident cases available for analyses including food diary data, which was collected at study follow-up, HES data rows were dropped where the event date preceded the date of follow-up questionnaire receipt (refer to figure 4.2). The number of data rows and individual cases for each diagnosis type occurring after follow-up questionnaire receipt are displayed in Table 4.9. There were 332 incident CHD cases and 173 incident stroke cases from follow-up questionnaire receipt. Note that case frequencies for CHD or stroke cases identified using all HES diagnosis codes are not presented as this follow-up incidence data will only be used in time-to-event analyses (Chapter 7). As the secondary diagnosis fields in HES may represent historical events, it is not appropriate to use the admission date for a potentially unrelated condition in calculating survival times for modelling disease risk.

Table 4.9 Row and case frequency since study follow-up for each CVD outcome, calculated using the primary HES diagnosis field

	Frequency of data rows in HES records*	Case frequency#
Total CHD	788	332
Myocardial Infarction	151	86
Acute coronary syndrome	309	161
Chronic cardiac events	458	267
Other cardiac events	21	12
Total stroke	405	173
Subarachnoid haemorrhage	46	16
Intracerebral haemorrhage	48	25
Intracranial haemorrhage	34	16
Cerebral Infarction	206	95
Other cerebrovascular events	71	41

*Includes multiple rows/events per participant
#Subsequent events per participant removed, no duplicate IDs

HES data for comparison with MINAP

In order to suitably compare the 'quality' of HES and MINAP and establish if both sources have identified the same cases, ACS cases from HES were restricted to the same time-frame as MINAP. In total there were 1,007 data rows relating to 462 English individuals, with event dates inside the date range 1st January 2003 to 30th June 2011. HES ACS cases were defined both using the main or any diagnosis fields.

4.3.5 Capture-recapture methodology

It is not necessary to count every case in a population because the recapture or source overlap information can be used to estimate this (Chao et al., 2003). The capture-recapture method attempts to estimate the total number of disease cases in a population and thus the completeness of different data sources can be calculated (Hook and Regal, 1995).

Log-linear modelling is typically used to model count data when there are three or more variables of interest and can be applied, among other uses, to capture-recapture count data. The expected cell counts are estimated in a similar way to the application of the Chi-squared method when there are two variables of interest. The method essentially compares the observed number of cases in each cell with the number that is expected to occur by chance (Cramer, 2003). "The log-linear approach models the logarithm of the expected value of each observable category" (Chao et al., 2003) and the dependant variable in this analysis is the difference between the observed and expected case frequency, expressed as a likelihood ratio chi-square (Cramer, 2003). Log-linear models are the method of choice for assessing capture-recapture count data (Hook and Regal, 1995) and have been extensively used to handle dependence among samples (Chao et al., 2003).

Using the log-linear approach, various models are fitted to observed cells. The fit of the various models can be assessed to identify which model includes the simplest or most parsimonious explanation of the case distribution (Chao et al., 2003, Cramer, 2003). The 'best' model can then be applied to predict the expected number of cases missing from all lists (Chao et al., 2003).

Three-source approach

Using the numbers of observed cases from each source individually and from all combinations of sources, log-linear analysis was used to calculate the number of expected incident cases not captured on any of the three lists. The number of cases not identified with any list is represented by '*h*' in Table 4.10.

Table 4.10 Three-way cross over table illustrating all potential combinations of case reporting from the three event data sources

		Case in mortality records?			
		YES		NO	
		Case in HES?		Case in HES?	
		YES	NO	YES	NO
Case in MINAP?	YES	a	b	c	d
	NO	e	f	g	h

In comparing mortality data and cases from HES and MINAP, the cases considered were IHD mortality cases received from NHSIC and ACS cases identified through both HES and in MINAP.

Lists may not be independent in the way cases are identified, for example, the characteristics that lead to being recorded in one list may mean cases are more likely to be identified in other lists. This list dependency was explored using eight (2^3) models, to identify the most parsimonious model for calculating the completeness of datasets and therefore estimate the number of missed cases.

Models with increasing dependency are presented in Table 4.11. The first model assumes all sources to be independent with no interactions between lists, three models include one two-way interaction, three include two pair-wise interactions and one model was saturated, with three pair-wise interactions. It is not possible to model a three way interaction as the number of missed cases is unknown so it is assumed that no three way interaction exists, in order to estimate the number of cases missed from all lists.

Table 4.11 Eight models fitted for three-source data, from independent to dependant lists

		Source interactions in each model
i	Unsaturated model, assumes independent lists	ONS, MINAP, HES
ii		ONS*MINAP, HES
iii		ONS*HES, MINAP
iv		MINAP*HES, ONS
v		ONS*MINAP, HES*ONS
vi		ONS*MINAP, MINAP*HES
vii		MINAP*HES, ONS*HES
viii	Saturated model, assumes dependent lists	ONS*MINAP, ONS*HES, MINAP*HES

*= interaction between sources

Models were reviewed for goodness-of-fit to identify whether the inclusion of interactions between sources improved estimates. Using the most appropriate model, completeness of each data source was estimated followed by the completeness of the three overlapped sources, calculated as the proportion of cases observed to the number expected.

As the MINAP dataset covered a shorter time-frame (1st January 2003 to 30th June 2011) than both HES and the mortality data, cases occurring outside these dates were excluded from analyses to ensure comparability with MINAP. Estimates of missing cases were generated twice, once including a diagnosis of ACS from the main HES field and again using any of the diagnosis fields within HES records.

The 'poisson' command was used in Stata which produces estimates comparable to the Stata user-written program 'recap' (an der Heiden, 2009). In order that symmetric behaviour of the likelihood function was not assumed, 95% CIs around the maximum likelihood ratio value were generated according to the goodness-of-fit method suggested by Regal and Hook (Regal and Hook, 1984). The Stata command for goodness-of-fit based 95% CIs was kindly provided by Dr Darren Greenwood.

Two-source approach

Stroke data

The completeness of stroke case capture was assessed using HES and mortality data spanning from 1st January 2003 to 30th June 2011. This analysis was conducted twice, once including a diagnosis of stroke in the main HES diagnosis field and again using any of the diagnosis fields within HES records.

HES vs. MINAP

HES and MINAP should essentially capture the same participants who are admitted into hospitals with ACS. The successful capture of cases in HES and MINAP has been examined over a comparable time frame for both datasets, 1st January 2003 to 30th June 2011. As with stroke analyses, both the main and all diagnoses fields within HES were separately considered in this comparison. Additionally, as the quality of data recording is likely to have improved over time, since MINAP has become fully established, the analysis was repeated looking at cases reported between 1st January 2003 to 31st December 2006 and from 1st January 2007 to 30th June 2011.

4.4 Results

4.4.1 Case cross-matches

Spanning from 1st January 2003 to 30th June 2011, 149 IHD mortality cases were observed, 198 ACS cases from the MINAP dataset and 339 or 419 ACS cases were identified within the HES dataset, using either the main or all diagnosis fields, respectively (Table 4.12). When the three sources were combined, and so including case overlap, a total of 516 cases were identified using the main HES diagnosis field and this increased to 573 coronary cases identified using any diagnosis field within HES. The degree of overlap between the three sources, using both the main or using all diagnosis fields within HES, is also displayed in Figures 4.4 and 4.5, respectively.

Table 4.12 CHD case capture and overlap within 3 sources

IHD mortality case	MINAP ACS case	HES ACS case	Source overlap frequency using main HES diagnosis field	Source overlap frequency using all HES diagnosis fields
-	-	-	?	?
-	-	✓	181	238
-	✓	-	51	37
-	✓	✓	135	149
✓	-	-	123	115
✓	-	✓	14	22
✓	✓	-	3	2
✓	✓	✓	9	10
n=149	n=198	Main diagnosis 339 Any diagnosis 419	Total n=516	Total n=573

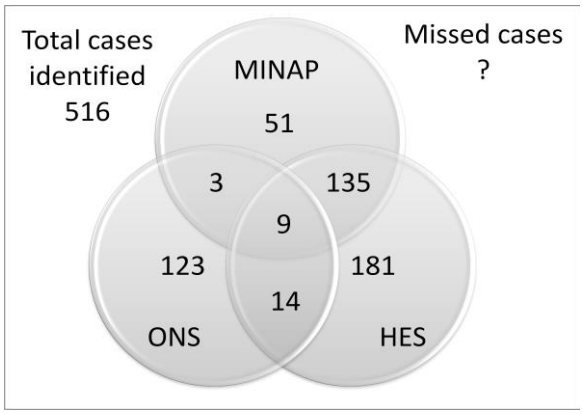


Figure 4.4 Source overlap including only the main HES diagnosis field

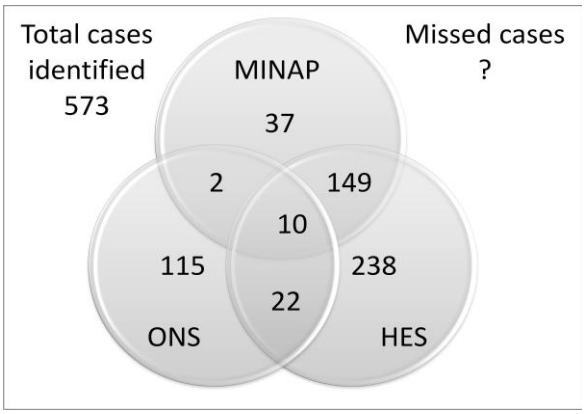


Figure 4.5 Source overlap including all HES diagnosis fields

Spanning from 1st January 2003 to 30th June 2011, 145 stroke mortality cases were observed and 368 or 406 stroke cases were identified within the HES dataset, using either the main or all diagnosis fields respectively (Table 4.13). When both sources were combined, and so including case overlap, a total of 433 cases were identified using the main field and this increased to 467 stroke cases identified with any diagnosis field in HES.

Table 4.13 Stroke case capture and overlap within 2 sources

Stroke mortality case	HES stroke case	Source overlap frequency using main HES diagnosis field	Source overlap frequency using all HES diagnosis fields
-	-	?	?
-	✓	288	322
✓	-	65	61
✓	✓	80	84
n=145	Main diagnosis field=368 All diagnosis fields=406	Total n=433	Total n=467

ACS cases from both HES and MINAP records were assessed spanning from 1st January 2003 to 30th June 2011 and also split by event date. Using the primary diagnosis field within HES, 393 ACS cases were observed in HES and MINAP and using all diagnosis fields this increased to 458 (Table 4.14).

Table 4.14 ACS case capture and overlap within HES and MINAP over different timeframes

ACS case observed in MINAP	ACS case observed in HES	Case frequency observed using main HES diagnosis field			Case frequency observed using all HES diagnosis fields		
		01/01/03 to 30/06/11	01/01/03 to 31/12/06	01/01/07 to 30/06/11	01/01/03 to 30/06/11	01/01/03 to 31/12/06	01/01/07 to 30/06/11
-	-	?	?	?	?	?	?
-	✓	195	85	112	260	100	162
✓	-	54	16	40	39	12	29
✓	✓	144	56	86	159	60	97
		n=393	n=157	n=238	n=458	n=172	n=288

4.4.2 Estimation of total CHD cases

Each of the eight list dependency models (Table 4.11) was applied to the CHD case data to identify the best fit. Estimates for expected missing cases, total case estimation and CHD ascertainment rates using the three combined sources in each of eight models which account for different source interactions, are displayed in Tables 4.15 and 4.16.

The model without source interaction terms (*i*) fit the data least well (refer to Table 4.11), having the highest values for goodness-of-fit and Akaike's Information Criterion (AIC), which is a measure of goodness-of-fit corrected for the number of parameters in the model. Of the models that included just one, two-way interaction between sources (models *ii*, *iii* and *iv*), models *ii* and *iii* which had an interaction term between mortality information and MINAP or HES, do not fit the data as well as model *iv* with an interaction between HES and MINAP sources. Models *vi*, *vii* and *viii* that include two and three interaction terms between sources also have comparably low goodness-of-fit and AIC values as model *iv* but are more complex with additional interactions and therefore do not represent the most parsimonious explanation of list dependence. Model *iv* was selected as the best or most parsimonious of the eight, having acceptable goodness-of-fit and AIC values.

Using model *iv*, the best model, the estimated number of CHD cases missed by all three sources is 1736 (95% CI: 1138 to 2741) when cases were identified with the primary field in HES (Table 4.17). This gives an estimate for total CHD cases of 2252 (95% CI: 1654 to 3257) and indicates that just 23% (95% CI: 16 to 31%) of cases were captured in total. Estimates are

similar when cases were identified using the broader HES definition (all diagnosis fields). The number of missing cases is estimated as slightly lower at 1434 (95% CI: 973 to 2162) and so in total, 2007 (95% CI: 1546 to 2734) cases are estimated. The total ascertainment rate for CHD events using death records, MINAP and all diagnoses codes in HES is 29% (95% CI: 21 to 37%).

The estimated proportion of total cases identified through death records was just 7% (95% CI: 5 to 9%), through MINAP was 9% (95% CI: 6 to 12%) and 15% (95% CI: 10 to 20%) of total cases were observed in HES. Numbers were similar when all HES diagnosis fields were included in the definition of ACS, but the proportion of total cases identified through HES increased to 21% (95% CI: 15 to 27%) (Table 4.17).

Table 4.15 Assessing best model fit for list dependency: Estimation of expected missing cases and total cases estimated from each of eight interaction models, using ACS cases identified from the main HES diagnosis field

Model	Terms and interactions in each model	Pearson goodness -of-fit	AIC	Expected missing (95% CI)*	Estimated total cases (95% CI)*	Total ascertainment rate (%)(95% CI)	Death ascertainment rate (%)(95% CI)	MINAP ascertainment rate (%)(95% CI)	HES ascertainment rate (%)(95% CI)
i)	D , M, H	182	173.2	275 (209, 357)	791 (725, 873)	65 (59, 71)	19 (17, 21)	25 (23, 27)	43 (39, 47)
ii)	D , M, H, D*M	127	136.94	203 (150, 270)	719 (666, 786)	72 (66, 77)	21 (19, 22)	28 (25, 30)	47 (43, 51)
iii)	D , M, H, D*H	68	85.75	110 (73, 159)	626 (589, 675)	82 (76, 87)	24 (22, 25)	32 (29, 34)	54 (50, 58)
iv)	D , M, H, M*H	0.23	-3.77	1736 (1138, 2741)	2252 (1654, 3257)	23 (16, 31)	7 (5, 9)	9 (6, 12)	15 (10, 20)
v)	D , M, H, D*M, D*H	35	22.32	68 (43, 102)	584 (559, 618)	88 (83, 92)	26 (24, 27)	34 (32, 35)	58 (55, 61)
vi)	D , M, H, D*M, M*H	0.03	-1.97	1590 (925, 2941)	2106 (1441, 3457)	25 (15, 36)	7 (4, 10)	9 (6, 14)	16 (10, 24)
vii)	D , M, H, D*H, M*H	0.11	-1.89	2091 (754, 8663)	2607 (1270, 9179)	20 (6, 41)	6 (2, 12)	8 (2, 16)	13 (4, 27)
viii)	D , M, H, D*M, D*H, M*H	0	0	1802 (453, 9184)	2318 (969, 9700)	22 (5, 53)	6 (2, 15)	9 (2, 20)	15 (3, 35)

Key as below

Table 4.16 Assessing best model fit for list dependency: Estimation of expected missing cases and total cases estimated from each of eight interaction models, using ACS cases identified from all HES diagnosis fields

Model	Terms and interactions in each model	Pearson goodness -of-fit	AIC	Expected missing (95% CI)*	Estimated total cases (95% CI)*	Total ascertainment rate (%)(95% CI)	Death ascertainment rate (%)(95% CI)	MINAP ascertainment rate (%)(95% CI)	HES ascertainment rate (%)(95% CI)
i)	D , M, H	181	171.87	256 (194, 332)	829 (767, 905)	69 (63, 75)	18 (16, 19)	24 (22, 26)	51 (46, 55)
ii)	D , M, H, D*M	135	141.09	202 (151, 267)	776 (724, 840)	74 (68, 79)	19 (18, 21)	26 (24, 27)	54 (50, 58)
iii)	D , M, H, D*H	58	75.35	86 (54, 129)	659 (627, 702)	87 (82, 91)	23 (21, 24)	30 (28, 32)	64 (60, 67)
iv)	D , M, H, M*H	1	-2.96	1434 (973, 2161)	2007 (1546, 2734)	29 (21, 37)	7 (5, 10)	10 (7, 13)	21 (15, 27)
v)	D , M, H, D*M, D*H	30	21.47	59 (36, 91)	632 (609, 664)	91 (86, 94)	24 (22, 24)	31 (30, 33)	66 (63, 69)
vi)	D , M, H, D*M, M*H	0.07	-1.92	1244 (786, 2049)	1817 (1359, 2622)	32 (22, 42)	8 (6, 11)	11 (8, 15)	23 (16, 31)
vii)	D , M, H, D*H, M*H	0.66	-1.32	2128 (639 13164)	2701 (1212, 13737)	32 (4, 47)	6 (1, 12)	7 (1, 16)	16 (3, 35)
viii)	D , M, H, D*M, D*H,M*H	1.14	0	1545 (354 10848)	2118 (927, 11421)	27 (5, 62)	7 (1, 16)	9 (2, 21)	20 (4, 45)

AIC Akaike Information Criterion; CI confidence intervals; D IHD death; H HES ACS case; M MINAP ACS case. *95% CIs calculated using goodness-of-fit based method (Regal and Hook, 1984)

4.4.3 Estimation of total stroke cases

Stroke cases identified from death records and through HES were combined to estimate missing cases. A total of 433 and 467 cases were observed using either the main or all HES diagnosis fields, respectively (Table 4.18). The total number of cases expected, using the main HES field, was 667 (95% CI: 593 to 769) and 701 (95% CI: 627 to 803) when all diagnosis codes were considered.

The total case ascertainment rate for mortality records was 22% (95% CI: 19 to 24%) or 21% (95% CI: 18 to 23%) and for HES was 55% (95% CI: 48 to 62%) or 58% (95% CI: 51 to 65%), using the main or all HES diagnosis fields, respectively.

4.4.4 Estimation of total ACS cases from HES and MINAP

ACS cases identified through MINAP and HES were combined and 393 or 458 cases were identified using the main or all HES diagnosis fields, respectively. An estimated 73 cases (95% CI: 47 to 108) were missed when just the primary HES diagnosis field was used and this reduced to 64 cases missed (95% CI: 39 to 97) with all diagnosis fields (Table 4.19).

The total ascertainment rate and separately, MINAP and HES ascertainment rates, appear to have improved from the early time period to the latter when considering ACS cases estimated from the main HES diagnosis field. The total ascertainment rate rose from 87% between 2003 and 2006 to 95% between 2007 and 2011. However, when considering the broader ACS definition within HES data (cases identified using all diagnosis fields), the total case ascertainment rate dropped slightly from 90% to 86% whilst the MINAP rate remained constant at 38% and HES ascertainment increased from 83% to 95%.

Table 4.17 Observed and expected cell counts and case ascertainment rates for CHD events reported by mortality records, MINAP and HES

	Total observed cases	Expected missing (95% CI)	Estimated total cases (95% CI)	Total ascertainment rate (%) (95% CI)	Death record ascertainment rate (%) (95% CI)	MINAP ascertainment rate (%) (95% CI)	HES ascertainment rate (%) (95% CI)
Comparison using main HES diagnosis field	516	1736 (1138 to 2741)	2252 (1654 to 3257)	23 (16 to 31)	7 (5 to 9)	9 (6 to 12)	15 (10 to 20)
Comparison using all HES diagnosis fields	573	1434 (973 to 2161)	2007 (1546 to 2734)	29 (21 to 37)	7 (5 to 10)	10 (7 to 13)	21 (15 to 27)

95% CIs calculated using the goodness-of-fit based method (Regal and Hook, 1984).

Table 4.18 Observed and expected cell counts and case ascertainment rates for stroke events reported by mortality records and from the primary and all HES diagnosis fields

Stroke death vs. HES stroke	Stroke deaths observed n	HES strokes observed n	Total observed strokes n	Expected missing (95% CIs)	Total expected (95% CIs)	Total ascertainment rate (%) (95% CIs)	Death records ascertainment rate (%) (95% CIs)	HES ascertainment rate (%) (95% CIs)
Comparison using main HES diagnosis field	145	368	433	234 (160 to 336)	667 (593 to 769)	65 (56 to 73)	22 (19 to 24)	55 (48 to 62)
Comparison using all HES diagnosis fields	145	406	467	234 (160 to 336)	701 (627 to 803)	67 (58 to 74)	21 (18 to 23)	58 (51 to 65)

95% CIs calculated using the goodness-of-fit based method (Regal and Hook, 1984).

Table 4.19 Observed and expected cell counts and case ascertainment rates for ACS in HES and MINAP

MINAP vs. HES	Comparison timeframe	MINAP observed n	HES observed n	Total observed n	Expected missing (95% CIs)	Total expected (95% CIs)	Total ascertainment rate (95% CIs)	MINAP ascertainment rate (95% CIs)	HES ascertainment rate (95% CIs)
Comparison using main HES diagnosis field	01/01/03 to 30/06/11	198	339	393	73 (47 to 108)	466 (440 to 501)	84 (78 to 89)	42 (40 to 45)	76 (68 to 77)
	01/01/03 to 31/12/06	72	141	157	24 (11 to 46)	181 (168 to 203)	87 (77 to 93)	40 (35 to 43)	78 (69 to 84)
	01/01/07 to 30/06/11	126	198	238	40 (22 to 65)	251 (233 to 276)	95 (86 to 102)	50 (46 to 54)	79 (72 to 85)
Comparison using all HES diagnosis fields	01/01/03 to 30/06/11	198	419	458	64 (39 to 97)	522 (497 to 555)	88 (83 to 92)	38 (36 to 40)	80 (75 to 84)
	01/01/03 to 31/12/06	72	160	172	20 (8 to 40)	192 (180 to 212)	90 (81 to 96)	38 (34 to 40)	83 (75 to 89)
	01/01/07 to 30/06/11	126	319	288	48 (27 to 79)	336 (215 to 367)	86 (78 to 134)	38 (34 to 59)	95 (87 to 148)

95% CIs calculated using the goodness-of-fit based method (Regal and Hook, 1984).

4.5 Discussion

4.5.1 Summary of findings

The most suitable model to explain the three source dependency included an interaction term between HES and MINAP. Dependency between these two sources is unsurprising as both record hospital inpatient events. Using this best model, an estimated 1736 (95% CI: 1138 to 2741) cases were missed from all three sources of CHD event data. The CIs are relatively wide and it is possible that this estimate for total case capture is not truly reflective of the number of cases. This estimate could be unreliable because of relatively small numbers of cases and the fact that for each interaction, there are even fewer cases available on which to base the estimate. This creates greater uncertainty around the estimates.

With the high estimate for total CHD cases, the total ascertainment rate is relatively low at just 23% with the death, MINAP and HES records respectively observing just 7%, 9% and 15% of cases. These findings suggest that there may be a substantial number of events occurring within the community which do not result in hospitalisation or death and thus are not recorded in the data sources used here. A recent study that included 4 sources of MI case data from English HES, MINAP, death registry and primary care datasets found that when just non-fatal events were considered from HES, MINAP and primary care data, 52.5% of all events were observed in MINAP, 67.9% in HES and 74.5% in the primary care dataset (Herrett et al., 2013). This study, published after the comparison of case data for the UKWCS was completed, supports the necessity of obtaining case data from multiple sources and in terms of MI demonstrates that primary care data have a valuable role to play (Herrett et al., 2013). A limitation of this study however was that no estimation of missing numbers of cases was made using a log-linear approach despite having large numbers of cases and comparable datasets in terms of outcomes and timeframes. The study, by contrast to the work presented in this chapter, assumes that no cases are missed when calculating the proportion of cases from each dataset. This may somewhat explain why the proportion of MI cases captured from HES and MINAP exceed the values identified here for the UKWCS case data. Additionally, the narrower outcome classification (MI) used by Herrett and colleagues may also account for the greater estimates of case capture. Some sensitivity may have been lost through comparing all CHD events, as was done in this chapter.

List dependency could not be modelled for estimating the number of missing and therefore total number of stroke cases because only two sources of event data were available. There were an estimated 234 (95% CI: 160 to 336) stroke cases for the UKWCS participants that were not captured either with HES or death records. The total ascertainment rate for both sources was estimated as 65%, with death records capturing 22% and the main HES diagnosis field capturing 55% of cases.

When just HES and MINAP cases were considered together, an estimated 73 (95% CI: 47 to 108) cases were missed over the study time period, with 24 (95% CI: 11 to 46) being missed between 2003 and 2006 and 40 (95% CI: 22 to 65) cases being missed between 2007 and 2011. The models estimate that a greater proportion of total ACS cases were recorded in the latter time period, even when taking into consideration that the later period is 6 months longer than the early period. Total ACS cases expected during 2003-2006 was 181 (95% CI: 168 to 203) and was considerably higher for the period 2007-2011, at 251 (95% CI: 233 to 276). During the years 1999 to 2007 it has been demonstrated that ACS admission rates decreased in women across all age groups (Pearson-Stuttard et al., 2012). The greater observed and total expected case numbers in this later period may therefore indicate more complete ACS capture and recording within HES and MINAP systems or could simply reflect the ageing profile of the UKWCS participants and greater risk of events with increasing age.

Participants who died were unlikely to appear in HES or MINAP, just 8% of death cases appeared in MINAP and 15% or 21% of death cases were listed in HES, using the main or all diagnosis fields, respectively. Similar observations were also made in the recent evaluation of the completeness and diagnostic validity of MI recording in four health record sources within England, discussed above. A total of 21,482 acute MI cases were observed and 36.7% of fatal MI cases were recorded in HES whilst just 17.1% were observed in MINAP. A far greater proportion of fatal MI cases, 55.9%, were however observed in primary care records (Herrett et al., 2013).

The three source case estimation results are informative but the application may somewhat over-estimate the degree of under-capture and therefore provide overly large estimates of the number of missed cases. This may occur because it is not possible to include the three-way source interaction in models, for this the number of missed cases would be needed and it is this that the model tries to estimate. Additionally, the two-way interactions are not estimated very precisely when applied to the case data for the UKWCS. Firstly, the model assumes there

is an equal chance of being captured by lists but fatalities that are not admitted to hospital would never be recorded in HES or MINAP. Secondly, there are relatively small numbers of cases available over the matching time periods and even fewer cases available to be used in source dependency models, where the number of overlapping cases is used. Thirdly, different case types are captured in the lists with only fatal events from one source, clinically confirmed ACS from MINAP and potentially unverified cases being recorded in HES. Additionally, less critical cases or silent MI cases are missed from all three sources. The recent work by Herrett and colleagues explores the less critical and retrospective case diagnosis within primary care settings in England and identified that many more cases can be identified with the use of this data (Herrett et al., 2013).

4.5.2 Strengths and limitations

Model assumptions

Principle assumptions should be met when applying capture-recapture methods for multiple lists. The first assumption, that individual identifiers are not lost, is not an issue where participants have unique identifiers. However, the second assumption, that lists are independent, is problematic in health applications (IWGDMF, 1995b). List or source dependency occurs because the nature of being on one list means participants are more likely to appear on others. In the context of the data available for the UKWCS, lists are likely to be dependent, simply being admitted to hospital and therefore appearing in HES means that participants are then available for entry in the MINAP system. Women who die before reaching hospital will not be identified in HES or MINAP and these records may therefore not be independent from HES or MINAP but partly mutually exclusive. Conversely, those women who die of cardiovascular disease whilst in hospital may be more likely to have accurate mortality records than those who die outside of hospital and therefore will appear in all three sources.

The assumption that homogeneity of capture probability should exist within lists is often violated in human populations (IWGDMF, 1995b). However the application of log-linear analysis permits modelling of this list dependency and can be appropriately applied to produce total case estimates where it is not possible to ensure list independence. Capture-recapture methods are widely used and are generally accepted as a practical way to estimate the degree of under-capture and to calculate the actual number of cases (Reintjes et al., 2007).

A further assumption made with capture-recapture methods is that the study population is closed (Hook and Regal, 1995), which essentially means that there are no additions or losses during a study period (Chao et al., 2003). Whilst there is a fixed list of participants for the UKWCS there is no guarantee that all cases would be captured, for example any non-fatal events outside of English hospitals would not be identified. However assuming that net migration between England and Wales or Scotland is roughly reciprocal, and given the small proportion of HES records received for Scottish or Welsh residents (Table 4.6), only a small proportion of events would be missed this way.

Sources of data

A major limitation is that there is no national register for all cardiovascular events in the UK and so sources must be combined for maximum case identification. Additionally, neither HES nor MINAP include cases spanning back to study inception and so any non-fatal events occurring before HES and MINAP began reliably collecting data cannot be identified.

HES data are principally collected for administrative purposes and so an obvious concern is regarding the accuracy of diagnostic code reporting and the validity of applying these for epidemiological research. Some authors caution careful interpretation of hospital activity data as it may not accurately reflect the underlying prevalence of disease (Hansell et al., 2001). Causes of data quality issues in HES include poor recording in patient notes or coding into the database system, failure to meet submission deadlines, leading to gaps in coverage and the fact that local systems vary greatly, which can lead to differences in data quality (HES, 2013). Although HES data quality has improved since the early 1990s, there is still reportedly wide variation between some health authority data especially for diagnostic codes (Hansell et al., 2001).

A comprehensive comparison study between HES records and GP records in England was carried out for random samples of participants from the Million Women Study (Wright et al., 2012). The reporting of vascular diseases between 1997 and 2005 was the main aim of the work to assess whether the reporting in HES is sufficiently reliable for epidemiological research. GPs were contacted for women presenting in HES with diagnoses of IHD, venous thromboembolism and cerebrovascular disease and also for a sample of women with no HES record of these diagnoses. For a diagnosis of IHD in HES, 92% of GP records had a matching diagnosis or for a closely related IHD diagnosis and for cerebrovascular disease this was 94%. The authors also identified that agreement was highest for the more severe outcomes which

are likely to be of greater epidemiological interest. By examining GP records for those women without vascular disease diagnosis in HES and finding very few contradicting diagnoses, the study found that recording of admissions for vascular disease in HES was virtually complete. The authors indicate that the Million Women Study is likely to be representative of middle-aged women in the general population. This validation work is therefore very informative and reassures that the likely quality of the data collected for UKWCS participants is acceptable for research, despite population differences between UKWCS participants and those in the Million Women Study or in the general population.

Furthermore, a recently published systematic review assessing the accuracy of routinely collected data from 25 British studies, published over two decades, where the comparison was against case notes found that 80% of diagnoses in routine data were confirmed in case notes (Burns et al., 2012). Over the time period assessed, coding practice has greatly changed with the introduction of payment by results and also changes to ICD classification (Burns et al., 2012). The studies were however heterogeneous with varying outcomes and methods and the historical nature of the data limits contemporary applicability somewhat (Burns et al., 2012). Despite this, and a lack of consensus on what level of data accuracy is acceptable, the authors identify that accuracy rates have improved and continue to do so (Burns et al., 2012).

Because of the limitations of using HES data, it was not relied on alone but used in conjunction with MINAP. The completeness of key fields in MINAP, including main discharge diagnosis, is closely monitored and has been found to be generally above 95%. Validation exercises using randomly selected MINAP records also indicate the median level of agreement between re-entered data and original has risen from 72% in 2003 to 89.5% in 2008 (Herrett et al., 2010). The scale of MINAP and the representativeness of MINAP data are also cited as key strengths which underpin its use as a research tool (Herrett et al., 2010).

4.6 Summary

This chapter has outlined the different sources of event data available for the UKWCS and how each dataset was prepared for use, with details of case definitions and exclusion of non-English participants.

Estimation of the validity of these datasets has been undertaken using both three and two source comparisons, as appropriate. It is likely that the three source comparisons, although informative, provide inflated estimates for total cases. The comparison of the matching case

type (ACS) through using MINAP and HES data may provide closer estimates of missing case numbers but list dependence cannot be accounted for in these two-source models. This validation exercise does however provide useful estimates for expected missing CHD, stroke or ACS cases for the UKWCS population despite limitations in either having only two sources of event data or limitations with case types not being exactly comparable in the three source estimations. The validation work indicates that none of the lists are complete and this finding supports the use of multiple sources of event data, as has been applied in Chapters 5, 6 and 7.

The next chapter utilises the mortality data, detailed above, and examines risk of IHD, stroke and total CVD mortality in association with dietary fibre intake. For this work all mortality cases since study baseline are utilised and not just those which were used in the validation exercise.

Subsequent chapters then build upon this work by assessing risk of total (fatal plus non-fatal) CVD risk in relation to fibre intake assessed using, firstly FFQs (Chapter 6), then food diaries (Chapter 7).

Chapter 5 Dietary fibre and fatal ischaemic heart disease, stroke and cardiovascular disease

5.1 Chapter overview

The current chapter explores dietary fibre and risk of fatal IHD, stroke and CVD. Fibre intakes assessed using FFQs contribute data on total fibre intakes, both soluble and insoluble fibre and also fibre from key food sources. Participants were followed from baseline for just over 14 years on average. Figure 5.1 details dietary and outcome data sources, as described in earlier chapters that have been used in the current chapter: diet assessed by FFQ and mortality data only.

This chapter details how potential confounders in the relationship between fibre and CVD risk were identified and handled in models. Survival analysis has been used to model the relationship between fibre and CVD and the approach described in the method section of this chapter applies to chapters 6 and 7 also.

Using survival analysis methods to explore exposures and outcomes, there was no evidence of an association between fibre intakes of any kind and IHD, stroke or CVD mortality risk in the full sample of women, after adjustments were applied. There was some evidence that greater cereal fibre intake may confer protection for stroke mortality in women who were classed as overweight or obese using their BMI at baseline. Greater fibre density of the diet was also protectively associated with fatal stroke risk in women who were free of hypertension or angina at study baseline.

An article was published from the work in this chapter (Threapleton et al., 2013b), in addition to two abstracts presented to the 2012 Meeting of the Society of Social Medicine (Threapleton et al., 2012b) and the winter meeting of the Nutrition Society 2012 (Threapleton et al., 2012a).

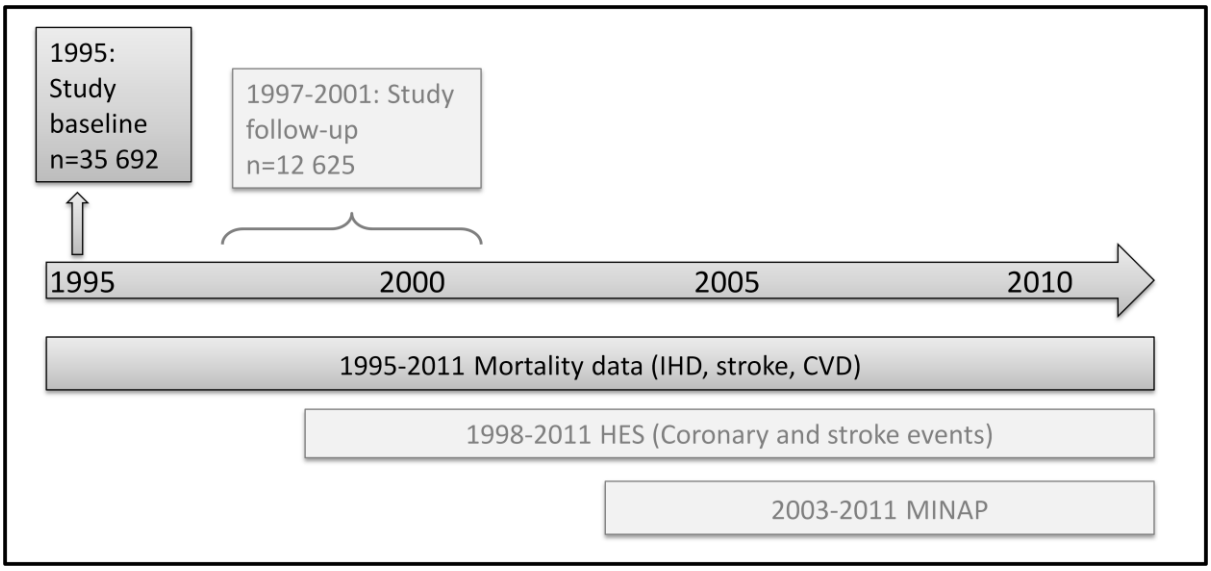


Figure 5.1 Data sources used in this chapter: Dietary data from FFQs and mortality (CHD, stroke and CVD) data spanning from study baseline

5.2 Background

CVD accounts for almost half of all deaths across Europe and is the main cause of disease burden (Allender et al., 2008). Although women typically experience CVD events later in life than men (Vaidya et al., 2011, Worrall-Carter et al., 2011), the annual mortality burden for CHD and stroke in women is estimated to be greater than that in men, at around 597,000 compared to 548,000 cases within the European Union (EU) (Nichols et al., 2012). In the UK specifically, around one in three deaths is attributed to CVD and for the year 2010 that is approximately 180,000 deaths. CHD was responsible for around 45% of the total CVD deaths and stroke for 28%, the rest being caused by a range of other circulatory diseases (Townsend et al., 2012). Rates of CVD are in decline in many developed European countries (Allender et al., 2008) and incidence rates are also declining in the UK, (Allender et al., 2008, Gale et al., 2012) a fall that has been attributed to improvements in risk factors for CVD, by means of lifestyle improvements (Unal et al., 2004).

A UK national target set in 1999 to reduce the death rate from CHD and stroke, among other diseases, by at least two fifths by 2010 was reached in 2009. In addition there has been progress towards reducing CVD inequalities for death rates in England between the population as a whole and the most deprived areas (Townsend et al., 2012). Although death from CVD related causes is likely to always be one of the primary causes of death in any population, key aims must include improving rates in premature mortality, typically classed as death under 75

years. Over the past 10 years premature mortality from CVD in the UK has fallen by 44% and in 2010 an estimated 28% of premature deaths in men and 19% in women were from CVD (Townsend et al., 2012).

Trends in CVD incidence reflect trends in mortality in that they have been declining over the past few decades. In England, over 10% of men and 15% of women who are admitted to hospital for MI die within 30 days and for stroke the figures are higher but are measured over a 60 day period. In England, an estimated 17% of men and 25% of women admitted to hospital with strokes die within 60 days but these mortality rates are substantially lower in individuals under 75 years (Townsend et al., 2012).

Gender differences in CVD mortality rates or life course disease trends exist (Nichols et al., 2012, Vaidya et al., 2011), indicating the importance in exploring preventative strategies separately between the sexes. Social inequalities in CVD mortality rates persist and are more striking in women than men. In the most recently available data from 2001/03 it was estimated that female workers with routine jobs had CHD death rates five times higher than their professional/managerial counterparts (Townsend et al., 2012). Given that 'most CVD in women is preventable' (Mosca et al., 2007, Worrall-Carter et al., 2011), the current work utilises dietary data from middle-aged women and explores fatal CVD risk in relation to total fibre intake as well as exploring major food sources of fibre, in order to characterise potentially beneficial dietary behaviour.

5.3 Method

5.3.1 Dietary data

A validated 217-item FFQ was used at baseline to assess typical intake over the previous 12 months (Calvert et al., 1997, Cade et al., 2004a). NSP intake values were estimated using data from McCance & Widdowson's *The Composition of Foods* (5th edition) (Holland et al., 1991) Fibre calculated using the AOAC method, in addition to soluble and insoluble fibre intakes were also used.

NSP estimates from specific food sources were also generated and include fibre from the following food groups: total cereal foods, breakfast cereals, vegetables (excluding potatoes), fruit (excluding juice), legumes and nuts/seeds.

Refer back to the 'Dietary assessment' section of Chapter 3 for additional details on dietary data collection or generation of exposure variables.

5.3.2 Mortality data

Mortality data were available from baseline for participants who provided sufficient information for their records to be traced through the NHSIC (98% of baseline participants were traced to allow linkage). Deaths were classified using ICD 9th edition and 10th edition codes. Fatal cerebrovascular events were identified with codes 430 to 438 or I60 to I698 and fatal IHD events with codes 410 to 4149 or I20 to I259. CVD cases were classed as either a cerebrovascular or heart disease case. Refer back to the 'Mortality Data' section of Chapter 4 for additional information.

5.3.3 Exclusions

Participants were excluded from analyses where the following criteria were met:

1) Not successfully tracked through the NHSIC (generally because NHS number or date of birth were incorrect) (n=695)
2) Tracked through national registers but did not provide both lifestyle and dietary information (n=318)
3) Daily calorie intake from the FFQ was outside a plausible range (500-6000 kcal/day) (n=405)
4) Reported personal history of cancer (n=2445), stroke (n=264), diabetes (n=646) or heart attack (n=498) at study baseline
5) Died within one year of returning FFQ (removed to limit reverse causality for any latent disease that may have caused diet to change) (n=98)
6) Requested their data not to be used in future studies (n=1)

5.3.4 Testing dose-response and non-linear associations

For full sample analyses (not subgroups), fibre intake was explored as a categorical exposure (fifths of intake) where each subsequent category was compared to the lowest intake category or where low and high consumers were compared to the middle intake group, in order to observe the shape of any associations. Models with continuous exposure variables were also used to determine whether any associations were linear, thus meeting one of the established criteria for causality (Hill, 1965).

Subgroup analyses were carried out only with the continuous exposure to minimise the effects of multiple testing and because case numbers were diminished, so splitting the samples into fifths would have resulted in small case numbers in each group.

5.3.5 Descriptive statistics

After exclusions, characteristics of high and low fibre consumers were explored when the sample was split into five categories based on NSP intake. The same dietary and lifestyle characteristics were also explored separately for women who reported personal history of stroke or heart attacks at baseline and were therefore excluded from analyses, those not classified as cases by the censor date (either died of other causes or were alive until study censor date) and those who became cases (study censor dates are discussed in Chapter 4). Stata version 11 (Statacorp, 2009) was used for all data manipulation and analyses in this chapter.

5.3.6 Survival analyses

Survival analyses were conducted using Cox regression (Cox and Oakes, 1984). The time variable used in survival analyses was time in the study (person years), calculated as the time from the date the questionnaire was completed until either a report of death or the censor date of the analysis, whichever comes first. Note that the censor date for those who died of other causes was their date of death. Censor date for surviving participants was set at February 2011, approximately the latest date of death received to that point for the cohort. Models were weighted by the inverse of the probability of being sampled to take into account the large proportion of vegetarians in the cohort and give less weight in models to data from vegetarian participants. The weighting variable had been generated by Dr Darren Greenwood for a previous study (Cade et al., 2007).

5.3.7 Checking proportional hazards

In order to ensure variables in the model were associated with proportional hazards over time, each variable was examined using log-log survival curves for each of the three outcomes (IHD, stroke, CVD) to ensure the survival function was constant over time and hazards were proportional in different groups. Here, the survival function [–ln(-ln(S))] was plotted over ln(time). Categorical variables were plotted with two or three categories and continuous variables were divided into two equal weight groups to explore whether risks were proportional in the high vs. low fibre consumers or younger vs. older women etc. Roughly

parallel lines on these plots indicate that hazards are proportional over time among different intake or lifestyle groups, thus meeting the requirement for the models.

Figure 5.2 displays smoothed (log-log) survival curves for IHD mortality in higher compared to lower NSP consumers. The closely parallel lines display how the hazards in low and high consumers are similar over the time frame of the study. This gives confidence that generating one hazard ratio for the association between NSP and risk is therefore sufficient to explain the association over the whole study duration because the risk association is constant in high compared to low level consumers.

This condition was met and hazard ratios were proportional in the case of each exposure and for all covariates used in models.

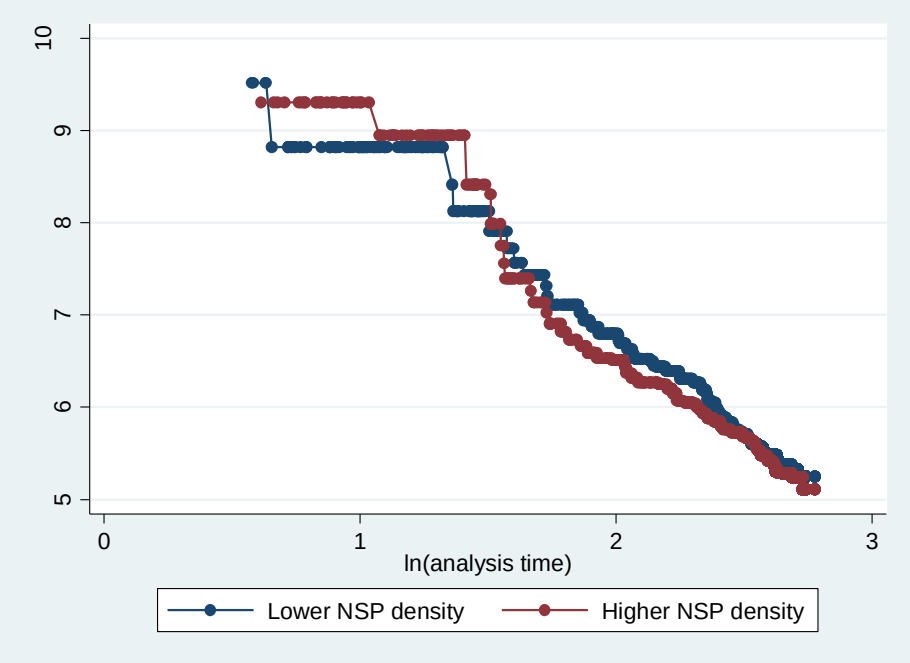


Figure 5.2 Survival function for IHD mortality in those consuming lower and higher NSP density diets

5.3.8 Confounder adjustment

Confounders are generally considered to be causal determinants of outcomes and are also associated with exposure (Wang, 2002). The concept of using causal diagrams to explore confounders has been formalised with the use of directed acyclic graphs (DAGs) and arguments supporting this approach are detailed by Greenland *et al.* (Greenland et al., 1999). DAGs allow the user to identify which potential confounders to adjust for in models to help limit over- or under-adjustment (Greenland et al., 1999).

Selection of confounding variables to adjust for in models was undertaken in the first instance using a DAG to identify the minimal number of adjustments needed, to avoid issues of over adjustment. Firstly, a diagram was drawn for all causal paths between exposure, confounders and outcome, then arrows originating from the exposure (fibre) were removed. Unblocked backdoor paths were then identified. These 'unblocked' paths are defined as routes between fibre and CVD using arrows of any direction that do not incorporate 'colliders'. Colliders are considered to 'block' a path and exist where the arrows of a certain path both point towards each other at a variable on the path.

These principles are detailed using the simplified DAG example below (Figure 5.3), unblocked backdoor paths exists between CVD and fibre via two confounder paths [CVD←C←B→Fibre] [CVD←C←A→Fibre], but not another [CVD ←E→D←B→Fibre], as arrows collide at 'D' on this path. Using this example, A, B and C would be adjusted for in the model but D and E would not.

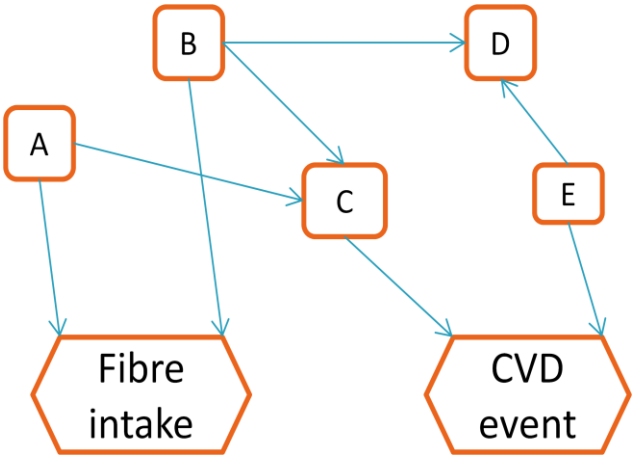


Figure 5.3 Example of simplified directed acyclic graph

A causal diagram for fibre, CVD and associated potential confounders has been generated for the UKWCS so the method of using DAGs can be applied here (Figure 5.4). In principle, it may only be necessary to adjust for a few confounders in a system or for one variable within an unblocked backdoor path because confounders on one path are causally related and adjustment for just one represents the minimal sufficiency set (Greenland et al., 1999). However, applying this principle here would mean adjusting only for SES, family history and age as all paths run via these three confounders. A problem with categorising participants into just three groups based on SES is that this simple classification is not likely to accurately capture the complexity of the relationship between many lifestyle characteristics and fibre or CVD and other confounding variables were therefore also used as adjustments as they existed on separate backdoor paths.

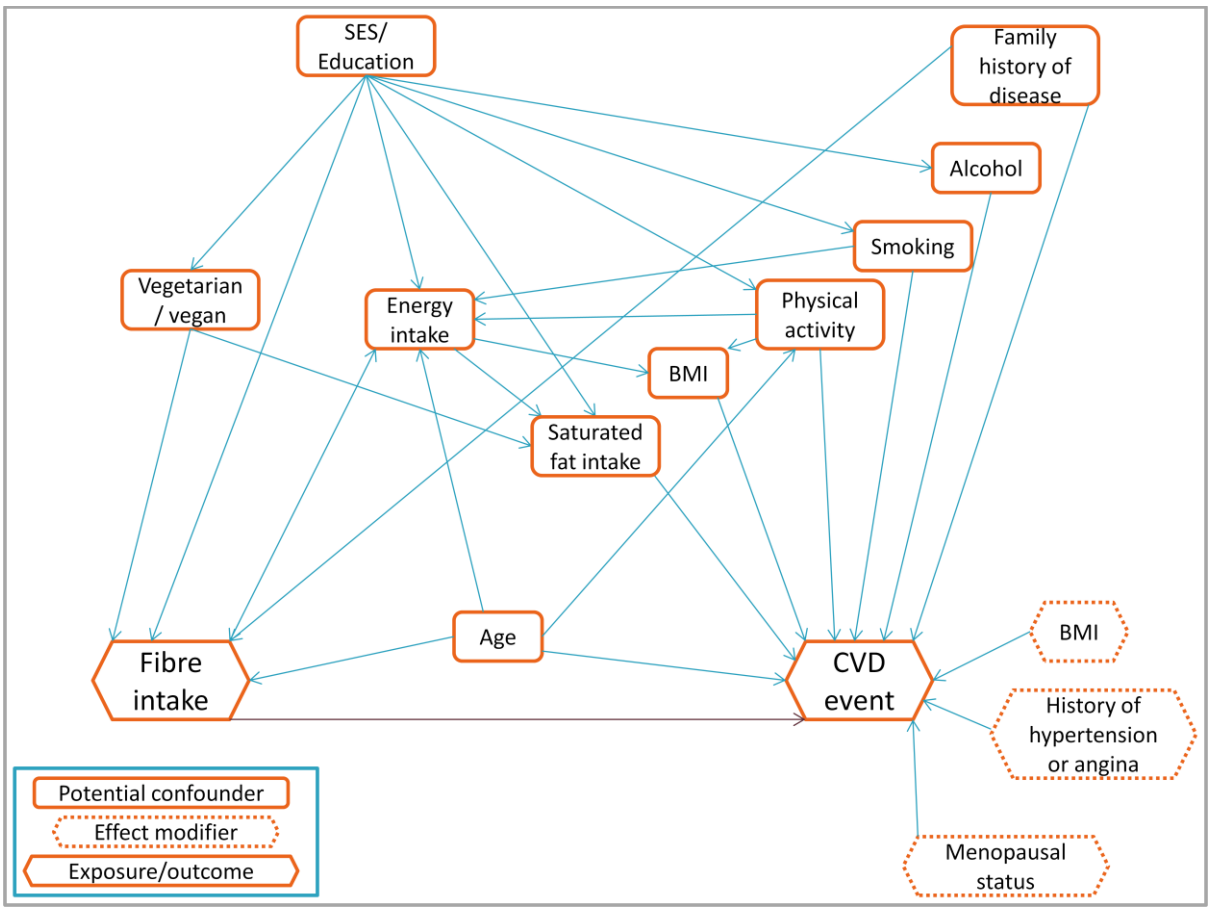


Figure 5.4 Causal diagram for the variables associated with fibre and CVD

Correlations between different potential confounders (continuous variables) were explored in order to identify collinear associations and avoid over-adjustment in models (Table 5.1). These

tests revealed that the selected variables were not so closely associated to preclude them from being used as adjustments in the same model.

Table 5.1 Correlation between continuous variables

	Age	Saturated fat	BMI	Physical activity	Ethanol	Fibre
Energy	0.00	0.76	-0.01	0.17	-0.03	0.68
Age		0.01	0.14	-0.04	-0.10	0.02
Saturated Fat			0.01	0.12	0.02	0.22
BMI				-0.04	-0.06	-0.08
Physical activity					-0.01	0.15
Ethanol						-0.06

Notes on selection of potential confounders:

- Vegetarian status was not included as a confounder since models had been weighted on this factor.
- An ANOVA test revealed significant differences ($p<0.001$) in the distribution of participants across categories for SES and education. As education data were less complete for this sample, SES was selected as the adjustment within in the model.
- Saturated fat intake correlated highly with energy intake (0.76) and this was therefore not included in models to avoid collinearity.
- Menopausal status was not included as an adjustment as it is functionally related to age.
- Ethnicity was not considered for adjustment because 99% of the cohort was classified as white and ethnicity did not appear to be related to fibre intake in the sample, likely because of the relatively small number of non-white participants.
- History of hypertension or angina was not used as an adjustment in models as both sit on the causal pathway for the development of CVD.
- BMI is considered both as a potential confounder and effect modifier in Figure 4.4 because of the many potential mechanisms through which fibre may affect CVD risk. If protective effects of fibre are mediated through body weight changes then this exists on the causal pathway and need not be adjusted for. If another mechanism is in play then BMI should be included as a potential confounder as BMI will dictate energy intake (and thus fibre) and is also causally related to CVD risk. It is also known that BMI is independently related to CVD risk and it may therefore modify any effect of fibre on CVD risk.
- Family history of disease data were not available for use.

The following three levels of adjustment were applied:

1) Age (years)
2) Age (years), alcohol (ethanol g/day), smoking status (non-smoker, current-smoker, ex-smoker), physical activity-metabolic equivalents (MET-hours/week) and SES (professional/managerial, intermediate or routine/manual).
3) As model 2 with the addition of energy intake (kcal/day)* and BMI (kg/m^2).

* By way of sensitivity analysis, when modelling CVD risk associated with fibre density of the diet, energy intake was additionally excluded in separate models, as suggested for nutrient density analyses by Willett (2013b). Results were not appreciably different with or without adjustment for energy intake in fibre density analyses (data not shown) and the results presented here are without adjustment for energy intake.

The intermediate model adjusted for potential confounders except for energy intake and BMI as the action of fibre on satiety, energy intake and ultimately BMI is one plausible mechanism for the action of fibre and adjustment for this could therefore mask potential effects. The inclusion of BMI and energy intake as adjustments did not greatly alter risk estimates. For brevity, results from model 2 are therefore not included in tables, as they offer little extra information but are discussed where relevant.

5.3.9 Subgroup analyses

Subgroup analyses were conducted for potential effect modifiers, where a biologically plausible mechanism exists for the different effect of fibre on CVD within these subgroups.

1) Menopausal status was explored through subgroup analyses because of its proposed independent association with CVD risk, possibly via influencing lipid changes (Matthews et al., 2009). Too few cases existed in the pre-menopausal women to allow analysis with this sub-group. Menopausal status groups were derived in a previous study (Cade et al., 2007) and briefly, women were classified as either pre-menopausal, post-menopausal or 'not applicable', which included women who were pregnant, taking contraceptive pills or using hormone replacement therapy.
2) BMI was explored through subgroup analyses since greater BMI has been independently associated with CHD risk (Nordestgaard et al., 2012, Logue et al., 2011) and may modify the effect of fibre. World Health Organisation cut-points were applied to group

participants on BMI; underweight BMI <18.5, healthy weight BMI 18.5-24.9, overweight BMI 25.0-29.9, obese BMI ≥30 kg/m^2. Overweight and obese participants (BMI ≥25kg/m^2) were grouped due to insufficient case numbers in either group separately. Subgroup analyses were therefore conducted in women who reported a healthy BMI at baseline and those who were classed as overweight or obese.

3) Personal history of hypertension or angina (cardiovascular event risk factors) were also considered potential effect modifiers as they may interact with any protective association of fibre on cardiovascular health. Analyses were conducted only on those not reporting history of angina or hypertension at baseline as there were insufficient event cases to run models for those with a history of these conditions.

5.3.10 Statistical significance

For primary analyses (full sample) a 2-sided p-value ≤0.05 was considered statistically significant but to acknowledge issues with multiple testing, for subgroup analyses the accepted p-value was reduced to ≤0.01, thus reducing the probability of observing false positive results.

5.4 Results

5.4.1 Descriptive statistics

After applying exclusions, 31,036 women remained from 35,690. Table 5.2 shows the characteristics and cardiovascular risk factors for those women who at baseline had reported a personal history of stroke or heart attacks, and therefore were excluded from the analysis and data for those women who subsequently developed CVD during the study period. As expected, the non-cases were younger, with a median age of 50 years (IQR 14), compared to 67 years (IQR 11) and 66 years (IQR 11) for the stroke and IHD cases, respectively (p<0.001 for comparisons of cases to non-cases). Unsurprisingly, menopausal status also was significantly different for cases and non-cases with 78% of cases being classified as postmenopausal compared to 35% of non-cases (p<0.001); a reflection of greater age in the case group.

BMI was lower in the non-cases compared to IHD cases (23.6 vs. 24.6 kg/m^2, p<0.001) but BMI was not significantly different for stroke cases and non-cases (24.1 vs. 23.6 kg/m^2, p=0.98). The proportion of women grouped by each SES category was also not different when comparing stroke cases and non-cases but differed significantly for IHD and CVD cases compared to non-cases. Educational achievement was significantly different for cases and non-cases for all disease comparisons, with a greater proportion of women grouped in the lower

educational achievement groups for stroke and IHD cases (36 and 37%) compared to non-cases (15%).

Ethanol intake was markedly lower for the IHD cases, 1.2g/day (IQR 7.2) compared to stroke cases, 5.0g/day (IQR 12.2) and the non-cases 5.5g/day (IQR 11.6). Smoking status classification was also different among cases and non-cases. Many other lifestyle characteristics however did not differ among cases and non-cases. Physical activity was 14.3 MET-hours/week (IQR 13.9) for IHD cases and 12.3 (IQR 12.6) in stroke cases and was similar in non-cases, 14.5 (IQR 13.2). Dietary characteristics such as energy intake, saturated fat intake and total dietary fibre also did not differ among cases and non-cases. The exceptions to this were fibre from breakfast cereals for IHD cases (2.1g/day) vs. non-cases (1.8g/day) ($p=0.05$) and fibre from legumes which was significantly different in IHD cases (0.8g/day) and stroke cases (0.9g/day) compared to non-cases (1.1g/day), $p<0.001$.

Table 5.3 shows how characteristics differ across increasing categories of fibre (NSP) intake. Age appears to vary little with increasing fibre intake but an ANOVA test indicated significant age differences among fibre quintiles ($p<0.01$). BMI clearly decreases across increasing categories ($p<0.001$), whilst physical activity and energy intake both increase across increasing fibre intake categories ($p<0.001$).

Unsurprisingly, fewer meat-eaters and more vegetarians were categorised in the higher fibre intake groups and a Chi^2 test revealed significant differences among groups ($p<0.001$). More smokers were classed within the lower intake groups, a difference that was also significant among the categories ($p<0.001$). Additionally, the education and socio-economic profile improved with increased levels of fibre intake.

Saturated fat intake was not significantly different in cases or non-cases but did increase with increasing fibre intake quintiles, the difference between quintiles was statistically significant ($p<0.001$). The increasing saturated fat intake level seen with increasing fibre intake, likely reflects greater overall consumption of food across increasing fibre intake groups.

Table 5.2 Baseline cross-sectional characteristics in those excluded because of history of stroke and heart-attacks, those subsequently suffering a fatal stoke or IHD and those women still alive at censor date or whose cause of death was not attributed to CVD

		History of stroke or IHD (excluded from analysis)	Fatal Stroke	Fatal IHD	Non-cases (No fatal stroke or IHD)	P-value stroke vs. non stroke cases†	P-value IHD vs. non IHD cases†	P-value CVD vs. non CVD cases†
N		565	130	128	30778			
Age, years		62.6(12.4)	67.3(11.4)	65.8(10.5)	50.3 (14.1)	<0.001	<0.001	<0.001
BMI, kg/m^2		25.5 (5.7)	24.1 (5.1)	24.6 (6.2)	23.6 (4.7)	0.98	<0.001	0.01
Smoking status (%)	Current	68 (13)	22 (18)	24 (20)	3228 (11)	0.01	<0.01	<0.01
	Former	222 (40)	27 (22)	43 (35)	9123 (30)			
	Never smoked	274 (47)	76 (60)	56 (45)	17545 (59)			
Diet group (%)	Meat-eaters	437 (77)	99 (76)	98 (77)	20478 (67)	0.10	0.07	<0.01
	Fish-eaters	61 (11)	12 (9)	13 (10)	3938 (13)			
	Vegetarian	67 (12)	19 (14)	17 (13)	6363 (21)			
Socio-economic status NS-SEC (%)	Professional/ managerial	316 (58)	71 (58)	63 (52)	19214 (64)	0.45	0.01	0.01
	Intermediate	161 (30)	38 (31)	48 (40)	8198 (27)			
	Routine and manual	66 (12)	13 (11)	10 (8)	2722 (9)			
Highest educational achieve-ment (%)	No formal record	166 (33)	37 (36)	38 (37)	4469 (15)	<0.001	<0.001	<0.001
	O-level	122 (25)	20 (20)	25 (24)	8909 (32)			
	A-level	115 (23)	22 (22)	17 (16)	7039 (25)			
	Degree	93 (19)	23 (22)	24 (23)	7834 (28)			
Menopause status (%)	Postmenopause	384 (70)	100 (78)	94 (78)	10657 (35)	<0.001	<0.001	<0.001
	Premenopause	64 (12)	9 (7)	6 (5)	12808 (42)			
	Not applicable‡	99 (18)	19 (15)	21(17)	6838 (23)			
History of angina at baseline (%)	Yes	162 (37)	6 (5)	11 (10)	374 (1)	<0.001	<0.001	<0.001
	No	277 (63)	104 (95)	99 (90)	28421 (99)			
History of hypertension at baseline (%)	Yes	245 (52)	48 (41)	46 (39)	4596 (16)	<0.001	<0.001	<0.001
	No	226 (48)	69 (59)	73 (61)	24768 (84)			
Ethanol g/day		2.1 (10.0)	5.0 (12.2)	1.2 (7.2)	5.5 (11.6)	0.39	<0.001	<0.01
Physical activity, MET-hrs/wk		14.4(14.3)	14.3(13.9)	12.3(12.6)	14.5 (13.2)	0.54	0.27	0.22
Energy intake, kcal/day		2202 (952)	2175 (1010)	2215 (864)	2187 (863)	0.52	0.40	0.29
Saturated fat intake g/day		26.9(16.3)	28.2(15.7)	26.7(16.4)	27.0 (16.0)	0.66	0.72	0.57
NSP, g/day		24.3(13.3)	23.0(14.5)	23.3(12.1)	23.8 (12.4)	0.30	0.39	0.18
NSP density, g/1000kcal/day		11.1 (4.3)	11.1 (4.6)	11.0 (5.0)	11.0 (4.2)	0.64	0.86	0.65
AOAC fibre, g/day		37.6(20.4)	36.1(22.0)	34.8(17.1)	36.7 (18.9)	0.36	0.37	0.20
AOAC fibre density, g/1000kcal/day		17.2 (6.6)	17.1 (6.9)	16.3 (6.9)	16.9 (6.2)	0.68	0.79	0.63
Soluble fibre, g/day		10.6 (5.7)	9.6 (5.6)	9.6 (4.7)	10.4 (5.0)	0.28	0.14	0.07
Insoluble fibre, g/day		15.6 (9.5)	14.9(10.7)	14.4 (9.0)	15.3 (8.9)	0.36	0.55	0.28
NSP within Foods, g/day	Total fruit	4.5 (4.4)	4.2 (4.8)	3.9 (3.5)	4.2 (3.9)	0.17	0.60	0.55
	Vegetables	5.2 (4.7)	4.5 (3.6)	5.0 (4.2)	4.9 (3.7)	0.40	0.77	0.42
	Total cereal foods	7.8 (7.3)	7.2 (7.0)	7.5 (7.3)	7.6 (7.0)	0.82	0.38	0.65
	Breakfast cereals	1.9 (4.5)	2.1 (3.3)	1.7 (3.8)	1.8 (3.6)	0.84	0.0	0.12
	Nuts & Seeds	0.07(0.26)	0.07(0.17)	0.06(0.22)	0.08 (0.29)	0.12	0.20	0.05
	Legumes	1.1 (1.2)	0.8 (1.0)	0.9 (0.9)	1.1 (1.3)	0.01	0.01	<0.001

Values are median (interquartile range) or frequency (percent)

† p values were generated using χ^2 for categorical variables and t-tests for continuous variables and the comparisons are between specific case types and all other participants except those with history of stroke or IHD

‡ Pregnant, taking the contraceptive pill or hormone replacement therapy

Table 5.3 Dietary and lifestyle characteristics across increasing NSP quintiles, after applying exclusions to the sample

		1st fifth	2nd fifth	3rd fifth	4th fifth	5th fifth	p-value†
N		6207	6207	6207	6207	6208	
NSP g/day		14.2 (3.9)	19.5 (2.3)	23.8 (2.3)	29.1 (3.1)	38.3 (8.6)	
NSP density g/1000kcal/day		8.2 (2.9)	10.0 (3.0)	11.0 (3.2)	12.1 (3.4)	13.8 (3.8)	
AOAC fibre g/day		21.9 (6.0)	30.1 (3.9)	36.7 (4.2)	44.6 (5.3)	58.8 (13.8)	
Soluble fibre g/day		6.5 (1.9)	8.8 (1.6)	10.5 (1.9)	12.5 (2.2)	16.3 (4.3)	
Insoluble fibre g/day		8.4 (2.7)	12.2 (2.0)	15.3(14.3)	19.0 (2.6)	25.6 (6.2)	
Age, years		50.0(13.8)	50.2(14.1)	50.6(14.0)	50.1(14.4)	50.7 (14.5)	0.005
BMI, kg/m^2		24.0 (5.1)	23.8 (4.8)	23.6 (4.5)	23.3 (4.5)	23.1 (4.3)	<0.001
Hypertension at baseline %		19	18	17	18	18	0.356
Angina at baseline %		1.3	1.1	1.1	1.3	1.5	0.266
Frequency of deaths	Stroke	37	21	21	22	24	0.156
	IHD	26	30	20	26	17	0.566
	CVD (stroke+IHD)	63	51	41	48	41	0.250
Smoking status (%)	Current smoker	1075 (18)	701 (12)	545 (9)	489 (8)	464 (8)	<0.001
	Former smoker	1751 (29)	1831 (30)	1843 (30)	1904 (31)	1864 (31)	
	Never smoker	3189 (53)	3507 (58)	3658 (61)	3652 (61)	3671 (61)	
Diet group (%)	Meat-eaters	4959 (80)	4525 (73)	4281 (69)	3718 (60)	3192 (51)	<0.001
	Fish-eaters	429 (7)	629 (10)	760 (12)	949 (15)	1196 (19)	
	Vegetarian	819 (13)	1053 (17)	1167 (19)	1540 (25)	1820 (29)	
Socio-economic status NSSEC %	Professional/ managerial	3711 (61)	3784 (62)	3889 (64)	3945 (65)	4019 (66)	<0.001
	Intermediate	1759 (29)	1708 (28)	1672 (27)	1607 (27)	1538 (25)	
	Routine and manual	589 (10)	597 (10)	525 (9)	514 (8)	520 (9)	
Highest educational achieve-ment %	No formal record	1096 (19)	911 (16)	862 (15)	791 (14)	884 (16)	<0.001
	O-level	1925 (34)	1788 (31)	1821 (32)	1715 (30)	1705 (30)	
	A-level	1283 (23)	1397 (25)	1436 (25)	1516 (27)	1446 (25)	
	Degree	1367 (24)	1566 (28)	1602 (28)	1690 (29)	1656 (29)	
Menopause status %	Postmenopausal	2123 (35)	2094 (34)	2198 (36)	2172 (36)	2264 (37)	0.019
	Premenopausal	2543 (42)	2653 (43)	2534 (41)	2598 (42)	2495 (41)	
	Not-applicable‡	1432 (23)	1375 (23)	1391 (23)	1338 (22)	1342 (22)	
Ethanol g/day		5.7 (13.1)	5.6 (12.2)	5.8 (11.6)	5.4 (11.0)	4.8 (10.7)	<0.001
Physical activity, MET-hrs/wk		12.3(12.2)	13.6(12.2)	14.4(12.6)	15.2(12.8)	17.0 (14.8)	<0.001
Energy intake, kcal/day		1642(573)	1946(578)	2173(623)	2407(682)	2877 (873)	<0.001
Saturated fat intake g/day		23.2(13.7)	25.5(14.8)	27.3(15.4)	28.5(16.1)	30.9 (18.1)	<0.001

Values are median (interquartile range) or frequency (percent)

† p values were generated using χ^2 for categorical variables and t-tests for continuous variables

‡ Pregnant, taking the contraceptive pill or hormone replacement therapy

Case numbers

At follow up, a total of 258 deaths attributable to CVD were observed. Nine of the 128 IHD cases and five of the 130 stroke cases had missing BMI data and SES data were missing from an additional 6 IHD and 8 stroke cases. Fully adjusted models therefore included 113 IHD and 117 stroke cases. Subgroup analyses included fewer cases and the numbers included in fully-adjusted models are listed in Table 5.4.

Table 5.4 IHD, stroke and CVD case numbers in full sample and subgroup analyses

	Sample number	IHD cases in fully-adjusted model	Stroke cases in fully-adjusted model	CVD cases in fully-adjusted model
Full sample	31036	113	117	230
Postmenopausal women	10851	85	89	174
No history of hypertension or angina	26143	67	69	136
BMI 18.5-24.9kg/m^2	19042	55	65	120
BMI ≥25kg/m^2	11331	55	45	100

5.4.2 Survival analysis

In total, 31,036 women free from personal history of stroke or heart attacks were followed for a median of 14.3 years (IQR 1.4). The cause of death was attributed to stroke in 130 participants and 128 fatal IHD cases were observed.

HRs and 95% CIs for IHD, stroke and CVD mortality in relation to increasing fifths of the various fibre exposures are presented along with HRs for the linear dose-response associations in age-adjusted and fully adjusted models (Table 5.5). Results from intermediate models (without adjustment for BMI and energy intake) have not been presented as they did not appreciably differ from age or fully-adjusted models. For example, HRs for the age-adjusted, intermediate and fully-adjusted models for fatal CVD risk and each 6g/day higher total fibre intake were respectively, 0.93 (95% CI: 0.85 to 1.02) p=0.13, 0.96 (95% CI: 0.88 to 1.06) p=0.45 and 0.92 (95% CI: 0.80 to 1.05) p=0.20.

Comparisons were also made between lower or higher intake groups using a middle reference category, rather than comparing all levels to the lowest intake category, as the lowest consumers may differ in other characteristics not accounted for through adjustments (results not displayed in tables).

Full sample results-overview

For analyses including the full sample, no statistically significant results were observed for models using the continuous exposures, indicating that there is no evidence of a linear dose-response relationship between the fibre exposures and risk of fatal IHD, stroke or CVD (Table 5.5). The risk estimates are all less than 1, which would indicate an inverse association, but CIs span 1 in each model.

The models examining quintiles of fibre intake and risk did indicate inverse relationships but only for one or two quintiles compared to the lowest intake group. In most cases, statistical significance in age-adjusted models was attenuated in the multivariate models. In multivariate models displaying some statistically significant results, significance was often seen in just one or two of the four intake comparisons with the lowest quintile, despite HRs all being indicative of a protective association (lower than 1). The results for stroke risk and fibre from nuts and seeds depict this well, HRs and 95% CIs for comparisons with Q1 were as follows: Q2 0.86 (0.52 to 1.43) p=0.57; Q3 0.63 (0.35 to 1.13) p=0.12; Q4 0.50 (0.26 to 0.95) p=0.03; Q5 0.45 (0.23 to 0.85) p=0.01. Here, risk estimates decrease with greater intake levels, but statistical significance was reached in only two of the four comparisons. This likely explains the close but non-significant result for the continuous exposure risk estimate: risk of fatal stroke was reduced by 8% per 0.2g/day increase in fibre from nuts and seeds but this was not statistically significant 0.92 (0.83 to 1.02) p=0.13 (Table 5.5).

Modelling risk in high and low intakes compared to the mid-intake group

Modelling risk in quintile 1, 2, 4 and 5 compared to the middle category (Q3) on the whole provided little extra insight into the association between fibre from different sources as none of the results proved statistically significant. For fatal IHD risk and total fibre intake, no associations were seen in those consuming high or low intakes compared to the middle-intake category.

Stroke risk was greater in low fibre (NSP or AOAC) consumers compared to the middle intake group, for the lowest NSP group, HR 2.10 (95% CI: 1.08 to 4.11) p=0.03 and lowest AOAC group HR 2.11 (95% CI: 1.11 to 3.99) p=0.02, compared to Q3.

Risk of fatal CVD was 62% higher in comparisons for both the lowest intake groups for AOAC fibre density and insoluble fibre compared to the middle intake group, AOAC density HR 1.62

(95% CI: 1.02 to 2.56) p=0.04; Insoluble fibre HR 1.62 (95% CI: 1.02 to 2.58) p=0.04 and also for Q2 vs. Q3 of insoluble fibre intake HR 1.60 (95% CI: 1.01 to 2.53) p=0.05.

Taken together, these findings indicate lower fibre density or insoluble fibre are associated with greater fatal CVD risk and lower total fibre is associated with stroke risk increase. However, there did not appear to be a clear risk trend across increasing quintiles of fibre intakes except for this greater risk in the lowest intake level.

Total NSP and AOAC fibre intake and fibre density

No apparent association was seen for total dietary fibre, assessed either as NSP or AOAC and IHD risk. For stroke events, estimates indicated roughly 50% reduction in risk across many of the intake categories compared to the lowest intake level. Those consuming approximately 45g/day AOAC fibre compared to 21g/day saw a 53% risk reduction HR 0.47 (95% CI: 0.25 to 0.90) p=0.02. This association was not observed in Q2 or Q5 compared to Q1, which likely explains why the examination of continuous fibre intake did not indicate a protective association.

As noted above, stroke risk was significantly increased in the lowest total fibre intake categories compared to mid-intake levels.

There did not appear to be an association between fibre density (assessed either as NSP or AOAC) and risk of fatal IHD, stroke or CVD.

Soluble and insoluble fibre

Although the majority of risk estimates were on the side of indicating a protective association, CIs were generally wide and no significant associations were observed for soluble or insoluble fibre intake and risk of fatal IHD, stroke or CVD, in the full sample.

Fibre from food sources

Similar to total fibre intake, no specific food sources of fibre were significantly associated with fatal IHD, stroke or CVD risk in fully adjusted models using the full sample. One or two quintile comparisons were statistically significant, but this did not carry through to the linear dose-response test. Stroke risk was reduced by 55% for the highest group compared to the lowest intake of fibre from nuts and seeds HR 0.45 (95% CI: 0.23 to 0.85) p=0.01, but for every

0.2g/day increase in intake, the association was not statistically significant HR 0.92 (95% CI: 0.83 to 1.02) p=0.13.

Subgroup analyses

No significant associations were observed with any of the fibre exposures and risk of IHD, stroke or CVD in the sample of postmenopausal women or those with BMI within the healthy range at baseline (Table 5.6 and 5.7).

In women whose baseline BMI was ≥25kg/m^2 (n=11,331) 45 stroke cases were observed. Risk of fatal stroke appeared to be significantly reduced with greater intake of fibre from nuts and seeds (Table 5.8). With every 0.2g/day increase, risk was reduced by 32%, HR 0.68 (95% CI: 0.48 to 0.98) p=0.04, but this result did not reach the 1% pre-specified significance criterion. However, stroke risk was significantly reduced with greater cereal fibre intake in women whose baseline BMI was ≥25kg/m^2 HR 0.80 (95% CI: 0.65 to 0.93) p<0.01. No other notable associations were seen in this subgroup for risk of IHD or CVD.

In a healthy sub-sample of women, who were free of hypertension or angina at baseline (n=26,143) 69 stroke cases were observed and a protective association was seen for both NSP and AOAC fibre density and stroke risk (Table 5.9). For every 2g/1000kcal/day increase in NSP fibre, risk was reduced by 17%, HR 0.83 (95% CI: 0.70 to 0.99) p=0.04 and by 18% for every 3g/1000kcal/day increase in AOAC fibre, HR 0.82 (95% CI: 0.68 to 0.99) p=0.04. However, these results did not reach the pre-specified 1% significance level and so must be interpreted with caution because of the greater chance for type I error or false positive findings. Surprisingly, in this sub-group, greater risk of fatal IHD was associated with increasing fibre from total cereal foods, although this 15% risk increase was not significant at the 1% level; risk was 1.15 (95% CI: 1.00 to 1.31) p=0.05, for each 3g/day increase in fibre from cereal foods.

Table 5.5 Cardiovascular mortality risk assessed using both categorised fibre intake and by fitting a linear dose-response trend, using fibre as a continuous variable

		Median intake (IQR)	CHD HR (95% CI) p-value			Stroke HR (95% CI) p-value			Total CVD HR (95% CI) p-value		
			Cases[1]	Age-adjusted	Fully-adjusted[2]	Cases[1]	Age-adjusted	Fully-adjusted[2]	Cases[1]	Age-adjusted	Fully-adjusted[2]
NSP (g/day)	Q1	14.1 (3.9)	24	1	1	34	1	1	58	1	1
	Q2	19.4 (2.3)	28	0.99 (0.58, 1.68)	1.30 (0.76, 2.24)	21	0.61 (0.35, 1.04)	0.63 (0.35, 1.12)	49	0.78 (0.54, 1.14)	0.90 (0.61, 1.34)
	Q3	23.8 (2.3)	18	0.75 (0.42, 1.33)	0.93 (0.48, 1.82)	19	0.48 (0.27, 0.86)	0.48 (0.24, 0.93)	37	0.60 (0.40, 0.90)	0.66 (0.41, 1.06)
	Q4	29.1 (3.1)	26	0.84 (0.49, 1.45)	1.15 (0.61, 2.17)	21	0.57 (0.33, 0.99)	0.55 (0.39, 1.04)	47	0.69 (0.47, 1.02)	0.79 (0.51, 1.24)
	Q5	38.3 (8.6)	17	0.63 (0.34, 1.15)	0.89 (0.38, 2.08)	22	0.74 (0.44, 1.25)	0.61 (0.27, 1.36)	39	0.69 (0.47, 1.03)	0.74 (0.41, 1.33)
	Per 6g/day		113	0.92 (0.80, 1.05)	0.96 (0.79, 1.17)	117	0.95 (0.84, 1.07)	0.87 (0.73, 1.04)	230	0.93 (0.85, 1.02)	0.92 (0.80, 1.05)
	pTrend			*0.22*	*0.69*		*0.37*	*0.13*		0.13	*0.20*
AOAC (g/day)	Q1	21.0 (5.9)	27	1	1	35	1	1	62	1	1
	Q2	30.0 (3.4)	26	0.89 (0.52, 1.50)	1.18 (0.68, 2.03)	20	0.57 (0.33, 0.99)	0.59 (0.32, 1.06)	46	0.72 (0.49, 1.05)	0.84 (0.56, 1.25)
	Q3	36.8 (3.5)	18	0.72 (0.42, 1.26)	0.82 (0.43, 1.58)	20	0.50 (0.28, 0.87)	0.47 (0.25, 0.90)	38	0.60 (0.41, 0.89)	0.62 (0.39, 0.98)
	Q4	44.8 (4.8)	26	0.76 (0.44, 1.29)	0.99 (0.52, 1.86)	21	0.58 (0.34, 1.00)	0.54 (0.29, 1.00)	47	0.66 (0.45, 0.97)	0.72 (0.46, 1.13)
	Q5	63.0 (13.5)	16	0.56 (0.30, 1.02)	0.72 (0.28, 1.83)	21	0.69 (0.40, 1.16)	0.51 (0.21, 1.26)	37	0.63 (0.42, 0.93)	0.61 (0.32, 1.17)
	Per 11g/day		113	0.90 (0.76, 1.07)	0.96 (0.73, 1.26)	117	0.94 (0.81, 1.09)	0.86 (0.68, 1.08)	230	0.92 (0.82, 1.03)	0.91 (0.76, 1.08)
	pTrend			*0.23*	*0.76*		*0.42*	*0.19*		0.15	*0.28*
NSP density g/1000 kcal/day	Q1	7.4 (1.5)	29	1	1	27	1	1	56	1	1
	Q2	9.4 (0.8)	20	0.66 (0.38, 1.16)	0.85 (0.47, 1.52)	23	0.75 (0.43, 1.30)	0.81 (0.46, 1.44)	43	0.70 (0.47, 1.04)	0.83 (0.55, 1.24)
	Q3	11.0 (0.8)	22	0.72 (0.41, 1.25)	0.91 (0.51, 1.63)	19	0.64 (0.36, 1.15)	0.71 (0.39, 1.29)	41	0.68 (0.46, 1.02)	0.80 (0.53, 1.22)
	Q4	12.7 (1.0)	19	0.70 (0.40, 1.22)	0.74 (0.40, 1.38)	23	0.80 (0.46, 1.39)	0.79 (0.44, 1.44)	42	0.75 (0.51, 1.11)	0.77 (0.50, 1.19)
	Q5	15.4 (2.3)	23	0.84 (0.49, 1.43)	0.99 (0.55, 1.76)	25	0.82 (0.47, 1.43)	0.89 (0.49, 1.62)	48	0.83 (0.57, 1.22)	0.94 (0.62, 1.42)
	Per 2g/1000 kcal/day		113	0.98 (0.85, 1.14)	0.95 (0.86, 1.06)	117	0.92 (0.81, 1.04)	0.92 (0.80, 1.05)	230	0.95 (0.86, 1.05)	0.95 (0.86, 1.06)
	pTrend			*0.81*	*0.89*		*0.17*	*0.21*		0.30	*0.37*
AOAC density g/1000 kcal/day	Q1	11.3 (2.1)	31	1	1	28	1	1	59	1	1
	Q2	14.6 (1.2)	22	0.63 (0.37, 1.09)	0.76 (0.43, 1.33)	20	0.62 (0.25, 1.09)	0.65 (0.36, 1.16)	42	0.62 (0.42, 0.92)	0.70 (0.47, 1.05)
	Q3	16.9 (1.1)	21	0.70 (0.41, 1.19)	0.81 (0.46, 1.44)	21	0.64 (0.36, 1.12)	0.68 (0.38, 1.21)	42	0.67 (0.45, 0.99)	0.74 (0.49, 1.11)
	Q4	19.4 (1.4)	18	0.59 (0.33, 1.03)	0.68 (0.37, 1.25)	25	0.76 (0.44, 1.31)	0.84 (0.47, 1.49)	43	0.67 (0.45, 0.99)	0.76 (0.50, 1.15)
	Q5	24.3 (3.6)	21	0.73 (0.42, 1.25)	0.81 (0.45, 1.47)	23	0.75 (0.43, 1.31)	0.76 (0.42, 1.40)	44	0.74 (0.50, 1.09)	0.79 (0.52, 1.21)
	Per 3g/1000 kcal/day		113	0.98 (0.84, 1.15)	0.99 (0.84, 1.17)	117	0.92 (0.80, 1.05)	0.92 (0.80, 1.06)	230	0.95 (0.86, 1.05)	0.96 (0.86, 1.07)
	pTrend			*0.81*	*0.92*		*0.20*	*0.24*		0.32	*0.42*

		Median intake (IQR)	CHD HR (95% CI) p-value			Stroke HR (95% CI) p-value			Total CVD HR (95% CI) p-value		
			Cases[1]	Age-adjusted	Fully-adjusted[2]	Cases[1]	Age-adjusted	Fully-adjusted[2]	Cases[1]	Age-adjusted	Fully-adjusted[2]
Soluble fibre (g/day)	Q1	6.4 (1.6)	29	1	1	31	1	1	60	1	1
	Q2	8.6 (0.9)	24	0.88 (0.52, 1.47)	0.91 (0.52, 1.61)	27	0.80 (0.47, 1.36)	0.89 (0.50, 1.60)	51	0.84 (0.58, 1.22)	0.90 (0.60, 1.35)
	Q3	10.4 (0.9)	20	0.65 (0.37, 1.14)	0.88 (0.47, 1.62)	15	0.56 (0.31, 1.01)	0.52 (0.26, 1.03)	35	0.60 (0.40, 0.91)	0.68 (0.43, 1.08)
	Q4	12.5 (1.2)	21	0.64 (0.37, 1.11)	0.74 (0.37, 1.49)	26	0.85 (0.50, 1.43)	0.78 (0.40, 1.51)	47	0.74 (0.51, 1.08)	0.76 (0.47, 1.23)
	Q5	16.4 (3.8)	19	0.62 (0.35, 1.10)	0.76 (0.32, 1.80)	18	0.70 (0.40, 1.24)	0.60 (0.25, 1.43)	37	0.66 (0.44, 0.99)	0.68 (0.37, 1.25)
	Per 3g/day		113	0.87 (0.74, 1.03)	0.91 (0.69, 1.19)	117	0.95 (0.83, 1.10)	0.88 (0.70, 1.11)	230	0.91 (0.82, 1.02)	0.89 (0.75, 1.07)
	pTrend			*0.12*	*0.47*		*0.52*	*0.29*		0.11	*0.22*
Insoluble fibre (g/day)	Q1	8.4 (2.6)	26	1	1	32	1	1	58	1	1
	Q2	12.4 (1.6)	29	1.00 (0.60, 1.68)	1.29 (0.76, 2.19)	21	0.68 (0.39, 1.17)	0.73 (0.41, 1.32)	50	0.83 (0.57, 1.21)	0.98 (0.66, 1.46)
	Q3	15.3 (1.6)	15	0.64 (0.36, 1.14)	0.67 (0.35, 1.31)	19	0.56 (0.32, 1.00)	0.57 (0.30, 1.07)	34	0.60 (0.40, 0.90)	0.62 (0.39, 0.98)
	Q4	19.1 (2.2)	27	0.80 (0.47, 1.36)	1.09 (0.58, 2.06)	22	0.63 (0.36, 1.10)	0.69 (0.37, 1.28)	49	0.71 (0.49, 1.04)	0.87 (0.56, 1.36)
	Q5	25.6 (6.0)	16	0.57 (0.31, 1.05)	0.75 (0.33, 1.71)	23	0.83 (0.49, 1.41)	0.78 (0.37, 1.67)	39	0.71 (0.48, 1.05)	0.78 (0.45, 1.36)
	Per 4g/day		113	0.94 (0.82, 1.08)	1.00 (0.82, 1.20)	117	0.95 (0.85, 1.07)	0.90 (0.78, 1.05)	230	0.95 (0.87, 1.03)	0.95 (0.84, 1.07)
	pTrend			*0.38*	*0.96*		*0.40*	*0.20*		0.23	*0.41*
Total cereal fibre (g/day)	Q1	2.8 (1.4)	25	1	1	27	1	1	52	1	1
	Q2	5.1 (1.1)	18	0.70 (0.39,1.26)	0.87 (0.46, 1.65)	21	0.86 (0.49, 1.53)	0.88 (0.48, 1.62)	39	0.78 (0.52, 1.17)	0.88 (0.56, 1.36)
	Q3	7.6 (1.4)	27	0.98 (0.58, 1.65)	1.11 (0.61, 2.04)	26	0.92 (0.53, 1.59)	0.95 (0.53, 1.71)	53	0.95 (0.65, 1.39)	1.04 (0.68, 1.58)
	Q4	10.7 (1.8)	19	0.59 (0.33, 1.06)	0.72 (0.36, 1.45)	21	0.72 (0.41, 1.28)	0.69 (0.37, 1.26)	40	0.65 (0.43, 0.98)	0.71 (0.44, 1.12)
	Q5	15.7 (4.5)	24	0.75 (0.43, 1.29)	1.06 (0.52, 2.15)	22	0.81 (0.46, 1.41)	0.77 (0.40, 1.50)	46	0.78 (0.52, 1.15)	0.91 (0.56, 1.48)
	Per 3g/day		113	1.00 (0.88, 1.14)	1.03 (0.89, 1.19)	117	0.96 (0.87, 1.07)	0.94 (0.83, 1.07)	230	0.98 (0.90, 1.07)	0.99 (0.89, 1.09)
	pTrend			*0.99*	*0.71*		*0.48*	*0.34*		0.67	*0.77*
Fibre from breakfast cereals (g/day)	Q1	0.05 (0.14)	23	1	1	20	1	1	43	1	1
	Q2	0.5 (0.4)	25	1.20 (0.68, 2.09)	1.11 (0.61, 2.03)	21	1.17 (0.63, 2.17)	1.24 (0.65, 2.35)	46	1.18 (0.78, 1.79)	1.16 (0.75, 1.80)
	Q3	1.8 (0.7)	16	0.75 (0.41, 1.40)	0.82 (0.42, 1.62)	23	1.04 (0.57, 1.92)	1.06 (0.56, 2.02)	39	0.89 (0.57, 1.36)	0.93 (0.80, 1.48)
	Q4	3.5 (0.7)	23	0.93 (0.52, 1.64)	1.04 (0.55, 1.69)	30	1.41 (0.80, 2.50)	1.45 (0.79, 2.68)	53	1.15 (0.77, 1.71)	1.22 (0.79, 1.90)
	Q5	7.6 (2.6)	26	0.91 (0.52, 1.59)	1.09 (0.58, 2.02)	23	1.06 (0.59, 1.91)	1.11 (0.59, 2.10)	49	0.98 (0.65, 1.46)	1.09 (0.70, 1.71)
	Per 2g/day		113	1.04 (0.93, 1.16)	1.04 (0.93, 1.17)	117	0.98 (0.89, 1.07)	0.97 (0.88, 1.07)	230	1.01 (0.93, 1.09)	1.01 (0.93, 1.09)
	pTrend			*0.52*	*0.46*		*0.58*	*0.52*		0.81	*0.83*

		Median intake (IQR)	CHD HR (95% CI) p-value			Stroke HR (95% CI) p-value			Total CVD HR (95% CI) p-value		
			Cases[1]	Age-adjusted	Fully-adjusted[2]	Cases[1]	Age-adjusted	Fully-adjusted[2]	Cases[1]	Age-adjusted	Fully-adjusted[2]
Fruit fibre (g/day)	Q1	1.4 (0.9)	26	1	1	27	1	1	53	1	1
	Q2	2.9 (0.7)	24	0.64 (0.38, 1.08)	0.83 (0.47, 1.47)	21	0.56 (0.32, 1.00)	0.69 (0.39, 1.23)	45	0.60 (0.41, 0.89)	0.76 (0.50, 1.14)
	Q3	4.2 (0.7)	23	0.55 (0.32, 0.95)	0.76 (0.42, 1.38)	25	0.69 (0.40, 1.17)	0.73 (0.41, 1.31)	48	0.61 (0.42, 0.90)	0.74 (0.49, 1.13)
	Q4	5.8 (1.1)	23	0.53 (0.31, 0.90)	0.74 (0.40, 1.36)	18	0.43 (0.23, 0.78)	0.50 (0.25, 0.97)	41	0.48 (0.32, 0.72)	0.61 (0.39, 0.96)
	Q5	9.5 (4.1)	17	0.38 (0.21, 0.69)	0.55 (0.28, 1.06)	26	0.72 (0.43, 1.22)	0.79 (0.42, 1.48)	43	0.54 (0.37, 0.80)	0.68 (0.43, 1.06)
	Per 2g/day		113	0.91 (0.75, 1.10)	0.98 (0.83, 1.17)	117	1.00 (0.90, 1.12)	1.02 (0.90, 1.15)	230	0.96 (0.87, 1.06)	1.00 (0.90, 1.11)
	pTrend			*0.32*	*0.85*		*0.98*	*0.78*		0.42	*0.97*
Vegetable fibre (g/day)	Q1	2.3 (0.9)	27	1	1	27	1	1	54	1	1
	Q2	3.7 (0.6)	23	0.79 (0.46, 1.36)	0.92 (0.52, 1.62)	25	1.11 (0.65, 1.90)	1.10 (0.63, 1.93)	48	0.94 (0.64, 1.37)	1.00 (0.67, 1.50)
	Q3	4.9 (0.7)	20	0.73 (0.42, 1.27)	0.82 (0.45, 1.50)	21	0.68 (0.37, 1.23)	0.76 (0.41, 1.41)	41	0.71 (0.47, 1.06)	0.78 (0.51, 1.20)
	Q4	6.6 (1.0)	17	0.57 (0.32, 1.03)	0.64 (0.35, 1.18)	27	0.90 (0.53, 1.55)	1.00 (0.57, 1.75)	44	0.73 (0.49, 1.08)	0.82 (0.54, 1.23)
	Q5	9.5 (3.0)	26	0.74 (0.44, 1.26)	0.95 (0.52, 1.74)	17	0.69 (0.39, 1.22)	0.58 (0.30, 1.11)	43	0.72 (0.49, 1.06)	0.75 (0.49, 1.17)
	Per 2g/day		113	0.92 (0.81, 1.05)	0.95 (0.83, 1.09)	117	0.94 (0.83, 1.06)	0.90 (0.79, 1.03)	230	0.93 (0.85, 1.01)	0.93 (0.84, 1.02)
	pTrend			*0.21*	*0.50*		*0.30*	*0.11*		0.10	*0.21*
Legume fibre (g/day)	Q1	0.2 (0.2)	25	1	1	31	1	1	56	1	1
	Q2	0.65 (0.20)	29	1.08 (0.65, 1.79)	1.35 (0.77, 2.37)	30	1.06 (0.64, 1.74)	1.11 (0.65, 1.89)	59	1.07 (0.75, 1.52)	1.22 (0.83, 1.79)
	Q3	1.11 (0.18)	28	1.12 (0.66, 1.89)	1.58 (0.90, 2.79)	23	0.93 (0.54, 1.61)	1.08 (0.62, 1.90)	51	1.02 (0.71, 1.49)	1.30 (0.87, 1.94)
	Q4	1.66 (0.39)	14	0.84 (0.46, 1.56)	1.09 (0.54, 2.21)	19	0.89 (0.50, 1.61)	1.05 (0.58, 1.90)	33	0.87 (0.57, 1.33)	1.07 (0.68, 1.69)
	Q5	3.6 (1.4)	17	1.01 (0.55, 1.85)	1.33 (0.65, 2.71)	14	1.01 (0.55, 1.86)	0.79 (0.37, 1.67)	31	1.01 (0.66, 1.55)	1.03 (0.62, 1.72)
	Per 1g/day		113	0.95 (0.82, 1.10)	1.01 (0.87, 1.17)	117	0.97 (0.82, 1.14)	0.87 (0.74, 1.04)	230	0.96 (0.86, 1.07)	0.94 (0.84, 1.06)
	pTrend			*0.47*	*0.92*		*0.73*	*0.12*		0.46	*0.32*
Fibre from nuts and seeds (g/day)	Q1	0 (0.01)	38	1	1	45	1	1	83	1	1
	Q2	0.06 (0.01)	25	0.83 (0.51, 1.36)	0.86 (0.51, 1.45)	26	0.96 (0.60, 1.53)	0.86 (0.52, 1.43)	51	0.89 (0.64, 1.26)	0.86 (0.60, 1.23)
	Q3	0.08 (0.05)	16	0.74 (0.42, 1.30)	0.84 (0.50, 1.53)	18	0.67 (0.38, 1.17)	0.63 (0.35, 1.13)	34	0.70 (0.47, 1.05)	0.72 (0.47, 1.10)
	Q4	0.27 (0.13)	17	0.67 (0.38, 1.21)	0.87 (0.47, 1.61)	14	0.49 (0.29, 0.94)	0.50 (0.26, 0.95)	31	0.58 (0.38, 0.90)	0.66 (0.42, 1.03)
	Q5	0.85 (0.91)	17	0.65 (0.37, 1.15)	0.84 (0.44, 1.59)	14	0.52 (0.29, 0.96)	0.45 (0.23, 0.85)	31	0.58 (0.39, 0.88)	0.61 (0.39, 0.96)
	Per 0.2g/day		113	0.97 (0.89, 1.06)	0.99 (0.90, 1.08)	117	0.93 (0.85, 1.02)	0.92 (0.83, 1.02)	230	0.95 (0.89, 1.02)	0.96 (0.89, 1.03)
	pTrend			*0.51*	*0.76*		*0.13*	*0.13*		0.16	*0.25*

[1]Case numbers apply to fully-adjusted models. [2]Adjustments include Age (years), BMI (kg/m^2), calories from carbohydrate, fat and protein (kcal/day), ethanol intake (g/day), MET (hours/week), smoking status (current vs. not current smoker), socio-economic status (professional or managerial/ intermediate/ routine or manual). Note, adjustment for energy intake was not included in fibre density models. Shading for CIs that do not span 1 in fully adjusted models.

Table 5.6 Risk of fatal IHD, stroke and CVD in postmenopausal women, using continuous fibre variables

	Fibre exposure model increment	IHD HR (95% CI) p-value		Stroke HR (95% CI) p-value		CVD HR (95% CI) p-value	
		Age-adjusted model	Fully-adjusted model	Age-adjusted model	Fully-adjusted model	Age-adjusted model	Fully-adjusted model
Case number		94	85	100	89	194	174
NSP	6 g/day	0.92 (0.80, 1.06) 0.25	1.06 (0.85, 1.31) 0.60	0.98 (0.87, 1.10) 0.73	0.88 (0.73, 1.07) 0.20	0.95 (0.87, 1.04) 0.27	0.96 (0.83, 1.12) 0.63
NSP density	2 g/1000kcal/day	0.97 (0.84, 1.13) 0.70	1.06 (0.90, 1.24) 0.49	0.94 (0.83, 1.07) 0.34	0.92 (0.79, 1.07) 0.26	0.96 (0.87, 1.06) 0.37	0.99 (0.88, 1.11) 0.84
AOAC	11 g/day	0.91 (0.77, 1.07) 0.26	1.09 (0.81, 1.47) 0.58	0.98 (0.85, 1.13) 0.78	0.87 (0.67, 1.14) 0.32	0.94 (0.84, 1.05) 0.30	0.98 (0.80, 1.19) 0.81
AOAC density	3 g/1000kcal/day	0.97 (0.83, 1.14) 0.70	1.06 (0.89, 1.27) 0.49	0.94 (0.82, 1.07) 0.37	0.93 (0.79, 1.09) 0.35	0.96 (0.86, 1.06) 0.39	1.00 (0.88, 1.13) 0.95
Soluble fibre	3 g/day	0.87 (0.74, 1.04) 0.12	0.99 (0.73, 1.36) 0.97	0.99 (0.86, 1.14) 0.89	0.93 (0.72, 1.21) 0.60	0.93 (0.83, 1.04) 0.21	0.96 (0.78, 1.18) 0.70
Insoluble fibre	4 g/day	0.95 (0.83, 1.08) 0.42	1.10 (0.90, 1.34) 0.35	0.98 (0.88, 1.10) 0.75	0.91 (0.76, 1.07) 0.25	0.96 (0.88, 1.05) 0.41	1.00 (0.87, 1.14) 0.99
Fibre from total cereals	3 g/day	1.01 (0.88, 1.14) 0.93	1.10 (0.96, 1.27) 0.18	0.98 (0.88, 1.09) 0.70	0.88 (0.77, 1.01) 0.08	0.99 (0.91, 1.08) 0.87	1.00 (0.89, 1.11) 0.93
Fibre from breakfast cereals	2 g/day	1.04 (0.93, 1.16) 0.51	1.09 (0.99, 1.20) 0.07	0.98 (0.90, 1.07) 0.68	0.94 (0.83, 1.06) 0.32	1.01 (0.94, 1.10) 0.75	1.03 (0.94, 1.13) 0.51
Fibre from fruit	2 g/day	0.91 (0.75, 1.10) 0.35	1.00 (0.81, 1.24) 1.00	1.02 (0.92, 1.13) 0.67	1.06 (0.95, 1.19) 0.31	0.97 (0.88, 1.07) 0.58	1.04 (0.93, 1.16) 0.53
Fibre from vegetables	2 g/day	0.92 (0.81, 1.04) 0.19	0.94 (0.79, 1.12) 0.49	0.96 (0.86, 1.08) 0.53	0.91 (0.78, 1.06) 0.23	0.94 (0.86, 1.03) 0.16	0.92 (0.82, 1.04) 0.18
Fibre from legumes	1 g/day	0.97 (0.84, 1.12) 0.66	1.09 (0.95, 1.26) 0.23	0.98 (0.82, 1.17) 0.82	0.94 (0.78, 1.12) 0.47	0.97 (0.87, 1.09) 0.64	1.02 (0.91, 1.14) 0.75
Fibre from nuts and seeds	0.2 g/day	0.97 (0.89, 1.06) 0.56	1.00 (0.91, 1.09) 0.92	0.94 (0.86, 1.03) 0.18	0.94 (0.85, 1.04) 0.25	0.96 (0.90, 1.02) 0.22	0.97 (0.90, 1.05) 0.43

Table 5.7 Risk of fatal IHD, stroke and CVD in those with healthy BMI at baseline (18.5-24.9kg/m^2), using continuous fibre variables

	Fibre exposure model increment	IHD HR (95% CI) p-value		Stroke HR (95% CI) p-value		CVD HR (95% CI) p-value	
		Age-adjusted model	Fully-adjusted model	Age-adjusted model	Fully-adjusted model	Age-adjusted model	Fully-adjusted model
Case number		58	55	71	65	129	120
NSP	6 g/day	0.98 (0.85, 1.13) 0.81	1.04 (0.84, 1.29) 0.71	0.98 (0.84, 1.15) 0.80	0.98 (0.80, 1.20) 0.83	0.98 (0.88, 1.09) 0.73	1.01 (0.87, 1.17) 0.93
NSP density	2 g/1000kcal/day	1.01 (0.83, 1.22) 0.94	1.07 (0.89, 1.29) 0.46	0.96 (0.81, 1.13) 0.61	0.99 (0.84, 1.17) 0.92	0.98 (0.86, 1.11) 0.75	1.03 (0.91, 1.17) 0.65
AOAC	11 g/day	0.97 (0.82, 1.15) 0.75	1.03 (0.78, 1.36) 0.85	0.98 (0.81, 1.18) 0.80	0.98 (0.75, 1.29) 0.89	0.97 (0.86, 1.11) 0.70	1.00 (0.82, 1.22) 0.98
AOAC density	3 g/1000kcal/day	1.00 (0.81, 1.22) 0.97	1.06 (0.87, 1.29) 0.58	0.95 (0.80, 1.13) 0.59	1.00 (0.84, 1.19) 0.98	0.97 (0.85, 1.11) 0.67	1.03 (0.90, 1.17) 0.71
Soluble fibre	3 g/day	0.96 (0.81, 1.15) 0.67	0.99 (0.73, 1.34) 0.96	0.96 (0.79, 1.17) 0.69	0.94 (0.70, 1.26) 0.68	0.96 (0.84, 1.10) 0.57	0.96 (0.78, 1.19) 0.73
Insoluble fibre	4 g/day	0.99 (0.87, 1.12) 0.84	1.05 (0.87, 1.25) 0.63	0.99 (0.86, 1.14) 0.91	1.01 (0.85, 1.20) 0.94	0.99 (0.90, 1.09) 0.84	1.02 (0.90, 1.16) 0.71
Fibre from total cereals	3 g/day	1.03 (0.93, 1.15) 0.57	1.07 (0.95, 1.20) 0.27	1.01 (0.91, 1.13) 0.80	1.03 (0.91, 1.15) 0.67	1.02 (0.95, 1.11) 0.58	1.05 (0.96, 1.14) 0.30
Fibre from breakfast cereals	2 g/day	0.99 (0.89, 1.11) 0.87	1.02 (0.92, 1.13) 0.77	1.00 (0.91, 1.10) 0.96	1.01 (0.93, 1.10) 0.82	1.00 (0.93, 1.07) 0.94	1.01 (0.95, 1.08) 0.71
Fibre from fruit	2 g/day	0.87 (0.71, 1.06) 0.16	0.91 (0.76, 1.09) 0.30	1.02 (0.89, 1.18) 0.75	1.05 (0.91, 1.21) 0.50	0.97 (0.85, 1.10) 0.62	1.00 (0.88, 1.13) 0.99
Fibre from vegetables	2 g/day	1.02 (0.88, 1.17) 0.84	1.04 (0.90, 1.20) 0.59	0.95 (0.80, 1.13) 0.53	0.93 (0.79, 1.10) 0.40	0.98 (0.87, 1.10) 0.71	0.98 (0.87, 1.11) 0.78
Fibre from legumes	1 g/day	1.05 (0.89, 1.23) 0.57	1.07 (0.91, 1.27) 0.40	0.86 (0.69, 1.09) 0.22	0.85 (0.66, 1.08) 0.18	0.95 (0.83, 1.10) 0.52	0.96 (0.83, 1.12) 0.64
Fibre from nuts and seeds	0.2 g/day	0.97 (0.84, 1.12) 0.66	0.97 (0.85, 1.12) 0.69	0.96 (0.88, 1.05) 0.36	0.96 (0.88, 1.05) 0.36	0.96 (0.89, 1.04) 0.37	0.97 (0.89, 1.05) 0.39

Adjustments: Age (years), BMI (kg/m^2), calories from carbohydrate, fat and protein (kcal/day), ethanol intake (g/day), MET (hours/week), smoking status (current vs. not current smoker), socio-economic status (professional or managerial/ intermediate/ routine or manual). Energy intake was not included in fibre density models.

Table 5.8 Risk of fatal IHD, stroke and CVD in those with BMI classed as overweight or obese at baseline (≥25kg/m^2), using continuous fibre variables

	Fibre exposure model increment	IHD HR (95% CI) p-value		Stroke HR (95% CI) p-value		CVD HR (95% CI) p-value	
		Age-adjusted model	Fully-adjusted model	Age-adjusted model	Fully-adjusted model	Age-adjusted model	Fully-adjusted model
Case number		67	55	50	45	117	100
NSP	6 g/day	0.88 (0.70, 1.12) 0.30	0.90 (0.63, 1.28) 0.56	0.95 (0.78, 1.14) 0.57	0.80 (0.62, 1.04) 0.10	0.91 (0.78, 1.06) 0.23	0.86 (0.68, 1.08) 0.19
NSP density	2 g/1000kcal/day	1.00 (0.80, 1.24) 1.00	0.94 (0.74, 1.19) 0.60	0.91 (0.75, 1.11) 0.35	0.89 (0.72, 1.09) 0.25	0.96 (0.83, 1.12) 0.63	0.92 (0.78, 1.08) 0.34
AOAC	11 g/day	0.87 (0.65, 1.16) 0.33	0.93 (0.58, 1.50) 0.76	0.93 (0.75, 1.17) 0.56	0.76 (0.53, 1.09) 0.14	0.89 (0.74, 1.08) 0.25	0.86 (0.62, 1.17) 0.33
AOAC density	3 g/1000kcal/day	1.01 (0.80, 1.28) 0.94	0.96 (0.74, 1.24) 0.75	0.90 (0.73, 1.11) 0.34	0.88 (0.70, 1.10) 0.26	0.97 (0.82, 1.14) 0.68	0.93 (0.78, 1.11) 0.40
Soluble fibre	3 g/day	0.82 (0.62, 1.09) 0.18	0.88 (0.56, 1.37) 0.57	1.00 (0.82, 1.22) 1.00	0.90 (0.63, 1.27) 0.54	0.90 (0.75, 1.07) 0.24	0.89 (0.67, 1.18) 0.42
Insoluble fibre	4 g/day	0.92 (0.73, 1.17) 0.50	0.96 (0.67, 1.35) 0.80	0.92 (0.76, 1.11) 0.38	0.79 (0.62, 1.01) 0.07	0.92 (0.78, 1.08) 0.31	0.88 (0.70, 1.11) 0.29
Fibre from total cereals	3 g/day	0.97 (0.75, 1.26) 0.83	0.95 (0.69, 1.29) 0.72	0.87 (0.74, 1.03) 0.10	0.80 (0.65, 0.93) <0.01	0.93 (0.78, 1.11) 0.43	0.87 (0.71, 1.07) 0.19
Fibre from breakfast cereals	2 g/day	1.10 (0.90, 1.34) 0.34	1.07 (0.85, 1.35) 0.57	0.91 (0.77, 1.08) 0.29	0.85 (0.69, 1.04) 0.11	1.04 (0.87, 1.24) 0.67	0.99 (0.81, 1.20) 0.90
Fibre from fruit	2 g/day	0.96 (0.74, 1.25) 0.77	1.06 (0.86, 1.31) 0.57	1.00 (0.87, 1.14) 0.95	1.02 (0.88, 1.18) 0.84	0.98 (0.83, 1.14) 0.76	1.04 (0.91, 1.20) 0.55
Fibre from vegetables	2 g/day	0.86 (0.72, 1.04) 0.12	0.90 (0.74, 1.09) 0.28	0.99 (0.84, 1.15) 0.87	0.91 (0.76, 1.11) 0.35	0.92 (0.81, 1.04) 0.18	0,90 (0.79, 1.04) 0.15
Fibre from legumes	1 g/day	0.85 (0.66, 1.10) 0.22	0.94 (0.72, 1.21) 0.62	1.09 (0.89, 1.35) 0.40	0.94 (0.74, 1.20) 0.63	0.97 (0.82, 1.15) 0.72	0.94 (0.79, 1.12) 0.50
Fibre from nuts and seeds	0.2 g/day	1.00 (0.92, 1.08) 1.00	1.01 (0.91, 1.11) 0.91	0.81 (0.59, 1.09) 0.17	0.68 (0.48, 0.98) 0.04	0.96 (0.86, 1.06) 0.41	0.93 (0.81, 1.07) 0.29

Table 5.9 Risk of fatal IHD, stroke and CVD in those without history of hypertension or angina at baseline, using continuous fibre variables

	Fibre exposure model increment	IHD HR (95% CI) p-value		Stroke HR (95% CI) p-value		CVD HR (95% CI) p-value	
		Age-adjusted model	Fully-adjusted model	Age-adjusted model	Fully-adjusted model	Age-adjusted model	Fully-adjusted model
Case number		75	67	81	69	156	136
NSP	6 g/day	0.91 (0.77, 1.07) 0.25	1.03 (0.81, 1.32) 0.79	0.94 (0.80, 1.10) 0.44	0.81 (0.65, 1.01) 0.07	0.92 (0.82, 1.04) 0.18	0.91 (0.77, 1.07) 0.27
NSP density	2 g/1000kcal/day	1.04 (0.87, 1.25) 0.64	1.07 (0.89, 1.30) 0.47	0.85 (0.72, 0.99) 0.04	0.83 (0.70, 0.99) 0.04	0.94 (0.83, 1.06) 0.31	0.95 (0.83, 1.08) 0.41
AOAC	11 g/day	0.86 (0.71, 1.05) 0.15	1.01 (0.73, 1.38) 0.98	0.92 (0.76, 1.12) 0.42	0.76 (0.56, 1.02) 0.07	0.89 (0.78, 1.03) 0.12	0.86 (0.70, 1.08) 0.19
AOAC density	3 g/1000kcal/day	1.03 (0.85, 1.24) 0.78	1.06 (0.87, 1.29) 0.58	0.83 (0.70, 0.99) 0.04	0.82 (0.68, 0.99) 0.04	0.92 (0.81, 1.05) 0.24	0.93 (0.81, 1.07) 0.34
Soluble fibre	3 g/day	0.82 (0.67, 1.01) 0.06	0.87 (0.64, 1.20) 0.40	0.93 (0.77, 1.13) 0.48	0.80 (0.57, 1.06) 0.11	0.88 (0.77, 1.02) 0.08	0.82 (0.66, 1.02) 0.08
Insoluble fibre	4 g/day	0.93 (0.79, 1.09) 0.36	0.99 (0.97, 1.02) 0.68	0.93 (0.80, 1.09) 0.38	0.83 (0.69, 1.01) 0.06	0.93 (0.83, 1.04) 0.21	0.94 (0.81, 1.10) 0.45
Fibre from total cereals	3 g/day	1.02 (0.89, 1.17) 0.75	1.15 (1.00, 1.31) 0.05	0.97 (0.86, 1.11) 0.69	0.93 (0.80, 1.09) 0.39	1.00 (0.91, 1.10) 0.96	1.04 (0.93, 1.16) 0.49
Fibre from breakfast cereals	2 g/day	1,04 (0.92, 1.17) 0.51	1.09 (0.98, 1.23) 0.12	1.00 (0.90, 1.10) 0.98	0.98 (0.88, 1.10) 0.78	1.02 (0.94, 1.11) 0.63	1.04 (0.95, 1.14) 0.36
Fibre from fruit	2 g/day	0.79 (0.65, 0.94) 0.01	0.86 (0.73, 1.02) 0.08	0.92 (0.74, 1.13) 0.42	0.92 (0.72, 1.18) 0.52	0.86 (0.74, 1.00) 0.05	0.89 (0.76, 1.06) 0.19
Fibre from vegetables	2 g/day	0.87 (0.73, 1.05) 0.15	0.88 (0.72, 1.09) 0.24	0.94 (0.79, 1.10) 0.43	0.87 (0.73, 1.04) 0.13	0.91 (0.80, 1.03) 0.12	0.88 (0.77, 1.01) 0.06
Fibre from legumes	1 g/day	0.97 (0.82, 1.13) 0.66	1.02 (0.87, 1.20) 0.76	0.97 (0.77, 1.21) 0.77	0.83 (0.65, 1.05) 0.12	0.97 (0.84, 1.11) 0.63	0.93 (0.81, 1.06) 0.28
Fibre from nuts and seeds	0.2 g/day	0.98 (0.86, 1.11) 0.71	1.00 (0.88, 1.13) 0.97	0.97 (0.89, 1.05) 0.44	0.96 (0.87, 1.06) 0.40	0.97 (0.90, 1.05) 0.46	0.98 (0.90, 1.06) 0.60

Adjustments: Age (years), BMI (kg/m^2), calories from carbohydrate, fat and protein (kcal/day), ethanol intake (g/day), MET (hours/week), smoking status (current vs. not current smoker), socio-economic status (professional or managerial/ intermediate/ routine or manual). Energy intake was not included in fibre density models. Shading in fully adjusted models where CIs do not span 1.

5.5 Discussion

5.5.1 Result summary

In this prospective study of healthy females, there was no evidence of any statistically significant associations between total fibre intake or fibre from certain food sources and risk of fatal IHD, stroke or CVD in analyses of the full sample. Results suggest that greater intake of cereal sources of fibre in those with higher BMI may be associated with reduced fatal stroke risk. Risk was reduced by 20% for every 3g/day increase in cereal fibre, this is roughly equivalent to fibre contained within a standard portion of brown rice or two slices of wholemeal bread (Holland et al., 1991). This specific observation relating to cereal fibre intake may be indicative of the greater insoluble fibre content of cereals compared to fruit and vegetables and is concordant with the protective associations observed for insoluble fibre, but not soluble fibre intake with CHD risk in Chapter 2 (Figure 2.4) and lower stroke risk with greater insoluble but not soluble fibre (Threapleton et al., 2013d). Cereal fibre may also be a surrogate for total cereal intake and the protective association observed in overweight women may reflect other beneficial components of cereal grains. Compounds within grains such as antioxidants, hormonally-active lignans, phytosterols, amylase inhibitors and saponins have all been shown to influence risk factors for CHD and it is likely that this combination of compounds within grains is responsible for their protective effect (Slavin, 2004). It has also been noted in this and other cohort studies that FFQs tend to overestimate intake of some foods such as vegetables (refer to Chapter 3, discussion) and the protective association evident only for cereal fibre may be because of fewer issues with measurement error in assessing cereal intake compared to other foods, especially vegetables.

For both fatal stroke and CVD HRs decreased with greater intake of fibre from nuts and seeds and for both outcomes, risk was significantly reduced in the highest group compared to the lowest consumers. However, the evidence for a linear dose-response relationship was lacking, perhaps because fibre intake from nut and seed sources was relatively low and protective associations may only become apparent with intakes at sufficiently high levels.

In the healthy subgroup of participants, that is those free of hypertension and angina, there appeared to be a protective association with fibre density and fatal stroke risk. This had not been observed when those with hypertension or angina were combined in the full sample, suggesting greater fibre density may prevent CVD development in those who are healthy,

rather than halt or reverse disease development in those already displaying risk factors. This result however did not reach the pre-specified 1% level of statistical significance that was set for subgroup analyses. The risk of Type I error, or false positive findings, is even greater in subgroup analyses, where sample sizes are diminished (Bowers et al., 2006a). The possible protective effect of fibre density in those without key CVD risk factors should be further explored in cohorts with larger case numbers.

The attenuation of associations between the age or mid-adjusted models and fully-adjusted models indicates that the variables identified as potential confounders explain some of the variation in risk and are associated with fibre intake. Differing results by BMI classification indicate that BMI may modify the effect of fibre on CVD risk despite no different risk estimates being observed between models with and without adjustment for BMI.

5.5.2 Comparison with other published studies

A recent pan-European EPIC study observed contrasting results to the findings in this chapter and report a protective association between total fibre intake and total CVD mortality (Chuang et al., 2012). However, the definition for CVD mortality differed from the UKWCS and included all cardiovascular-related death rather than coronary plus stroke events (this study is included in meta-analyses and has been discussed in Chapter 2). Another observation from this EPIC study was for cereal fibre intake; greater intake was associated with risk reduction, as was seen for the UKWCS in obese women. Similarly, both the EPIC study and the work in this chapter indicate no evidence of protective associations for fruit or vegetable fibre intake (Chuang et al., 2012). Fatal CHD risk within EPIC was explored in a separate publication and a significant risk reduction was also reported with greater intake of total fibre (Crowe et al., 2012). However, as discussed in Chapter 2, one study focusing on just the UK data from the pooled EPIC study found that fibre assessed using food diaries was protectively associated with risk, but this was not the case for fibre calculated from FFQs (Ward et al., 2012), mirroring the non significant results observed here. As discussed in Chapter 3, the limitations in assessing diet using FFQ are potentially greater than with food diaries and this observation from Ward and colleagues may be attributed to this.

Systematic reviews and data pooling projects for dietary fibre and CVD or CHD report protective associations for dietary fibre intake (Pereira et al., 2004, Mente et al., 2009, Ye et al., 2012) as was seen in the systematic review and meta-analyses reported in Chapter 2. These reviews examined both incidence and mortality data together but it is possible the underlying

pathology for non-fatal events differs from fatal events. However, other prospective studies reporting just on fatal CVD or CHD events also tend to observe protective associations both in men and women (Eshak et al., 2010, Park et al., 2011, Pietinen et al., 1996, Rimm et al., 1996, Streppel et al., 2008, Wolk et al., 1999) unlike the results observed here. One exception to this trend is the Australian Blue Mountain Eye study, which did not report a protective association for total fibre and fatal CVD (Buyken et al., 2010).

The picture for total (fatal plus non-fatal) stroke risk is less consistent than CHD and CVD, whereby some studies report no evidence of protective associations with increased fibre intake (Oh et al., 2005, Bazzano et al., 2003) and several observe protective associations (Kokubo et al., 2011, Larsson et al., 2009, Ascherio et al., 1998). Few studies report fatal stroke events and total fibre intake but a lack of association was reported in one Japanese cohort for both men and women (Eshak et al., 2010) and a cohort from the US (Bazzano et al., 2003). Additionally fatal stroke risk was not associated with greater cereal fibre intake in an Australian cohort (Kaushik et al., 2009).

Considering the existing evidence from observational studies and meta-analyses together, there appears to be an inverse association between total dietary fibre and both total (fatal plus non-fatal) and fatal CVD or CHD risk, contrasting observations in the UKWCS which were not statistically significant despite being in the same direction. The general lack of evidence for a protective association of total fibre and stoke mortality risk does mirror observations from some other cohorts (discussed above) but does not elucidate possible reasons for the protective associations which were observed for stroke risk in some subgroup analyses.

5.5.3 Strengths and limitations

Strengths of this work include that data are from a large prospective cohort that has been followed up for a relatively long period of time. The cohort was designed to allow exploration of disease in relation to healthy dietary characteristics by recruiting a large proportion of health conscious individuals. Diet was also assessed using a tool that had been validated for use in the sample.

Results were weighted to reduce the impact of data from vegetarian participants as a much greater proportion exist in this sample than the general population, meaning risk estimates are more applicable to the general population. The inclusion of this weighting factor actually had relatively little impact on the estimates but does account somewhat for the oversampling of

vegetarians. However, the sample does include women who are generally well-educated, middle-class and are reasonably healthy and therefore, the generalisablility of findings to other populations is unknown.

One limitation in dietary assessment here is that diets may change over time but only diet assessed at baseline was considered in this analysis. However, some work using a sub-sample of cohort participants assessed 5-years after baseline indicated moderate stability in dietary pattern classification (Greenwood et al., 2003). Other shortcomings in dietary assessment using FFQs are issues such as measurement error and the tendency to over-estimate consumption of certain foods like fruit and vegetables (Cade et al., 2002). Refer to Chapter 3 where strengths and limitation of dietary assessment methods are discussed.

Uncontrolled confounding is a limitation with observational work (Willett, 1998a), meaning some other lifestyle or dietary factor, not adequately controlled for or accounted for in models, could explain observations. Evidence from RCTs would be required to confirm associations as causal but given the long progression of CVD, trials are unlikely to be feasible. Another problem with observational work of this kind is the inability to distinguish single nutrient specific end-points from other nutrients that are highly correlated (Bingham et al., 1994). Here it is not possible to identify whether fibre from a specific food is related to end-points or whether intake of the whole food, with associated nutrients, is responsible.

Further limitations include the imperfect measurement of confounding variables, BMI was derived from self-reported weight and height, SES in the model was based on three broad groupings and physical activity expenditure was calculated from a series of questions which asked participants to estimate the time spent on usual activities, all of which introduce error into calculations. The use of only mortality data is a limitation as non-fatal cases are unidentified and are therefore misclassified as non-cases. Case numbers are also limited by using only mortality data, especially for sensitivity analyses. However, given that different associations have been reported in some studies for mortality or non-fatal CVD outcomes and fibre intake (Bazzano et al., 2003, Pietinen et al., 1996), it is plausible that the underlying pathology for the two outcomes is distinct and combining events may cloud rather than elucidate associations. The lack of consideration for the time-frame of exposure and disease development in prospective work has also been criticised (Willett, 1998a) but case numbers here were too few to explore this.

There are reported inconsistencies in the coding of the 'original underlying cause' of death in the UK, with 78% of cases reportedly matching to the fourth ICD code level and 90% matching to the ICD10 chapter, in a sample of 7,914 deaths which were coded in duplicate. The national death data are however routinely checked and pass through a series of validation checks and processes to highlight potential errors (ONS, 2013). Mortality data have been recorded in a consistent way through national registry services for decades in the UK and registers were well established before the start of the UKWCS.

5.6 Summary

Greater total dietary fibre intake may confer no additional cardiovascular benefit in already health-conscious women but may contribute to lower fatal stroke risk in those free of cardiovascular risk factors (hypertension/angina). Cereal fibre may contribute to lower fatal stroke risk specifically in women with higher BMI and there are suggestions that fibre from nuts and seeds may contribute to lower stroke mortality risk in women free of CVD risk factors. There was no suggestion of protective associations for other sources of fibre for fatal stroke risk reduction or for any food source of fibre with fatal CHD or CVD.

Further observational work will ideally include incidence data to boost statistical power. Experimental studies should contribute towards explaining observations from this and other cohort studies through exploring possible mechanisms underlying the relationship between CVD risk factors, BMI and the different types of fibre or sources of fibre.

The principles used to explore CVD mortality and fibre intake, assessed with FFQs, in this chapter are extended in the next chapter (Chapter 6) to explore fatal plus non-fatal CVD risk in relation to fibre intake. Risk of specific types of CHD and stroke subtypes in relation to fibre intake are also explored in Chapter 6.

Chapter 6 Dietary fibre intake and risk of total CVD, non-fatal CVD and CVD subtypes

6.1 Chapter overview

The current chapter explores dietary fibre assessed with FFQs and risk of non-fatal or total CHD, stroke and CVD. As in the last chapter, FFQs provide data on total fibre intake, both soluble and insoluble fibre and also fibre from key food sources. Mortality event data are combined with HES and MINAP cases to estimate total CVD incidence. For non-fatal events, data from HES and MINAP were combined. Figure 6.1 highlights the sources of dietary and outcome data that are used and presented in this chapter.

The previous chapter detailed the methods used to select confounding variables and survival analysis methods. These methods are also applied in this chapter and are therefore only described here briefly, with additional detail to highlight any differences from the work in the previous chapter.

Few of the fibre exposures were associated with risk of non-fatal or total incident CHD in this chapter but protective associations were observed for fibre density and total fibre assessed as AOAC and insoluble fibre intake when MI was assessed separately from other CHD outcomes. Each 11g/day greater AOAC fibre intake was associated with lower MI risk, HR 0.86 (95% CI: 0.73 to 1.00) p=0.04. Higher fibre intake was also protectively associated with risk of total stroke (fatal plus non-fatal) but this association only extended to cereal fibre when non-fatal strokes were assessed. The associations observed with fibre intake and stroke risk were apparent for ischaemic, rather than haemorrhagic strokes but there were comparatively fewer haemorrhagic stroke cases observed over the 14 year follow up. Lower overall CVD risk was associated with greater fibre density, with every 2g/1000kcal/day higher NSP density, risk was 0.95 (95% CI: 0.91 to 1.00) p=0.03.

Two abstracts presenting results from work in this chapter were accepted for presentation at the European Congress of Epidemiology 2013 (Threapleton et al., 2013a) and the UK Society for Social Medicine annual meeting 2013 (Threapleton et al., 2013c) and have now been published.

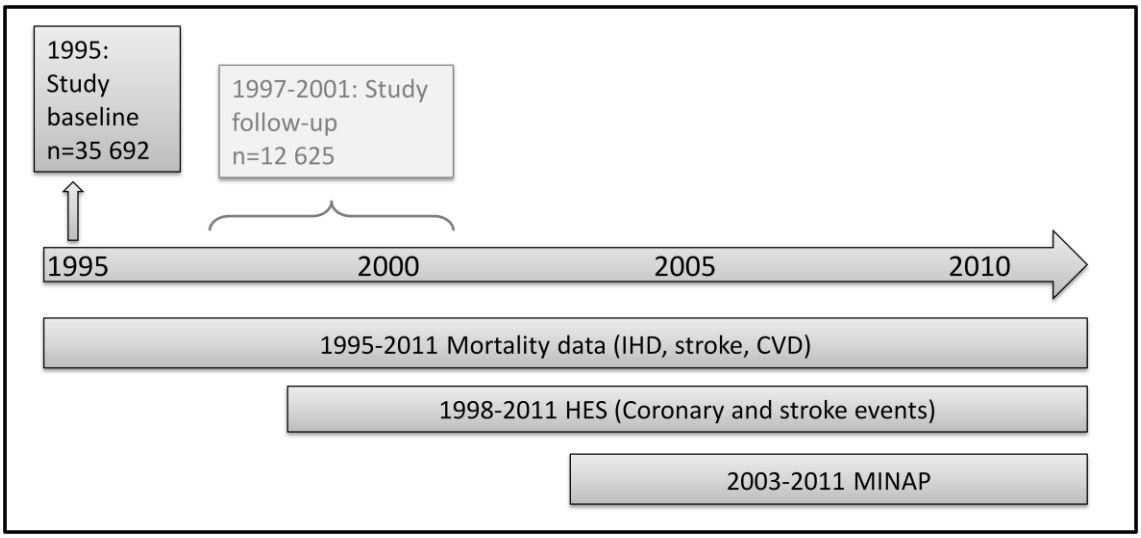


Figure 6.1 Data sources used in this chapter: Dietary data from FFQs, mortality data spanning from study baseline, HES data since 1998 and MINAP since 2003

6.2 Background

There are an estimated 103,000 heart attacks in the UK each year, with approximately 50,000 of these in English men and 32,000 in English women (Townsend et al., 2012). Although incidence rates for CHD and strokes are declining, CVD remains an enormous social and financial burden for the UK, as noted in Chapter 1.

Reductions in MI over the past few decades have been driven by improvements in risk factors (Unal et al., 2004) and the incidence rate in England has decreased by a third between 2002 and 2010. Generally, MI incidence increases with age and it is this characteristic that means women experience more events in total than men, because they are living longer (Townsend et al., 2012).

Stroke incidence has also decreased over the past few decades but there remains a greater number of stroke events in women, again, because they live longer in general (Townsend et al., 2012). In England there are an estimated 68,000 strokes in women annually and 57,000 in men. For the whole UK, there are an estimated 152,000 strokes annually (Townsend et al., 2012).

Unlike mortality, which is a clearly defined outcome and where the style of recording events on death certificates has been largely consistent over many decades, non-fatal CVD events were unrecorded, or recorded in a non centralised system, during the early phase of the UKWCS. Still today, unlike cancer events, there is no national database from which to draw

complete CVD event data. In addition, the severity of non-fatal events will differ meaning that only the most severe types lead to hospitalisation, clear diagnosis and event recording whereas many minor events or risk factors for CVD may go unrecorded. Despite limitations in the availability of historical CVD event data for the participants of the UKWCS, compiling cases from multiple available sources, even though they span different timeframes, has much potential benefit. This approach allows not only case numbers to be increased for improved statistical power but also provides sufficient case numbers to allow exploration of stroke and coronary event sub-types as the aetiology of the different event types may differ.

As noted in Chapter 2 (systematic review of literature) other cohorts studies have separately assessed fibre intake in relation to risk of fatal, non-fatal or total CVD events and found different observations. Two of the identified studies explored fatal events in addition to total (fatal plus non-fatal events combined) and found contrasting results (Bazzano et al., 2003, Pietinen et al., 1996). In the NHANES I study, a significant protective association was seen for fibre intake and total CVD or CHD risk but the association was not apparent when examining only risk of fatal CVD or CHD, possibly because of fewer events being available (Bazzano et al., 2003). The Finnish ATBC study observed the opposite, with a protective association between total fibre intake and fatal CHD but there was no evidence of an association when non-fatal MI events were combined with fatal CHD (Pietinen et al., 1996).

The US Health Professionals' Follow-up Study (HPFS) and Nurses' Health Study assessed risk of fatal, non-fatal or total CHD (Rimm et al., 1996, Wolk et al., 1999). The HPFS found protective associations for risk with total fibre intake in all three outcomes but protective associations for both fatal and non-fatal CHD were seen in the Nurses' Health Study but not when the two events combined (Wolk et al., 1999). The relatively small number of other studies which have assessed risk of non-fatal events in addition to fatal and total, only one of which reported stroke events (Bazzano et al., 2003), and the inconsistent results observed in these studies indicates the need for further research in this area. This issue is addressed in this chapter where risk of non fatal CVD is explored along with total CVD and CVD subtypes.

6.3 Method

6.3.1 Dietary data

As discussed in previous chapters, FFQs were used to assess typical intake over the previous 12 months. Fibre calculated as both NSP and using the AOAC method were examined. Fibre

density of the diet in addition to soluble and insoluble fibre intakes and NSP from key food sources were explored, as in the previous chapter.

6.3.2 Mortality data

The same IHD and stroke definitions were applied to mortality records to identify cases, as discussed in Chapters 4 and 5. The inclusion of case event dates was however extended, from February 2011, to 30th June 2011 to reflect newly available data and to match with the latest available case information from the two other sources (HES and MINAP).

6.3.3 HES records

CHD cases include those with records listing ICD10 codes I20 to I25.9 and I46 to I59.0 (Table 4.4). Stroke cases were identified as those with records listing any stroke event ICD10 I60.0 to I64X (Table 4.5).

Stroke or CHD events were identified from HES using only the primary diagnosis field in the dataset and not using any of the other diagnosis fields. This approach was taken for survival analyses as events listed in secondary diagnosis fields may relate to old conditions, not necessarily to the reason for inpatient admission. As it is not clear when events in secondary diagnosis fields occurred, accurate survival times could not be generated.

6.3.4 MINAP records

Coronary events from MINAP were those where a final diagnosis of MI, threatened MI or ACS were identified (refer to Chapter 4 for details).

6.3.5 Censor date

MINAP events were examined closely to observe if a lag-time in event reporting and therefore drop-off in events was visible. If this was the case, a censor date earlier than the date of the latest record would be set, to reflect a month where event reporting appeared complete. No lag-time and therefore drop off in MINAP event reporting was observed because the numbers of cases each month was small and there was a large degree of variation in case numbers between months (refer to Figure 4.3 in Chapter 4). The date of the most recent event in MINAP was therefore chosen as the censor date for all outcome sources. Survival times were calculated in years from receipt of baseline questionnaire until either, earliest CVD event date, date of death for any other cause or until 30th June 2011, whichever came first. The censor

date of 30th June 2011 was applied to HES and mortality events, where records extended after this time.

6.3.6 Exclusions

In addition to the exclusions applied for mortality analyses (detailed below and in Chapter 5), participants whose baseline address was in Scotland, Wales or Northern Ireland were excluded. This decision was taken because MINAP records relate to English and Welsh hospitals but HES data were only obtained for English hospitals. Women were also excluded where they reported personal history of angina, unlike in the previous chapter, because this was included as an outcome for ACS, chronic and total CHD events.

For analyses examining risk of non-fatal events, those who died, of any cause, within 1 month of the earliest CHD, stroke or CVD event, were excluded. For example, a participant with non-fatal stroke occurring in one year, followed by a fatal stroke in the next year would remain in the analysis but a CHD event followed by death within 30 days, from any cause, would be excluded.

Participants were excluded from analyses where the following criteria were met:

1) Not successfully tracked through the ONS/ NHSIC for record linkage (n=695)
2) Tracked through national registers but did not provide both lifestyle and dietary information (n=318)
3) Daily calorie intake from the FFQ was outside a plausible range (500-6000 kcal/day) (n=405)
4) Reported personal history of cancer (n=2445), stroke (n=264), diabetes (n=646), angina (n=718) or heart attack (n=498) at study baseline
5) Died (any cause) or experienced CVD event within one year of returning FFQ (n=129)
6) Requested their data not to be used in future studies (n=1)
7) Address not in England (n=3874)
8) Died within 30 days of CHD (n=107), stroke (n=99) or CVD (n=192) event [this criteria was applied separately in each analysis using non-fatal data]

6.3.7 Testing dose-response and non-linear associations

CVD risk was assessed with fibre as both a categorical exposure (fifths of intake) where each subsequent category was compared to the lowest intake category and as a continuous

exposure to explore potential dose-response associations. Exposure increments used in dose-response models were the same as those detailed in Chapter 3. Categorical exposures were re-generated specifically for this sample of women, where only English participants were included and are not identical to those in the previous chapter.

6.3.8 Survival analyses

Survival analyses were conducted using Cox proportional hazard regression (Cox and Oakes, 1984) and again models were weighted by the inverse of the probability of being sampled, to take into account the large proportion of vegetarians in the cohort (as noted in Chapter 5, methods).

Assumptions for proportional hazards were once again checked with the use of log-log survival curves for each outcome with all exposure and confounding variables. This condition was met for each exposure and with all covariates used in models.

For primary analyses (full sample) a 2-sided p-value ≤0.05 was considered statistically significant and for subgroup analyses, this was reduced to ≤0.01. In later sections of this chapter, results from subgroup analyses are presented where CIs indicate an association but where the p-value did not reach the 1% level. These associations are of importance but should be interpreted with caution because of the greater chance of false positive findings with multiple tests.

Microsoft Access was used for identifying earliest CHD, stroke or CVD events within each participant ID, thus condensing data from long (multiple row) format to wide (one row per participant) with the help of a database manager. Survival times were then calculated and all other data manipulation and analyses were conducted by myself, using Stata version 12 (StataCorp, 2011).

A recent study indicates that vegetarians and meat-eaters may possess different intestinal bacteria types and that the bacteria present to a larger degree in meat-eaters may contribute to increased CVD risk (Koeth et al., 2013). Since fibre is understood to affect gut micro-flora, post-hoc analysis of key associations between total fibre intake and CHD or stroke risk in vegetarian and non-vegetarian participants was undertaken to explore any potential different associations. The same exclusions and adjustments were applied as with other models but the factor that weighted on vegetarian status was removed.

6.3.9 Confounder adjustment

Potential confounding variables chosen as adjustments for models were the same as in the previous chapter. Please refer back to Chapter 5 for details surrounding the selection of these. The following three levels of adjustment were applied:

1) Age (years)
2) Age (years), alcohol (ethanol g/day), smoking status (non-smoker, current-smoker, ex-smoker), physical activity-metabolic equivalents (MET- hours/week) and SES (professional/managerial, intermediate or routine/manual).
3) As model 2 with the addition of energy intake (kcal/day)* and BMI (kg/m^2)

* As noted in the previous chapter, by way of sensitivity analysis, when modelling CVD risk associated with fibre density of the diet, energy intake was additionally excluded in separate models, as suggested for nutrient density analyses by Willett (2013b). Results were not appreciably different with or without adjustment for energy intake in fibre density analyses (data not shown) and the results presented here are without adjustment for energy intake.

In the intermediate model, the absence of BMI and energy intake as adjustments did not greatly alter risk estimates compared to model 3. In the interest of brevity, results from model 2 are therefore not included in tables as they offer little extra information but are discussed where relevant.

6.3.10 Cohort subgroup analyses

Secondary analyses were conducted using the same subgroups as detailed in Chapter 5; menopausal status, BMI classification and self-reported hypertension at baseline. Subgroups analysis was only conducted where, for each outcome type, a minimum of 50 cases existed within the subgroup (case numbers are detailed in Tables 6.1 to 6.4). Analyses restricted to subgroups were only carried out for linear dose-response associations and not where exposures were split into 5 intake levels as case numbers tended to be quite small in the various categories.

6.3.11 Case subtype analyses

In addition to examining total and non-fatal coronary and stroke events, fibre intake was examined in relation to risk of specific classifications of CHD and stroke (refer to Chapter 4, Tables 4.4 and 4.5 for details of which ICD10 codes are included in each outcome):

- Myocardial infarction
- Acute coronary syndromes (ACS) (including MI and other acute coronary events)
- Chronic CHD (all chronic and not acute events)
- Haemorrhagic stroke
- Ischaemic stroke
- Stroke, type not specified
- Ischaemic and unspecified stroke

Stroke events were initially grouped as haemorrhagic, ischaemic or those cases where event type was not reported (unspecified). Because of strongly protective risk associations observed only for the unspecified stroke cases, ischaemic and unspecified strokes were grouped together to represent 'mostly ischaemic' type stroke, for post-hoc exploration. As the majority of first stroke events are ischaemic, grouping the ischaemic and unspecified will approximately represent ischaemic strokes. Estimates from other studies of the proportion of first stroke events that are ischaemic in type varies from 78% in a collaborative study including 22 countries around the world (O'Donnell et al., 2010) to 90% of first stroke events being recorded as ischaemic in a Danish study (Andersen et al., 2009). Grouping cases in this way provides a larger sample of ischaemic type strokes, despite the possibility of including a small proportion of haemorrhagic stroke cases.

6.4 Results

6.4.1 Sample and case numbers

After applying exclusions to the sample, 27,400 women remained from 35,692. Up to 30th June 2011, a total of 821 incident CHD and 388 incident stroke cases were observed. For each participant, the first CHD or stroke event was identified, disregarding later events. When combined and either the first CHD or stroke event was identified, a total of 1162 incident CVD cases were observed (Table 6.1). After removal of participants with missing data for any of the selected covariates, 760 incident CHD, 347 incident stroke and 1067 incident CVD cases were available in the fully-adjusted model analyses.

In fully adjusted models, 125 haemorrhagic, 160 ischaemic and 125 unspecified stroke cases were available (Table 6.2) and 217 MI, 361 ACS and 535 chronic heart disease cases were available (Table 6.3). For non-fatal CHD, stroke and CVD analyses, 668, 258 and 914 cases were available in fully-adjusted models, respectively (Table 6.4).

Table 6.1 Cohort participant and total (fatal plus non-fatal) CHD, stroke and CVD case frequency (%) included in unadjusted and fully adjusted models

	UKWCS participant numbers †		Total CHD case frequency		Total stroke case frequency		Total CVD case frequency	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Full sample	27400	25810	821	760	388	347	1162	1067
Premenopausal	11447 (42)	10908 (42)	118 (14)	111 (15)	48 (12)	38 (11)	163 (14)	146 (14)
Postmenopausal	9475 (35)	8816 (34)	476 (58)	440 (58)	248 (64)	225 (65)	692 (60)	638 (60)
Underweight*	578 (2)	557 (2)	11 (1)	10 (1)	16 (4)	14 (4)	27 (2)	24 (2)
Healthy weight*	16965 (62)	16589 (64)	428 (52)	412 (54)	211 (54)	203 (59)	614 (53)	592 (55)
Overweight*	6457 (24)	6318 (24)	234 (29)	223 (29)	97 (25)	92 (27)	320 (28)	305 (29)
Obese*	3371 (12)	2318 (9)	148 (18)	115 (15)	64 (16)	38 (11)	201 (17)	146 (14)
Hypertension	4011 (15)	3745 (15)	242 (29)	224 (29)	117 (30)	108 (31)	347 (30)	323 (30)
No hypertension	23389 (85)	22065 (85)	579 (71)	536 (71)	271 (70)	239 (69)	815 (70)	744 (70)

Key as below. Note there are fewer CVD cases than CHD plus stroke numbers combined as CVD represents the first occurrence of either a CHD or stroke event for each participant

Table 6.2 Cohort participant and stroke sub-type case frequency (%) included in unadjusted and fully adjusted models

	UKWCS participant numbers †		Haemorrhagic stroke case frequency		Ischaemic stroke case frequency		Unspecified stroke case frequency		Ischaemic/ unspecified stroke case frequency	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Full sample	27400	25810	135	125	184	160	138	125	284	251
Premenopausal	11447 (42)	10908 (42)	30 (22)	25 (20)	22 (12)	17 (11)	6 (4)	6 (5)	23 (8)	18 (7)
Postmenopausal	9475 (35)	8816 (34)	66 (49)	62 (50)	119 (65)	106 (66)	106 (77)	96 (77)	199 (70)	179 (71)
Underweight*	578 (2)	557 (2)	6 (4)	6 (5)	3 (2)	2 (1)	8 (6)	7 (6)	11 (4)	9 (4)
Healthy weight*	16965 (62)	16589 (64)	74 (55)	71 (57)	98 (53)	95 (59)	76 (55)	72 (58)	154 (54)	149 (59)
Overweight*	6457 (24)	6318 (24)	33 (24)	11 (9)	53 (29)	49 (31)	32 (23)	31 (25)	71 (25)	66 (26)
Obese*	3371 (12)	2318 (9)	22 (16)	15 (12)	30 (16)	14 (9)	22 (16)	15 (12)	48 (17)	27 (11)
Hypertension	4011 (15)	3745 (15)	37 (27)	36 (29)	50 (27)	43 (27)	51 (37)	48 (38)	91 (32)	82 (33)
No hypertension	23389 (85)	22065 (85)	98 (73)	89 (71)	134 (73)	117 (73)	87 (63)	77 (62)	193 (68)	169 (67)

Unadjusted: sample/case frequency (%) available in unadjusted models; Adjusted: sample/case frequency (%) available in fully adjusted models

*using WHO cut-points for BMI <18.5, 18.5-24.9, 25.0-29.9, ≥30 kg/m^2 † Numbers include cases

Table 6.3 Participant and CHD sub-type case frequency (%) included in unadjusted and fully adjusted models

	UKWCS participant numbers †		MI case frequency		ACS case frequency		Chronic heart disease case frequency	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Full sample	27400	25810	236	217	392	361	573	535
Premenopausal	11447 (42)	10908 (42)	28 (12)	27 (12)	49 (13)	46 (13)	77 (13)	74 (14)
Postmenopausal	9475 (35)	8816 (34)	156 (66)	142 (65)	241 (61)	222 (61)	328 (57)	302 (56)
Underweight*	578 (2)	557 (2)	5 (2)	4 (2)	6 (2)	5 (1)	6 (1)	5 (1)
Healthy weight*	16965 (62)	16589 (64)	134 (57)	127 (59)	214 (55)	202 (56)	291 (51)	283 (53)
Overweight*	6457 (24)	6318 (24)	54 (23)	51 (24)	107 (27)	99 (27)	173 (30)	167 (31)
Obese*	3371 (12)	2318 (9)	43 (18)	35 (16)	65 (17)	55 (15)	103 (18)	80 (15)
Hypertension	4011 (15)	3745 (15)	70 (30)	66 (30)	116 (30)	106 (29)	180 (31)	171 (32)
No hypertension	23389 (85)	22065 (85)	166 (70)	151 (70)	276 (70)	255 (71)	393 (69)	364 (68)

Key as below.

Table 6.4 Participant and non-fatal CHD, stroke and CVD case frequency (%) included in unadjusted and fully adjusted models

	UKWCS participant numbers †		Non-fatal CHD case frequency		UKWCS participant numbers		Non-fatal stroke case frequency		UKWCS participant numbers		Non-fatal CVD case frequency	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Full sample	27287	25606	714	668	27302	25721	290	258	27287	25609	990	914
Premenopausal	11444 (42)	10893 (43)	112 (16)	107 (16)	11439 (42)	10901 (42)	40 (14)	31 (12)	11435 (42)	10893 (43)	151 (15)	137 (15)
Postmenopausal	9394 (34)	8674 (34)	399 (56)	372 (56)	9406 (34)	8754 (34)	179 (62)	163 (63)	9337 (34)	8677 (34)	569 (57)	527 (58)
Underweight*	575 (2)	548 (2)	8 (1)	7 (1)	570 (2)	551 (2)	8 (3)	8 (3)	567 (2)	548 (2)	16 (2)	15 (2)
Healthy weight*	16898 (62)	16483 (64)	380 (53)	368 (55)	16914 (62)	16541 (64)	160 (55)	155 (60)	16855 (62)	16484 (64)	535 (54)	518 (57)
Overweight*	6415 (24)	6257 (24)	202 (28)	192 (29)	6435 (24)	6296 (24)	75 (26)	70 (27)	6395 (23)	6258 (24)	272 (27)	257 (28)
Obese*	3370 (12)	2290 (9)	124 (17)	101 (15)	3354 (12)	2305 (9)	47 (16)	25 (10)	3356 (12)	2291 (9)	167 (17)	124 ()14
Hypertension	3960 (15)	3669 (14)	205 (29)	192 (29)	3976 (15)	3711 (14)	82 (28)	74 (29)	3929 (14)	3670 (14)	283 (29)	263 (29)
No hypertension	23327 (85)	21937 (86)	509 (71)	476 (71)	23326 (85)	22010 (86)	208 (72)	184 (71)	23273 (85)	21939 (86)	707 (71)	651 (71)

Unadjusted: sample/case frequency (%) available in unadjusted models; Adjusted: sample/case frequency (%) available in fully adjusted models

*using WHO cut-points for BMI <18.5, 18.5-24.9, 25.0-29.9, ≥30 kg/m^2 † Numbers include cases.

Note there are fewer CVD cases than CHD plus stroke numbers combined as CVD represents the first occurrence of either a CHD or stroke event for each participant

6.4.2 Descriptive statistics

Stroke cases were unsurprisingly older than non-cases, median age was 62.7 (IQR 13.7) years compared to 50.1 (IQR 13.9) years (Table 6.5). Stroke cases also reported slightly higher BMI 24.0 kg/m^2 (IQR 4.9) compared to non-cases 23.4 kg/m^2 (IQR 4.6). A greater proportion of stroke cases were also meat eaters 74% compared to non-cases 66% and educational achievement levels were lower in cases. Haemorrhagic stroke cases tended to be younger than other stroke types, median age was 58.4 years (IQR 17.3) for haemorrhagic stroke and was 63.1 years (IQR 13.5) and 65.9 years (IQR 9.3) for ischaemic and unspecified stroke, respectively. Haemorrhagic stroke cases also reported slightly greater fibre intake levels than other cases, although this may relate to the marginally higher energy intake reported for the haemorrhagic compared to other cases. Stroke cases were also much more likely to report history of hypertension compared to non-cases 30% vs. 14%. Intake levels for total fibre and fibre from different sources were roughly comparable for different stroke cases types and non-cases.

Similar to stroke, CHD cases were almost 10 years older than non-cases. At baseline, median age was 59.4 years (IQR 13.5) vs. 50.0 years (IQR 13.9) and median BMI was over one point higher in the cases 24.6 kg/m^2 (IQR 5.6) compared to non-cases 23.4 kg/m^2 (IQR 4.6) (Table 6.6). Most characteristics were similar between acute and chronic heart disease cases but some differences exist between these and the small number of 'other' heart disease cases, namely the 'other CHD' cases were younger and had BMI closer to the non-cases. Heart disease cases also tended to be meat-eaters, smokers and have lower educational achievement compared to the non-cases. Personal history of hypertension was higher in CHD cases, with 30% of participants reporting this at baseline, compared to the non-cases where just 14% had existing hypertension.

CVD cases were more than 10 years older than non-cases 60.4 years (IQR 13.8) compared to 49.9 years (IQR 13.7) and BMI was one unit higher 24.4 kg/m^2 (IQR 5.1) compared to non-cases 23.4 kg/m^2 (IQR 4.6) (Table 6.7). Similar to the stroke and CHD cases and non-cases, CVD cases tended to have a lower education or employment profile and tended to include more meat-eaters and smokers compared to the non-cases. Physical activity, energy intake and other dietary characteristics were however largely comparable in the CVD cases and non-cases.

Table 6.5 Baseline cross-sectional characteristics for fatal plus non-fatal stroke and stroke subtype cases, non-fatal stroke cases and non cases

		Haemorrhagic stroke (fatal plus non-fatal cases)	*Ischaemic stroke (fatal plus non-fatal cases)*	*Unspecified stroke (fatal plus non-fatal cases)*	*Ischaemic or unspecified stroke (fatal plus non-fatal cases)*	*Total stroke (all fatal plus non-fatal cases)*	*Non-fatal stroke*	*Non stroke cases*
N		*135*	*184*	*138*	*284*	*388*	*355*	*27012*
Age, years		*58.4 (17.3)*	*63.1 (13.5)*	*65.9 (9.3)*	*64.5 (12.1)*	*62.7 (13.6)*	*61.9 (13.6)*	*50.1 (13.9)*
BMI, kg/m^2		*23.7 (4.8)*	*24.4 (5.2)*	*24.0 (5.3)*	*24.1 (5.0)*	*24.0 (4.9)*	*24.0 (4.8)*	*23.4 (4.6)*
Smoking status (%)	Current	*22 (16)*	*25 (14)*	*21 (15)*	*38 (13)*	*52 (13)*	*35 (12)*	*2848 (11)*
	Former	*37 (27)*	*43 (24)*	*38 (28)*	*71 (25)*	*105 (27)*	*89 (31)*	*8144 (30)*
	Never smoked	*76 (56)*	*115 (63)*	*79 (57)*	*174 (61)*	*230 (60)*	*165 (57)*	*15975 (59)*
Diet group (%)	Meat-eaters	*98 (73)*	*138 (75)*	*108 (78)*	*215 (76)*	*289 (74)*	*211 (59)*	*17836 (66)*
	Fish-eaters	*20 (15)*	*17 (9)*	*10 (7)*	*25 (9)*	*41 (11)*	*31 (11)*	*3518 (13)*
	Vegetarian	*17 (13)*	*29 (16)*	*20 (14)*	*44 (15)*	*58 (15)*	*48 (17)*	*5658 (21)*
Socio-economic status NS-SEC (%)	Professional/ managerial	*78 (59)*	*96 (54)*	*64 (49)*	*143 (53)*	*199 (54)*	*146 (52)*	*16787 (63)*
	Intermediate	*38 (29)*	*61 (35)*	*51 (39)*	*100 (37)*	*132 (35)*	*103 (37)*	*7297 (28)*
	Routine and manual	*16 (12)*	*18 (10)*	*16 (12)*	*27 (10)*	*41 (11)*	*30 (11)*	*2378 (9)*
Highest educational achievement (%)	No formal record	*30 (24)*	*49 (33)*	*46 (42)*	*80 (35)*	*105 (32)*	*73 (30)*	*3868 (16)*
	O-level	*31 (24)*	*37 (25)*	*22 (20)*	*54 (24)*	*80 (25)*	*65 (27)*	*8081 (32)*
	A-level	*34 (27)*	*34 (23)*	*27 (24)*	*55 (24)*	*80 (24)*	*62 (25)*	*6057 (24)*
	Degree	*32 (25)*	*28 (19)*	*14 (13)*	*41 (18)*	*62 (19)*	*45 (18)*	*6820 (27)*
Menopause status (%)	Post	*66 (51)*	*119 (66)*	*106 (80)*	*199 (72)*	*248 (66)*	*179 (64)*	*9227 (35)*
	Pre	*30 (23)*	*22 (12)*	*6 (5)*	*23 (8)*	*48 (13)*	*40 (14)*	*11399 (43)*
	Not applicable*	*34 (26)*	*40 (22)*	*21 (16)*	*54 (20)*	*80 (21)*	*62 (22)*	*5986 (22)*
History of hypertension at baseline (%)	Yes	*37 (27)*	*50 (27)*	*51 (37)*	*91 (32)*	*117 (30)*	*82 (28)*	*3894 (14)*
	No	*98 (73)*	*134 (72)*	*87 (63)*	*193 (68)*	*271 (70)*	*208 (72)*	*23118 (86)*
Ethanol g/day		*4.9 (10.8)*	*3.3 (10.7)*	*3.5 (9.6)*	*3.3 (10.4)*	*3.8 (10.5)*	*3.4 (10.1)*	*5.6 (11.7)*
Physical activity, MET-hours/week		*14.0 (14.4)*	*14.3 (13.8)*	*16.8 (14.0)*	*15.3 (13.5)*	*14.6 (13.7)*	*15.0 (14.0)*	*14.5 (13.1)*
Energy intake, kcal/day		*2237 (831)*	*2176 (935)*	*2012 (952)*	*2143 (933)*	*2153 (896)*	*2169 (915)*	*2179 (857)*
Saturated fat intake g/day		*28.1 (15.1)*	*26.9 (20.1)*	*26.1 (17.1)*	*27.2 (17.9)*	*27.5 (17.6)*	*27.5 (19.2)*	*26.8 (15.8)*

		Haemorrhagic stroke (fatal plus non-fatal cases)	*Ischaemic stroke (fatal plus non-fatal cases)*	*Unspecified stroke (fatal plus non-fatal cases)*	*Ischaemic or unspecified stroke (fatal plus non-fatal cases)*	*Total stroke (all fatal plus non-fatal cases)*	*Non-fatal stroke*	*Non stroke cases*
NSP, g/day		*24.3 (13.3)*	*23.1 (13.1)*	*20.8 (13.1)*	*23.0 (13.7)*	*23.1 (13.4)*	*23.6 (13.3)*	*23.8 (12.3)*
NSP density, g/1000kcal/day		*11.1 (4.2)*	*10.9 (4.9)*	*10.1 (4.8)*	*10.6 (4.9)*	*10.8 (4.6)*	*10.9 (4.5)*	*11.0 (4.2)*
AOAC fibre, g/day		*36.6 (20.7)*	*35.7 (19.5)*	*32.1 (21.8)*	*35.7 (21.1)*	*35.9 (20.8)*	*36.4 (20.7)*	*36.6 (18.7)*
AOAC fibre density, g/1000kcal/day		*16.3 (6.4)*	*16.7 (6.7)*	*16.4 (7.4)*	*16.6 (7.2)*	*16.6 (7.0)*	*16.5 (6.7)*	*16.9 (6.1)*
Soluble fibre, g/day		*10.3 (5.5)*	*10.1 (4.9)*	*9.1 (6.2)*	*10.0 (5.3)*	*10.1 (5.5)*	*10.3 (5.5)*	*10.4 (5.0)*
Insoluble fibre, g/day		*15.3 (8.8)*	*14.6 (9.2)*	*13.5 (9.4)*	*14.7 (9.8)*	*14.7 (9.4)*	*14.9 (9.3)*	*15.3 (8.8)*
NSP within Foods, g/day	Total fruit	*4.0 (3.5)*	*3.7 (4.0)*	*4.0 (4.2)*	*3.8 (4.4)*	*4.0 (4.2)*	*3.9 (3.8)*	*4.2 (3.9)*
	Vegetables	*5.5 (3.8)*	*4.9 (4.4)*	*4.6 (3.3)*	*4.8 (3.8)*	*5.0 (3.8)*	*5.3 (3.9)*	*5.0 (3.7)*
	Total cereal foods	*7.4 (7.3)*	*6.8 (8.0)*	*6.8 (6.5)*	*6.9 (7.4)*	*7.1 (7.3)*	*6.8 (7.9)*	*7.6 (7.0)*
	Breakfast cereals	*1.7 (3.8)*	*1.8 (3.8)*	*2.1 (3.9)*	*2.0 (3.9)*	*1.9 (3.8)*	*1.8 (3.9)*	*1.8 (3.6)*
	Nuts & Seeds	*0.07 (0.2)*	*0.07 (0.3)*	*0.07 (0.2)*	*0.07 (0.2)*	*0.07 (0.2)*	*0.08 (0.3)*	*0.09 (0.3)*
	Legumes	*1.0 (1.1)*	*1.1(1.0)*	*0.8 (1.2)*	*0.9 (1.1)*	*0.9 (1.0)*	*1.1 (1.0)*	*1.1(1.3)*

* Pregnant, taking the contraceptive pill/ hormone replacement therapy.

Values are median (IQR) or numbers (percentages).

Table 6.6 Baseline cross-sectional characteristics for fatal plus non-fatal CHD and CHD subtype cases, non-fatal CHD cases and non cases

		Total MI (fatal plus non-fatal cases)	*Total ACS (fatal plus non-fatal cases)*	*Chronic heart disease (fatal plus non-fatal cases)*	*Other heart disease (fatal plus non-fatal cases)*	*Total CHD incidence†(fatal plus non-fatal cases)*	*Non-fatal CHD*	*Non CHD cases*
N		*236*	*392*	*573*	*38*	*812*	*714*	*26573*
Age, years		*61.2 (13.2)*	*60.3 (13.4)*	*59.2 (13.0)*	*54.2 (15.2)*	*59.4 (13.5)*	*58.6 (13.4)*	*50.0 (13.9)*
BMI, kg/m^2		*24.1 (5.1)*	*24.4 (5.4)*	*24.6 (5.7)*	*23.6 (5.0)*	*24.6 (5.6)*	*24.5 (5.6)*	*23.4 (4.6)*
Smoking status (%)	Current	*53 (23)*	*68 (17)*	*78 (14)*	*4 (11)*	*126 (16)*	*100 (14)*	*2773 (10)*
	Former	*72 (31)*	*120 (31)*	*181 (32)*	*14 (37)*	*261(32)*	*227 (23)*	*7986 (30)*
	Never smoked	*110 (47)*	*201 (52)*	*313 (54)*	*20 (53)*	*431 (53)*	*384 (54)*	*15771 (59)*
Diet group (%)	Meat	*184 (78)*	*297 (76)*	*433 (76)*	*24 (63)*	*607 (75)*	*530 (74)*	*17508 (66)*
	Fish	*26 (11)*	*45 (11)*	*66 (12)*	*6 (16)*	*100 (12)*	*86 (12)*	*3460 (13)*
	Vegetarian	*26 (11)*	*50 (13)*	*74 (13)*	*8 (21)*	*114 (14)*	*98 (14)*	*5605 (21)*
Socio-economic status NS-SEC (%)	Professional/ managerial	*124 (56)*	*213 (58)*	*313 (56)*	*20 (53)*	*445 (55)*	*384 (54)*	*16536 (62)*
	Intermediate	*79 (35)*	*126 (34)*	*179 (32)*	*4 (11)*	*264 (33)*	*231 (32)*	*7167 (27)*
	Routine and manual	*20 (9)*	*31 (8)*	*64 (12)*	*14 (37)*	*80 (10)*	*74 (10)*	*2337 (9)*
Highest educational achieve-ment (%)	No formal record	*62 (30)*	*100 (29)*	*138 (28)*	*9 (28)*	*201 (25)*	*171 (24)*	*3770 (14)*
	O-level	*57 (28)*	*102 (30)*	*149 (30)*	*13 (41)*	*221 (27)*	*198 (28)*	*7937 (30)*
	A-level	*44 (21)*	*69 (20)*	*103 (21)*	*9 (28)*	*153 (19)*	*138 (19)*	*5985 (23)*
	Degree	*43 (21)*	*73 (21)*	*100 (20)*	*1 (3)*	*135 (17)*	*115 (16)*	*6747(25)*
Menopause status (%)	Post	*156 (68)*	*241 (63)*	*328 (59)*	*17 (45)*	*476 (59)*	*399 (56)*	*8995 (34)*
	Pre	*28 (12)*	*49 (13)*	*77 (14)*	*12 (32)*	*118 (15)*	*112 (16)*	*11332 (43)*
	NA*	*47 (20)*	*92 (24)*	*152 (27)*	*9 (24)*	*204 (25)*	*187 (26)*	*5858 (22)*
History of hypertension at baseline (%)	Yes	*70 (30)*	*116 (30)*	*180 (31)*	*9 (24)*	*242 (30)*	*205 (29)*	*3755 (14)*
	No	*166 (70)*	*276 (70)*	*393 (69)*	*29 (76)*	*579 (71)*	*509 (71)*	*22818 (86)*
Ethanol g/day		*1.8 (10.5)*	*3.5 (10.6)*	*3.6 (9.8)*	*3.6 (11.4)*	*3.6 (10.4)*	*3.9 (10.9)*	*5.6 (11.7)*
Physical activity, MET-hours/week		*14.0 (13.0)*	*14.3 (14.0)*	*14.8 (13.4)*	*13.3 (11.3)*	*14.4 (13.2)*	*14.8 (13.3)*	*14.5 (13.1)*
Energy intake, kcal/day		*2215 (862)*	*2181 (876)*	*2202 (856)*	*2089 (886)*	*2198 (875)*	*2200 (881)*	*2178 (858)*
Saturated fat intake g/day		*27.8 (15.8)*	*26.7 (14.5)*	*26.9 (16.1)*	*26.9 (15.6)*	*26.7 (15.7)*	*26.6 (15.7)*	*26.8 (15.9)*

		Total MI (fatal plus non-fatal cases)	*Total ACS (fatal plus non-fatal cases)*	*Chronic heart disease (fatal plus non-fatal cases)*	*Other heart disease (fatal plus non-fatal cases)*	*Total CHD incidence†(fatal plus non-fatal cases)*	*Non-fatal CHD*	*Non CHD cases*
NSP, g/day		*21.5 (11.6)*	*23.1 (11.9)*	*24.2 (13.2)*	*23.0 (10.1)*	*24.0 (12.9)*	*24.2 (12.9)*	*23.9 (12.3)*
NSP density, g/1000kcal/day		*10.7 (4.6)*	*11.0 (4.5)*	*11.1 (4.2)*	*11.0 (3.5)*	*11.0 (4.3)*	*11.1 (4.2)*	*11.0 (4.2)*
AOAC fibre, g/day		*34.0 (17.5)*	*35.3 (17.9)*	*37.3 (19.7)*	*33.9 (13.3)*	*36.6 (19.2)*	*37.0 (19.6)*	*36.6 (18.7)*
AOAC fibre density, g/1000kcal/day		*16.3 (6.1)*	*16.6 (6.1)*	*16.9 (6.1)*	*17.0 (5.3)*	*16.9 (6.0)*	*17.0 (6.0)*	*16.9 (6.2)*
Soluble fibre, g/day		*9.8 (4.9)*	*10.2 (4.8)*	*10.4 (5.3)*	*10.3 (4.9)*	*10.4 (5.2)*	*10.5 (5.3)*	*10.4 (5.0)*
Insoluble fibre, g/day		*13.8 (8.3)*	*14.9 (9.0)*	*15.6 (9.2)*	*14.6 (7.4)*	*15.3 (9.0)*	*15.5 (9.1)*	*15.3 (8.8)*
NSP within Foods, g/day	Total fruit	*4.3 (3.6)*	*4.3 (4.1)*	*4.3 (3.9)*	*5.6 (4.2)*	*4.3 (4.2)*	*4.3 (4.3)*	*4.2 (3.9)*
	Vegetables	*5.0 (3.9)*	*5.1 (3.9)*	*5.3 (3.9)*	*5.3 (4.5)*	*5.2 (3.9)*	*5.3 (3.9)*	*4.9 (3.7)*
	Total cereal foods	*7.1 (6.8)*	*7.6 (7.4)*	*7.7 (7.1)*	*8.7 (4.6)*	*7.7 (7.1)*	*8.0 (7.2)*	*7.6 (7.0)*
	Breakfast cereals	*1.7 (3.6)*	*1.9 (4.4)*	*1.8 (4.4)*	*3.5 (5.2)*	*1.9 (4.4)*	*1.9 (4.4)*	*1.8 (3.6)*
	Nuts & Seeds	*0.08 (0.2)*	*0.08 (0.2)*	*0.08 (0.2)*	*0.07 (0.2)*	*0.08 (0.2)*	*0.08 (0.2)*	*0.09 (0.3)*
	Legumes	*0.9 (1.1)*	*0.9 (1.1)*	*0.9 (1.2)*	*1.2 (1.9)*	*1.0 (1.2)*	*1.0 (1.2)*	*1.1 (1.3)*

Values are median (IQR) or numbers (percentages).

* Pregnant, taking the contraceptive pill/ hormone replacement therapy.

†Total CHD incidence includes ACS, Chronic and Other cases.

Table 6.7 Baseline cross-sectional characteristics for non-fatal and total CVD (fatal plus non-fatal) cases in addition to non-cases

			Non-fatal CVD	*Total CVD (Fatal plus non-fatal events) †*	*Non-CVD cases*
N			*990*	*1162*	*26232*
Age, years			*59.3 (13.9)*	*60.4 (13.8)*	*49.9 (13.7)*
BMI, kg/m^2			*24.4 (5.1)*	*24.4 (5.1)*	*23.4 (4.6)*
Smoking status (%)	Current		*133 (13)*	*168 (14)*	*2731 (10)*
	Former		*312 (32)*	*352 (30)*	*7895 (30)*
	Never smoked		*541 (55)*	*639 (55)*	*15563 (59)*
Diet group (%)	Meat-eaters		*733 (74)*	*863 (74)*	*17252 (66)*
	Fish-eaters		*116 (12)*	*135 (12)*	*3425 (13)*
	Vegetarian		*141 (14)*	*164 (14)*	*5555 (21)*
Socio-economic status NS-SEC (%)	Professional/ managerial		*522 (53)*	*622 (54)*	*16359 (62)*
	Intermediate		*331 (33)*	*378 (33)*	*7053 (27)*
	Routine and manual		*101 (10)*	*117 (10)*	*2300 (9)*
Highest educational achievement (%)		No formal record	*242 (24)*	*292 (25)*	*3679 (14)*
		O-level	*256 (26)*	*292 (25)*	*7866 (30)*
		A-level	*198 (20)*	*221 (19)*	*5917 (23)*
		Degree	*158 (16)*	*193 (17)*	*6689 (25)*
Menopause status (%)		Post	*569 (57)*	*692 (60)*	*8779 (33)*
		Pre	*151 (15)*	*163 (14)*	*11287 (43)*
		Not applicable*	*246 (25)*	*274 (24)*	*5788 (22)*
History of hypertension at baseline (%)		Yes	*283 (29)*	*347 (30)*	*3650 (14)*
		No	*707 (71)*	*815 (70)*	*22582 (86)*
Ethanol g/day			*3.8 (10.6)*	*3.6 (10.5)*	*5.6 (11.7)*
Physical activity, MET-hours/week			*14.7 (13.1)*	*14.4 (13.3)*	*14.5 (13.1)*
Energy intake, kcal/day			*2187 (892)*	*2177 (885)*	*2178 (857)*
Saturated fat intake g/day			*26.8 (16.7)*	*26.8 (16.5)*	*26.8 (15.8)*
NSP, g/day			*24.0 (13.0)*	*23.7 (12.9)*	*23.9 (12.3)*
NSP density, g/1000kcal/day			*11.0 (4.4)*	*11.0 (4.3)*	*11.0 (4.2)*
AOAC fibre, g/day			*36.6 (19.9)*	*36.2 (19.8)*	*36.7 (18.7)*
AOAC fibre density, g/1000kcal/day			*16.9 (6.2)*	*16.7 (6.3)*	*16.9 (6.2)*
Soluble fibre, g/day			*10.4 (5.4)*	*10.2 (5.3)*	*10.4 (5.0)*
Insoluble fibre, g/day			*15.2 (9.3)*	*15.0 (9.2)*	*15.3 (8.8)*
NSP within Foods, g/day	Total fruit		*4.2 (4.2)*	*4.2 (4.1)*	*4.2 (3.9)*
	Vegetables		*5.2 (3.9)*	*5.2 (3.8)*	*5.0 (3.7)*
	Total cereal foods		*7.4 (7.3)*	*7.4 (7.1)*	*7.6 (7.0)*
	Breakfast cereals		*1.9 (4.3)*	*1.9 (4.1)*	*1.8 (3.6)*
	Nuts & Seeds		*0.08 (0.3)*	*0.08 (0.2)*	*0.10 (0.3)*
	Legumes		*1.0 (1.2)*	*1.0 (1.1)*	*1.1(1.3)*

Values are median (IQR) or numbers (percentages).

* Pregnant, taking the contraceptive pill/ hormone replacement therapy.

† CVD cases include heart disease or stroke cases combined.

6.4.3 Survival analysis

In total, 27,400 English residents, free from personal history of stroke or heart attacks were followed for over 14 years. During follow-up, and after excluding non eligible women, 821 incident (fatal plus non-fatal) CHD and 388 incident stroke cases were observed, with 760 CHD and 347 stroke cases available in fully adjusted models. When first stroke and CHD events were considered together and after median follow up of 14.3 years, 1162 CVD cases (1067 available in fully adjusted models) were observed.

HRs (95% CIs) are presented for categorised exposures and by fitting a linear trend over the categories for assessing dose-response trends in both age-adjusted and fully adjusted models. Total (fatal plus non-fatal) CHD, stroke or CVD events are presented in Table 6.8 and non-fatal events in Table 6.9. Risk estimates for stroke sub-types are presented in Tables 6.10 and 6.11 and for sub-types of CHD in Table 6.12.

Total CHD, stroke and CVD incidence (fatal plus non-fatal events)

All participants

Greater intake of total dietary fibre, assessed as NSP or using the AOAC method, higher fibre density and greater intake of soluble fibre, insoluble fibre and fibre from cereals were all associated with significantly lower risk for total stroke (Table 6.8). With each 6g/day increase in total NSP fibre, risk of total stroke reduced by 11%: HR 0.89 (95% CI: 0.81 to 0.99) p=0.03. For these different fibre exposures, risks tended to decrease with increasing intake categories and risk was significantly lower in each increasing intake group compared to the lowest consumers. For example with total NSP intake, risk of total stroke compared to the lowest consumers in Q2 was 0.65 (95% CI: 0.46 to 0.92), Q3 0.63 (95% CI: 0.44 to 0.90), Q4 0.63 (95% CI: 0.44 to 0.91) and in Q5 was 0.54 (95% CI: 0.35 to 0.85). With each 4g/day greater insoluble fibre intake the HR for total stroke was 0.90 (95% CI: 0.82 to 0.99) p=0.03 and with each 3g/day increase in cereal fibre HR 0.92 (95% CI: 0.85 to 1.00) p=0.04.

For total CVD, only fibre density assessed either using the NSP or AOAC method was associated with significantly lower risk: with every 2g/1000kcal/day higher NSP density, risk was 0.95 (95% CI: 0.91 to 1.00) p=0.03.

No significant associations were observed for any of the different fibre exposures and risk for total CHD.

Continuous risk estimates for models just adjusted for age, SES, alcohol, smoking status and MET-hours/week, without adjustment for BMI and energy intake, were not appreciably different from risk estimates for the fully-adjusted models (BMI and energy intake included). In most cases, statistical significance in multivariate models was similar to or slightly strengthened compared to the age-adjusted models.

Subgroups

In subgroups where a minimum of 50 cases were available for fully adjusted analyses, only one significant association, at the pre specified 1% significance level, was observed for stroke. With each 1g/day increase in fibre from legumes, there was 40% decreased risk of a total stroke HR 0.60 (95% CI: 0.41 to 0.87) p<0.01 in women who were classed as obese at baseline.

In women free of hypertension, legume fibre (per 1g/day) was associated with lower CHD risk 0.92 (95% CI: 0.85 to 0.99) p=0.02 in the fully adjusted model and this association was also apparent in the model without adjustment for energy intake and BMI, 0.93 (95% CI: 0.87 to 0.99) p=0.03. In the same subgroup, many of the different fibre exposures were inversely associated with CVD risk in both the mid and fully adjusted models. In those without hypertension, and in the fully adjusted models, CVD risk per 6g/day increase in total NSP was 0.94 (95% CI: 0.88 to 1.00) p=0.05; per 2g/1000kcal/day greater intake in NSP density 0.94 (95% CI: 0.89 to 1.00) p=0.02; per 11g/day greater AOAC fibre intake 0.91 (95% CI: 0.84 to 0.99) p=0.03; per 3g/1000kcal/day increase in AOAC density 0.93 (95% CI: 0.88 to 0.98) p=0.01; per 3g/day increase in soluble fibre 0.91 (95% CI: 0.84 to 0.99) p=0.03 and per 1g/day increase in legume fibre 0.94 (0.88 to 1.00) p=0.04.

In 7,723 participants that identified themselves as vegetarian at baseline, risk of total CHD (174 cases) and total stroke (77 cases) per 6g/day greater NSP intake was 0.96 (95% CI 0.85 to 1.09) and 1.08 (95% CI 0.90 to 1.30), respectively. In 18,794 participants that were reported meat-eaters at baseline, risk of total CHD (612 cases) and total stroke (297 cases) per 6g/day greater NSP intake was 0.98 (95% CI 0.91 to 1.05) and 0.88 (95% CI 0.79 to 0.98), respectively. No significant model interaction was observed by vegetarian status for CHD outcomes (p=0.56) or stroke outcomes (p=0.77) using the likelihood ratio test though power was low for this test of interaction.

Table 6.8 Dietary fibre intake and associated risk for total CHD, total stroke and total CVD

		Median intake (IQR)	Total CHD Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Total stroke Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Total CVD Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend
NSP (g/day)	Q1	14.1 (3.9)	163	1	1	87	1	1	243	1	1
	Q2	19.5 (2.2)	156	0.93 (0.74, 1.16)	0.99 (0.78, 1.25)	63	0.75 (0.55, 1.03)	0.65 (0.46, 0.92)	213	0.87 (0.73, 1.05)	0.87 (0.72, 1.06)
	Q3	23.8 (2.3)	138	0.81 (0.64, 1.01)	0.89 (0.68, 1.14)	63	0.69 (0.50, 0.95)	0.63 (0.44, 0.90)	192	0.75 (0.62, 0.91)	0.77 (0.62, 0.96)
	Q4	29.1 (3.1)	149	0.85 (0.68, 1.06)	0.88 (0.67, 1.14)	69	0.70 (0.51, 0.97)	0.63 (0.44, 0.91)	211	0.80 (0.66, 0.96)	0.79 (0.64, 0.98)
	Q5	38.1 (8.5)	154	0.90 (0.72, 1.13)	0.96 (0.71, 1.29)	65	0.71 (0.52, 0.98)	0.54 (0.35, 0.85)	208	0.82 (0.68, 0.99)	0.79 (0.61, 1.01)
	Per 6g/day		760	0.98 (0.94, 1.03) 0.44	0.98 (0.92, 1.05) 0.59	347	0.94 (0.88, 1.02) 0.13	0.89 (0.81, 0.99) 0.03	1067	0.97 (0.93, 1.01) 0.09	0.95 (0.90, 1.01) 0.09
AOAC (g/day)	Q1	21.8 (5.9)	166	1	1	90	1	1	250	1	1
	Q2	30.0 (3.4)	154	0.92 (0.74, 1.14)	0.96 (0.76, 1.22)	63	0.72 (0.52, 0.98)	0.64 (0.46, 0.91)	214	0.86 (0.72, 1.03)	0.86 (0.70, 1.04)
	Q3	36.6 (3.5)	131	0.75 (0.60, 0.95)	0.81 (0.62, 1.05)	58	0.57 (0.41, 0.79)	0.50 (0.34, 0.71)	180	0.67 (0.56, 0.82)	0.67 (0.54, 0.84)
	Q4	44.5 (4.7)	156	0.85 (0.68, 1.06)	0.89 (0.68, 1.17)	72	0.75 (0.55, 1.02)	0.66 (0.46, 0.95)	217	0.80 (0.67, 0.96)	0.79 (0.63, 0.99)
	Q5	58.5 (13.2)	153	0.92 (0.73, 1.15)	0.94 (0.69, 1.29)	64	0.65 (0.47, 0.90)	0.46 (0.29, 0.73)	206	0.80 (0.67, 0.97)	0.74 (0.56, 0.96)
	Per 11g/day		760	0.98 (0.93, 1.04) 0.48	0.98 (0.90, 1.06) 0.62	347	0.94 (0.86, 1.03) 0.16	0.87 (0.76, 0.99) 0.03	1067	0.96 (0.91, 1.01) 0.11	0.94 (0.88, 1.01) 0.09
NSP density g/1000 kcal/day	Q1	7.5 (1.5)	165	1	1	85	1	1	242	1	1
	Q2	9.4 (0.8)	148	0.84 (0.67, 1.05)	0.93 (0.74, 1.17)	75	0.89 (0.66, 1.20)	0.90 (0.66, 1.24)	216	0.86 (0.71, 1.03)	0.92 (0.76, 1.12)
	Q3	11.0 (0.8)	150	0.91 (0.73, 1.13)	0.97 (0.77, 1.23)	53	0.56 (0.40, 0.79)	0.61 (0.43, 0.88)	194	0.78 (0.64, 0.94)	0.84 (0.69, 1.03)
	Q4	12.7 (1.0)	141	0.81 (0.65, 1.01)	0.88 (0.70, 1.13)	58	0.64 (0.46, 0.88)	0.62 (0.43, 0.88)	193	0.76 (0.63, 0.91)	0.80 (0.65, 0.97)
	Q5	15.4 (2.4)	156	0.88 (0.71, 1.10)	0.97 (0.76, 1.24)	76	0.74 (0.54, 1.01)	0.78 (0.56, 1.09)	222	0.82 (0.68, 0.99)	0.90 (0.74, 1.10)
	2g/1000kcal/d		760	0.96 (0.91, 1.01) 0.08	0.98 (0.93, 1.03) 0.35	347	0.91 (0.85, 0.98) 0.01	0.92 (0.85, 0.99) 0.03	1067	0.94 (0.90, 0.98) <0.01	0.95 (0.91, 1.00) 0.04
AOAC density g/1000 kcal/day	Q1	11.7 (2.1)	158	1	1	87	1	1	238	1	1
	Q2	14.6 (1.2)	163	0.99 (0.80, 1.24)	1.05 (0.84, 1.31)	69	0.73 (0.54, 1.00)	0.73 (0.53, 1.01)	224	0.89 (0.74, 1.07)	0.93 (0.79, 1.12)
	Q3	16.9 (1.1)	141	0.88 (0.71, 1.11)	0.93 (0.73, 1.17)	54	0.58 (0.42, 0.81)	0.59 (0.42, 0.83)	188	0.76 (0.63, 0.92)	0.79 (0.65, 0.97)
	Q4	19.4 (1.4)	158	0.95 (0.76, 1.19)	1.03 (0.81, 1.30)	64	0.60 (0.43, 0.83)	0.65 (0.46, 0.91)	212	0.82 (0.68, 0.99)	0.88 (0.72, 1.07)
	Q5	23.4 (3.5)	140	0.84 (0.66, 1.06)	0.90 (0.70, 1.16)	73	0.69 (0.50, 0.94)	0.69 (0.49, 0.96)	205	0.77 (0.64, 0.93)	0.82 (0.67, 1.01)
	3g/1000kcal/d		760	0.95 (0.91, 1.00) 0.07	0.97 (0.93, 1.03) 0.35	347	0.91 (0.84, 0.98) 0.01	0.91 (0.84, 0.99) 0.03	1067	0.94 (0.90, 0.98) <0.01	0.95 (0.91, 1.00) 0.03

		Median intake (IQR)	Total CHD Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Total stroke Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Total CVD Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend
Soluble fibre (g/day)	Q1	6.4 (1.6)	168	1	1	88	1	1	250	1	1
	Q2	8.6 (0.9)	141	0.84 (0.67, 1.05)	0.86 (0.67, 1.09)	62	0.71 (0.52, 0.98)	0.67 (0.47, 0.95)	197	0.79 (0.65, 0.95)	0.78 (0.64, 0.95)
	Q3	10.4 (0.9)	142	0.84 (0.67, 1.05)	0.88 (0.69, 1.14)	59	0.73 (0.53, 1.00)	0.62 (0.43, 0.89)	195	0.79 (0.66, 0.95)	0.78 (0.63, 0.97)
	Q4	12.5 (1.2)	143	0.78 (0.62, 0.98)	0.85 (0.65, 1.11)	73	0.75 (0.55, 1.03)	0.68 (0.47, 0.99)	209	0.76 (0.63, 0.92)	0.78 (0.62, 0.97)
	Q5	16.3 (3.8)	166	0.97 (0.78, 1.21)	1.01 (0.75, 1.36)	65	0.71 (0.52, 0.98)	0.57 (0.36, 0.90)	216	0.84 (0.70, 1.01)	0.81 (0.62, 1.04)
	Per 3g/day		760	0.98 (0.93, 1.04) 0.55	0.97 (0.90, 1.05) 0.45	347	0.95 (0.87, 1.03) 0.22	0.88 (0.77, 1.00) 0.05	1067	0.97 (0.92, 1.01) 0.16	0.94 (0.88, 1.01) 0.09
Insoluble fibre (g/day)	Q1	8.4 (2.6)	165	1	1	89	1	1	248	1	1
	Q2	12.2 (1.6)	154	0.92 (0.74, 1.14)	0.96 (0.76, 1.21)	66	0.81 (0.60, 1.11)	0.71 (0.50, 0.99)	214	0.88 (0.74, 1.06)	0.87 (0.72, 1.05)
	Q3	15.3 (1.6)	143	0.82 (0.66, 1.03)	0.90 (0.70, 1.16)	60	0.68 (0.49, 0.94)	0.60 (0.42, 0.85)	197	0.77 (0.64, 0.93)	0.78 (0.64, 0.97)
	Q4	19.0 (2.2)	149	0.87 (0.69, 1.08)	0.90 (0.70, 1.16)	67	0.68 (0.49, 0.94)	0.64 (0.45, 0.91)	206	0.79 (0.66, 0.95)	0.79 (0.64, 0.97)
	Q5	25.5 (5.9)	149	0.88 (0.70, 1.11)	0.92 (0.69, 1.23)	65	0.72 (0.52, 0.98)	0.53 (0.34, 0.81)	202	0.80 (0.66, 0.96)	0.75 (0.59, 0.96)
	Per 4g/day		760	0.99 (0.94, 1.03) 0.51	0.99 (0.94, 1.05) 0.85	347	0.94 (0.88, 1.01) 0.10	0.90 (0.82, 0.99) 0.03	1067	0.97 (0.93, 1.00) 0.08	0.96 (0.91, 1.01) 0.12
Total cereal fibre (g/day)	Q1	2.8 (1.4)	165	1	1	84	1	1	240	1	1
	Q2	5.1 (1.1)	139	0.85 (0.68, 1.06)	0.90 (0.71, 1.14)	72	0.87 (0.63, 1.19)	0.83 (0.59, 1.16)	207	0.86 (0.71, 1.04)	0.89 (0.73, 1.08)
	Q3	7.6 (1.4)	146	0.85 (0.68, 1.06)	0.86 (0.67, 1.10)	60	0.66 (0.47, 0.92)	0.60 (0.41, 0.87)	200	0.78 (0.65, 0.94)	0.77 (0.62, 0.94)
	Q4	10.6 (1.8)	166	0.92 (0.74, 1.15)	1.01 (0.79, 1.29)	61	0.68 (0.49, 0.94)	0.56 (0.39, 0.81)	217	0.82 (0.68, 0.99)	0.85 (0.69, 1.04)
	Q5	15.6 (4.5)	144	0.84 (0.67, 1.05)	0.89 (0.68, 1.16)	70	0.76 (0.55, 1.04)	0.68 (0.46, 0.99)	203	0.80 (0.66, 0.96)	0.81 (0.65, 1.00)
	Per 3g/day		760	1.00 (0.95, 1.04) 0.84	1.01 (0.96, 1.06) 0.62	347	0.95 (0.89, 1.01) 0.11	0.92 (0.85, 1.00) 0.04	1067	0.98 (0.94, 1.01) 0.18	0.98 (0.94, 1.02) 0.37
Fibre from breakfast cereals (g/day)	Q1	0.05 (0.1)	152	1	1	71	1	1	219	1	1
	Q2	0.5 (0.4)	157	1.06 (0.85, 1.33)	1.02 (0.81, 1.30)	67	1.19 (0.85, 1.66)	1.11 (0.78, 1.59)	217	1.08 (0.89, 1.31)	1.04 (0.85, 1.27)
	Q3	1.8 (0.7)	132	0.85 (0.67, 1.08)	0.84 (0.65, 1.07)	64	0.96 (0.68, 1.36)	0.94 (0.65, 1.35)	188	0.86 (0.70, 1.05)	0.85 (0.69, 1.05)
	Q4	3.5 (0.7)	140	0.88 (0.70, 1.12)	0.89 (0.70, 1.14)	77	1.10 (0.79, 1.53)	1.06 (0.75, 1.51)	206	0.90 (0.74, 1.10)	0.92 (0.75, 1.12)
	Q5	7.6 (2.6)	179	1.04 (0.83, 1.29)	1.06 (0.84, 1.34)	68	0.93 (0.67, 1.30)	0.84 (0.59, 1.20)	237	0.98 (0.81, 1.18)	0.97 (0.80, 1.19)
	Per 2g/day		760	1.00 (0.96, 1.05) 0.84	1.02 (0.97, 1.06) 0.46	347	0.97 (0.92, 1.03) 0.36	0.96 (0.90, 1.02) 0.19	1067	0.99 (0.95, 1.02) 0.46	0.99 (0.95, 1.03) 0.70

		Median intake (IQR)	Total CHD Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Total stroke Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Total CVD Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend
Fruit fibre (g/day)	Q1	1.4 (0.9)	157	1	1	77	1	1	228	1	1
	Q2	2.9 (0.7)	149	0.82 (0.65, 1.03)	0.88 (0.70, 1.12)	66	0.69 (0.49, 0.96)	0.75 (0.53, 1.06)	209	0.80 (0.66, 0.97)	0.86 (0.70, 1.04)
	Q3	4.2 (0.7)	146	0.74 (0.58, 0.93)	0.82 (0.64, 1.04)	68	0.79 (0.58, 1.09)	0.76 (0.53, 1.07)	204	0.75 (0.62, 0.90)	0.78 (0.64, 0.96)
	Q4	5.8 (1.1)	143	0.78 (0.62, 0.97)	0.84 (0.65, 1.07)	66	0.63 (0.45, 0.87)	0.67 (0.47, 0.96)	202	0.74 (0.61, 0.89)	0.78 (0.63, 0.96)
	Q5	9.4 (4.0)	168	0.83 (0.67, 1.04)	0.94 (0.73, 1.19)	70	0.62 (0.45, 0.86)	0.65 (0.45, 0.93)	224	0.74 (0.61, 0.90)	0.81 (0.66, 1.00)
	Per 2g/day		760	0.98 (0.94, 1.02) 0.38	0.99 (0.95, 1.04) 0.79	347	0.96 (0.90, 1.04) 0.32	0.97 (0.90, 1.05) 0.47	1067	0.97 (0.94, 1.01) 0.12	0.98 (0.95, 1.02) 0.38
Vegetable fibre (g/day)	Q1	2.3 (0.9)	155	1	1	80	1	1	225	1	1
	Q2	3.7 (0.6)	127	0.78 (0.61, 0.99)	0.82 (0.64, 1.05)	50	0.69 (0.49, 0.97)	0.68 (0.48, 0.99)	173	0.76 (0.62, 0.92)	0.78 (0.63, 0.96)
	Q3	5.0 (0.7)	156	0.93 (0.74, 1.16)	0.99 (0.78, 1.26)	67	0.84 (0.61, 1.15)	0.85 (0.60, 1.20)	217	0.91 (0.75, 1.09)	0.96 (0.78, 1.17)
	Q4	6.6 (1.0)	159	0.89 (0.71, 1.12)	0.96 (0.76, 1.22)	82	0.84 (0.62, 1.15)	0.94 (0.67, 1.30)	234	0.88 (0.73, 1.06)	0.97 (0.79, 1.18)
	Q5	9.6 (3.0)	163	0.92 (0.74, 1.15)	0.94 (0.73, 1.20)	68	0.71 (0.51, 0.97)	0.67 (0.48, 0.95)	218	0.84 (0.70, 1.01)	0.85 (0.69, 1.05)
	Per 2g/day		760	0.99 (0.95, 1.04) 0.78	0.99 (0.94, 1.03) 0.57	347	0.97 (0.90, 1.04) 0.40	0.97 (0.90, 1.04) 0.40	1067	0.98 (0.95, 1.02) 0.35	0.98 (0.94, 1.02) 0.34
Legume fibre (g/day)	Q1	0.2 (0.2)	184	1	1	82	1	1	253	1	1
	Q2	0.7 (0.2)	162	0.90 (0.73, 1.11)	0.86 (0.69, 1.08)	87	1.06 (0.79, 1.43)	1.14 (0.83, 1.56)	242	0.97 (0.81, 1.15)	0.97 (0.81, 1.16)
	Q3	1.1 (0.2)	146	0.88 (0.71, 1.10)	0.87 (0.70, 1.10)	70	1.00 (0.74, 1.37)	1.03 (0.73, 1.44)	213	0.95 (0.79, 1.14)	0.96 (0.79, 1.16)
	Q4	1.6 (0.4)	142	0.98 (0.78, 1.22)	0.96 (0.76, 1.22)	58	0.84 (0.60, 1.19)	0.96 (0.67, 1.39)	188	0.93 (0.76, 1.12)	0.96 (0.78, 1.17)
	Q5	3.6 (1.4)	126	0.91 (0.72, 1.16)	0.87 (0.68, 1.12)	50	1.03 (0.73, 1.45)	0.94 (0.63, 1.41)	171	0.97 (0.79, 1.18)	0.92 (0.74, 1.15)
	Per 1g/day		760	0.97 (0.92, 1.02) 0.22	0.96 (0.91, 1.02) 0.17	347	1.00 (0.92, 1.08) 0.93	0.96 (0.88, 1.06) 0.46	1067	0.98 (0.94, 1.03) 0.39	0.97 (0.92, 1.02) 0.18
Fibre from nuts and seeds (g/day)	Q1	0 (0.01)	186	1	1	115	1	1	288	1	1
	Q2	0.06 (0.01)	173	1.07 (0.87, 1.31)	1.09 (0.88, 1.36)	63	0.68 (0.51, 0.92)	0.66 (0.48, 0.92)	226	0.93 (0.78, 1.10)	0.94 (0.78, 1.13)
	Q3	0.08 (0.05)	146	0.98 (0.79, 1.23)	1.05 (0.84, 1.33)	59	0.64 (0.46, 0.88)	0.70 (0.50, 0.99)	200	0.86 (0.72, 1.04)	0.94 (0.78, 1.14)
	Q4	0.28 (0.12)	131	0.95 (0.75, 1.19)	1.01 (0.79, 1.28)	47	0.59 (0.42, 0.83)	0.63 (0.44, 0.91)	174	0.82 (0.68, 1.00)	0.88 (0.72, 1.08)
	Q5	0.87 (0.92)	124	0.79 (0.62, 1.00)	0.87 (0.67, 1.13)	63	0.65 (0.47, 0.89)	0.70 (0.49, 1.00)	179	0.74 (0.61, 0.89)	0.81 (0.65, 1.00)
	Per 0.2g/day		760	0.98 (0.96, 1.01) 0.19	0.99 (0.97, 1.02) 0.46	347	0.96 (0.93, 1.00) 0.06	0.97 (0.94, 1.01) 0.17	1067	0.98 (0.96, 1.00) 0.06	0.99 (0.97, 1.01) 0.24

[1]Case numbers apply to fully-adjusted models. [2]Adjustments include Age (years), BMI (kg/m^2), calories from carbohydrate, fat and protein (kcal/day), ethanol intake (g/day), MET (hours/week), smoking status (current vs. not current smoker), socio-economic status (professional or managerial/ intermediate/ routine or manual). Note: adjustment for energy intake was not included in fibre density models. Highlight=CIs do not span 1 in fully adjusted model.

Non-fatal CHD, non-fatal stroke and non-fatal CVD

All participants

Risk of non-fatal coronary or stroke events in relation to the different fibre exposures in the full sample of women are shown in Table 6.9. After median follow up of 14.7 years and excluding non-eligible participants, 668 non-fatal CHD cases were available for fully-adjusted models. Slightly fewer stroke cases were available for fully adjusted models (258) and when CHD and stroke cases were considered together, 914 non-fatal CVD cases were available for fully adjusted models.

For stroke events, estimates indicated roughly 20 to 50% lower risk across many of the intake categories compared to the lowest intake level for total fibre, insoluble and soluble fibre. However, only total cereal fibre intake appeared significantly associated with lower stroke risk in the dose-response model HR 0.93 (95% CI: 0.83 to 0.99) p=0.03 per 3 g/day increase.

As with total (fatal plus non-fatal) CHD, no apparent association was seen with many exposures for CHD risk, the only exception being legume fibre. Non-fatal CHD risk per 1g/day greater legume fibre was 0.94 (95% CI: 0.88 to 1.00) p=0.05.

There was no evidence of an association between fibre intake and non-fatal CVD risk in the linear dose-response models and risk was significantly associated with greater fibre intake in only a handful of the category comparisons. Non-fatal CVD was significantly associated with higher fibre density, both assessed as NSP and AOAC, in the intermediate model (energy intake and BMI not included as covariates). With each 2g/1000kcal/day greater NSP density, risk was 0.95 (95% CI: 0.91 to 1.00) p=0.03 and the risk per 3g/1000kcal/day AOAC fibre density 0.95 (95% CI: 0.91 to 1.00) p=0.04.

Subgroups

In the various subgroups of women few associations existed for risk of non-fatal events in dose-response models. No associations were apparent for non-fatal CHD and for non-fatal stroke just a few of the risk estimates did not span the line of no effect, suggesting some evidence of an association, but did not reach the 1% significance criterion and are described below.

As in the full sample, total cereal fibre intake remained protectively associated with non-fatal stroke in women whose baseline BMI was in the healthy range (18.5-25kg/m^2). In women with history of hypertension, lower non-fatal stroke risk was associated with greater intake of cereal fibre (per 3g/day) 0.81 (95% CI: 0.68 to 0.98) p=0.03.

Several of the fibre exposures were significantly associated with lower non-fatal CVD risk in women with no history of hypertension: NSP (per 6/day) 0.92 (95% CI: 0.85 to 0.98) p=0.01; NSP density (per 2g/1000kcal/day) 0.91 (95% CI: 0.87 to 0.97) p<0.01; AOAC (per 11g/day) 0.89 (95% CI: 0.82 to 0.97) p=0.01; AOAC density (per3g/1000kcal/day) 0.91 (95% CI: 0.86 to 0.96) p<0.01; soluble fibre (per 3g/day) 0.91 (95% CI: 0.84 to 0.99) p=0.04; insoluble fibre (per 4g/day) 0.92 (95% CI: 0.86 to 0.98) p=0.01; legume fibre (per 1g/day) 0.92 (95% CI: 0.86 to 0.99) p=0.02.

Table 6.9 Dietary fibre intake and associated risk for non-fatal CHD, non-fatal stroke and non-fatal CVD

		Median intake (IQR)	Non-fatal CHD			Non-fatal stroke			Non fatal CVD		
			Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend
NSP (g/day)	Q1	14.1 (3.9)	143	1	1	61	1	1	202	1	1
	Q2	19.5 (2.2)	132	0.91 (0.72, 1.16)	0.94 (0.73, 1.21)	45	0.76 (0.53, 1.11)	0.63 (0.42, 0.95)	176	0.87 (0.71, 1.06)	0.86 (0.69, 1.06)
	Q3	23.8 (2.3)	122	0.83 (0.65, 1.05)	0.86 (0.66, 1.14)	47	0.77 (0.53, 1.12)	0.67 (0.44, 1.00)	162	0.78 (0.63, 0.96)	0.79 (0.62, 0.99)
	Q4	29.1 (3.1)	131	0.88 (0.69, 1.12)	0.86 (0.65, 1.14)	55	0.79 (0.55, 1.15)	0.69 (0.46, 1.04)	185	0.86 (0.70, 1.05)	0.84 (0.66, 1.06)
	Q5	38.1 (8.5)	140	0.94 (0.74, 1.19)	0.94 (0.68, 1.29)	50	0.73 (0.50, 1.06)	0.50 (0.31, 0.83)	189	0.87 (0.71, 1.07)	0.82 (0.63, 1.07)
	Per 6g/day		668	1.00 (0.95, 1.05) 0.90	0.99 (0.93, 1.05) 0.73	184	0.96 (0.88, 1.04) 0.31	0.90 (0.81, 1.01) 0.08	914	0.98 (0.94, 1.03) 0.44	0.97 (0.92, 1.03) 0.30
AOAC (g/day)	Q1	21.8 (5.9)	143	1	1	64	1	1	207	1	1
	Q2	30.0 (3.4)	131	0.90 (0.71, 1.15)	0.92 (0.71, 1.19)	42	0.68 (0.46, 0.99)	0.58 (0.39, 0.89)	172	0.83 (0.68, 1.02)	0.82 (0.66, 1.02)
	Q3	36.6 (3.5)	119	0.79 (0.62, 1.01)	0.82 (0.62, 1.09)	42	0.60 (0.41, 0.88)	0.49 (0.32, 0.75)	156	0.71 (0.57, 0.87)	0.70 (0.55, 0.88)
	Q4	44.5 (4.7)	135	0.89 (0.70, 1.13)	0.89 (0.66, 1.18)	61	0.89 (0.63, 1.26)	0.76 (0.50, 1.15)	192	0.87 (0.71, 1.06)	0.85 (0.67, 1.08)
	Q5	58.5 (13.2)	140	0.97 (0.76, 1.23)	0.95 (0.68, 1.32)	49	0.65 (0.45, 0.95)	0.43 (0.25, 0.72)	187	0.85 (0.70, 1.05)	0.78 (0.59, 1.03)
	Per 11g/day		668	1.00 (0.94, 1.06) 0.96	0.99 (0.91, 1.07) 0.76	184	0.96 (0.87, 1.06) 0.44	0.89 (0.77, 1.03) 0.12	914	0.98 (0.93, 1.04) 0.52	0.96 (0.90, 1.04) 0.33
NSP density g/1000 kcal/day	Q1	7.5 (1.5)	139	1	1	63	1	1	199	1	1
	Q2	9.4 (0.8)	134	0.90 (0.71, 1.14)	0.98 (0.76, 1.25)	56	0.95 (0.67, 1.34)	0.95 (0.65, 1.37)	189	0.93 (0.76, 1.13)	0.98 (0.80, 1.20)
	Q3	11.0 (0.8)	131	0.93 (0.73, 1.18)	0.99 (0.77, 1.27)	40	0.39 (0.39, 0.86)	0.63 (0.41, 0.95)	164	0.80 (0.65, 0.99)	0.86 (0.69, 1.07)
	Q4	12.7 (1.0)	126	0.85 (0.67, 1.09)	0.94 (0.72, 1.21)	43	0.45 (0.45, 0.95)	0.63 (0.42, 0.96)	167	0.79 (0.64, 0.97)	0.84 (0.68, 1.05)
	Q5	15.4 (2.4)	138	0.91 (0.72, 1.16)	1.00 (0.77, 1.29)	56	0.49 (0.49, 1.02)	0.73 (0.50, 1.08)	195	0.86 (0.70, 1.06)	0.94 (0.76, 1.17)
	2g/1000kcal/d		668	0.96 (0.91, 1.01) 0.10	0.98 (0.93, 1.03) 0.41	184	0.91 (0.84, 0.99) 0.03	0.92 (0.84, 1.00) 0.06	914	0.95 (0.91,0.99) 0.01	0.96 (0.92, 1.01) 0.12
AOAC density g/1000 kcal/day	Q1	11.7 (2.1)	131	1	1	63	1	1	193	1	1
	Q2	14.6 (1.2)	147	1.07 (0.85, 1.36)	1.11 (0.87, 1.41)	52	0.79 (0.56, 1.13)	0.77 (0.53, 1.13)	194	0.96 (0.79, 1.16)	0.98 (0.79, 1.20)
	Q3	16.9 (1.1)	121	0.90 (0.70, 1.15)	0.94 (0.73, 1.22)	41	0.63 (0.43, 0.92)	0.64 (0.43, 0.95)	159	0.80 (0.65, 0.98)	0.83 (0.67, 1.03)
	Q4	19.4 (1.4)	146	1.04 (0.82, 1.33)	1.12 (0.87, 1.44)	47	0.61 (0.42, 0.89)	0.64 (0.43, 0.96)	189	0.89 (0.72, 1.08)	0.95 (0.77, 1.18)
	Q5	23.4 (3.5)	123	0.86 (0.67, 1.11)	0.93 (0.71, 1.22)	55	0.68 (0.47, 0.98)	0.67 (0.46, 0.99)	179	0.81 (0.65, 0.99)	0.86 (0.69, 1.08)
	3g/1000kcal/d		668	0.96 (0.91, 1.01) 0.09	0.98 (0.93, 1.03) 0.38	184	0.92 (0.84, 1.00) 0.05	0.92 (0.84, 1.01) 0.09	914	0.94 (0.90, 0.99) 0.01	0.96 (0.92, 1.01) 0.12

		Median intake (IQR)	Non-fatal CHD Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend		Non-fatal stroke Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend		Non fatal CVD Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend	
Soluble fibre (g/day)	Q1	6.4 (1.6)	144	1		1		64	1		1		207	1		1	
	Q2	8.6 (0.9)	122	0.85 (0.66, 1.08)		0.85 (0.66, 1.10)		38	0.64 (0.43, 0.94)		0.56 (0.37, 0.86)		160	0.78 (0.63, 0.96)		0.76 (0.61, 0.95)	
	Q3	10.4 (0.9)	126	0.87 (0.69, 1.11)		0.89 (0.68, 1.16)		46	0.77 (0.53, 1.11)		0.64 (0.42, 0.95)		168	0.82 (0.67, 1.01)		0.80 (0.64, 1.00)	
	Q4	12.5 (1.2)	126	0.81 (0.64, 1.04)		0.85 (0.64, 1.13)		58	0.83 (0.58, 1.19)		0.74 (0.49, 1.12)		181	0.80 (0.65, 0.98)		0.81 (0.64, 1.03)	
	Q5	16.3 (3.8)	150	1.02 (0.81, 1.29)		1.00 (0.73, 1.37)		52	0.72 (0.50, 1.05)		0.56 (0.34, 0.92)		198	0.90 (0.74, 1.10)		0.85 (0.65, 1.12)	
	Per 3g/day		668	1.00 (0.95, 1.06)	0.94	0.98 (0.91, 1.06)	0.68	184	0.97 (0.87, 1.07)	0.50	0.90 (0.78, 1.04)	0.17	914	0.99 (0.94, 1.04)	0.59	0.96 (0.90, 1.03)	0.30
Insoluble fibre (g/day)	Q1	8.4 (2.6)	143	1		1		63	1		1		206	1		1	
	Q2	12.2 (1.6)	130	0.90 (0.70, 1.14)		0.92 (0.71, 1.18)		48	0.84 (0.58, 1.20)		0.68 (0.46, 1.01)		175	0.87 (0.71, 1.06)		0.84 (0.68, 1.04)	
	Q3	15.3 (1.6)	132	0.88 (0.69, 1.12)		0.94 (0.72, 1.23)		44	0.72 (0.50, 1.06)		0.61 (0.40, 0.92)		172	0.81 (0.66, 0.99)		0.82 (0.66, 1.03)	
	Q4	19.0 (2.2)	128	0.90 (0.71, 1.15)		0.89 (0.67, 1.16)		53	0.77 (0.53, 1.12)		0.68 (0.46, 1.02)		177	0.84 (0.69, 1.03)		0.82 (0.65, 1.03)	
	Q5	25.5 (5.9)	135	0.92 (0.72, 1.17)		0.91 (0.67, 1.24)		50	0.73 (0.51, 1.07)		0.49 (0.30, 0.79)		184	0.85 (0.70, 1.05)		0.79 (0.61, 1.02)	
	Per 4g/day		668	1.00 (0.95, 1.05)	0.97	1.00 (0.94, 1.06)	0.95	184	0.95 (0.88, 1.03)	0.25	0.91 (0.82, 1.02)	0.07	914	0.98 (0.94, 1.02)	0.42	0.98 (0.92, 1.03)	0.36
Total cereal fibre (g/day)	Q1	2.8 (1.4)	144	1		1		63	1		1		203	1		1	
	Q2	5.1 (1.1)	125	0.87 (0.68, 1.11)		0.91 (0.71, 1.18)		58	0.90 (0.63, 1.28)		0.85 (0.58, 1.25)		182	0.88 (0.72, 1.08)		0.91 (0.74, 1.13)	
	Q3	7.6 (1.4)	126	0.84 (0.66, 1.07)		0.84 (0.65, 1.09)		37	0.54 (0.36, 0.81)		0.46 (0.29, 0.73)		164	0.76 (0.62, 0.94)		0.74 (0.59, 0.93)	
	Q4	10.6 (1.8)	144	0.96 (0.76, 1.21)		1.02 (0.79, 1.32)		45	0.65 (0.44, 0.95)		0.53 (0.34, 0.81)		183	0.84 (0.69, 1.03)		0.85 (0.68, 1.07)	
	Q5	15.6 (4.5)	129	0.88 (0.69, 1.12)		0.88 (0.67, 1.17)		55	0.77 (0.54, 1.11)		0.64 (0.41, 0.99)		182	0.85 (0.70, 1.04)		0.83 (0.66, 1.06)	
	Per 3g/day		668	1.01 (0.96, 1.05)	0.90	1.01 (0.96, 1.06)	0.77	184	0.94 (0.87, 1.02)	0.13	0.91 (0.83, 0.99)	0.03	914	0.99 (0.95, 1.03)	0.49	0.98 (0.94, 1.03)	0.49
Fibre from breakfast cereals (g/day)	Q1	0.05 (0.1)	131	1		1		56	1		1		185	1		1	
	Q2	0.5 (0.4)	133	1.02 (0.79, 1.30)		1.00 (0.77, 1.29)		51	1.18 (0.80, 1.73)		1.07 (0.71, 1.62)		184	1.07 (0.87, 1.32)		1.03 (0.83, 1.29)	
	Q3	1.8 (0.7)	119	0.85 (0.66, 1.10)		0.85 (0.65, 1.11)		46	0.90 (0.60, 1.34)		0.87 (0.57, 1.33)		160	0.85 (0.69, 1.06)		0.85 (0.67, 1.06)	
	Q4	3.5 (0.7)	129	0.91 (0.71, 1.16)		0.94 (0.73, 1.22)		54	0.99 (0.67, 1.45)		0.95 (0.63, 1.42)		182	0.93 (0.76, 1.15)		0.95 (0.77, 1.19)	
	Q5	7.6 (2.6)	156	1.05 (0.83, 1.33)		1.05 (0.82, 1.35)		51	0.88 (0.60, 1.28)		0.75 (0.50, 1.12)		203	0.99 (0.81, 1.22)		0.97 (0.78, 1.20)	
	Per 2g/day		668	1.01 (0.96, 1.05)	0.82	1.01 (0.97, 1.06)	0.62	184	0.97 (0.90, 1.04)	0.34	0.95 (0.87, 1.02)	0.16	914	0.99 (0.96, 1.03)	0.77	1.00 (0.96, 1.04)	0.84

		Median intake (IQR)	Non-fatal CHD Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend		Non-fatal stroke Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend		Non fatal CVD Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend	
Fruit fibre (g/day)	Q1	1.4 (0.9)	134	1		1		55	1		1		188	1		1	
	Q2	2.9 (0.7)	127	0.89 (0.69, 1.13)		0.90 (0.70, 1.17)		54	0.82 (0.56, 1.19)		0.87 (0.59, 1.31)		179	0.86 (0.70, 1.06)		0.90 (0.72, 1.11)	
	Q3	4.2 (0.7)	125	0.76 (0.59, 0.98)		0.82 (0.63, 1.07)		48	0.85 (0.59, 1.23)		0.76 (0.50, 1.14)		170	0.78 (0.63, 0.96)		0.80 (0.64, 0.99)	
	Q4	5.8 (1.1)	127	0.86 (0.67, 1.10)		0.88 (0.68, 1.15)		52	0.71 (0.48, 1.04)		0.71 (0.47, 1.07)		178	0.81 (0.66, 1.00)		0.83 (0.67, 1.04)	
	Q5	9.4 (4.0)	155	0.95 (0.75, 1.21)		1.01 (0.78, 1.32)		49	0.59 (0.39, 0.87)		0.58 (0.38, 0.90)		199	0.81 (0.66, 1.00)		0.87 (0.69, 1.09)	
	Per 2g/day		668	1.00 (0.96, 1.04)	0.90	1.00 (0.96, 1.05)	0.84	184	0.97 (0.89, 1.06)	0.46	0.97 (0.88, 1.06)	0.47	914	0.99 (0.95, 1.02)	0.46	0.99 (0.95, 1.03)	0.76
Vegetable fibre (g/day)	Q1	2.3 (0.9)	132	1		1		56	1		1		188	1		1	
	Q2	3.7 (0.6)	190	0.78 (0.61, 1.02)		0.82 (0.62, 1.07)		29	0.56 (0.37, 0.87)		0.57 (0.36, 0.91)		135	0.69 (0.56, 0.87)		0.72 (0.57, 0.91)	
	Q3	5.0 (0.7)	140	1.00 (0.78, 1.27)		1.03 (0.80, 1.33)		53	0.99 (0.69, 1.43)		0.99 (0.66, 1.47)		192	0.98 (0.80, 1.20)		1.00 (0.81, 1.24)	
	Q4	6.6 (1.0)	142	0.94 (0.74, 1.20)		1.00 (0.77, 1.29)		62	0.90 (0.62, 1.29)		0.99 (0.67, 1.45)		202	0.91 (0.74, 1.11)		0.98 (0.79, 1.21)	
	Q5	9.6 (3.0)	145	0.95 (0.75, 1.21)		0.95 (0.73, 1.25)		58	0.82 (0.57, 1.19)		0.82 (0.55, 1.20)		197	0.88 (0.72, 1.08)		0.88 (0.70, 1.02)	
	Per 2g/day		668	1.00 (0.96, 1.05)	0.86	0.99 (0.95, 1.04)	0.83	184	1.00 (0.93, 1.08)	0.99	1.01 (0.93, 1.09)	0.85	914	1.00 (0.96, 1.04)	0.94	1.00 (0.96, 1.04)	0.83
Legume fibre (g/day)	Q1	0.2 (0.2)	164	1		1		54	1		1		213	1		1	
	Q2	0.7 (0.2)	144	0.90 (0.72, 1.12)		0.84 (0.66, 1.05)		63	1.15 (0.81, 1.63)		1.22 (0.84, 1.79)		205	0.98 (0.81, 1.18)		0.94 (0.77, 1.15)	
	Q3	1.1 (0.2)	122	0.84 (0.66, 1.06)		0.78 (0.62, 1.00)		55	1.13 (0.79, 1.63)		1.14 (0.76, 1.71)		176	0.92 (0.75, 1.12)		0.89 (0.72, 1.09)	
	Q4	1.6 (0.4)	129	0.96 (0.76, 1.22)		0.91 (0.71, 1.16)		45	0.93 (0.62, 1.39)		1.05 (0.68, 1.62)		169	0.95 (0.77, 1.17)		0.95 (0.76, 1.18)	
	Q5	3.6 (1.4)	109	0.87 (0.68, 1.12)		0.79 (0.61, 1.03)		41	1.18 (0.79, 1.75)		1.13 (0.71, 1.81)		151	0.97 (0.79, 1.21)		0.91 (0.72, 1.14)	
	Per 1g/day		668	0.96 (0.90, 1.02)	0.15	0.94 (0.88, 1.00)	0.05	184	1.01 (0.93, 1.10)	0.75	1.00 (0.90, 1.10)	0.96	914	0.98 (0.93, 1.03)	0.40	0.97 (0.92, 1.02)	0.19
Fibre from nuts and seeds (g/day)	Q1	0 (0.01)	154	1		1		79	1		1		231	1		1	
	Q2	0.06 (0.01)	153	1.15 (0.92, 1.44)		1.16 (0.92, 1.47)		45	0.67 (0.47, 0.96)		0.66 (0.45, 0.97)		194	0.97 (0.80, 1.17)		0.98 (0.80, 1.19)	
	Q3	0.08 (0.05)	135	1.07 (0.84, 1.35)		1.13 (0.88, 1.45)		46	0.72 (0.50, 1.03)		0.80 (0.54, 1.19)		178	0.93 (0.76, 1.14)		1.01 (0.82, 1.24)	
	Q4	0.28 (0.12)	117	1.02 (0.80, 1.31)		1.06 (0.81, 1.37)		36	0.64 (0.43, 0.96)		0.69 (0.45, 1.05)		152	0.88 (0.71, 1.08)		0.93 (0.74, 1.16)	
	Q5	0.87 (0.92)	109	0.84 (0.65, 1.09)		0.89 (0.68, 1.19)		52	0.74 (0.51, 1.06)		0.81 (0.54, 1.22)		159	0.80 (0.65, 0.98)		0.86 (0.69, 1.09)	
	Per 0.2g/day		668	0.99 (0.96, 1.01)	0.41	0.99 (0.97, 1.02)	0.66	184	0.98 (0.94, 1.02)	0.27	0.99 (0.95, 1.02)	0.46	914	0.99 (0.97, 1.01)	0.22	0.99 (0.97, 1.01)	0.52

[1]Case numbers apply to fully-adjusted models. [2]Adjustments include Age (years), BMI (kg/m^2), calories from carbohydrate, fat and protein (kcal/day), ethanol intake (g/day), MET (hours/week), smoking status (current vs. not current smoker), socio-economic status (professional or managerial/ intermediate/ routine or manual). Note: adjustment for energy intake was not included in fibre density models. Highlight=CIs do not span 1 in fully adjusted model.

Haemorrhagic, ischaemic and unspecified stroke incidence (fatal plus non-fatal events)

All participants

After a median follow up of 14.4 years, 135 haemorrhagic and 184 ischaemic strokes were observed. There were 138 cases where the type of stroke was not detailed in records.

Total fibre intake, insoluble fibre, soluble fibre and vegetable fibre were all significantly associated with lower risk of unspecified strokes in the fully-adjusted dose-response models (Table 6.10). Each 6g/day increase in NSP was associated with 24% lower risk HR 0.76 (95% CI: 0.63 to 0.92) p<0.01 and with each 2g/day increase in vegetable fibre HR 0.80 (95% CI: 0.68 to 0.92) p<0.01.

The majority of risk estimates for haemorrhagic and ischaemic stroke indicated a protective association but CIs were generally wide and no significant associations were observed in the fully-adjusted models for dose-response associations except with cereal fibre. For each 3g/day increase in cereal fibre, risk of ischaemic stroke was 0.89 (95% CI: 0.80 to 1.00) p=0.05.

When ischaemic and unspecified strokes were combined 284 cases were observed, with 251 cases being available for use in the fully adjusted models. Risk estimates for 'mostly ischaemic' type stroke, where ischaemic and unspecified cases were combined, are presented in Table 6.11. Risk estimates from the dose-response models in this larger subgroup of stroke cases largely reflect those seen for the 'unspecified type' stroke. The risk estimates for 'mostly ischaemic' strokes tended to be slightly weaker compared to those seen for the unspecified strokes only, but CIs were narrower on the whole in this larger category. For example, with total fibre intake, assessed as AOAC, risk was 0.74 (95% CI: 0.57 to 0.94) with each 11g/day greater intake for unspecified stroke and 0.80 (95% CI: 0.68 to 0.95) for 'mostly ischaemic' strokes.

There was lower risk for 'mostly ischaemic' type stroke of between 12 and 16% with greater intake of total fibre (NSP and AOAC), fibre density, soluble and insoluble fibre. A protective association was also seen for 'mostly ischaemic' stroke risk with greater cereal fibre intake (per 3g/day) HR 0.88 (95% CI: 0.80 to 0.96) p<0.01. An association was also seen with cereal fibre and ischaemic stroke risk, 0.89 (95% CI: 0.80 to 1.00) p=0.05 but with unspecified stroke type CIs span just over the line of no effect, 0.89 (95% CI: 0.78 to 1.02) p=0.11.

A protective association observed with greater vegetable fibre intake in the unspecified stroke cases (per 2g/day) 0.80 (95% CI: 0.68 to 0.92) p<0.01, disappeared when combined with the ischaemic cases 0.92 (95% CI: 0.84 to 1.02) p=0.13.

Subgroups

In the postmenopausal subgroup, vegetable fibre per 2g/day increase was associated with an increased risk of haemorrhagic stroke 1.08 (95% CI: 1.02 to 1.14) p=0.01 but a decreased risk of unspecified stroke 0.80 (95% CI: 0.66 to 0.96) p=0.02 in fully adjusted models. Total NSP and NSP density were also associated with lower risk of unspecified stroke (per 6g/day) 0.77 (95% CI: 0.63 to 0.95) p=0.02 and (per 2g/1000kcal/day) 0.86 (95% CI: 0.73 to 0.99) p=0.04, respectively in fully adjusted models. In this subgroup of postmenopausal women, these protective associations were not the same with fibre assessed as AOAC but were present for both soluble (per 3g/day) 0.76 (95% CI: 0.57 to 1.00) p=0.05 and insoluble fibre (per 4g/day) 0.82 (95% CI: 0.69 to 0.99) p=0.04.

Protective associations observed with greater NSP or AOAC fibre, greater insoluble fibre and greater cereal fibre with risk for 'mostly ischaemic' stroke in the full sample remained protective in the postmenopausal group. However, soluble fibre was no longer significantly associated with risk of 'mostly ischaemic' stroke in this subgroup 0.87 (95% CI: 0.73 to 1.04) p=0.14, per 3g/day increase.

In healthy weight women, protective associations for the various fibre exposures (total fibre, fibre density, insoluble fibre, cereal fibre) and risk of 'mostly ischaemic' stroke remained, as in the full sample of participants. For each 4g/day increase in insoluble fibre, risk was 0.81 (95% CI: 0.70 to 0.95) p<0.01, however the protective association observed in the full sample for soluble fibre did not remain in this healthy weight subgroup. With each 3g/day increase in soluble fibre, risk of 'mostly ischaemic' stroke was 0.83 (95% CI: 0.67 to 1.02) p=0.08.

In healthy weight women, each 1g/day increase in legume fibre was associated with an increased risk of haemorrhagic stroke 1.11 (95% CI: 1.00 to 1.24) p=0.05 in the fully adjusted models. Additionally in fully adjusted models, total cereal fibre (per 3g/day), fibre from breakfast cereals (per 2g/day), AOAC fibre density (per 3g/1000kcal/day) and insoluble fibre (per 4g/day) were all associated with ischaemic stroke risk reduction, 0.83 (95% CI: 0.71 to 0.98) p=0.03; 0.78 (95% CI: 0.66 to 0.92) p<0.01; 0.85 (95% CI: 0.73 to 1.00) p=0.05; 0.82 (95% CI: 0.67 to 0.99) p=0.04, respectively. Total fibre (NSP) (per 6g/day), vegetable fibre (per

2g/day) and soluble fibre (per 3g/day) were associated with lower risk of unspecified type stroke in healthy weight women 0.78 (95% CI: 0.61 to 0.99) p=0.04; 0.80 (95% CI: 0.68 to 0.95) p=0.01 and 0.74 (95% CI: 0.55 to 1.00) p=0.05, respectively.

Those classified as overweight or obese were combined because of small case numbers in the two groups separately. Haemorrhagic stroke risk reductions were observed with four exposures in this overweight subgroup, total cereal fibre (per 3g/day) 0.85 (95% CI: 0.72 to 1.00) p=0.05; breakfast cereal fibre (per 2g/day) 0.83 (95% CI: 0.69 to 1.00) p=0.05; AOAC fibre (per11g/day) 0.76 (95% CI: 0.59 to 0.97) p=0.03 and insoluble fibre (per 4g/day) 0.81 (95% CI: 0.69 to 0.97) p=0.02. Fibre from nuts or seeds was additionally associated with lower risk of unspecified stroke, for each 0.2g increase in daily intake risk was 0.78 (95% CI: 0.62 to 0.97) p=0.03.

In participant subgroups, there were only sufficient cases to explore associations for 'mostly ischaemic' stroke risk in those reporting to have hypertension at baseline. In this subgroup, only greater cereal fibre intake, per 3g/day increase, was associated with lower risk 0.84 (95% CI: 0.70 to 1.00) p=0.05.

Women without hypertension had lower risk of ischaemic stroke with greater NSP density (per 2g/1000kcal/day) 0.88 (95% CI: 0.77 to 1.00) p=0.05, AOAC density (per 3g/1000kcal/day) 0.86 (95% CI: 0.75 to 0.98) p=0.02 and fibre from breakfast cereals (per 2g/day) 0.81 (95% CI: 0.71 to 0.93) p<0.01. Lower risk for unspecified stroke was observed for NSP (per 6g/day) 0.71 (95% CI: 0.56 to 0.90) p<0.01; NSP density (per 2g/1000kcal/day) 0.78 (95% CI: 0.60 to 0.92) p<0.01; AOAC (per 11g/day) 0.67 (95% CI: 0.48 to 0.93) p=0.02; AOAC density (per 3g/1000kcal/day) 0.78 (95% CI: 0.65 to 0.93) p<0.01; soluble fibre (per 3g/day) 0.64 (95% CI: 0.46 to 0.89) p<0.01; insoluble fibre (per 4g/day) 0.77 (95% CI: 0.62 to 0.96) p=0.02 and vegetable fibre (per 2g/day) 0.76 (95% CI: 0.61 to 0.96) 0.02, in fully adjusted models.

For 'mostly ischaemic' stroke in women without hypertension at baseline, the following items were protectively associated with risk: NSP (per 6g/day) 0.81 (95% CI: 0.69 to 0.94) p<0.01; NSP density (per 2g/1000kcal/day) 0.82 (95% CI: 0.73 to 0.91) p<0.01; AOAC (per 11g/day) 0.76 (95% CI: 0.63 to 0.92) p<0.01; AOAC density (per 3g/1000kcal/day) 0.81 (95% CI: 0.72 to 0.91) p<0.01; insoluble fibre (per 4g/day) 0.81 (95% CI: 0.71 to 0.93) p<0.01; soluble fibre (per 3g/day) 0.79 (95% CI: 0.65 to 0.97) p=0.02; cereal fibre (per 3g/day) 0.90 (95% CI: 0.80 to 1.00) p=0.05 and fibre from breakfast cereals (per 2g/day) 0.89 (95% CI: 0.80 to 0.99) p=0.03.

Table 6.10 Dietary fibre intake and associated risk for haemorrhagic, ischaemic and 'unspecified' stroke

		Median intake (IQR)	Total haemorrhagic stroke				Total ischaemic stroke				Total unspecified stroke			
			Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend	Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend	Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend
NSP (g/day)	Q1	14.1 (3.9)	27	1		1	40	1		1	43	1		1
	Q2	19.5 (2.2)	24	0.74 (0.43, 1.29)		0.74 (0.41, 1.34)	26	0.79 (0.50, 1.24)		0.60 (0.35, 1.01)	24	0.57 (0.35, 0.94)		0.56 (0.33, 0.96)
	Q3	23.8 (2.3)	21	0.66 (0.37, 1.17)		0.68 (0.37, 1.27)	35	0.88 (0.56, 1.38)		0.76 (0.46, 1.26)	19	0.44 (0.26, 0.76)		0.42 (0.23, 0.79)
	Q4	29.1 (3.1)	25	0.82 (0.48, 1.41)		0.81 (0.45, 1.45)	32	0.67 (0.41, 1.07)		0.59 (0.34, 1.03)	19	0.40 (0.29, 0.69)		0.37 (0.20, 0.66)
	Q5	38.1 (8.5)	28	0.88 (0.52, 1.49)		0.86 (0.42, 1.77)	27	0.65 (0.40, 1.05)		0.40 (0.21, 0.74)	20	0.53 (0.32, 0.89)		0.43 (0.20, 0.91)
	Per 6g/day		125	0.95 (0.85, 1.06)	0.37	0.92 (0.79, 1.05)	0.22	160	0.96 (0.85, 1.07)	0.46	0.90 (0.77, 1.06)	0.20	125	0.84 (0.73, 0.98)	0.02	0.76 (0.63, 0.92)	<0.01
AOAC (g/day)	Q1	21.8 (5.9)	30	1		1	42	1		1	43	1		1
	Q2	30.0 (3.4)	22	0.65 (0.37, 1.12)		0.61 (0.33, 1.10)	25	0.66 (0.41, 1.05)		0.54 (0.32, 0.91)	26	0.66 (0.40, 1.06)		0.67 (0.40, 1.13)
	Q3	36.6 (3.5)	20	0.58 (0.33, 1.02)		0.54 (0.30, 0.99)	33	0.70 (0.45, 1.09)		0.57 (0.34, 0.96)	16	0.34 (0.19, 0.61)		0.32 (0.16, 0.62)
	Q4	44.5 (4.7)	31	0.92 (0.56, 1.52)		0.80 (0.46, 1.41)	31	0.67 (0.42, 1.05)		0.56 (0.32, 0.99)	19	0.42 (0.25, 0.73)		0.42 (0.22, 0.78)
	Q5	58.5 (13.2)	22	0.63 (0.36, 1.09)		0.47 (0.22, 1.02)	29	0.59 (0.37, 0.94)		0.35 (0.18, 0.70)	21	0.59 (0.35, 0.97)		0.45 (0.20, 1.00)
	Per 11g/day		125	0.93 (0.82, 1.07)	0.32	0.87 (0.73, 1.05)	0.15	160	0.96 (0.83, 1.10)	0.52	0.88 (0.72, 1.07)	0.20	125	0.83 (0.70, 0.99)	0.04	0.74 (0.57, 0.94)	0.02
NSP density g/1000 kcal/day	Q1	7.5 (1.5)	27	1		1	45	1		1	32	1		1
	Q2	9.4 (0.8)	26	0.99 (0.59, 1.68)		1.06 (0.61, 1.84)	28	0.67 (0.43, 1.04)		0.64 (0.39, 1.04)	34	0.99 (0.62, 1.60)		1.08 (0.66, 1.75)
	Q3	11.0 (0.8)	25	0.74 (0.42, 1.29)		0.85 (0.49, 1.50)	26	0.56 (0.35, 0.90)		0.62 (0.37, 1.01)	14	0.46 (0.25, 0.85)		0.50 (0.26, 0.95)
	Q4	12.7 (1.0)	24	0.83 (0.48, 1.43)		0.82 (0.46, 1.48)	25	0.48 (0.30, 0.78)		0.48 (0.28, 0.81)	20	0.55 (0.31, 0.96)		0.56 (0.30, 1.02)
	Q5	15.4 (2.4)	23	0.77 (0.44, 1.36)		0.84 (0.47, 1.51)	36	0.63 (0.41, 0.99)		0.71 (0.44, 1.14)	25	0.70 (0.41, 1.19)		0.69 (0.39, 1.21)
	2g/1000kcal/d		125	0.93 (0.82, 1.04)	0.21	0.94 (0.83, 1.06)	0.31	160	0.89 (0.80, 0.99)	0.03	0.91 (0.81, 1.02)	0.12	125	0.86 (0.76, 0.98)	0.02	0.86 (0.75, 0.98)	0.02
AOAC density g/1000 kcal/day	Q1	11.7 (2.1)	29	1		1	41	1		1	36	1		1
	Q2	14.6 (1.2)	27	0.93 (0.56, 1.56)		0.97 (0.57, 1.64)	30	0.67 (0.43, 1.05)		0.65 (0.40, 1.06)	23	0.59 (0.35, 0.98)		0.59 (0.35, 1.01)
	Q3	16.9 (1.1)	22	0.65 (0.37, 1.14)		0.71 (0.40, 1.26)	26	0.61 (0.38, 0.97)		0.63 (0.38, 1.04)	18	0.49 (0.28, 0.84)		0.51 (0.29, 0.91)
	Q4	19.4 (1.4)	22	0.63 (0.36, 1.11)		0.64 (0.35, 1.15)	30	0.57 (0.36, 0.92)		0.65 (0.39, 1.07)	22	0.51 (0.30, 0.87)		0.55 (0.31, 0.97)
	Q5	23.4 (3.5)	25	0.76 (0.44, 1.30)		0.81 (0.46, 1.43)	33	0.61 (0.39, 0.96)		0.64 (0.39, 1.04)	26	0.66 (0.40, 1.10)		0.64 (0.37, 1.11)
	3g/1000kcal/d		125	0.91 (0.80, 1.03)	0.15	0.92 (0.81, 1.05)	0.22	160	0.88 (0.79, 0.98)	0.02	0.90 (0.80, 1.01)	0.08	125	0.87 (0.75, 0.99)	0.04	0.86 (0.74, 0.99)	0.04

		Median intake (IQR)	Total haemorrhagic stroke				Total ischaemic stroke				Total unspecified stroke			
			Cases[1]	Age-adjusted HR (95% CI)	p-trend	Fully-adjusted[2] HR (95% CI)	p-trend	Cases[1]	Age-adjusted HR (95% CI)	p-trend	Fully-adjusted[2] HR (95% CI)	p-trend	Cases[1]	Age-adjusted HR (95% CI)	p-trend	Fully-adjusted[2] HR (95% CI)	p-trend
Soluble fibre (g/day)	Q1	6.4 (1.6)	27	1		1		38	1		1		44	1		1	
	Q2	8.6 (0.9)	24	0.85 (0.49, 1.48)		0.90 (0.51, 1.61)		29	0.78 (0.49, 1.25)		0.71 (0.42, 1.20)		25	0.55 (0.33, 0.90)		0.57 (0.33, 0.97)	
	Q3	10.4 (0.9)	20	0.70 (0.40, 1.25)		0.67 (0.36, 1.26)		30	0.88 (0.56, 1.39)		0.72 (0.42, 1.23)		15	0.37 (0.21, 0.66)		0.37 (0.20, 0.69)	
	Q4	12.5 (1.2)	27	0.93 (0.54, 1.58)		0.86 (0.48, 1.56)		34	0.76 (0.48, 1.22)		0.75 (0.42, 1.31)		22	0.48 (0.29, 0.80)		0.44 (0.23, 0.82)	
	Q5	16.3 (3.8)	27	0.85 (0.49, 1.46)		0.83 (0.39, 1.78)		29	0.73 (0.45, 1.18)		0.54 (0.28, 1.05)		19	0.51 (0.30, 0.86)		0.44 (0.22, 0.88)	
	Per 3g/day		125	0.94 (0.83, 1.07)	0.37	0.89 (0.74, 1.07)	0.21	160	0.98 (0.86, 1.12)	0.79	0.93 (0.77, 1.13)	0.48	125	0.82 (0.69, 0.98)	0.03	0.72 (0.57, 0.92)	<0.01
Insoluble fibre (g/day)	Q1	8.4 (2.6)	28	1		1		42	1		1		43	1		1	
	Q2	12.2 (1.6)	27	0.82 (0.48, 1.40)		0.81 (0.46, 1.42)		26	0.82 (0.52, 1.30)		0.61 (0.37, 1.01)		22	0.62 (0.37, 1.02)		0.59 (0.34, 1.01)	
	Q3	15.3 (1.6)	18	0.57 (0.31, 1.02)		0.56 (0.30, 1.04)		32	0.81 (0.52, 1.28)		0.68 (0.41, 1.13)		22	0.55 (0.33, 0.92)		0.51 (0.29, 0.89)	
	Q4	19.0 (2.2)	27	0.88 (0.52, 1.49)		0.84 (0.48, 1.45)		31	0.65 (0.40, 1.04)		0.59 (0.34, 1.00)		19	0.40 (0.23, 0.71)		0.44 (0.25, 0.78)	
	Q5	25.5 (5.9)	25	0.76 (0.44, 1.31)		0.67 (0.34, 1.34)		29	0.67 (0.42, 1.08)		0.42 (0.22, 0.79)		19	0.56 (0.33, 0.94)		0.43 (0.20, 0.92)	
			125														
	Per 4g/day			0.95 (0.85, 1.05)	0.32	0.92 (0.81, 1.04)	0.20	160	0.95 (0.85, 1.06)	0.34	0.90 (0.78, 1.03)	0.13	125	0.86 (0.75, 0.98)	0.03	0.80 (0.68, 0.96)	0.01
Total cereal fibre (g/day)	Q1	2.8 (1.4)	27	1		1		38	1		1		34	1		1	
	Q2	5.1 (1.1)	25	1.01 (0.58, 1.73)		0.93 (0.52, 1.66)		42	1.05 (0.68, 1.62)		1.05 (0.65, 1.71)		24	0.68 (0.40, 1.17)		0.77 (0.44, 1.35)	
	Q3	7.6 (1.4)	23	0.78 (0.44, 1.38)		0.75 (0.40, 1.38)		19	0.47 (0.27, 0.79)		0.37 (0.20, 0.70)		30	0.82 (0.50, 1.35)		0.93 (0.54, 1.58)	
	Q4	10.6 (1.8)	27	0.81 (0.47, 1.41)		0.75 (0.42, 1.34)		21	0.55 (0.34, 0.92)		0.40 (0.21, 0.73)		19	0.59 (0.34, 1.02)		0.60 (0.33, 1.10)	
	Q5	15.6 (4.5)	23	0.81 (0.47, 1.42)		0.74 (0.39, 1.41)		40	0.84 (0.54, 1.30)		0.74 (0.42, 1.30)		18	0.60 (0.35, 1.03)		0.57 (0.29, 1.11)	
	Per 3g/day		125	0.96 (0.86, 1.06)	0.40	0.95 (0.84, 1.07)	0.40	160	0.94 (0.86, 1.04)	0.23	0.89 (0.80, 1.00)	0.05	125	0.91 (0.81, 1.03)	0.12	0.89 (0.78, 1.02)	0.11
Fibre from breakfast cereals (g/day)	Q1	0.05 (0.1)	24	1		1		37	1		1		27	1		1	
	Q2	0.5 (0.4)	26	1.10 (0.62, 1.93)		1.02 (0.57, 1.84)		31	1.05 (0.65, 1.71)		0.97 (0.58, 1.64)		21	1.18 (0.67, 2.09)		1.18 (0.65, 2.14)	
	Q3	1.8 (0.7)	22	0.85 (0.47, 1.53)		0.84 (0.45, 1.56)		29	0.94 (0.58, 1.53)		0.89 (0.53, 1.50)		23	0.88 (0.49, 1.59)		0.99 (0.54, 1.80)	
	Q4	3.5 (0.7)	29	1.07 (0.61, 1.86)		1.03 (0.57, 1.85)		32	0.93 (0.58, 1.49)		0.85 (0.50, 1.42)		29	1.25 (0.73, 2.14)		1.41 (0.80, 2.50)	
	Q5	7.6 (2.6)	24	0.86 (0.49, 1.52)		0.81 (0.45, 1.46)		31	0.81 (0.50, 1.30)		0.66 (0.39, 1.11)		25	1.05 (0.61, 1.80)		1.09 (0.61, 1.96)	
	Per 2g/day		125	0.98 (0.88, 1.08)	0.68	0.98 (0.88, 1.09)	0.65	160	0.94 (0.86, 1.02)	0.15	0.89 (0.80, 0.99)	0.03	125	1.00 (0.92, 1.09)	0.97	1.00 (0.91, 1.10)	0.95

		Median intake (IQR)	Total haemorrhagic stroke				Total ischaemic stroke				Total unspecified stroke			
			Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend	Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend	Cases[1]	Age-adjusted HR (95% CI) p-trend		Fully-adjusted[2] HR (95% CI) p-trend
Fruit fibre (g/day)	Q1	1.4 (0.9)	26	1		1	39	1		1	37	1		1
	Q2	2.9 (0.7)	23	0.61 (0.34, 1.08)		0.72 (0.40, 1.31)	33	0.74 (0.47, 1.18)		0.77 (0.47, 1.28)	20	0.40 (0.23, 0.70)		0.49 (0.28, 0.86)
	Q3	4.2 (0.7)	29	0.95 (0.57, 1.59)		0.96 (0.55, 1.68)	32	0.81 (0.52, 1.26)		0.74 (0.44, 1.22)	20	0.45 (0.26, 0.76)		0.47 (0.26, 0.84)
	Q4	5.8 (1.1)	26	0.68 (0.40, 1.18)		0.77 (0.42, 1.40)	22	0.43 (0.26, 0.73)		0.48 (0.28, 0.84)	27	0.52 (0.32, 0.86)		0.58 (0.34, 0.98)
	Q5	9.4 (4.0)	21	0.57 (0.32, 1.01)		0.63 (0.33, 1.19)	34	0.59 (0.37, 0.94)		0.59 (0.35, 1.00)	31	0.40 (0.23, 0.68)		0.44 (0.24, 0.78)
	Per 2g/day		125	0.94 (0.84, 1.06)	0.33	0.94 (0.83, 1.08) 0.39	160	0.96 (0.85, 1.07)	0.46	0.97 (0.86, 1.09) 0.62	125	0.92 (0.79, 1.07)	0.29	0.95 (0.80, 1.12) 0.52
Vegetable fibre (g/day)	Q1	2.3 (0.9)	26	1		1	39	1		1	34	1		1
	Q2	3.7 (0.6)	15	0.68 (0.36, 1.27)		0.64 (0.33, 1.25)	24	0.62 (0.38, 1.02)		0.70 (0.41, 1.19)	22	0.71 (0.42, 1.20)		0.71 (0.41, 1.21)
	Q3	5.0 (0.7)	23	0.96 (0.55, 1.70)		0.99 (0.55, 1.77)	28	0.73 (0.46, 1.17)		0.78 (0.46, 1.32)	27	0.73 (0.44, 1.21)		0.74 (0.44, 1.27)
	Q4	6.6 (1.0)	34	1.21 (0.71, 2.05)		1.33 (0.78, 2.27)	34	0.67 (0.42, 1.07)		0.81 (0.49, 1.33)	27	0.70 (0.42, 1.15)		0.69 (0.40, 1.18)
	Q5	9.6 (3.0)	27	1.01 (0.58, 1.74)		0.93 (0.52, 1.66)	35	0.76 (0.49, 1.19)		0.80 (0.50, 1.27)	15	0.36 (0.20, 0.66)		0.33 (0.18, 0.62)
	Per 2g/day		125	1.02 (0.94, 1.11)	0.65	1.02 (0.93, 1.11) 0.71	160	1.00 (0.90, 1.11)	0.97	1.02 (0.92, 1.13) 0.75	125	0.83 (0.72, 0.96)	0.01	0.80 (0.68, 0.92) <0.01
Legume fibre (g/day)	Q1	0.2 (0.2)	24	1		1	38	1		1	42	1		1
	Q2	0.7 (0.2)	33	1.24 (0.75, 2.09)		1.26 (0.73, 2.20)	37	0.95 (0.62, 1.48)		1.10 (0.68, 1.77)	28	0.75 (0.47, 1.20)		0.81 (0.50, 1.32)
	Q3	1.1 (0.2)	26	1.05 (0.60, 1.83)		1.19 (0.66, 2.12)	32	1.05 (0.67, 1.64)		1.02 (0.61, 1.70)	21	0.60 (0.35, 1.02)		0.68 (0.39, 1.18)
	Q4	1.6 (0.4)	22	0.98 (0.55, 1.73)		1.07 (0.58, 1.97)	26	0.82 (0.49, 1.36)		1.00 (0.57, 1.73)	21	0.65 (0.37, 1.12)		0.79 (0.44, 1.40)
	Q5	3.6 (1.4)	20	1.13 (0.63, 2.04)		1.08 (0.56, 2.08)	27	1.10 (0.67, 1.79)		1.09 (0.60, 1.99)	13	0.70 (0.38, 1.28)		0.62 (0.30, 1.27)
	Per 1g/day		125	1.05 (0.93, 1.18)	0.43	1.05 (0.92, 1.19) 0.50	160	1.00 (0.90, 1.11)	0.99	0.99 (0.87, 1.12) 0.82	125	0.90 (0.73, 1.11)	0.33	0.85 (0.68, 1.06) 0.14
Fibre from nuts and seeds (g/day)	Q1	0 (0.01)	38	1		1	59	1		1	46	1		1
	Q2	0.06 (0.01)	26	0.88 (0.53, 1.44)		0.82 (0.49, 1.39)	31	0.63 (0.42, 0.95)		0.60 (0.38, 0.94)	25	0.72 (0.44, 1.16)		0.76 (0.45, 1.27)
	Q3	0.08 (0.05)	23	0.72 (0.42, 1.24)		0.72 (0.41, 1.27)	21	0.44 (0.27, 0.73)		0.47 (0.28, 0.81)	19	0.64 (0.38, 1.09)		0.70 (0.39, 1.23)
	Q4	0.28 (0.12)	16	0.54 (0.29, 1.01)		0.57 (0.31, 1.07)	21	0.48 (0.29, 0.80)		0.50 (0.29, 0.86)	19	0.74 (0.44, 1.26)		0.80 (0.45, 1.42)
	Q5	0.87 (0.92)	22	0.75 (0.43, 1.30)		0.68 (0.38, 1.22)	28	0.53 (0.33, 0.85)		0.60 (0.35, 1.02)	16	0.41 (0.22, 0.76)		0.50 (0.26, 0.95)
	Per 0.2g/day		125	0.98 (0.92, 1.04)	0.44	0.97 (0.91, 1.04) 0.40	160	0.95 (0.89, 1.02)	0.14	0.97 (0.91, 1.03) 0.30	125	0.89 (0.81, 0.99)	0.03	0.92 (0.84, 1.01) 0.08

[1]Case numbers apply to fully-adjusted models. [2]Adjustments include Age (years), BMI (kg/m^2), calories from carbohydrate, fat and protein (kcal/day), ethanol intake (g/day), MET (hours/week), smoking status (current vs. not current smoker), socio-economic status (professional or managerial/ intermediate/ routine or manual). Note, adjustment for energy intake was not included in fibre density models. Highlight=CIs do not span 1 in fully adjusted model.

Table 6.11 Dietary fibre intake and associated risk for 'mostly ischaemic' type stroke (ischaemic type plus unspecified stroke)

		Median intake (IQR)	Cases[1]	Age-adjusted HR (95% CI)	p-trend	Fully-adjusted[2] HR (95% CI)	p-trend
NSP (g/day)	Q1	14.1 (3.9)	68	1		1	
	Q2	19.5 (2.2)	43	0.70 (0.49, 1.02)		0.56 (0.38, 0.84)	
	Q3	23.8 (2.3)	47	0.68 (0.47, 0.99)		0.56 (0.37, 0.84)	
	Q4	29.1 (3.1)	49	0.60 (0.41, 0.87)		0.49 (0.32, 0.75)	
	Q5	38.1 (8.5)	44	0.61 (0.42, 0.88)		0.35 (0.20, 0.60)	
	Per 6g/day		251	0.92 (0.84, 1.01)	0.08	0.83 (0.73, 0.94)	<0.01
AOAC (g/day)	Q1	21.8 (5.9)	69	1		1	
	Q2	30.0 (3.4)	44	0.69 (0.48, 0.99)		0.60 (0.40, 0.89)	
	Q3	36.6 (3.5)	42	0.55 (0.38, 0.81)		0.43 (0.28, 0.67)	
	Q4	44.5 (4.7)	49	0.64 (0.44, 0.92)		0.53 (0.34, 0.81)	
	Q5	58.5 (13.2)	47	0.61 (0.42, 0.88)		0.35 (0.19, 0.61)	
	Per 11g/day		251	0.92 (0.83, 1.03)	0.14	0.80 (0.68, 0.95)	<0.01
NSP density g/1000 kcal/day	Q1	7.5 (1.5)	68	1		1	
	Q2	9.4 (0.8)	53	0.77 (0.55, 1.10)		0.77 (0.53, 1.11)	
	Q3	11.0 (0.8)	35	0.49 (0.33, 0.72)		0.51 (0.34, 0.78)	
	Q4	12.7 (1.0)	38	0.47 (0.32, 0.69)		0.44 (0.29, 0.68)	
	Q5	15.4 (2.4)	57	0.58 (0.41, 0.84)		0.61 (0.42, 0.90)	
	2g/1000kcal/d		251	0.86 (0.79, 0.94)	<0.01	0.86 (0.79, 0.95)	<0.01
AOAC density g/1000 kcal/day	Q1	11.7 (2.1)	67	1		1	
	Q2	14.6 (1.2)	46	0.62 (0.43, 0.89)		0.59 (0.40, 0.87)	
	Q3	16.9 (1.1)	38	0.54 (0.37, 0.78)		0.54 (0.36, 0.80)	
	Q4	19.4 (1.4)	46	0.52 (0.36, 0.75)		0.55 (0.37, 0.83)	
	Q5	23.4 (3.5)	54	0.55 (0.38, 0.80)		0.55 (0.37, 0.81)	
	3g/1000kcal/d		251	0.86 (0.79, 0.94)	<0.01	0.86 (0.78, 0.95)	<0.01
Soluble fibre (g/day)	Q1	6.4 (1.6)	67	1		1	
	Q2	8.6 (0.9)	45	0.69 (0.48, 1.01)		0.63 (0.42, 0.95)	
	Q3	10.4 (0.9)	41	0.69 (0.48, 1.00)		0.56 (0.37, 0.85)	
	Q4	12.5 (1.2)	53	0.70 (0.48, 1.00)		0.61 (0.39, 0.94)	
	Q5	16.3 (3.8)	45	0.66 (0.45, 0.96)		0.46 (0.27, 0.77)	
	Per 3g/day		251	0.94 (0.84, 1.04)	0.23	0.84 (0.72, 0.98)	0.03
Insoluble fibre (g/day)	Q1	8.4 (2.6)	70	1		1	
	Q2	12.2 (1.6)	42	0.73 (0.50, 1.05)		0.58 (0.39, 0.87)	
	Q3	15.3 (1.6)	47	0.69 (0.48, 0.99)		0.55 (0.37, 0.82)	
	Q4	19.0 (2.2)	46	0.55 (0.37, 0.81)		0.49 (0.32, 0.75)	
	Q5	25.5 (5.9)	46	0.62 (0.43, 0.90)		0.36 (0.22, 0.61)	
	Per 4g/day		251	0.92 (0.85, 1.00)	0.06	0.84 (0.75, 0.94)	<0.01
Total cereal fibre (g/day)	Q1	2.8 (1.4)	64	1		1	
	Q2	5.1 (1.1)	54	0.80 (0.56, 1.16)		0.77 (0.52, 1.15)	
	Q3	7.6 (1.4)	42	0.60 (0.41, 0.88)		0.52 (0.34, 0.81)	
	Q4	10.6 (1.8)	37	0.55 (0.37, 0.81)		0.42 (0.27, 0.66)	
	Q5	15.6 (4.5)	54	0.71 (0.50, 1.02)		0.57 (0.36, 0.89)	
	Per 3g/day		251	0.93 (0.86, 1.01)	0.07	0.88 (0.80, 0.96)	<0.01
Fibre from breakfast cereals (g/day)	Q1	0.05 (0.1)	54	1		1	
	Q2	0.5 (0.4)	47	1.19 (0.81, 1.77)		1.11 (0.73, 1.69)	
	Q3	1.8 (0.7)	45	0.95 (0.63, 1.42)		0.91 (0.59, 1.39)	
	Q4	3.5 (0.7)	55	1.11 (0.76, 1.63)		1.05 (0.69, 1.57)	
	Q5	7.6 (2.6)	50	0.91 (0.62, 1.34)		0.79 (0.52, 1.20)	
	Per 2g/day		251	0.96 (0.90, 1.03)	0.25	0.93 (0.86, 1.01)	0.09
Fruit fibre (g/day)	Q1	1.4 (0.9)	59	1		1	
	Q2	2.9 (0.7)	47	0.65 (0.45, 0.96)		0.70 (0.46, 1.05)	
	Q3	4.2 (0.7)	49	0.73 (0.51, 1.05)		0.69 (0.46, 1.04)	
	Q4	5.8 (1.1)	42	0.53 (0.36, 0.79)		0.55 (0.36, 0.84)	
	Q5	9.0 (4.0)	54	0.57 (0.39, 0.82)		0.57 (0.38, 0.87)	
	Per 2g/day		251	0.96 (0.88, 1.05)	0.38	0.97 (0.88, 1.07)	0.53
Vegetable fibre (g/day)	Q1	2.3 (0.9)	62	1		1	
	Q2	3.7 (0.6)	39	0.69 (0.47, 1.02)		0.69 (0.46, 1.05)	
	Q3	5.0 (0.7)	49	0.78 (0.54, 1.13)		0.77 (0.51, 1.14)	
	Q4	6.6 (1.0)	55	0.72 (0.50, 1.03)		0.74 (0.50, 1.09)	
	Q5	9.6 (3.0)	46	0.59 (0.40, 0.86)		0.53 (0.36, 0.79)	
	Per 2g/day		251	0.94 (0.86, 1.03)	0.19	0.92 (0.84, 1.02)	0.13
Legume fibre (g/day)	Q1	0.2 (0.2)	66	1		1	
	Q2	0.7 (0.2)	59	0.98 (0.69, 1.38)		1.05 (0.72, 1.51)	
	Q3	1.1 (0.2)	46	0.93 (0.65, 1.33)		0.89 (0.60, 1.34)	
	Q4	1.6 (0.4)	42	0.80 (0.53, 1.19)		0.92 (0.60, 1.40)	
	Q5	3.6 (1.4)	38	0.98 (0.66, 1.46)		0.89 (0.55, 1.42)	
	Per 1g/day		251	0.97 (0.88, 1.08)	0.62	0.93 (0.83, 1.05)	0.23
Fibre from nuts and seeds (g/day)	Q1	0 (0.01)	87	1		1	
	Q2	0.06 (0.01)	47	0.70 (0.50, 0.98)		0.70 (0.48, 1.02)	
	Q3	0.08 (0.05)	39	0.60 (0.41, 0.87)		0.66 (0.44, 0.99)	
	Q4	0.28 (0.12)	34	0.59 (0.39, 0.87)		0.62 (0.40, 0.95)	
	Q5	0.87 (0.92)	44	0.54 (0.37, 0.79)		0.62 (0.41, 0.95)	
	Per 0.2g/day		251	0.94 (0.89, 0.99)	0.03	0.96 (0.91, 1.01)	0.09

Key as on previous page

Risk of MI, ACS and chronic heart disease (fatal plus non-fatal events)

All participants

After median follow-up of 14.4 years, 236 MI, 392 ACS and 573 chronic heart disease cases were observed. Greater total and insoluble fibre intakes were associated with lower risk of MI. For each 11g/day increase in fibre (AOAC), there was 14% lower risk: 0.86 (95% CI: 0.73 to 1.00) p=0.04 and for each 3g/1000kal/day increase in AOAC density of the diet, risk was 0.89 (95% CI: 0.80 to 0.98) p=0.02. Each 4g/day greater intake of insoluble fibre was associated with 12% lower risk for MI 0.88 (95% CI: 0.79 to 0.99) p=0.03. These protective associations were not evident when total ACS or chronic heart disease cases were considered. Greater legume fibre (per 1g/day) however was protectively associated with lower risk of chronic heart disease, 0.92 (95% CI: 0.86 to 0.99) p=0.03 but for MI and ACS, risk estimates were close to 1, indicating no evidence of any associations with greater legume fibre intake.

Subgroups

As with the analysis using the full sample of women, no protective associations were apparent in any of the subgroups for ACS and fibre intake. For chronic heart disease risk, legume fibre (per 1g/day) remained protectively associated in both the women with healthy BMI at baseline 0.88 (95% CI: 0.78 to 0.99) p=0.03 and those without hypertension 0.87 (95% CI: 0.79 to 0.96) p=0.004. Soluble fibre (per 3g/day) was also associated with 12% lower risk of chronic disease risk in women without hypertension, 0.88 (95% CI: 0.78 to 0.99) p=0.04.

The protective associations observed with total fibre and insoluble fibre for the full sample remained in the postmenopausal subgroup: per 6g/day increase in NSP, 0.86 (95% CI: 0.73 to 1.00) p=0.05, per 2g/1000kcal/day increase in NSP density, 0.87 (95% CI: 0.77 to 0.99) p=0.03, per 3g/1000kcal/day increase in AOAC density, 0.88 (95% CI: 0.77 to 0.99) p=0.04 and per 4g/day increase in insoluble fibre, risk was 0.86 (95% CI: 0.75 to 1.00) p=0.05. Fruit fibre intake, per 2g/day increase was also associated with lower MI risk 0.90 (95% CI: 0.81 to 1.00) p=0.05.

When considering women without hypertension at baseline, the protective associations for MI risk seen in the full sample, did not remain. They were however apparent in those with hypertension; for each 2g/1000kcal/day increase in NSP fibre density, risk of MI was 0.83 (95% CI: 0.69 to 0.99) p=0.04 and for AOAC density per 3g/1000kcal/day risk was 0.81 (95% CI: 0.67 to 0.98) p=0.03.

Table 6.12 Dietary fibre intake and associated risk for MI, ACS or chronic heart disease

		Median intake (IQR)	Incident MI (fatal plus non-fatal)			Incident ACS (fatal plus non-fatal)			Incident chronic heart disease (no acute events) (fatal plus non-fatal)		
			Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend
NSP (g/day)	Q1	14.1 (3.9)	47	1	1	78	1	1	117	1	1
	Q2	19.5 (2.2)	61	1.27 (0.88, 1.85)	1.33 (0.88, 2.00)	88	1.11 (0.82, 1.51)	1.16 (0.84, 1.60)	105	0.84 (0.65, 1.10)	0.90 (0.68, 1.19)
	Q3	23.8 (2.3)	30	0.63 (0.41, 0.99)	0.63 (0.38, 1.06)	55	0.71 (0.50, 1.00)	0.73 (0.50, 1.08)	91	0.73 (0.55, 0.96)	0.81 (0.59, 1.10)
	Q4	29.1 (3.1)	45	0.79 (0.52, 1.19)	0.79 (0.48, 1.29)	76	0.93 (0.67, 1.27)	0.93 (0.64, 1.37)	108	0.83 (0.64, 1.09)	0.88 (0.65, 1.21)
	Q5	38.1 (8.5)	34	0.70 (0.45, 1.08)	0.55 (0.31, 0.99)	64	0.80 (0.57, 1.12)	0.76 (0.48, 1.19)	114	0.94 (0.72, 1.22)	1.05 (0.74, 1.49)
	Per 6g/day		217	0.93 (0.84, 1.02) 0.14	0.89 (0.79, 1.01) 0.06	361	0.96 (0.90, 1.03) 0.31	0.95 (0.87, 1.05) 0.34	535	0.98 (0.93, 1.04) 0.54	0.99 (0.91, 1.07) 0.73
AOAC (g/day)	Q1	21.8 (5.9)	46	1	1	76	1	1	118	1	1
	Q2	30.0 (3.4)	58	1.29 (0.88, 1.89)	1.38 (0.91, 2.09)	86	1.16 (0.85, 1.58)	1.24 (0.89, 1.72)	104	0.82 (0.63, 1.07)	0.86 (0.65, 1.15)
	Q3	36.6 (3.5)	36	0.76 (0.49, 1.17)	0.76 (0.46, 1.25)	59	0.78 (0.56, 1.10)	0.83 (0.56, 1.22)	90	0.70 (0.54, 0.93)	0.76 (0.56, 1.04)
	Q4	44.5 (4.7)	45	0.82 (0.54, 1.25)	0.80 (0.48, 1.35)	78	1.01 (0.73, 1.39)	1.01 (0.68, 1.50)	110	0.81 (0.62, 1.05)	0.88 (0.64, 1.21)
	Q5	58.5 (13.2)	32	0.72 (0.46, 1.13)	0.55 (0.30, 1.02)	62	0.82 (0.58, 1.15)	0.78 (0.48, 1.25)	113	0.95 (0.73, 1.23)	1.03 (0.71, 1.49)
	Per 11g/day		217	0.92 (0.82, 1.03) 0.15	0.86 (0.73, 1.00) 0.04	361	0.96 (0.88, 1.04) 0.33	0.94 (0.83, 1.06) 0.30	535	0.98 (0.91, 1.05) 0.53	0.98 (0.89, 1.08) 0.71
NSP density g/1000 kcal/day	Q1	7.5 (1.5)	57	1	1	85	1	1	116	1	1
	Q2	9.4 (0.8)	41	0.66 (0.44, 0.98)	0.76 (0.50, 1.14)	67	0.76 (0.55, 1.04)	0.82 (0.59, 1.13)	106	0.86 (0.66, 1.12)	0.93 (0.71, 1.22)
	Q3	11.0 (0.8)	46	0.83 (0.57, 1.21)	0.88 (0.59, 1.31)	72	0.88 (0.65, 1.20)	0.91 (0.66, 1.26)	101	0.83 (0.64, 1.09)	0.89 (0.67, 1.18)
	Q4	12.7 (1.0)	36	0.58 (0.38, 0.88)	0.66 (0.43, 1.03)	62	0.66 (0.47, 0.92)	0.74 (0.53, 1.04)	107	0.89 (0.69, 1.16)	0.95 (0.72, 1.26)
	Q5	15.4 (2.4)	37	0.57 (0.37, 0.87)	0.67 (0.42, 1.05)	75	0.83 (0.60, 1.13)	0.90 (0.64, 1.25)	105	0.87 (0.66, 1.14)	0.94 (0.71, 1.24)
	2g/1000kcal/d		217	0.86 (0.78, 0.95) <.01	0.89 (0.81, 0.98) 0.02	361	0.93 (0.87, 1.00) 0.87	0.95 (0.88, 1.02) 0.18	535	0.96 (0.90, 1.01) 0.14	0.97 (0.92, 1.04) 0.40
AOAC density g/1000 kcal/day	Q1	11.7 (2.1)	54	1	1	78	1	1	113	1	1
	Q2	14.6 (1.2)	53	0.90 (0.62, 1.31)	1.00 (0.68, 1.48)	82	1.00 (0.73, 1.36)	1.06 (0.77, 1.45)	110	0.95 (0.73, 1.23)	0.98 (0.75, 1.28)
	Q3	16.9 (1.1)	37	0.67 (0.45, 1.02)	0.74 (0.48, 1.14)	62	0.82 (0.59, 1.13)	0.85 (0.60, 1.19)	102	0.85 (0.65, 1.11)	0.89 (0.67, 1.17)
	Q4	19.4 (1.4)	44	0.81 (0.55, 1.20)	0.87 (0.57, 1.32)	75	0.92 (0.67, 1.26)	0.97 (0.69, 1.35)	111	0.93 (0.71, 1.22)	1.00 (0.76, 1.32)
	Q5	23.4 (3.5)	29	0.46 (0.28, 0.73)	0.54 (0.33, 0.89)	64	0.76 (0.54, 1.06)	0.83 (0.58, 1.19)	99	0.85 (0.65, 1.12)	0.90 (0.68, 1.21)
	3g/1000kcal/d		217	0.86 (0.78, 0.95) <.01	0.89 (0.80, 0.98) 0.02	361	0.93 (0.87, 1.00) 0.07	0.95 (0.88, 1.03) 0.20	535	0.95 (0.90, 1.01) 0.12	0.97 (0.91, 1.04) 0.38

		Median intake (IQR)	Incident MI (fatal plus non-fatal)			Incident ACS (fatal plus non-fatal)			Incident chronic heart disease (no acute events) (fatal plus non-fatal)		
			Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend
Soluble fibre (g/day)	Q1	6.4 (1.6)	52	1	1	79	1	1	121	1	1
	Q2	8.6 (0.9)	45	0.94 (0.63, 1.39)	0.93 (0.61, 1.42)	72	0.96 (0.70, 1.32)	0.99 (0.70, 1.39)	100	0.78 (0.60, 1.01)	0.80 (0.61, 1.07)
	Q3	10.4 (0.9)	40	0.83 (0.55, 1.25)	0.84 (0.53, 1.35)	68	0.98 (0.71, 1.35)	1.00 (0.69, 1.45)	98	0.75 (0.57, 0.98)	0.78 (0.58, 1.05)
	Q4	12.5 (1.2)	39	0.67 (0.44, 1.03)	0.69 (0.41, 1.15)	71	0.88 (0.63, 1.22)	0.94 (0.63, 1.39)	99	0.72 (0.55, 0.95)	0.78 (0.56, 1.07)
	Q5	16.3 (3.8)	41	0.79 (0.52, 1.19)	0.66 (0.37, 1.17)	71	0.88 (0.64, 1.23)	0.90 (0.57, 1.43)	117	0.95 (0.74, 1.23)	0.97 (0.68, 1.38)
	Per 3g/day		217	0.95 (0.85, 1.07) 0.40	0.90 (0.77, 1.04) 0.16	361	0.97 (0.89, 1.05) 0.40	0.94 (0.84, 1.05) 0.28	535	0.97 (0.91, 1.04) 0.44	0.95 (0.86, 1.05) 0.34
Insouble fibre (g/day)	Q1	8.4 (2.6)	50	1	1	82	1	1	118	1	1
	Q2	12.2 (1.6)	56	1.14 (0.78, 1.65)	1.18 (0.79, 1.76)	81	1.01 (0.74, 1.37)	1.05 (0.76, 1.45)	105	0.83 (0.63, 1.08)	0.88 (0.66, 1.16)
	Q3	15.3 (1.6)	34	0.64 (0.42, 1.00)	0.69 (0.42, 1.13)	60	0.72 (0.51, 1.00)	0.77 (0.53, 1.12)	103	0.80 (0.61, 1.04)	0.90 (0.67, 1.20)
	Q4	19.0 (2.2)	45	0.78 (0.52, 1.17)	0.76 (0.47, 1.23)	75	0.91 (0.67, 1.25)	0.91 (0.63, 1.31)	101	0.79 (0.61, 1.03)	0.84 (0.62, 1.14)
	Q5	25.5 (5.9)	32	0.66 (0.43, 1.03)	0.55 (0.31, 0.95)	63	0.78 (0.56, 1.09)	0.75 (0.49, 1.16)	108	0.90 (0.68, 1.16)	0.96 (0.68, 1.35)
	Per 4g/day		217	0.92 (0.84, 1.00) 0.06	0.88 (0.79, 0.99) 0.03	361	0.97 (0.90, 1.03) 0.33	0.96 (0.88, 1.05) 0.40	535	0.99 (0.94, 1.04) 0.65	1.00 (0.93, 1.08) 0.93
Total cereal fibre (g/day)	Q1	2.8 (1.4)	55	1	1	84	1	1	112	1	1
	Q2	5.1 (1.1)	39	0.78 (0.52, 1.17)	0.83 (0.54, 1.26)	66	0.81 (0.59, 1.12)	0.87 (0.62, 1.21)	100	0.86 (0.65, 1.13)	0.93 (0.70, 1.23)
	Q3	7.6 (1.4)	41	0.73 (0.49, 1.09)	0.72 (0.46, 1.11)	65	0.74 (0.53, 1.02)	0.72 (0.51, 1.03)	103	0.84 (0.64, 1.10)	0.89 (0.66, 1.19)
	Q4	10.6 (1.8)	43	0.70 (0.46, 1.05)	0.78 (0.49, 1.23)	77	0.84 (0.61, 1.15)	0.94 (0.67, 1.33)	115	0.90 (0.69, 1.17)	1.00 (0.75, 1.35)
	Q5	15.6 (4.5)	39	0.65 (0.43, 0.98)	0.68 (0.42, 1.09)	69	0.83 (0.60, 1.13)	0.87 (0.60, 1.25)	105	0.85 (0.65, 1.12)	0.95 (0.69, 1.30)
	Per 3g/day		217	0.93 (0.86, 1.02) 0.11	0.94 (0.86, 1.04) 0.22	361	1.00 (0.93, 1.06) 0.91	1.02 (0.95, 1.09) 0.65	535	0.99 (0.94, 1.05) 0.80	1.01 (0.95, 1.07) 0.67
Fibre from breakfast cereals (g/day)	Q1	0.05 (0.1)	51	1	1	76	1	1	108	1	1
	Q2	0.5 (0.4)	50	1.06 (0.71, 1.59)	1.02 (0.67, 1.54)	77	1.03 (0.75, 1.43)	1.02 (0.73, 1.42)	107	1.00 (0.76, 1.31)	0.95 (0.72, 1.27)
	Q3	1.8 (0.7)	39	0.73 (0.48, 1.13)	0.74 (0.47, 1.16)	65	0.85 (0.60, 1.18)	0.82 (0.58, 1.16)	96	0.82 (0.62, 1.09)	0.82 (0.61, 1.09)
	Q4	3.5 (0.7)	33	0.68 (0.44, 1.04)	0.64 (0.40, 1.02)	59	0.75 (0.53, 1.06)	0.74 (0.52, 1.06)	96	0.81 (0.61, 1.07)	0.84 (0.63, 1.13)
	Q5	7.6 (2.6)	44	0.81 (0.54, 1.20)	0.78 (0.50, 1.20)	84	0.98 (0.72, 1.34)	0.98 (0.71, 1.36)	128	0.99 (0.76, 1.29)	1.04 (0.79, 1.38)
	Per 2g/day		217	0.93 (0.85, 1.02) 0.11	0.93 (0.85, 1.02) 0.13	361	1.01 (0.94, 1.07) 0.87	1.02 (0.95, 1.09) 0.62	535	1.00 (0.95, 1.06) 0.96	1.02 (0.97, 1.08) 0.48

		Median intake (IQR)	Incident MI (fatal plus non-fatal) Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Incident ACS (fatal plus non-fatal) Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend	Incident chronic heart disease (no acute events) (fatal plus non-fatal) Cases[1]	Age-adjusted HR (95% CI) p-trend	Fully-adjusted[2] HR (95% CI) p-trend
Fruit fibre (g/day)	Q1	1.4 (0.9)	44	1	1	71	1	1	111	1	1
	Q2	2.9 (0.7)	46	0.80 (0.53, 1.21)	0.99 (0.64, 1.54)	73	0.92 (0.66, 1.27)	1.02 (0.72, 1.45)	103	0.81 (0.62, 1.06)	0.85 (0.64, 1.13)
	Q3	4.2 (0.7)	40	0.70 (0.46, 1.07)	0.85 (0.54, 1.33)	68	0.81 (0.58, 1.13)	0.93 (0.65, 1.32)	107	0.74 (0.56, 0.97)	0.80 (0.60, 1.06)
	Q4	5.8 (1.1)	48	0.78 (0.52, 1.17)	0.98 (0.62, 1.54)	74	0.85 (0.61, 1.18)	0.98 (0.68, 1.41)	101	0.79 (0.60, 1.03)	0.81 (0.61, 1.09)
	Q5	9.4 (4.0)	39	0.61 (0.39, 0.93)	0.72 (0.45, 1.15)	75	0.81 (0.58, 1.13)	0.93 (0.64, 1.35)	113	0.81 (0.62, 1.06)	0.89 (0.66, 1.19)
	Per 2g/day		217	0.93 (0.86, 1.01) 0.08	0.94 (0.87, 1.01) 0.11	361	0.96 (0.91, 1.02) 0.17	0.97 (0.92, 1.03) 0.36	535	0.99 (0.94, 1.04) 0.62	1.00 (0.95, 1.05) 0.97
Vegetable fibre (g/day)	Q1	2.3 (0.9)	49	1	1	80	1	1	108	1	1
	Q2	3.7 (0.6)	32	0.71 (0.46, 1.09)	0.73 (0.46, 1.17)	55	0.76 (0.54, 1.06)	0.77 (0.53, 1.10)	91	0.74 (0.56, 0.99)	0.80 (0.60, 1.08)
	Q3	5.0 (0.7)	44	0.80 (0.53, 1.20)	0.92 (0.60, 1.43)	74	0.88 (0.64, 1.20)	0.90 (0.63, 1.27)	108	0.87 (0.66, 1.14)	0.93 (0.70, 1.24)
	Q4	6.6 (1.0)	49	0.82 (0.55, 1.22)	0.94 (0.62, 1.45)	79	0.85 (0.62, 1.17)	0.92 (0.66, 1.30)	113	0.89 (0.68, 1.16)	0.96 (0.72, 1.27)
	Q5	9.6 (3.0)	43	0.71 (0.47, 1.06)	0.72 (0.45, 1.15)	73	0.81 (0.59, 1.11)	0.80 (0.55, 1.16)	115	0.95 (0.73, 1.23)	0.97 (0.72, 1.30)
	Per 2g/day		217	0.96 (0.88, 1.05) 0.40	0.95 (0.87, 1.05) 0.32	361	0.97 (0.91, 1.03) 0.31	0.95 (0.89, 1.03) 0.21	535	1.00 (0.96, 1.05) 0.85	1.00 (0.95, 1.05) 0.98
Legume fibre (g/day)	Q1	0.2 (0.2)	52	1	1	87	1	1	133	1	1
	Q2	0.7 (0.2)	47	0.96 (0.65, 1.41)	0.94 (0.62, 1.42)	82	0.98 (0.73, 1.33)	0.94 (0.68, 1.28)	125	0.93 (0.73, 1.19)	0.90 (0.70, 1.16)
	Q3	1.1 (0.2)	42	0.92 (0.61, 1.39)	1.00 (0.65, 1.53)	75	0.95 (0.70, 1.30)	0.97 (0.70, 1.34)	96	0.83 (0.64, 1.07)	0.79 (0.60, 1.04)
	Q4	1.6 (0.4)	43	1.22 (0.82, 1.83)	1.21 (0.78, 1.88)	62	1.06 (0.77, 1.46)	0.99 (0.70, 1.39)	94	0.85 (0.65, 1.12)	0.83 (0.63, 1.11)
	Q5	3.6 (1.4)	33	1.04 (0.66, 1.63)	1.03 (0.64, 1.65)	55	0.94 (0.66, 1.34)	0.92 (0.64, 1.33)	87	0.87 (0.66, 1.15)	0.80 (0.59, 1.08)
	Per 1g/day		217	1.00 (0.91, 1.10) 0.93	1.00 (0.90, 1.10) 0.96	361	0.97 (0.90, 1.05) 0.45	0.97 (0.89, 1.05) 0.43	535	0.94 (0.88, 1.01) 0.07	0.92 (0.86, 0.99) 0.03
Fibre from nuts and seeds (g/day)	Q1	0 (0.01)	52	1	1	91	1	1	130	1	1
	Q2	0.06 (0.01)	53	1.17 (0.80, 1.70)	1.21 (0.82, 1.79)	82	1.02 (0.76, 1.37)	1.03 (0.76, 1.41)	130	1.12 (0.88, 1.43)	1.17 (0.91, 1.51)
	Q3	0.08 (0.05)	36	0.92 (0.60, 1.41)	1.04 (0.67, 1.61)	70	0.99 (0.72, 1.35)	1.02 (0.73, 1.42)	104	1.00 (0.77, 1.29)	1.08 (0.82, 1.42)
	Q4	0.28 (0.12)	37	1.03 (0.67, 1.57)	1.10 (0.71, 1.71)	60	0.93 (0.66, 1.30)	0.96 (0.68, 1.36)	87	0.87 (0.66, 1.15)	0.94 (0.70, 1.25)
	Q5	0.87 (0.92)	39	0.92 (0.60, 1.41)	0.98 (0.62, 1.57)	58	0.77 (0.55, 1.08)	0.78 (0.54, 1.15)	84	0.75 (0.56, 0.99)	0.86 (0.63, 1.18)
	Per 0.2g/day		217	1.00 (0.96, 1.04) 0.96	1.01 (0.97, 1.04) 0.76	361	0.99 (0.95, 1.02) 0.52	0.99 (0.95, 1.03) 0.59	535	0.98 (0.94, 1.01) 0.14	0.99 (0.96, 1.02) 0.55

[1]Case numbers apply to fully-adjusted models. [2]Adjustments include Age (years), BMI (kg/m^2), calories from carbohydrate, fat and protein (kcal/day), ethanol intake (g/day), MET (hours/week), smoking status (current vs. not current smoker), socio-economic status (professional or managerial/ intermediate/ routine or manual). Note, adjustment for energy intake was not included in fibre density models. Highlight=CIs do not span 1 in fully adjusted model.

6.5 Discussion

6.5.1 Total CVD, CHD and stroke

Unlike key findings from the previous chapter which focused on CVD mortality and where only cereal fibre intake was significantly associated with fatal stroke risk reduction, many more associations were apparent when non-fatal events were combined with these fatal events. Reduced risk of total stroke was associated with greater intakes of total fibre, fibre density, soluble and insoluble fibre. CHD risk reduction was associated with greater legume fibre intake and lower CVD risk with higher fibre density.

The estimated 13% risk reduction observed here with total stroke (fatal plus non fatal) and total dietary fibre intake, assessed as AOAC (per 11g/day increase) 0.87 (95% CI: 0.76 to 0.99) p=0.03 is of a similar magnitude to the 7% reduction per 7g/day seen in the recent systematic review and meta-analysis of other prospective cohort studies (discussed in Chapter 2) (Threapleton et al., 2013d). A clearer association was also observed for soluble fibre intake and stroke in the UKWCS, compared to the systematic review and meta-analysis. In the UKWCS, stroke risk decreased by 12% for each 3g/day higher soluble fibre intake 0.88 (95% CI: 0.77 to 1.00) p=0.05 and in the systematic review there was an indication of 6% reduction in risk 0.94 (0.88 to 1.01) for each 4g/day higher intake but in fact this result did not reach statistical significance (Threapleton et al., 2013d). This finding may be attributed to study population differences, namely the greater variation in dietary intakes in the UKWCS, compared to other studies identified in Chapter 2. This feature may allow associations with soluble fibre to be fully explored here, whereas no evidence of an association may be apparent in those studies with relatively few participants consuming adequate fibre intakes.

As seen for fatal stroke risk and cereal fibre intake (Chapter 5) (Threapleton et al., 2013b) there was a protective association with total stroke risk in this chapter. The Finnish ATBC Study of male smokers, identified during the systematic review, reported no protective associations with either insoluble or cereal fibre for stroke risk (Larsson et al., 2009). These findings contrast those observed in the UKWCS here and the other studies identified during the systematic review, which observed protective associations when examining cereal fibre intake (Oh et al., 2005, Kaushik et al., 2009) or insoluble fibre (Eshak et al., 2010, Kokubo et al., 2011).

The systematic review also identified just two other cohort studies reporting fruit or vegetable fibre intake in relation to stroke risk. As observed here for the UKWCS, there were no protective associations with either fruit or vegetable fibre intake (Larsson et al., 2009, Oh et al., 2005). As discussed in the previous chapter, the protective associations with cereal and not fruit or vegetable fibre may reflect protective benefits of cereal grains generally (Slavin, 2003), the greater relative proportion of insoluble to soluble type fibre (Lunn and Buttriss, 2007) or may simply reflect better measurement of cereal foods compared to fruit and especially vegetables as there is some evidence of over-reporting of vegetables in other British cohort studies (Bingham et al., 1997, Brunner et al., 2001). Additionally, a systematic review of cohort studies examining stroke risk with whole fruit and vegetable intake observed lower risk for each portion increase in fruit HR 0.95 (95% CI: 0.92 to 0.97) but not for vegetables HR 0.97 (95% CI: 0.92 to 1.02) (Dauchet et al., 2005) and another meta-analysis found lower risk of stroke with greater combined fruit and vegetable intake (He et al., 2006). These associations for reduced stroke with whole fruit and vegetable intake contrast the lack of any association observed in this and other studies (Larsson et al., 2009, Oh et al., 2005) suggesting that the protective benefits of fruit and vegetables for stroke may not be mediated via fibre intake but other micronutrients. Alternatively, the opposing findings may reflect differences in measurement error in estimating whole food and nutrient intakes, which may lead to reduced effect sizes.

Again, as in the previous chapter, there were no apparent associations with many fibre exposures and CHD risk in analyses of the full sample, the exception being an inverse association between total CHD risk and legume fibre. The general lack of associations directly contrast the pooled results generated from other cohort studies reported in Chapter 2, where lower CHD risk was associated with higher fibre intake. Several possible explanations exist for the null association seen with total CHD risk and total fibre intake: insufficient dietary variation, where the intakes of the sample only sit within a ‘flat’ portion of the dose-response curve may explain the lack of observations (Willett, 2013d). Despite participants here having a wide range of intake levels it is possible that associations are stronger in populations with lower intakes. The intake levels of the majority of UKWCS participants may be at or greater than a threshold level where benefits are seen. Another potential explanation for null associations is that the method of measuring diet was insufficiently accurate to measure differences that truly existed (Willett, 2013d). While this is possible, it is likely that a great deal of variation in fibre intakes do exist in this population and it is accepted that FFQs are able to

discriminate between and rank participants well with good agreement with food diaries (Brunner et al., 2001, Willett and Lenart, 2013) even though overestimation of nutrients tends to occur with FFQs (Bingham et al., 1994, Willett and Lenart, 2013).

Potential associations may also have been missed because of low statistical power and low case numbers (Willett and Lenart, 2013), however this is likely not the culprit as strong associations were apparent for stroke risk, where fewer cases existed than for CHD. Other explanations include that the assessment of diet did not encompass the true latent period where diet influences disease risk. There is also the possibility of an opposing variable, whereby another factor is associated with greater intake of fibre but is detrimental and creates negative confounding (Willett, 2013d).

The protective associations observed for greater fibre and stroke risk in the full sample were also observed in the obese but not healthy weight or overweight subsamples. Obesity is a well established risk factor for stroke (Goldstein et al., 2011) and results in systemic inflammation which is thought may work to initiate and mediate the development of vascular damage (Berg and Scherer, 2005) (discussed in Chapter 1). Additional fibre intake may confer no additional benefit in those who are at lower risk of stroke (i.e. not obese) but could be particularly beneficial where risk is greater because of higher BMI and inflammation.

Hypertension is a strong risk factor for stroke (O'Donnell et al., 2010) and inverse associations for stroke risk with greater fruit fibre intake became apparent when women with hypertension were excluded from the analysis. Additional fruit fibre intake may have no influence on risk in this already higher risk group but when these participants were excluded, general protective associations became apparent for the rest of the sample who had no history of hypertension.

The protective associations observed for fibre density and total CVD may simply be a reflection of the protective association with stroke rather than risk reduction for all events per se. The risk estimates for CVD tend to be weaker than with stroke and may simply reflect a dilution with the addition of CHD cases rather than any protective association separately for CHD. However, as with stroke events, inverse association for CVD became apparent when the sample was restricted to women without history of hypertension. Again, this finding indicates that greater fibre intake may not be additionally beneficial for those with hypertension, a risk factor for CVD, but may be helpful in healthy individuals.

6.5.2 Non-fatal cardiovascular disease

Similar to the inverse association for total stroke risk with higher cereal fibre, a 9% lower risk for non-fatal stroke with each 3g/day greater intake was observed in the dose-response analysis including all participants. Again, there were indications that total fibre, insoluble fibre and cereal fibre were protective in many of the category comparisons, primarily for risk of non-fatal stroke.

With non-fatal stroke, but not CVD and CHD, there were significant associations for many exposure categories compared to the lowest intake group but not for the dose-response models. For example for non-fatal stroke risk in each group compared to the lowest intake group of soluble fibre, risk in Q2 was 0.56 (95% CI: 0.37 to 0.86), Q3 0.64 (95% CI: 0.42 to 0.95), Q4 0.74 (95% CI: 0.49 to 1.12) and Q5 0.56 (95% CI: 0.34 to 0.92) but per 3g/day increase, risk was 0.90 (95% CI: 0.78 to 1.04). These results suggest that the association between fibre intake and stroke risk may not be linear in nature. Traditionally, the key criteria for identification of causal relationships in epidemiology has been the presence of a linear trend (Hill, 1965). However it has been suggested that a biological gradient may not be appropriate in contemporary epidemiology where diseases have multifactorial pathogenesis and where associations may be U-shaped or reach a threshold (Lucas and McMichael, 2005). It may well be that high intakes of some specific nutrients or elements are detrimental to health (Willett, 2013d) and it is unclear whether the association between fibre and CVD risk is likely to be linear or not. Certainly it is unlikely that high intakes of fibre would prove detrimental to health, as in the case of some other nutrients and it would seem logical that greater consumption, at least within plausible population intake ranges would likely have greater impact on lowering risk profile, although if associations are causal, beneficial effects may reach a threshold. An additional explanation for the protective associations that were more often observed in lower intake categories may be residual confounding and it may be some other aspect of diet or lifestyle in these lower fibre consumers that is conferring greater risk reduction.

Total fibre, fibre density, soluble, insoluble and fibre from legumes were associated with lower risk of non-fatal CVD in women without history of hypertension. As noted above, it may be that any beneficial action of fibre can stall disease development in healthier women but not offer benefit for those with this cardiovascular risk factors. The inclusion of participants with this risk factor in the analysis of the full sample may be clouding any associations and explain

why evidence is lacking when all participants were included in analysis together. This explanation may also extend to the indication of non-fatal stroke risk reduction seen with greater cereal fibre intake only in women with healthy BMI. The beneficial association seen in the full sample did not remain for overweight or obese women, who are at increased cardiovascular risk, and it may be that greater benefit can be seen in preventing the disease in those with fewer risk factors. However the smaller number of cases observed in these smaller subgroups may also explain null associations here. Additionally non-fatal stroke risk reduction was also observed for greater cereal fibre intake in women with history of hypertension.

6.5.3 Haemorrhagic vs. ischaemic stroke

Only one significant association was observed in the full sample analyses for haemorrhagic or ischaemic types of stroke; an 11% lower risk for ischaemic stroke was observed for each 3g/day greater cereal fibre intake. The protective associations observed with total fibre, soluble, insoluble and cereal fibre in the unspecified type stroke were also apparent when ischaemic cases were combined with the unspecified strokes. Combining cases in this way tended to slightly attenuate the protective associations but CIs were generally tighter in this larger sample of cases. The narrowing of CIs gives greater certainty to the estimates quantifying the degree of risk reduction seen with each specified fibre type.

Total fibre intake, soluble fibre, insoluble fibre and fibre from cereals were all associated with lower risk of 'mostly ischaemic' stroke in the full sample analysis. Four other cohorts identified during the systematic review of literature (Chapter 2) had also considered the associations between fibre and stroke sub-types (Oh et al., 2005, Larsson et al., 2009, Wallstrom et al., 2012, Kokubo et al., 2011). Findings from these studies do not help to explain observations seen for the UKWCS as they are not consistent between the studies. The Malmo Diet and Cancer Cohort Study identified an inverse association between fibre intake and ischaemic stroke in men but not women (Wallstrom et al., 2012). The Nurses' Health Study also saw no evidence of an association for ischaemic stroke risk with total, cereal, fruit or vegetable fibre but did observe a beneficial association for haemorrhagic stroke with cereal fibre intake (Oh et al., 2005). A Finnish cohort of male smokers found that only vegetable fibre was associated with reduced risk of ischaemic stroke but not haemorrhagic stroke when examining total fibre, soluble, insoluble, cereal, fruit and vegetable sources of fibre (Larsson et al., 2009). The fourth study, a Japanese cohort reported a protective association for women and not men with total fibre and insoluble fibre for both ischaemic and haemorrhagic stroke but there was no

evidence of associations with soluble fibre (Kokubo et al., 2011). The different findings mean formulating a consensus on risk of different types of stroke in relation to fibre intake is challenging. The contrasting observations may result from measurement error in assessing fibre intake from different foods in the various assessment tools or reflect the likely large variation in diets and variation in sources of fibre between the UK, US, Finland, Japan and Sweden.

In subgroup analyses for the UKWCS, results for the various fibre exposures do not tend to give a clear impression of the associations with haemorrhagic, ischaemic, unspecified or 'mostly ischaemic' stroke type. This issue may result from the greater uncertainty around estimates which comes from including fewer cases in these sub-group analyses or issues with multiple testing. When conducting multiple tests and using small sample sizes, there is greater chance of false positive results (Bowers et al., 2006a). In postmenopausal women, greater vegetable fibre was associated with increased risk for haemorrhagic stroke but decreased risk of ischaemic stroke. In the same subgroup, fibre intake and fibre density assessed as NSP were associated with risk reduction for unspecified type stroke but this was not the case when fibre intake or fibre density were assessed using the AOAC method.

In women without history of hypertension, more of the fibre exposures (total fibre, soluble fibre, insoluble fibre, cereal fibre, vegetable fibre, fibre from breakfast cereals) were associated with lower risk of ischaemic, unclassified and 'mostly ischaemic' strokes compared to those with personal history of hypertension. In those with this risk factor, just vegetable fibre and fibre from nut and seed sources were associated with risk reductions for either unspecified or 'mostly ischaemic' strokes. As discussed previously, the action of fibre may have greater effect in preventing disease rather than reverse disease progression in those already with risk factors. Oddly, greater total fibre and fibre from breakfast cereals were both associated with increased risk of ischaemic stroke in this subgroup. As discussed above, these results may be 'false positives' caused by multiple testing and smaller case numbers. Or, alternatively, residual confounding might somewhat explain these observations whereby women with knowledge of hypertension are both at greater risk of CVD and are following healthier diets. Excluding more cases occurring in the years after dietary assessment may resolve the issue of residual confounding and help to investigate this, however there are too few cases available to explore this here.

Protective associations for haemorrhagic stroke became apparent when examining overweight or obese women. For these women, greater intake of cereal fibre, fibre from breakfast cereals, insoluble fibre and total fibre, assessed as AOAC, were associated with risk reduction. These associations were not seen for haemorrhagic stroke in women with healthy BMI suggesting that the effect of fibre on risk is modified with greater BMI.

6.5.4 Acute vs. chronic heart disease

Protective associations were only apparent for fibre intake with MI and not the broader acute syndrome category. Unstable angina and other ischaemic heart diseases (ICD10 I24) were additionally included in the ACS category with MI. It is possible that the conditions included in the broader acute category and conditions included in the chronic disease category have different pathogenesis to MI. These other conditions may be influenced by different risk factors and any beneficial effect of fibre may do little to influence overall disease risk.

For chronic heart disease, just fibre from legumes was protectively associated with risk. This association remained in subgroup analyses for healthy BMI and those without history of hypertension. However, this lone observation, without any associations evident for total fibre intake, seems more likely a result of residual confounding and may reflect healthy behaviours of legume consumers rather than a specific benefit from fibre within legumes.

6.5.5 Strengths and limitations

As discussed in the previous chapter, strengths of this work include that there has been a relatively long period of follow up from this large prospective study and with the addition of non-fatal events here, case numbers are boosted. A benefit of this is that there are sufficient cases of each type of stroke or CHD event, allowing exploration of fibre intake in relation to the different types of events which is especially important for stroke because risk factors for the two main types (ischaemic and haemorrhagic), differ (Andersen et al., 2009). The pathology of disease development may be different for the two main stroke types and having sufficient cases of each allows exploration of this. While combining ischaemic with the unknown type stroke cases increased the number of strokes, which are likely to mostly be ischaemic, some sensitivity may be lost through including a small number of unidentified haemorrhagic stroke cases into this category.

A further unique strength of this cohort study is the use of a validated FFQ in a sample that includes diverse dietary intakes and this allows exploration of dose-response associations between very different levels of fibre intake with CVD risk. However, despite an early validation study with the FFQ (discussed in previous chapters) that indicated relatively stable dietary habits in participants for the five years since baseline (Greenwood et al., 2003) there are naturally limitations in assessing diet through any method and specific limitations with the use of FFQs, (Cade et al., 2002, Willett and Lenart, 2013) as discussed in Chapter 3.

Although the UKWCS includes women with a range of different education and socioeconomic classifications, it is a clear limitation that results from the UKWCS may not directly relate to the general population as participants are likely to be better educated and healthier than the UK population on the whole. Also, being a cohort of women means that the applicability of results to men of similar ages is unclear.

As highlighted in the previous chapter discussion, a major limitation with analysis of data from prospective observational studies is the potential for uncontrolled confounding, either via another lifestyle variable not considered in models or via an included confounder that has been imperfectly measured. It is conceivable that fibre itself is not directly acting to influence CVD risk, despite plausible mechanisms for its action (see Chapter 1), but another closely correlated nutrient or food component, or maybe both, may elicit the effect (Bingham et al., 1994).

Excluding participants with prevalent disease at baseline (cancer, diabetes, stroke, angina and heart attacks) removes those participants where knowledge of disease presence may have affected diet, thus allowing examination of disease incidence. However, hypertension is a risk factor for CHD and the inclusion of these women in the sample may serve to dilute risk estimates as associations are not consistent in those with or without hypertension. Another key risk factor for CVD development is being obese. These women were also retained in the full sample analysis to ensure sufficient case numbers and were then explored separately through sub-group analysis. Examining hypertension and overweight in this way allows more specific exploration of the relationship between fibre and these potential effect modifiers.

A limitation however is that relying on self-reported disease prevalence and BMI may lead to mis-classification of women with existing CVD conditions who will remain in analysis of total events or who are not included in the correct sub-group analyses. Various cohort studies have

attempted to estimate the validity of self-reported disease prevalence through comparison with medical records (Colditz et al., 1986, Okura et al., 2004, Britton et al., 2012). MI and cerebrovascular disease ascertainment rates in an American study of women, the Nurses' Health Study, were estimated as 68% and 66% respectively, although the authors ascribe these low rates to the application of strict criteria for case definitions (Colditz et al., 1986). From a cohort study of middle aged, mostly white participants in the US, good rates of reporting sensitivity were identified for self-reported hypertension (82%), stroke (78%) and MI (90%) when medical records were consulted, and similarly for specificity for the same conditions respectively, 92%, 99% and 98% (Okura et al., 2004). In the British Whitehall II cohort study, the validity of self reported stroke events was found to be high with almost 90% being validated and confirmed by medical records and just a small number of false positives or false negatives being identified (Britton et al., 2012).

6.6 Summary

Greater fibre intake, both in the form of soluble or insoluble fibre and particularly from cereal sources is associated with reduced CVD risk in the UKWCS. Patterns of association were clearer for strokes, especially ischaemic type and in women who did not have hypertension but relatively few associations were observed for CHD.

Exploration of effect modifiers (BMI, hypertension and menopausal status) has revealed some key differences in the risk profile for the various exposure and outcome combinations in these different groups.

In this and the previous chapter it has been possible to utilise huge quantities of dietary data, available because administering and processing nutrients from FFQs is relatively speedy and was carried out many years ago for the UKWCS. In the next chapter a different approach is taken and diet is assessed from four-day weighed food diaries. For practical reasons only case and control diaries are available for use and the sample numbers for controls are therefore greatly reduced. However, weighed food diary assessment of dietary intake is considered to be the gold standard in a field where no method is perfect (Willett and Lenart, 2013). The comparison of results generated with these two key methods will hopefully provide further insight into the association between dietary fibre and CVD risk.

Chapter 7 Dietary fibre intake and risk of cardiovascular disease, a case-cohort approach

7.1 Chapter overview

In this chapter a case-cohort approach is used to assess associations between dietary fibre intake, recorded from food diaries, and risk of CVD.

Risks for different CVD outcomes were estimated in association with higher intake of total fibre (NSP) and fibre from food sources. Fatal IHD, fatal stroke and fatal CVD (IHD plus stroke events) were explored along with risk of fatal plus non-fatal MI and ACS. The different sources of dietary and CVD event data used and presented in this chapter are displayed in Figure 7.1.

Findings, using the case-cohort method, include an inverse association between ACS and cereal fibre intake and between fatal stroke risk and total fibre intake. Positive associations were also observed with higher fibre intake and increased risk of fatal IHD and fatal stroke, although case numbers for the fatal outcomes were particularly small.

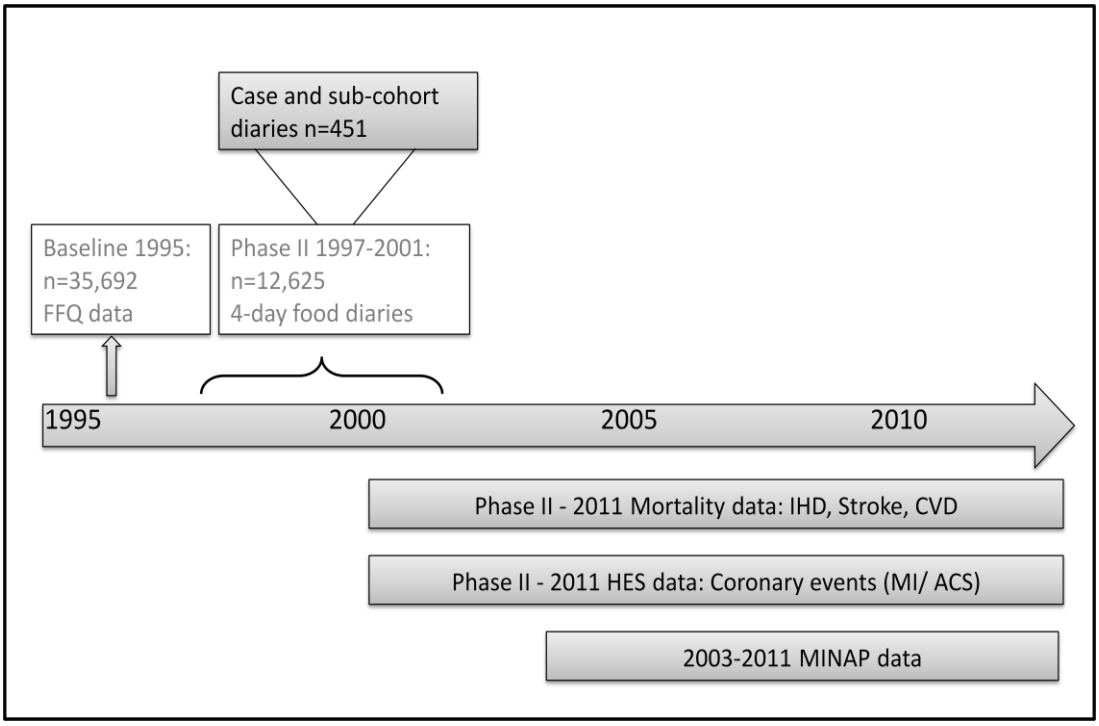


Figure 7.1 Data sources used in this chapter: Dietary data from case and sub-cohort food diaries, mortality and HES cases identified after receipt of food diaries plus MINAP cases identified since 2003

7.2 Background

The case-cohort approach, a variation on the nested case-control study, was proposed for failure time analysis by Prentice in the mid 1980's to deal with situations where it is not possible or practical to assess data from whole cohorts (Prentice, 1986). While it can be challenging to find participants willing to record their dietary intake in detail, over several days, processing the collected data for whole studies can be extremely time consuming and costly. This case-cohort method, as with case-control designs, only uses data for participants who become cases plus other members of the full cohort. In case-cohort analysis the controls are unmatched, are selected at random from the whole cohort and are known as the 'sub-cohort'. The design has the added advantage that sub-cohorts offer flexibility as they can be shared among different outcome classifications and timeframes (Barlow et al., 1999).

The systematic review of studies published in this area (reported in Chapter 2) (Threapleton et al., 2013d, Threapleton et al., 2013e) only identified a handful of studies that had assessed diet using measures other than FFQs (details are presented in Chapter 2 and the discussion section of the current chapter). Given that no dietary assessment method is without some error, the methods employed by studies that allow open ended responses and portion sizes to be estimated, undoubtedly capture a greater impression of actual intakes during the period of observation, be it 24-hours or 7 days in length (refer the discussion section of Chapter 3 for detailed description and comparison of dietary assessment approaches). Just one of these publications had explored CVD risk associated with fibre intake estimated using two different approaches and found different associations depending on the method used (Ward et al., 2012). In view of the different observations found in the EPIC Norfolk study and that two different diet assessment methods had been used in the UKWCS, associations explored in Chapters 5 and 6, using FFQ data, were extended here using the case-cohort approach.

7.3 Methods

7.3.1 Dietary data

As already detailed in Chapter 3, at the second point of data collection for the UKWCS, diet was assessed using weighed four-day food diaries. Dietary data from the diaries was analysed using our in house nutrient analysis software DANTE. It was only possible to estimate fibre intake from the diaries as NSP because British food tables do not include the AOAC values for the majority of food items. Similarly, values for soluble and insoluble fibre could not be accurately calculated for food diary data as many missing values exist in the food tables

(Holland et al., 1991). There are estimated factors for standard conversion between NSP and AOAC fibre but these are not useful to apply to total fibre intake in this context because the intake value would be similarly inflated for all participants. It is also not appropriate to apply a standard conversion factor for fibre assessed from specific food groups, because the proportions of lignin and resistant starch, the components not included in NSP values, differ from one food to another (Buttriss and Stokes, 2008).

Fibre density of the diet was calculated using the estimated total daily energy intake from the diaries. NSP from key food sources were also calculated to enable better comparison of these results with those presented in Chapter 5 (CVD mortality risk) and Chapter 6 (total CVD risk). CVD risk is presented in relation to NSP from total cereals, breakfast cereals, fruit (excluding juice), vegetables (excluding potatoes), legumes and nuts/seeds and details of which foods are grouped in each category are presented in Chapter 3.

7.3.2 Mortality data

The same IHD and stroke definitions were applied to mortality records to identify cases here, as presented and used in previous chapters (case definitions are presented in Chapter 4; mortality data are utilised in Chapters 5 and 6). In order to account for the potential influence of any latent disease, cases were only included where food diary information was received at least 12 months prior to the date of death. Cases used in this analysis extended from 1 year after questionnaire receipt to 3rd October 2012. In analyses where mortality data were combined with non-fatal cases, events extend only to 30th June 2011 to be consistent with the latest available case information from the two other sources of event data, HES and MINAP.

7.3.3 HES records

As detailed in prior chapters, CHD cases were identified using only the primary diagnosis field within the HES dataset. Because of limited resources, it was not possible to code dietary data for stroke cases or chronic heart disease identified in HES and a narrower CHD class was explored. ACS cases were defined using ICD10 codes I20.0 and I21.0-I24.9 and MI cases using I21.0 to I22.0 (refer to Chapter 4 for full details of event case classification).

7.3.4 MINAP records

Coronary events identified from MINAP were those participants where a final diagnosis of MI, threatened MI or ACS were recorded in the MINAP dataset (see Chapter 4 for details).

7.3.5 Censor date

Survival times were calculated, in years, from receipt of the follow-up questionnaire and food diary until either, earliest CVD event date, date of death from any other cause or the censor date, whichever came first.

When fatal and non-fatal events were combined, the censor date of 30th June 2011 was applied to HES and mortality events. Where only fatal events were included in analyses, the censor date was set as 3rd October 2012.

7.3.6 Exclusions

Of the 647 participants with follow-up data who were identified either as sub-cohort or case subjects, women were excluded where the following criteria were met:

1) Did not provide accurate NHS number or GP information at study baseline and were therefore not successfully traceable via the NHSIC (n=5).
2) Did not provide both lifestyle and dietary information (n=19).
3) Reported personal history of cancer, stroke, diabetes, angina or heart attack at study baseline or Phase II (n=167).
4) Died (any cause) or experienced CVD event within one year of receipt of Phase II data (n=10).
5) Food diary recording was less than 3 days (n=2).

7.3.7 Testing dose-response and non-linear associations

Unlike previous Chapters (5 and 6), associations between CVD risk and fibre intakes were assessed only using linear, dose-response models. Categorical exposures were not explored here because the limited case numbers for the different outcomes would result in extremely small numbers of cases in each intake category if risk was explored in this way and result in unreliable estimates. However, as detailed in Chapter 3 the fibre intake increments used in dose-response models were generated by calculating the mean intake difference between the five categories.

7.3.8 Case-cohort analysis and selection of sub-cohort participants

Cases with completed food diaries were identified from the 12,625 available participants and a corresponding number of non-case (sub-cohort or control) diaries were selected to be coded

1:1 with the cases. This enabled pairs of diaries, with one case and one control diary to be given to coders, guaranteeing coder blinding to the disease status of participants. Ensuring each coder had an equal number of both cases and controls also helped to minimise the effect of individual coder bias on the case and sub-cohort diaries.

No restrictions were placed on the sub-cohort random selection so that case diaries were also eligible to act within the sub-cohort, thus ensuring the sub-cohort was not artificially free of CVD cases. This is a main benefit of the case-cohort method (Prentice, 1986), in that sub-cohort diaries have been selected without restriction i.e. are not matched to specific cases, and may therefore act as controls in any future work. This sampling strategy, whereby some cases act as sub-cohort controls is accounted for in statistical analysis so estimates are not biased (discussed below). Strengths and weaknesses of the case-cohort design are discussed, at length, in the discussion of this chapter.

7.3.9 Survival analyses using the case-cohort method

Cox proportional hazards regression (Cox and Oakes, 1984) was used to explore the association between fibre intake assessed from diaries and CVD risk, with some modifications to allow for the sampling (Barlow et al., 1999). The Cox approach was modified for the case-cohort design according to the Prentice method (Prentice, 1986). Briefly, applying this method involves separating the dataset into participants selected to act as the sub-cohort (including any cases) and in a separate dataset, the additional cases. The two sets are separately processed using survival time and case information, before being rejoined.

The usual approach for preparing data for the proportional hazard models involves specifying failures (or cases) within the dataset. A modified approach, including probability weighting, was therefore taken, to ensure all participants in the sub-cohort set were classed as ‘non-failures’. Any cases in the sub-cohort set are additionally flagged as ‘failures’. The case and sub-cohort overlap in these analyses were n=2 for ACS outcomes and n=6 for CVD mortality. Participants in the case dataset were identified as ‘failures’ (Coviello, 2001). The two prepared datasets were then appended and case weights were specified to reflect the proportion of the total available diaries that had been selected as sub-cohort (i.e. 314 of 12,625) and thus attempt to estimate findings that would result from a full cohort analysis (Barlow, 1994, Barlow et al., 1999). Proportional hazard models were then run, as in previous chapters, though this modified method does not allow the inclusion of weighting factors. However,

including the inverse probability weighting for vegetarian status made little difference to risk estimates in Chapters 5 and 6.

Assumptions for proportional hazards were once again checked with the use of log-log survival curves for each outcome with all exposure and confounding variables. This condition was met for each exposure and with all covariates used in models.

A two-sided p-value ≤0.05 was considered statistically significant. Analyses were conducted using Stata version 12 (StataCorp, 2011).

7.3.10 Confounder adjustment

Adjustment for confounding variables was the same as Chapters 5 and 6 and the following three levels of adjustment were applied:

1) Age (years)
2) Age (years), alcohol (ethanol g/day), smoking status (non-smoker, current-smoker, ex-smoker), physical activity-metabolic equivalents (MET-hours/week) and socio-economic status (SES) (professional/managerial, intermediate or routine/manual).
3) As model 2 with the addition of energy intake from carbohydrate, fat and protein (excluding calories from ethanol) (kcal/day)* and BMI (kg/m^2).

* As in Chapters 5 and 6, results presented in here for fibre density models do not include adjustment for energy intake. Results were not appreciably different with or without adjustment for energy intake in fibre density models (data not shown).

7.3.11 Cohort subgroup analyses

Too few cases existed in the sample to explore associations by menopausal status, BMI category or history of hypertension.

7.4 Results

7.4.1 Sample and case numbers

There were 12,625 participants who responded to Phase II contact and provided both dietary and lifestyle information. Of these women, there were 186 ACS cases (102 MI cases) in English participants identified since receipt of Phase II data up to 30th June 2011. There were also 53 fatal IHD cases, 63 fatal stroke cases and therefore 116 fatal CVD cases identified from

questionnaire receipt up to 3rd October 2012. From the total 12,625 diaries, 373 sub-cohort diaries were randomly selected as 'controls'.

After applying exclusions to the sample of 283 case diaries, 144 were eligible to be used in analyses. These included 53 MI cases, 88 ACS cases, 25 fatal IHD cases, 42 fatal stroke cases and 67 fatal CVD cases (Table 7.1). After exclusions, 314 of the 373 diaries were eligible for use in the sub-cohort.

Table 7.1 Frequency of available participants and mean follow-up duration in both sub-cohort and case groups

	Available participants after exclusions applied to sample		Median (IQR) participant follow-up time in the study, years
	Unadjusted model	Fully-adjusted model	
Sub-cohort (control diaries)	314	291	10.6 (1.5)
MI case	53	49	6.1 (5.3)
ACS case	88	77	6.5 (4.5)
IHD mortality case	25	22	7.0 (6.2)
Stroke mortality case	42	39	7.8 (4.5)
CVD mortality case	67	61	7.7 (4.9)

7.4.2 Descriptive statistics

ACS cases were around 8 years older than sub-cohort participants at baseline, mean age was 59.9 years (SD 8.2) for ACS cases and 51.9 years (SD 9.1) in the sub-cohort (Table 7.2). Mortality cases were 62.4 years (SD 7.8) old at baseline. All case groups, except IHD mortality cases (n=25), had BMIs at least half a unit greater than the sub-cohort, where median BMI was 23.4 kg/m^2 (IQR 4.1).

The greatest proportion of current smokers was in the IHD mortality group (24%) but the case groups tended to include more reports of being a current smoker, at follow-up assessment (9 to 24%) than the sub-cohort, where only 7% of participants had indicated they were current smokers. Unsurprisingly, more cases also reported history of hypertension (33 to 48%) compared to the sub-cohort (19%). The sub-cohort participants also reported higher activity

levels and higher energy intake, which may be a reflection of the age difference between cases and the sub-cohort women.

In terms of social class, the case groups were not dissimilar to the sub-cohort in the proportion of women in each group; however the education profile of sub-cohort women was marginally better than the various case groups. In the sub-cohort, 37% of women reported being educated to degree level but the majority of the case groups had listed their highest educational achievement as A-levels.

Mean daily NSP intake in the sub-cohort was 17.5g (SD 6.3), higher than the ACS cases at 16.7g (SD 6.0) or the fatal CVD cases 16.6g (SD 6.3). These differences between the cases and sub-cohort are less apparent when comparing NSP density. The mean NSP density for the fatal CVD cases was 9.9g/1000kcal/day (SD 3.7) and for the ACS cases was 10.1g/1000kcal/day (SD 3.9) but this was marginally lower in the sub-cohort participants 9.8g/1000kcal/day (SD 3.2).

Descriptive characteristics of all participants (cases and sub-cohort) based on fibre intake, as assessed from the diaries, are discussed and presented in Chapter 3 (Tables 3.6, 3.7 and 3.8). In summary, lower fibre consumers tended to report lower energy intake, marginally higher BMI and lower levels of physical activity. In fibre density groups, the differences in energy intake were less distinct, with similar reported energy intakes across groups, except the highest fibre density group where energy intake was markedly lower. There was also little difference in BMI between these fibre density groups.

Lower fibre consumers tended to report lower levels of academic achievement and a higher proportion were grouped in the lower socio-economic class.

Table 7.2 Follow-up assessment cross-sectional characteristic for MI, ACS, IHD mortality, stroke mortality cases and randomly selected sub-cohort participants

		MI case	*ACS case*	*IHD mortality case*	*Stroke mortality case*	*CVD mortality case*	*Sub-cohort*
N		53	88	25	42	67	314
Age at baseline, years		61.1 (12.4)	60.5 (12.6)	65.0 (10.4)	63.9 (12.2)	64.3 (10.9)	50.2 (13.3)
BMI at Phase II, kg/m^2		24.0 (4.0)	24.2 (4.3)	23.1 (6.6)	24.1 (3.5)	24.0 (4.2)	23.4 (4.1)
Smoking status at Phase II (%)	Not a current smoker	58 (91)	79 (90)	19 (76)	37 (88)	56 (84)	291 (93)
	Current smoker	5 (9)	9 (10)	6 (24)	5 (12)	11 (16)	23 (7)
Diet group at baseline (%)	Meat-eaters	40 (75)	64 (73)	14 (56)	30 (71)	44 (66)	183 (58)
	Fish-eaters	5 (9)	10 (11)	5 (20)	5 (12)	10 (15)	46 (15)
	Vegetarian	8 (15)	14 (16)	6 (24)	7 (17)	13 (19)	85 (27)
Socio-economic status NS-SEC at baseline (%)	Professional/ managerial	33 (63)	57 (66)	17 (68)	25 (63)	42 (65)	207 (67)
	Intermediate	18 (35)	25 (29)	8 (32)	12 (30)	20 (31)	85 (27)
	Routine and manual	1 (2)	4 (5)	0	3 (8)	3 (5)	18 (6)
Highest educational achievement at baseline (%)	No formal record	13 (26)	20 (24)	1 (4)	11 (31)	12 (21)	33 (11)
	O-level	10 (20)	17 (20)	8 (35)	4 (11)	12 (21)	82 (28)
	A-level	15 (30)	26 (31)	7 (30)	11 (31)	18 (31)	71 (24)
	Degree	12 (24)	20 (24)	7 (30)	9 (26)	16 (28)	107 (37)
Menopause status Phase II (%)	Post	51 (96)	84 (95)	25 (96)	38 (90)	63 (93)	195 (62)
	Pre	2 (4)	4 (5)	1 (4)	4 (10)	5 (7)	120 (38)
History of hypertension at Phase II (%)	Yes	20 (43)	26 (34)	8 (33)	20 (48)	28 (42)	55 (19)
	No	27 (57)	51 (66)	16 (67)	22 (52)	38 (58)	241 (81)
Physical activity at Phase II (%)	No weekly activity	8 (16)	12 (16)	5 (23)	3 (7)	8 (13)	20 (7)
	Light/moderate activity	28 (56)	42 (53)	13 (59)	23 (56)	36 (57)	136 (46)
	Vigorous activity 1-2 times per week	11 (22)	17 (22)	3 (14)	9 (22)	12 (19)	78 (26)
	Vigorous activity 3 or more times per week	3 (6)	8 (10)	1 (5)	6 (15)	7 (11)	61 (21)
Ethanol at Phase II from diary, g/day mean (SD)		6.4 (12.3)	7.2 (11.0)	5.6 (11.9)	11.3 (15.3)	9.2 (14.3)	10.2 (13.0)
Energy intake at Phase II from diary, kcal/day		1757 (545)	1715 (68)	1758 (529)	1712 (385)	1758 (626)	1838 (583)
Protein intake from food diary g/day		68.8 (17.0)	68.8 (17.5)	69.8 (20.2)	65.5 (15.9)	66.0 (17.9)	69.7 (23.6)
Total fat intake from food diary g/day		62.8 (30.3)	62.1 (27.0)	67.5 (22.7)	57.0 (27.4)	61.9 (32.8)	66.8 (34.6)
Carbohydrate intake from food diary g/day		217.8 (46.1)	218.1 (62.8)	211.2 (52.3)	226.0 (96.0)	222.0 (79.6)	226.8 (80.7)

		MI case	*ACS case*	*IHD mortality case*	*Stroke mortality case*	*CVD mortality case*	*Sub-cohort*
Saturated fat intake at Phase II from diary g/day		21.7 (12.8)	22.0 (12.5)	24.5 (14.8)	19.8 (9.1)	21.8 (12.7)	22.5 (13.9)
NSP at Phase II from diary, g/day		16.7 (8.7)	16.7 (8.4)	18.1 (11.6)	15.3 (7.9)	15.8 (9.7)	17.1 (8.4)
NSP density at Phase II from diary, g/1000kcal/day		9.5 (5.3)	9.5 (4.9)	11.0 (5.8)	8.7 (3.5)	8.8 (5.0)	9.6 (4.4)
NSP within Foods from food diary, g/day	Cereals	2.0 (1.1)	1.7 (1.6)	2.3 (1.4)	1.9 (1.3)	2.1 (1.3)	2.0 (1.8)
	Breakfast Cereals	0.5 (1.2)	0.4 (1.2)	1.0 (1.9)	0.4 (1.3)	0.5 (1.8)	0.6 (1.8)
	Fruit	2.7 (3.0)	3.1 (3.1)	2.9 (3.2)	2.7 (2.0)	2.7 (2.6)	2.9 (2.7)
	Vegetables	3.3 (3.0)	3.3 (2.6)	3.2 (4.6)	2.7 (2.1)	3.0 (2.9)	3.4 (2.6)
	Legumes	0.9 (1.6)	0.8 (1.9)	0.9 (1.6)	0.7 (1.9)	0.8 (1.7)	0.8 (1.8)
	Nuts/Seeds	0 (0.3)	0 (0.2)	0 (0.4)	0 (0.3)	0 (0.3)	0 (0.2)

Values are median (IQR) or numbers (percentages) unless otherwise stated.

7.4.3 Survival analysis

In total, 314 participants acted as non-cases and were followed for 10.6 years (IQR 1.5) on average (median follow-up length). The median study time between receipt of Phase II diary and ACS events (n=88) was 6.5 years (IQR 4.5) and for all CVD mortality cases (n=67) was 7.7 years (IQR 4.9) (Table 7.1).

HRs and 95% CIs for each outcome were generated by assessing dose-response trends and are presented for age-adjusted, mid adjusted and fully adjusted models (as detailed above) in Tables 7.3 to 7.6, for the various dietary fibre exposures.

Incident MI or ACS

Higher intake of fibre from cereal foods was associated with lower risk of ACS, with each 3g/day increase, risk was 0.76 (95% CI: 0.58 to 1.00) in the fully adjusted model. None of the other fibre exposures were associated with risk for MI or ACS in any of the models. Risk estimates tended to sit below 1, with wide CIs when assessing risk in relation to total fibre intake, cereal fibre, fibre from breakfast cereals or fruit. However, for the other fibre exposures, risk estimates were close to or just higher than 1, again with wide CIs.

Fatal IHD, Stroke and CVD

Higher intake of total fibre was associated with lower fatal stroke risk in the age-adjusted model and also Model 2, HR 0.76 (95% CI: 0.58 to 1.00). However, this association was attenuated when BMI and energy intake were also included as covariates and the HR was 0.80 (95% CI: 0.60 to 1.06).

The risk for fatal stroke was significantly increased with higher intake of fibre from nuts/seeds HR 1.43 (95% CI: 1.04 to 1.98), in the fully adjusted model. Additionally, in the fully adjusted models higher risk for fatal IHD was associated with higher intake of fibre from vegetable, HR 1.21 (95% CI: 1.03 to 1.43) and legume sources, HR 1.18 (95% CI: 1.06 to 1.31).

Fatal CVD risk estimates tended to be lower than 1 for total fibre, cereals, fruit and vegetables and of a similar magnitude to observations in previous chapters, around 5-15% risk reduction, but CIs remained wide here and ultimately none of the exposures were significantly associated with fatal CVD risk.

Table 7.3 Total dietary fibre and fibre density of the diet and associated risk of total MI/ACS or fatal IHD, Stroke and CVD

	Cases in fully adjusted model	HR (95% CI) p-value, with each 4g/day higher NSP intake			HR (95% CI) p-value, with each 2g/1000kcal/day higher NSP density		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Total MI	49	0.96 (0.77, 1.18) 0.68	0.95 (0.75, 1.21) 0.70	0.98 (0.74, 1.30) 0.90	1.06 (0.87, 1.31) 0.55	1.02 (0.81, 1.27) 0.90	1.00 (0.79, 1.26) 0.98
Total ACS	77	0.91 (0.77, 1.09) 0.31	0.91 (0.74, 1.11) 0.34	0.95 (0.75, 1.20) 0.66	1.02 (0.86, 1.22) 0.79	0.99 (0.81, 1.21) 0.92	0.96 (0.78, 1.18) 0.71
Fatal IHD	22	1.11 (0.79, 1.56) 0.54	1.15 (0.82, 1.61) 0.43	1.16 (0.81, 1.66) 0.41	1.17 (0.86, 1.59) 0.30	1.11 (0.85, 1.44) 0.45	1.13 (0.85, 1.52) 0.40
Fatal Stroke	39	0.75 (0.58, 0.97) 0.03	0.76 (0.58, 1.00) 0.05	0.80 (0.60, 1.06) 0.12	0.84 (0.67, 1.06) 0.15	0.88 (0.69, 1.11) 0.27	0.81 (0.63, 1.05) 0.12
Fatal CVD	61	0.87 (0.70, 1.10) 0.25	0.91 (0.71, 1.16) 0.43	0.94 (0.72, 1.23) 0.66	0.96 (0.78, 1.18) 0.68	0.98 (0.79, 1.21) 0.85	0.94 (0.75, 1.19) 0.61

Key as below

Table 7.4 Fibre from cereals and breakfast cereals and associated risk of total MI/ACS or fatal IHD, Stroke and CVD

	Cases in fully adjusted model	HR (95% CI) p-value, with each 3g/day higher total cereal fibre intake			HR (95% CI) p-value, with each 2g/day higher breakfast cereal fibre		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Total MI	49	0.82 (0.61, 1.08) 0.16	0.80 (0.58, 1.09) 0.16	0.80 (0.58, 1.11) 0.18	0.90 (0.70, 1.16) 0.43	0.85 (0.63, 1.16) 0.31	0.85 (0.62, 1.16) 0.30
Total ACS	77	0.77 (0.61, 0.97) 0.03	0.75 (0.58, 0.96) 0.03	0.76 (0.58, 1.00) 0.05	0.87 (0.71, 1.07) 0.19	0.83 (0.65, 1.07) 0.15	0.83 (0.64, 1.07) 0.15
Fatal IHD	22	0.99 (0.69, 1.42) 0.95	0.91 (0.60, 1.39) 0.67	0.90 (0.59, 1.35) 0.60	1.04 (0.81, 1.33) 0.78	0.98 (0.73, 1.32) 0.89	0.98 (0.72, 1.33) 0.90
Fatal Stroke	39	0.78 (0.57, 1.05) 0.10	0.77 (0.56, 1.06) 0.11	0.80 (0.58, 1.11) 0.17	0.80 (0.60, 1.06) 0.12	0.79 (0.58, 1.09) 0.15	0.78 (0.57, 1.08) 0.14
Fatal CVD	61	0.85 (0.66, 1.10) 0.21	0.83 (0.63, 1.10) 0.20	0.85 (0.63, 1.13) 0.25	0.90 (0.73, 1.10) 0.30	0.87 (0.68, 1.11) 0.27	0.85 (0.66, 1.11) 0.23

Model 1: Adjusted only for age (years).

Model 2: Adjusted for age (years), socio-economic status (professional or managerial/ intermediate/ routine or manual), smoking (current vs. noncurrent smoker), physical activity level (no weekly physical activity/ light or moderate physical activity/ vigorous activity 1-2 times per week/ vigorous activity 3 or more times per week) and alcohol intake (g/day).

Model 3: As Model 2 and additionally adjusted BMI (kg/m^2) and energy intake from carbohydrates, fat and protein (kcal/day).

Shading indicates results where CIs do not span 1.

Table 7.5 Fibre from fruit and vegetables and associated risk of total MI/ACS or fatal IHD, Stroke and CVD

	Cases in fully adjusted model	HR (95% CI) p-value, with each 1g/day higher fruit fibre intake			HR (95% CI) p-value, with each 1g/day higher vegetable fibre intake		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Total MI	49	0.93 (0.79, 1.10) 0.41	0.94 (0.79, 1.13) 0.53	0.95 (0.79, 1.15) 0.62	1.11 (0.98, 1.26) 0.10	1.09 (0.97, 1.23) 0.16	1.10 (0.97, 1.24) 0.14
Total ACS	77	0.98 (0.87, 1.11) 0.79	0.99 (0.86, 1.13) 0.86	1.01 (0.87, 1.16) 0.93	1.08 (0.97, 1.20) 0.17	1.06 (0.96, 1.18) 0.26	1.07 (0.96, 1.19) 0.20
Fatal IHD	22	0.87 (0.69, 1.11) 0.26	0.86 (0.65, 1.12) 0.27	0.83 (0.64, 1.10) 0.19	1.20 (0.99, 1.45) 0.06	1.21 (1.03, 1.43) 0.02	1.21 (1.03, 1.43) 0.02
Fatal Stroke	39	0.89 (0.70, 1.13) 0.33	0.89 (0.67, 1.18) 0.41	0.91 (0.67, 1.22) 0.52	0.83 (0.67, 1.01) 0.07	0.83 (0.65, 1.04) 0.11	0.84 (0.66, 1.06) 0.15
Fatal CVD	61	0.88 (0.73, 1.07) 0.20	0.88 (0.70, 1.10) 0.27	0.89 (0.70, 1.13) 0.34	0.98 (0.83, 1.15) 0.79	1.00 (0.85, 1.18) 0.98	1.01 (0.87, 1.19) 0.85

Key as below

Table 7.6 Fibre from legumes and nuts/seeds and associated risk of MI/ACS or fatal IHD, Stroke and CVD

	Cases in fully adjusted model	HR (95% CI) p-value, with each 1g/day higher legume fibre intake			HR (95% CI) p-value, with each 1g/day higher nut/seed fibre intake		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Total MI	49	1.03 (0.84, 1.67) 0.76	1.03 (0.83, 1.27) 0.81	1.04 (0.84, 1.29) 0.71	0.91 (0.55, 1.53) 0.73	1.05 (0.71, 1.58) 0.80	1.11 (0.73, 1.68) 0.63
Total ACS	77	1.01 (0.85, 1.21) 0.91	0.99 (0.81, 1.20) 0.88	1.01 (0.83, 1.22) 0.95	0.82 (0.50, 1.34) 0.43	0.91 (0.58, 1.42) 0.67	1.01 (0.66, 1.54) 0.97
Fatal IHD	22	1.23 (1.07, 1.42) 0.01	1.17 (1.03, 1.33) 0.02	1.18(1.06, 1.31)<0.01	0.87 (0.44, 1.71) 0.69	1.06 (0.58, 1.93) 0.85	1.05 (0.47, 2.36) 0.91
Fatal Stroke	39	0.96 (0.75, 1.23) 0.73	0.93 (0.72, 1.20) 0.58	0.96 (0.74, 1.25) 0.77	1.16 (0.73, 1.84) 0.53	1.31 (0.90, 1.90) 0.16	1.43 (1.04, 1.98) 0.03
Fatal CVD	61	1.10 (0.87, 1.40) 0.43	1.08 (0.89, 1.31) 0.42	1.11 (0.92, 1.33) 0.27	1.08 (0.68, 1.74) 0.74	1.25 (0.87, 1.80) 0.22	1.36 (0.99, 1.87) 0.06

Model 1: Adjusted only for age (years).

Model 2: Adjusted for age (years), socio-economic status (professional or managerial/ intermediate/ routine or manual), smoking (current vs. noncurrent smoker), physical activity level (no weekly physical activity/ light or moderate physical activity/ vigorous activity 1-2 times per week/ vigorous activity 3 or more times per week) and alcohol intake (g/day).

Model 3: As Model 2 and additionally adjusted BMI (kg/m^2) and energy intake from carbohydrates, fat and protein (kcal/day).

Shading indicates results where CIs do not span 1.

7.5 Discussion

7.5.1 Result summary

A protective association was observed between cereal fibre intake and ACS, where 77 cases were available in fully-adjusted models but no associations were observed for MI, where 49 cases were included in fully adjusted models. Since MI cases make up the majority of the ACS cases, with other acute coronary events being included here, the different observations could indicate that cereal fibre has specific associations with ACS events that are not classified as MI. However, a clear limitation in this work is the small number of available cases that may account for the wide CIs seen with many of the risk estimates, such as with MI. For this reason, it would be unwise to place too great an emphasis on the precise risk estimates or indeed the specific lack of association with MI, as this may well be related to the wider CIs that come with fewer cases in analyses. However, given the small number of cases, many of the risk estimate CIs were not extremely wide and were similar in width to results in Chapters 5 and 6. The food diary approach, a potentially better assessment tool and the efficient case-cohort design may be mitigating the otherwise very wide CIs which could result from having few cases for some outcomes.

Higher intake of total fibre was associated with lower risk of fatal stroke in the age-adjusted and mid-adjusted model, but this association became non-significant, with wider CIs when BMI and energy intake were included as covariates. The attenuation may reflect the influence of BMI or energy intake on the association; the protective association may be mediated via an influence of fibre on energy intake and ultimately BMI. However the widened CIs for the fully-adjusted model may reflect the greater uncertainty around the estimate that comes from fewer cases being available to contribute data in the fully-adjusted model (n=39) than the age adjusted model (n=42). Two cases had missing information on socioeconomic classification and one for physical activity level.

In contrast to the protective association observed between fatal stroke risk and total fibre intake, an increased risk of fatal stroke was also associated with higher intake of fibre from nuts and seeds. It is worth bearing in mind that fibre from nuts and seeds contribute only a small fraction to the total intake of fibre. Additionally, unlike the FFQ which gives an impression of long-term intake, the diary may not be as good at assessing intake of fibre from nuts and seeds as they may be more sporadically consumed than other general foods like cereals, fruit and vegetables. This positive association, as with any other negative associations,

may be the result of residual confounding or indeed reverse causality, where participants at greater risk of stroke, for example with a family history, have consumed higher levels of nuts and seeds and also experience higher mortality rates.

The other positive associations between fatal IHD with higher intake of fibre from legumes or vegetables were unexpected but because of the even smaller number of cases in this case group (n=22), results may be particularly unreliable. An additional issue is the far greater proportion of case participants with hypertension (33-48% for the various outcomes), compared to the sub-cohort participants (19%). Because of the small case numbers available here, it was not possible to explore subgroup associations, as in Chapters 5 and 6. It is therefore not possible to know whether the inclusion of participants with hypertension is attenuating associations in the same manner as was seen in Chapter 5; associations between fatal stroke and fibre density became apparent only when hypertensive participants were removed. The positive associations here could be the result of reverse causality, where women with hypertension are adhering to healthier dietary practices whilst also being at greater risk for fatal IHD.

7.5.2 Findings compared with previous chapters

In Chapter 5, no associations were observed in the full sample when exploring fatal CVD risk and fibre estimated from the FFQs but protective associations were apparent in different subgroups. For example in overweight women, fatal stroke risk was reduced with higher intake of fibre from total cereals and nuts/seeds but there were no associations in the full sample. Conceivably, any protective effect of fibre on CVD development and risk may well be moderated by personal history of hypertension or BMI and the inability to explore these associations in this chapter is certainly a limitation since associations may not be visible when all participants are considered together.

As few protective associations were observed between coronary events and fibre in the previous chapter it is difficult to draw conclusions in light of the one positive association for ACS and cereal fibre seen in this chapter. In Chapter 6, lower MI risk was associated with higher fibre density and insoluble fibre intake but with too few MI cases available here, again it is challenging to draw meaningful conclusions.

7.5.3 Results compared to other studies

Few other cohort studies report fibre intake calculated using methods other than FFQs, in relation to risk of CVD. The systematic review and update searches (reported in Chapter 2) (Threapleton et al., 2013d, Threapleton et al., 2013e) indentified 4 studies, reported in 6 publications that used methods other than FFQs to assess diet (Bazzano et al., 2003, Streppel et al., 2008, Wallstrom et al., 2012, Ward et al., 2012, Crowe et al., 2012, Chuang et al., 2012). Diet was assessed using 24 hour recalls, diet histories or interviews in three of the studies (Bazzano et al., 2003, Streppel et al., 2008, Wallstrom et al., 2012), with 7 day weighed food diaries and FFQs in the Norfolk branch of the EPIC study (Ward et al., 2012) and by various methods in the pan-European EPIC studies (Crowe et al., 2012, Chuang et al., 2012). Of the identified studies, the most closely related in terms of population and assessment method was the work of Ward and colleagues using data from the Norfolk arm of the EPIC study. The authors report protective associations between fibre and CVD risk when using fibre estimated from food diaries and not with FFQs. The authors suggest that the additional information collected with the diary, in terms of the important contributors to fibre intake such as detail on the fibre source and portion size, may explain their differing observations by instrument (Ward et al., 2012).

In both UKWCS data and findings from EPIC Norfolk (Ward et al., 2012) limited associations were observed for CHD risk with fibre intake assessed using FFQs. In the EPIC study the protective association with fibre assessed from diaries contrasts the null association seen in this chapter for total fibre intake. Additionally the null associations with FFQ data in both studies contrast findings in Chapter 2 from many other cohorts. The different findings in the EPIC study with the two assessment tools and also between the two UK studies and pooled results in Chapter 2 is somewhat similar to the contrasting results found in studies of breast cancer and fat intake in the past decade. Two studies reported increased breast cancer risk with fat intake assessed using food diaries (Bingham et al., 2003, Freedman et al., 2006), where previously no associations had been seen with fat calculated from FFQs (Smith-Warner et al., 2001). This prompted a large collaborative pooling study to explore risk of breast cancer with fat intake, also assessed using food diaries (Key et al., 2011). The pooling study did not confirm the findings of previous work and suggested that lower case numbers in the earlier studies may be responsible for the different associations.

Applying this principle here, the larger case numbers reported in the EPIC Norfolk study for diet assessed using food diaries compared to the limited cases available in this chapter, suggests the observed protective association in the larger study is more reliable. Additionally, given the weight of data contributed by other studies which did assess diet using FFQs and the protective association that was seen for CHD risk with each 7g/day higher fibre intake HR 0.90 (95% CI: 0.87 to 0.94) (reported in Chapter 2), it seems likely that an association exists. The generally null association for total fibre intake and CHD risk in the UKWCS may therefore be explained by fewer cases or different population characteristics. As noted in earlier chapters, participants may be generally healthier than other study populations and may consume sufficiently high intakes of fibre, and may experience little additive benefit of greater intakes.

Dietary data from the UKWCS was incorporated into the pooling project (discussed above) but was identified as being most different from the other studies (Key et al., 2011). Differences between the studies were attributed to possible real differences, due to chance or were related to the different dietary coding methods employed in this cohort, compared to others (Key et al., 2011). These explanations of different results for the UKWCS are valid in this context and the contrasting results observed in this chapter, compared to those reported for Ward and colleagues (Ward et al., 2012) and in the meta-analyses of studies (reported in Chapter 2) (Threapleton et al., 2013d, Threapleton et al., 2013e). Opposing findings may therefore be due to real differences between the sample populations, chance, or variation in dietary assessment.

UKWCS participants consume higher levels of fibre than the general population (discussed and presented in Chapter 3). Since the FFQ was not so dissimilar to those used in other studies, it suggests that the different observations may be due to actual differences in fibre intakes rather than the tool used. The beneficial effect of fibre may have greater influence at lower fibre intakes and therefore associations may be less apparent in a sample such as the UKWCS, where many participants meet current recommended intake levels.

It is fair to consider that the null associations observed here may, in fact, be true and the influence of fibre on CVD risk is negligible. If this is the case, residual confounding of other dietary or lifestyle factors may be responsible for apparent associations in other study populations where there are relatively fewer well educated or health-conscious individuals. However, given the wealth of contrasting observations (Threapleton et al., 2013d, Threapleton et al., 2013e), this is unlikely to be the case.

7.5.4 Strengths and limitations of the case-cohort approach

The case-cohort design for failure time analysis was proposed by Prentice as a means of efficiently assessing exposure-disease associations in large studies (Prentice, 1986). Because of the usually low disease occurrence in large studies, much of the covariate data in the disease free subjects is redundant and would be costly and time-consuming to process (Prentice, 1986). The practical solution is therefore to process data for only a proportion of the disease-free participants, as is done in nested case-control or case-cohort designs.

The unique design of the case-cohort approach has key strengths and weaknesses. The fact that cases and controls are drawn from the same population, including disease-free participants at baseline, is a real strength. This attribute ensures the 'study base principle' and 'comparable accuracy principle' are not violated. These principles, outlined by Wacholder and colleagues concern the selection bias that may be introduced by selecting control subjects from a different population as cases and a form of exposure information bias that may be introduced when the assessment of exposures, in cases or controls, differs in accuracy (Wacholder et al., 1992a).

A significant benefit of the case-cohort method is the flexibility that comes from having unrestricted or randomly selected controls. This efficiency means that the sub-cohort may be shared among different outcome groups as no diagnosis restriction was used to identify the control participants (Self and Prentice, 1988, Barlow et al., 1999). The availability of a larger number of controls serves to improve the precision of risk estimates (Wacholder et al., 1992c). However, this method of control selection can lead to an issue with time comparability (Wacholder et al., 1992c) as cases contribute fewer study years than the sub-cohort, who likely remain in the study until censor date. This is especially a concern when exposures vary over time (Wacholder et al., 1992b). However this issue may not be such a concern for fibre intake in the UKWCS since reasonable stability was identified in dietary habits during a validation study (discussed in Chapter 5) (Greenwood et al., 2003).

Barlow and colleagues discuss the relative strengths and weaknesses of applying case-control and case-cohort methods within an example dataset (Barlow et al., 1999). They identify that the case-cohort method, while having some distinct advantages over case-control designs, is used infrequently in practice. The authors cite reasons such as perceived analytic complexity, perceived difficulty in variance computation and lack of appropriate software for case-cohort

methods as possible explanations for the infrequent use but recommend the design to be used in situations where flexibility is desired (Barlow et al., 1999).

In applying this method to assess CVD in the UKWCS, a major advantage has been the ability to use the larger set of sub-cohort diary information in analysing each of the different cardiovascular outcomes. Aside from the advantages brought with using a larger control dataset irrespective of the number of cases, a limitation is the relatively small numbers of case participants for some of the outcomes. The small numbers of cases may be responsible for greater uncertainty around estimates and therefore lack of clear associations. This makes it challenging to determine if associations do exist but are masked by the wider CIs resulting from limited case numbers.

7.5.5 Strengths and limitations of dietary and covariate assessment

Measurement error related to dietary assessment is a persistent problem when exploring relationships between diet and diseases. This error is a real concern and can have large effects on risk estimates. For example, the day to day variation in diet could mean that assessing intake over just a few days gives greater chance that people are misclassified because the true long term intake could be misreported in a short time-frame (Willett, 2013c). This within-person variation or random fluctuation above and below the true long term intake average could substantially distort RRs in epidemiologic associations. In general, the effect would be to reduce the strength of associations rather than exaggerate them; findings from a validation study of dietary assessment methods, using biomarkers, suggests that measurement error in FFQs can severely attenuate risk estimates (Kipnis et al., 2003). In this validation study and using protein intake as an example, it was found that a true association of 2 would appear as 1.1 or lower. This study compared 24 hour recalls with FFQs but given that diaries are considered the gold standard in terms of dietary assessment (Willett and Lenart, 2013) this problem of measurement error may be substantially lower for findings from food diaries. However in general, readers should place less emphasis on precise risk estimates and rather, should consider the general direction and relative magnitude of associations.

In addition to dietary measurement error, error and bias in assessing covariates is also a concern. For example, energy expenditure varies greatly between persons and fidgeting may contribute to energy expenditure of hundreds of calories per day and conventional assessment methods may not pick up this detail (Willett, 2013b). Where errors exist in the measurement

of confounders such as energy expenditure, the bias may be in either direction, not just to dilute the effect (Kipnis and Freedman, 2008, Freedman et al., 2011).

With diet diaries, the ability for participants to record diet in such an open-ended method and specify or estimate portion sizes is a significant advantage, over other methods like FFQs where food items and portion sizes are pre-specified. The strengths and limitations of different dietary assessment methods are discussed in depth in the discussion section of Chapter 3.

Dietary intake in the UKWCS has a wide range and is reflective of the recruitment of many health-conscious and vegetarian women. The wide variation allows assessment of potentially beneficial nutrients, such as fibre, since there is a greater proportion of the sample meeting dietary recommendations, than would be seen in a sample of the general population (discussed in Chapter 3). Although the recruitment here offers advantages, findings may be generally biased with respect to the general population. It is possible that protective associations observed in these women or, conversely, the lack of associations may not apply to other groups. For example, most of the influence of fibre on CVD risk may occur at lower intake levels and therefore be best displayed in populations where participants consume lower levels of fibre. This particular hypothesis may explain why findings here, especially for coronary outcomes do not reflect those generally observed in other studies (Threapleton et al., 2013d, Threapleton et al., 2013e) (reported in Chapter 2).

7.6 Summary

The case-cohort method was applied to assess associations between fibre and CVD risk, with fibre estimated using diaries in cases and a random sample of cohort participants. This method offers the primary benefit that all sub-cohort data can be used with each different outcome type.

Few associations were seen here and protective associations were observed for risk of ACS and higher intake of cereal fibre and for fatal stroke risk and total fibre intake. However, increased risks were also observed for fatal IHD and fatal stroke with higher intake of some of the other fibre exposures. Low case numbers are a major limitation here and contribute to the greater uncertainty around risk estimates.

It is possible that due to the many potential types of measurement error with quantifying diet and other confounders (discussed in Chapter 3 and above), true associations are attenuated and may even appear non-significant where they were weak.

With the exception of the significant association seen in the largest case group (88 ACS cases) with only cereal fibre, findings here do not reflect those seen in another, larger British study which used similar dietary assessment methods (Ward et al., 2012) or the combined findings of other prospective studies (Threapleton et al., 2013e, Threapleton et al., 2013d) (reported in Chapter 2).

Chapter 8 Summary and Conclusions

8.1 What was already known about fibre and CVD risk?

- Numerous observational studies have suggested that greater fibre intake is associated with lower risk of CVD, but many also report no evidence of any associations (Threapleton et al., 2013e).
- Previous literature reviews have been unsystematic, have only explored total dietary fibre rather than major food sources of fibre, or have not quantified the dose-response association between fibre and risk of CVD (Threapleton et al., 2013e)
- Meta-analyses of RCTs have identified links between intake of soluble fibre or barley and lower circulating lipid levels (Brown et al., 1999, Talati et al., 2009) and between greater fibre intake and lower blood pressure (Whelton et al., 2005, Streppel et al., 2005) (refer to Chapter 1).
- In the UK, just three previous studies have reported fibre intake and CHD risk (Appleby et al., 1999, Todd et al., 1999, Ward et al., 2012) with only one of these reporting on types of fibre (Ward et al., 2012). One existing UK study has previously reported fibre intake and total CVD risk (Akbaraly et al., 2011). These studies report inconsistent associations between CHD or CVD risk and fibre.

8.2 What this work has added

- The comprehensive systematic review and meta-analysis for CHD and CVD outcomes suggests that greater intake of total dietary fibre; insoluble type fibre; and fibre from cereal, fruit, or vegetable sources are associated with a lower risk of cardiovascular disease and coronary heart disease in healthy populations (Threapleton et al., 2013e).
- The systematic review and meta-analysis for stroke outcomes indicates that greater intake of total fibre is associated with lower stroke risk in healthy populations. Soluble fibre was not associated with lower stroke risk but only three studies were available to contribute data on this (Threapleton et al., 2013d).
- A limited study base and therefore gap in current knowledge about the association between total fibre and ischaemic or haemorrhagic stroke or key types of fibre and stroke risk has been identified.

- Work in Chapter 4 indicates that many important CVD events are not identified through one or two case sources and as many resources as are available should be recruited in order to fully ascertain case numbers in similar UK studies.
- Results from the UKWCS relating fibre intake and CHD or CVD risk add to the currently limited data on this topic in the UK.
- The results presented on fibre intake and stroke risk are the first of this type to be reported in a UK population.
- Examining fibre intake and CVD risk using two dietary assessment methods has only been reported by one other study (Ward et al., 2012). These results for the UKWCS therefore add to the limited existing evidence.

8.3 Thesis result summary

The analyses presented in this thesis have used data from the UKWCS and have combined previously assessed dietary information with CVD event records, for the first time in this cohort.

The objectives stated in Chapter 1 have been met and the association of dietary fibre intake in relation to CVD risk has been thoroughly examined using data from the UKWCS:

1) A systematic literature review was conducted to identify literature from observational studies reporting dietary fibre intake and risk of CHD, stroke and CVD (Chapter 2).
    - This work, which combines data from over 24 studies published over two decades, identified that each 7g/day greater total dietary fibre intake was associated with lower risk for CHD HR 0.90 (95% CI: 0.87to 0.94), stroke HR 0.93 (95% CI: 0.88 to 0.98) and also for total CVD HR 0.91 (95% CI: 0.88 to 0.94).
2) Characteristics of high and low fibre consumers, as estimated using FFQs and four day food diaries are presented in Chapter 3.
3) Section 251 approval and ethical approval was obtained specifically to permit linkage of CVD event data with the existing data held for the UKWCS (described in Chapter 4). The event records received from HES and MINAP were processed in order to reflect comparable timeframes, locations and outcomes and the quality, or completeness, of the CVD event data was explored using a capture-recapture approach and log-linear modelling.

- The work presented in Chapter 4 indicates the completeness of CVD event data for assessing acute events but other researchers have shown that primary care data may also be an important source of case information.

4) Associations between fatal CVD risk and fibre intake estimated using FFQs are assessed using survival analysis and presented in Chapter 5.
   - In total, 258 fatal CVD cases were reported since baseline and only cereal fibre intake was associated with significantly lower risk (at the 1% level) for fatal stroke in models including overweight and obese women. Subgroup analyses indicated greater fibre density is associated with lower risk of fatal stroke in women who were free of angina and hypertension at baseline.
5) The work in Chapter 5 is extended in Chapter 6 and incorporates non-fatal event data obtained from HES and MINAP. Dietary fibre intake, as assessed using FFQs, is explored in relation to risk of total CVD (fatal plus non-fatal events), non-fatal CVD and subtypes of CHD (MI, ACS, chronic events) and stroke (haemorrhage, ischaemia or unclassified stroke events).
   - A total of 821 CHD and 388 stroke fatal and non-fatal cases were reported for cohort participants. Total dietary fibre and insoluble fibre were associated with lower MI risk in the full sample of women. Lower total and non-fatal CVD risk was observed with many of the fibre exposures in women without history of hypertension. Protective associations were also seen with fibre intakes from different sources and stroke events, particularly ischaemic-type stroke.
6) Finally, detailed dietary information collected from four-day food diaries was processed for CVD cases and a random selection of non-case diaries. Risk of fatal stroke, fatal CHD and fatal CVD in addition to total MI and total ACS was explored using fibre assessed from these food diaries by applying a case-cohort approach and findings are presented in Chapter 7.
   - A total of 88 ACS cases and 67 fatal CVD cases, occurring after the follow-up dietary assessment, had completed and returned food diaries. Greater cereal fibre intake was associated with lower risk for ACS and greater intake of total fibre was also associated with lower fatal stroke risk. However, some positive associations were observed with some of the exposures for fatal IHD and fatal stroke risk, where case numbers were especially small.

8.4 Summary discussion

Drawing together the available evidence, fibre intake appears to be associated with CVD risk but the associations tended to stronger with stroke and total CVD rather than coronary outcomes alone. Also, cereal type fibre stood out among the different sources as being more consistently associated with lower CVD event risk in this cohort. The differential associations between fibre intake and the three main CVD outcomes supports the notion mentioned in Chapter 1 that the pathology of these conditions may differ and fibre may exert different influences on disease development processes.

As mentioned in the discussion sections of Chapters 5, 6 and 7 there are many explanations for the inverse associations observed between risk and fibre intake. The plausible mechanisms for the action of fibre (discussed in Chapter 1) may act to influence vascular health over the life course and ultimately lower risk. Alternatively, fibre intake may be a surrogate marker for other healthy lifestyle or dietary practices that were not measured or accounted for in models. For example, there are other beneficial components of cereal grain foods which may confer protection or it may be the combination of these compounds which serve to reduce risk of CVD (Slavin, 2004) (Section 5.5.1). In addition to the other beneficial components of cereals, the clearer associations with this source of fibre may reflect lower measurement error in assessing cereals, as compared with fruit and vegetables (discussed in section 5.5.1 and 6.5.1). Other explanations for the associations include that dietary fibre is a marker of available carbohydrate within foods and this has been associated with insulin and lipid profiles in clinical studies (Liu et al., 2000). Lower glycaemic load diets are also associated with improved levels of C-reactive protein, a risk factor for heart disease, and for CHD risk (Liu et al., 2000, Liu et al., 2002b).

Cereal fibre stood out as being more consistently associated with lower CVD risk compared to fruit or vegetables sources of fibre (Chapter 6). This observation is supported by findings from the Nurses' Health Study (discussed in Chapter 2), which were particularly similar to the UKWCS, with lower stroke risk associated with greater cereal fibre intake but not fruit or vegetable fibre (Oh et al., 2005). When breakfast cereal fibre was considered in the UKWCS, some of the same significant associations were observed as for cereal fibre for ischemic stroke risk (Table 6.10). Similarly, risk estimates and 95% CIs for both cereal fibre and breakfast cereal fibre were alike and were not associated with reduced risk of total CHD, non-fatal CHD or CHD sub-types (Tables 6.8, 6.9 and 6.12). In other analyses for total stroke, non-fatal stroke and

ischaemic plus unspecified stroke, the risk estimates for breakfast cereal-derived fibre were weaker than for total cereal fibre, with CIs that spanned the line of no effect despite risk estimates being in the same direction. For example with total stroke, the risk per 3g/day cereal fibre was 0.92 (95% CI 0.85 to 1.00) and per 1g/day greater intake of fibre from breakfast cereals 0.96 (95% CI 0.90 to 1.02) (Table 6.8). The few weaker associations may be indicative of the protective effects of different types of fibre, in that breakfast cereals generally contain added bran, which is principally insoluble in type. The composition of fibre types in many processed breakfast cereals, which contain greater insoluble fibre, may be less beneficial than whole cereal intake. However, these weaker associations with breakfast cereal fibre in some stroke analyses may reflect the greater measurement error that comes from assessing smaller nutrient intakes. In addition, as noted in Chapter 6, both soluble and insoluble fibres were associated with lower stroke risk (Table 6.8).

Given the current limitations in assessing self-reported diet in nutritional epidemiology studies it is rare when we can regard findings from single studies as definitive and associations are considered persuasive when data from many studies, including varied populations is combined (Kipnis and Freedman, 2008). The work in this thesis has allowed important examination of different sources and types of fibre and the associated risk with sub-types of CVD events in this sample of women. However, the combined study risk estimates reported in Chapter 2 may be more reliable when considering total fibre intake and CVD risk across a wide range of populations. The lack of many statistically significant associations seen for total fibre intake and CHD risk in the UKWCS may be explained by fewer cases or differences in study populations. Participants in this study may be generally healthier and consume sufficiently high intakes of fibre, and may experience lower additive benefit of greater intakes. Furthermore, the majority of studies included in Chapter 2 meta-analyses included men and it could be this principle difference that explains the somewhat weaker associations for some CHD outcomes in the UKWCS. However, when participant gender was explored through meta-regression of the separate study results, similar observations were seen for CHD outcomes for male and female study results (Table 2.4). Similarly for stroke outcomes, the estimated association with fibre intake was similar between the studies including men only and the one study including women only (Table 2.4).

Although many of the estimates observed for various exposures in this cohort were not statistically significant, they were all in the same direction as previous studies, and the confidence intervals often included the pooled estimates from the meta-analyses reported in

Chapter 2 (Threapleton et al., 2013e). So, in the context of the other work, results from the UKWCS are not entirely inconsistent. Furthermore, when study results for the UKWCS were included in the updated meta-analysis, the overall estimates and confidence intervals changed little from those reported in Chapter 2 (Threapleton et al., 2013e).

The several potential explanations for null or weaker observations in this study by comparison to other study populations identified during the systematic review are discussed in Section 6.5.1. Issues include insufficient dietary variation, insufficient accuracy to measure dietary differences, low case numbers, assessment not encompassing the true latent period of disease or an opposing variable in this study population that may create negative confounding (Willett, 2013d, Willett and Lenart, 2013). Imprecision in event reporting (discussed in Chapter 4) may also contribute towards apparent null associations in this study by comparison to others. The quality of MINAP data is well established (Herrett et al., 2010) but the lack of clinical case confirmation of events recorded in HES may contribute to error in case classification.

The quality of outcome reporting and type of event recorded by each data source is an important factor to consider. Mortality outcomes were examined in Chapter 5, separately from non-fatal or total incident events in Chapters 6 and 7. Whilst both approaches attempt to measure the burden of CVD, I have presented them separately for several reasons. Firstly, other cohort studies identified through the systematic review had reported different associations between fibre intake and risk for fatal or total CVD events and the findings were not consistent between studies (Bazzano et al., 2003, Pietinen et al., 1996). Secondly, analyses using HES and MINAP data (Chapter 6 and 7) had been restricted to English participants and a greater number of fatal events could therefore be examined by looking separately at mortality data. Additionally, the mortality event data represent a more complete outcome for this cohort in that the data span from baseline, whereas HES are available from 1998 and MINAP from 2003. As discussed in Chapter 4, the quality of mortality event reporting has been consistent for many years while the other datasets have been improving over recent years and examining mortality events separately allows different features of event reporting to be considered.

Understanding the quality of event data is a foundation for interpreting results and there are many potential reasons why CHD and stroke cases for cohort participants may be unidentified or not captured using the mortality, HES and MINAP datasets obtained. Only the mortality events span to study baseline so any non-fatal events would be unidentified during the early

phase of study follow-up. Also, any events in participants living in England at baseline who have moved to Scotland or overseas would not be identified. Recent work from other studies comparing dataset completeness indicates many CHD events are recorded in primary care data but not in hospital records (Herrett et al., 2013). Many less serious CHD events for cohort participants are likely to exist in these primary care records and these cases have not been identified during this study.

The potential for missing cases has been explored through the work in Chapter 4, the cross-comparisons using log-linear analysis provided estimates of the missing number of cases and capture quality for each dataset. The work was limited in that the different datasets had not captured the same type of case events and so the estimates of potential missing cases are inflated. For example, one dataset identifies only fatal cases. The model assumes that all of the non-fatal cases identified by the other sources are missed by this source, so over-estimating the estimate of total missing cases. However, excluding this source could lead to an under-estimate of the number missing, because some cases are only identified through death certificates, and missed by the other sources. Additionally, having case data contributed by three sources allowed for list dependency to be modelled, which is not possible with just two sources of case information, and list dependency is an important consideration in capture-recapture work (discussed in detail in Chapter 4).

8.5 Strengths and limitations

The challenges faced in epidemiology, to accurately assess exposure to various factors is possibly greatest in nutritional epidemiology where diet is a daily 'exposure', may be so varied and reporting or recording of diet may be influenced by so many factors. Assessing diet using multiple methods makes best use of imprecise and biased estimates and contributes better to our understanding of the relationship between diet and disease. Addressing the question of how fibre intake may influence CVD risk using two dietary assessment methods is therefore a key strength in this work.

As noted in a recent editorial, in order to gain better understanding of what mechanisms and which components of fibre might underlie protective associations, studies should investigate specific fibres and food sources (Landberg, 2012). This is a key strength of the work undertaken here, with a view to identifying the most beneficial sources of fibre. Additionally, the use of a validated FFQ in such a large sample of women is a real strength. Utilising CVD

event data from different sources has also made best use of available case information, given the lack of a standard clinical register for CVD events in the UK. This approach has allowed a greater number of cases to be identified, thus allowing exploration of CVD sub-types in view of potentially different pathology. Given the inconsistency in international definitions of fibre and the potential challenges in comparing findings from studies which have used different methods (discussed in Chapter 1 and below), an additional strength of analyses presented here is that risk has been presented in relation to fibre estimated both as NSP and AOAC.

Strengths and limitations of the systematic review and meta-analyses are discussed in depth in earlier sections (2.6.3, 2.6.4). A key limitation in the work reported in Chapter 2, meta-analyses using observational study data, concerns the correction of confounding. It is challenging to judge whether confounding has been adequately dealt with in each individual study although the majority of studies had included adjustment for the most relevant confounders, which were explored in turn through meta-regression (Table 2.4). Hypothesised causal links between fibre and CVD risk cannot be proven using observational studies but owing to the likely long pathogenesis of CVD, it is unlikely that trials of adequate duration and adherence would be feasible (Threapleton et al., 2013e). If the individual participant data for all the included cohorts were available, then this would provide a better means of deriving a pooled estimate. Firstly, because additional confounding could be adjusted for, if they had been collected by the cohort, and secondly because there could be consistency in how confounding was addressed thus reducing potential heterogeneity. The most suitable approach for answering this question appears to be via observational data such as that presented in this thesis in conjunction with clearly defined RCTs that examine specific types of fibre in relation with known risk factors for CHD and stroke.

An important limitation with this work is that the applicability of findings to other UK populations is unknown. However, it seems evident that greater fibre intake is associated with lower risk, especially with stroke and this information is relevant to other British women. Many study populations are limited in terms of dietary variation in intakes and results therefore depend more on accurate assessment of dietary exposures (Kipnis and Freedman, 2008). The unique strength of the UKWCS sample recruitment was the large proportion of vegetarian participants and thus a sample with a good range in terms of fibre intake. As noted in Chapter 3 discussion, the cohort participants on average consume a greater level of fibre than the general UK population. It is this characteristic that adds to limited evidence from cohorts based in the UK and lends itself to exploring healthy diets, while still having large numbers of

participants that have diets and fibre intakes similar to the national average. The unique characteristics of the cohort such as high proportion of vegetarians and high mean fibre intakes, while lending strength to examining a range of fibre intakes with CVD risk, also make this study distinct from others reported in Chapter 2. As noted in the discussion of Chapter 7 this population heterogeneity may somewhat explain weaker risk estimates observed for the UKWCS although most risk estimates here do lie on the protective side and CIs encompass a wide range of realistic risk reductions for this cohort.

To account for the oversampling of vegetarian participants, inverse probability weighting was used in models to address potential lack of generalisability of results. This method gives less weight to data from vegetarian participants while still benefitting from the full range of data provided by this large sub-sample. So, whilst the UKWCS is strengthened by including a large proportion of participants with high fibre intakes, estimates are still comparable to other population cohorts. In addition, the sensitivity analyses undertaken when this weighting factor was removed from the models indicated that it had relatively little impact on relative risk estimates, as noted in methods sections of Chapters 5 and 6.

8.6 Recommendations

8.6.1 In the cohort

Further work could focus on exploring food groups rather than specific nutrients in terms of CVD risk. Examining total cereal, fruit or vegetable intake might throw light on potential mechanisms and indicate whether any associations with these foods relate to the fibre content, other micronutrients or other associated behaviour.

As family history is a strong risk factor for CVD, this could be explored in those with and without family history of the disease to indicate whether fibre intake has a differential association on risk.

Participants were excluded from analyses where death or any CVD event was recorded within 1 year of dietary assessment to account for the influence of CVD and associated ill health that may have influenced dietary intake and thus bias associations through reverse causality. Whilst valuable cases may be missed from analyses, it is quite possible that many of these would have been latent, pre-existing cases, whose inclusion could lead to substantial bias in estimates. In addition, after other exclusions were applied to the sample, 59 participants were excluded

because of death or CVD event within 1 year of baseline assessment. Of the 1,193 total CVD cases, just 31 were removed in this process. Given the long pathogenesis of CVD discussed in Chapter 1, sensitivity analysis could be conducted to exclude CVD cases occurring within several years of baseline assessment, rather than one year. This longer exclusion time would account for any latent disease which may have influenced dietary choices during the assessment period.

Cardiovascular events examined in this work were principally ischaemic, with smaller proportions of the total cases classified as cardiac arrest or haemorrhagic stroke (Tables 4.4 and 4.5). Other CVD types could also be examined to explore whether the associations, or lack of associations in some comparisons, extend to other forms of CVD such as hypertensive CVD (see Figure 1.1). Additionally, the approach could be taken to examine all CVD by using all ICD10 codes in group 'I', such as was done in the recent EPIC study (Chuang et al., 2012).

It may also be beneficial to further explore stroke risk by accessing hospital records to determine if the type of stroke had been recorded. Identifying whether the unclassified strokes were ischaemic or haemorrhagic in nature may clarify why associations were seen in the unclassified strokes and if they are, as suspected, mainly ischaemic.

8.6.2 Other studies

As seen in Chapter 2, there are many existing studies that have assessed the association between total fibre intake and CVD, with fewer focusing on sources of fibre and CVD subtypes. The majority of these are conducted in the US and Europe with some work been done in Australia and Japan. Data from other population groups would therefore be a welcome addition to this body of work.

The question of whether total dietary fibre intake is associated with CVD risk seems largely resolved (Chapter 2) with greater intake being associated with lower risk in a dose-response fashion. As already noted, it is not possible to definitively distinguish whether these associations result from fibre per se or other associated nutrients or dietary and lifestyle habits. However, as discussed in Chapter 1, different types of fibre are demonstrated to have different physiological effects but fewer studies have examined this issue. Exploring sources of fibre in future studies, particularly for stroke outcomes, where this has been examined less, may elucidate the specific mechanisms by which fibre affects long-term vascular health. A further issue surrounds the inconsistency in definition and measurement of the fibre content

of foods, which this serves to limit the emphasis of fibre in labelling schemes (Buttriss and Stokes, 2008). International consensus is needed on fibre definitions and the categorisation of the various components of fibre and this especially important within research so findings may be comparable.

A recent systematic review identified a risk reduction for CVD of around 20% for high fibre consumers and separately for high whole-grain consumers, compared to the lowest consumers of each (Ye et al., 2012). Given that fibre and whole-grain consumption are likely to correlate highly, it remains a challenge to identify whether fibre intake is a surrogate marker for whole-grain intake and the potential beneficial compounds within grains or whether it is the fibre component of whole-grains which confers the protective associations seen with greater intake. Well conducted clinical trials may be useful to illuminate the many potential mechanisms through which both whole-grains and fibre from different foods may act.

Results from clinical trials, mentioned in Chapter 1 give the overall impression that the multiple risk factors for CVD may be influenced by fibre intake and it is the cross-linking between all these risk factors that makes pinpointing one key mechanism of action challenging. For example, one challenge with dietary intervention studies to explore potential mechanisms of action is that body weight is often not maintained and since this is related to other risk factors such as hypertension, dyslipidaemia, poor glucose control and inflammation it is not possible to distinguish the mechanism of action. Intervention studies should therefore address issues such as the interrelated nature of risk factors by attempting to maintain body weight changes equitably between intervention groups.

Measurement error is a major threat in nutritional epidemiology, a problem compounded by the fact that it is impossible to know whether the misclassification is random or differential and thus whether the precision or the validity of the study is affected (Michels, 2001). This issue is therefore crucial to address and with the advent of electronically collected data, error, cost and participant burden will reduce and this technology should be widely taken up (Baranowski, 2013). As noted in Chapter 3, it is well recognised that FFQs tend to universally inflate intake values as was seen for both fibre and energy intakes in the UKWCS. This imprecision means FFQs are not be ideal for assessing individual diet-disease risk associations be but FFQs are practical when ranking individuals and exploring trends in associations across specific populations (Cade et al., 2004b).

The unique feature of having such a large proportion of non meat-eaters in the UKWCS is ideal for exploring, in an observational setting, newly identified potential mechanisms behind CVD risk identified through metabolomic, animal and human studies. Recent work proposes CVD risk is linked to L-carnitine consumption, principally from red meat in the diet. The hypothesis set out by Koeth and colleagues suggests that gut microbiota are responsible for the digestion of L-carnitine that ultimately produces trimethylamine-N-oxide, a proatherogenic compound (Koeth et al., 2013). The weaker associations observed for the UKWCS compared to others in Chapter 2 could suggest a differential effect or interaction between the detrimental effects of L-carnitine digestion by the gut microflora of meat-eaters and the potential protective effects that high fibre intake may elicit on the proliferation of beneficial gut flora.

With this hypothesis in mind, that meat-eaters would experience greater beneficial effects of fibre through mitigating L-carnitine damage, key analyses from Chapter 6 were repeated separately for meat-eaters and vegetarians (Section 6.4.3-Subgroups). Whilst there was a suggestion that the association between NSP intake and stroke was slightly more protective in meat-eaters than in vegetarians, potentially supporting this hypothesis, there was no evidence to support this when formally tested by including an interaction in the model. The test indicated that the subgroup differences could easily be just due to chance. Neither was there any evidence to support this hypothesis for CHD, and with no reason to suggest that CHD and stroke would be influenced differently, these findings collectively do not support the hypothesis that meat-eaters have differential risk or that this is because of the presence of L-carnitine metabolites generated through their diet (Koeth et al., 2013). However, findings are inevitably inconclusive because subgroup numbers were small (e.g. 77 stroke cases in the vegetarian sub-group) and therefore power for the tests of interaction was low. Further observational study work to explore CVD risk by vegetarian status would therefore be useful to explore this novel hypothesis.

8.6.3 Public health messages

The various components or types of fibre are thought to have distinct effects and in order to maximise the health benefits from fibre in the diet it is thought important to come from a range of sources (Lunn and Buttriss, 2007). Having said that, cereal sources of fibre seem particularly beneficial for lowering risk and are widely consumed in the UK so recommendations based on cereal fibre intake may be well accepted.

Because of the different methods used to estimate dietary fibre content of foods, recommendations differ between countries (discussed in Chapter 1). The findings here indicate that increasing fibre intake in especially low consumers may have the most benefit as fewer associations were observed in this cohort compared to other studies. Messages should continue to promote higher fibre intake in the population but an issue may surround the fact that NSP is used for the recommendation, whilst back-of-pack values in the UK present AOAC values, potentially leading consumers to overestimate their fibre intake. In the very least, high intakes of fibre are unlikely to produce substantial deleterious effects and any potential negative effects of high fibre intake are much outweighed by the potential beneficial effects on risk for ill health (Lunn and Buttriss, 2007). Additionally, eating fibre-rich whole foods have the added benefit of being rich sources of micronutrients and other beneficial compounds and many are lower in energy (Slavin, 2003).

Messages to increase fibre intake have been in place for some time but fibre intake in the UK is well below the recommendation (18g/day) (COMA, 1991) and has only marginally increased over the past decade. Using data from adults in the NDNS, in 2000/1 women (19-64 years) were consuming 12.6 (SD 5.4) g/day (7.7 g/1000kcal/day) and in 2008/9 the mean intake was 12.8 (SD 4.5) g/day (7.8 g/1000kcal/day). Reported daily intakes in men were higher than the women but in terms of relative consumption, fibre density of the diet was lower. In 2000/1 men (19-64 years) were consuming 15.5 (SD 6.6) g/day (6.7 g/1000kcal/day) and in 2008/9 the mean intake was 14.9 (SD 5.6) g/day (6.8 g/1000kcal/day) (NDNS, 2011).

A key strategy for increasing intake must therefore address some of the important barriers. A recent qualitative study explored perceived barriers to whole-grain intake in the UK and found that the most prominent barriers were negative perception of sensory qualities, lack of knowledge about what whole-grains are, where to locate them and how to incorporate them into meals. Higher whole-grain consumers identified similar factors as facilitators for their consumption and preferred the taste, had good level of knowledge about whole-grains and good understanding of the health benefits (McMackin et al., 2013). Further education, specifically for low fibre consumers, of the health benefits, types and preparation methods of whole-grains could be very useful to increase intake of this rich source of fibre.

Lecturing individuals has little effect on behaviour change but bigger gains can potentially come from structural policy and legislative changes at regional, national and international levels (Capewell et al., 2008). The last decade has seen significant efforts by the food industry

to assist in consumer choice of whole-grains and increase the availability of whole-grain foods (McMackin et al., 2013). Components of foods that resist digestion may be one way to increase fibre intake 'by stealth' since the key barrier to whole-grain uptake is negative perception of sensory qualities (Buttriss and Stokes, 2008) and as noted in Chapter 1, resistant starch is fermented in the colon in the same way as soluble and some insoluble fibres and thus confers protective benefit. Using resistant starch in foods does not typically change the taste or texture of foods and these may therefore be accepted by consumers (Buttriss and Stokes, 2008). Manufacturers could also modify preparations for existing products by adding bran and this may be more acceptable for consumers as is less disruptive than whole dietary changes and also adds little by way of energy intake (Nicklas et al., 2011).

Fibre content of foods must be clearly displayed on pack labels if we hope or expect the general population to have an interest in or some knowledge of their own consumption. In 1999 the Food Standards Agency issued food back-of-pack labelling guidelines and it was recommended that fibre intake should be listed on the nutritional composition of foods, although this was only on a voluntary basis. The recommendation was initially to include NSP but in 2000, this was amended to AOAC fibre, in an effort to harmonise labels across Europe (FSA, 1999). Current recommendations still include listing the fibre composition of foods only on a voluntary basis (DEFRA, 2012). Since it is likely that most consumers are unaware of more recent research linking the properties of resistant carbohydrates to health benefits, a targeted education campaign with unambiguous public health messages is long overdue (Buttriss and Stokes, 2008).

8.7 Conclusions

- Total dietary fibre intake, in addition to insoluble fibre and from cereals and vegetables, are inversely associated with CHD, stroke and CVD risk when examining findings from many developed countries.
- Few studies report different stroke types in association with total fibre or fibre from key food sources.
- In a sample of health-conscious, middle-aged, English women, higher fibre intakes were associated with lower stroke risk and this association was particularly apparent with ischaemic stroke and with cereal fibre intake.
- Fewer associations were apparent between fibre intake and CHD risk, as compared to stroke and this finding, which is different to meta-analyses although in the same

direction, may be attributed to the relatively high fibre intake in the UKWCS by comparison to other studies and the general UK population.

- In women without hypertension greater fibre intake was associated with lower overall CVD risk indicating potential protective actions of fibre for those without existing risk factors for CVD.

References

References

AHA 2011. American Heart Association. Cardiovascular Disease Statistics. [Online] Available at: http://www.americanheart.org/presenter.jhtml?identifier=4478 [Accessed 09/02/2011].

AKBARALY, T. N., FERRIE, J. E., BERR, C., BRUNNER, E. J., HEAD, J., MARMOT, M. G., SINGH-MANOUX, A., RITCHIE, K., SHIPLEY, M. J. & KIVIMAKI, M. 2011. Alternative healthy eating index and mortality over 18 y of follow-up: Results from the Whitehall II cohort. *American Journal of Clinical Nutrition,* 94**,** 247-253.

ALLENDER, S., SCARBOROUGH, P., PETO, V., RAYNER, M., LEAL, J., LUENGO-FERNANDEZ, R. & GRAY, A. 2008. European cardiovascular disease statistics. [Online] Available at: http://www.heartstats.org/datapage.asp?id=7683 [Accessed 11/02/2011].

AN DER HEIDEN, M. 2009. RECAP: Stata module to perform capture-recapture analysis for three sources with Goodness-of-Fit based confidence intervals. Boston College Department of Economics. [Online]Accessed at http://ideas.repec.org/c/boc/bocode/s456859.html [Accessed March 2011].

ANDERSEN, K. K., OLSEN, T. S., DEHLENDORFF, C. & KAMMERSGAARD, L. P. 2009. Hemorrhagic and ischemic strokes compared: stroke severity, mortality, and risk factors. *Stroke,* 40**,** 2068-72.

APPLEBY, P. N., THOROGOOD, M., MANN, J. I. & KEY, T. J. 1999. The Oxford Vegetarian Study: an overview. *American Journal of Clinical Nutrition,* 70**,** Suppl-531S.

ASCHERIO, A., HENNEKENS, C., WILLETT, W. C., SACKS, F., ROSNER, B., MANSON, J., WITTEMAN, J. & STAMPFER, M. J. 1996. Prospective study of nutritional factors, blood pressure, and hypertension among US women. *Hypertension,* 27**,** 1065-72.

ASCHERIO, A., RIMM, E. B., GIOVANNUCCI, E. L., COLDITZ, G. A., ROSNER, B., WILLETT, W. C., SACKS, F. & STAMPFER, M. J. 1992. A prospective study of nutritional factors and hypertension among US men. *Circulation,* 86**,** 1475-1484.

ASCHERIO, A., RIMM, E. B., HERNAN, M. A., GIOVANNUCCI, E. L., KAWACHI, I., STAMPFER, M. J. & WILLETT, W. C. 1998. Intake of potassium, magnesium, calcium, and fiber and risk of stroke among US men. *Circulation,* 98**,** 1198-1204.

ATSMA, F., BARTELINK, M. L., GROBBEE, D. E. & VAN DER SCHOUW, Y. T. 2006. Postmenopausal status and early menopause as independent risk factors for cardiovascular disease: a meta-analysis. *Menopause,* 13**,** 265-79.

BACH, A., SERRA-MAJEM, L., CARRASCO, J. L., ROMAN, B., NGO, J., BERTOMEU, I. & OBRADOR, B. 2006. The use of indexes evaluating the adherence to the Mediterranean diet in epidemiological studies: a review. *Public Health Nutrition,* 9**,** 132-46.

BAER, H. J., GLYNN, R. J., HU, F. B., HANKINSON, S. E., WILLETT, W. C., COLDITZ, G. A., STAMPFER, M. & ROSNER, B. 2011. Risk factors for mortality in the Nurses' Health Study: a competing risks analysis. *American Journal of Epidemiology,* 173**,** 319-29.

BARANOWSKI, T. 2013. 24-Hour Recall and Deit Record Methods. *In:* HOFMAN A, M. M., SAMET J, SAVITZ DZ (ed.) *Nutritional Epidemiology.* Third ed. Oxford: Oxford University Press.

BARLOW, W. E. 1994. Robust variance estimation for the case-cohort design. *Biometrics,* 50**,** 1064-72.

BARLOW, W. E., ICHIKAWA, L., ROSNER, D. & IZUMI, S. 1999. Analysis of case-cohort designs. *Journal of Clinical Epidemiology,* 52**,** 1165-72.

BARRETT-CONNOR, E. 2013. Menopause, atherosclerosis, and coronary artery disease. *Curr Opin Pharmacol*.

BATES, B., LENNOX, A., PRENTICE, A., BATES, C. & SWAN, G. 2012. National Diet and Nutrition Survey: Headline results from Years 1, 2 and 3 (combined) of the Rolling Programme (2008/2009 – 2010/2011). [Online]. Available at: http://webarchive.nationalarchives.gov.uk/20130402145952/http://media.dh.gov.uk/network/261/files/2012/07/ndns-y3-report_all-text-docs-combined.pdf. [Accessed November 2013]

BATES, B., LENNOX, A. & SWAN, G. 2009. National Diet and Nutrition Survey: Headline results from Year 1 of the Rolling Programme. 2008/2009. [Online]. Available at: http://www.food.gov.uk/multimedia/pdfs/publication/ndnsreport0809year1results.pdf [Accessed 2011].

BAZZANO, L. A., HE, J., OGDEN, L. G., LORIA, C. M., WHELTON, P. K. & NATIONAL HEALTH NUTRITION EXAMINATION SURVEY 2003. Dietary fiber intake and reduced risk of coronary heart disease in US men and women: the National Health and Nutrition Examination Survey I Epidemiologic Follow-up Study. *Archives of Internal Medicine,* 163**,** 1897-1904.

BERG, A. H. & SCHERER, P. E. 2005. Adipose tissue, inflammation, and cardiovascular disease. *Circ Res,* 96**,** 939-49.

BERNSTEIN, A. M., ROSNER, B. A. & WILLETT, W. C. 2011. Cereal fiber and coronary heart disease: a comparison of modeling approaches for repeated dietary measurements, intermediate outcomes, and long follow-up. *European Journal of Epidemiology,* 26**,** 877-86.

BHUPATHIRAJU, S. N. & TUCKER, K. L. 2011. Coronary heart disease prevention: nutrients, foods, and dietary patterns. *Clin Chim Acta,* 412**,** 1493-514.

BINGHAM, S. 1987. The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. *Nutrition Abstracts and Reviews (Series A),* 57**,** 706-742.

BINGHAM, S. A., GILL, C., WELCH, A., CASSIDY, A., RUNSWICK, S. A., OAKES, S., LUBIN, R., THURNHAM, D. I., KEY, T. J., ROE, L., KHAW, K. T. & DAY, N. E. 1997. Validation of dietary assessment methods in the UK arm of EPIC using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin C and carotenoids as biomarkers. *International Journal of Epidemiology,* 26 Suppl 1**,** S137-51.

BINGHAM, S. A., GILL, C., WELCH, A., DAY, K., CASSIDY, A., KHAW, K. T., SNEYD, M. J., KEY, T. J., ROE, L. & DAY, N. E. 1994. Comparison of dietary assessment methods in nutritional epidemiology: weighed records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. *British Journal of Nutrition,* 72**,** 619-43.

BINGHAM, S. A., LUBEN, R., WELCH, A., WAREHAM, N., KHAW, K. T. & DAY, N. 2003. Are imprecise methods obscuring a relation between fat and breast cancer? *Lancet,* 362**,** 212-4.

BOWERS, D., HOUSE, A. & OWENS, D. 2006a. Chapter 21 Testing Hypotheses - the p-value. *Understanding Clinical Papers.* Second ed. Chichester: John Wiley & Sons Ltd.

BOWERS, D., HOUSE, A. & OWENS, D. 2006b. Chapter 26 Systematic review and meta-analysis. *Understanding Clinical Papers.* Second ed. Chichester: John Wiley & Sons Ltd.

BRITISH NUTRITION FOUNDATION 2005. *Cardiovascular Disease: Diet, Nutrition and Emerging Risk Factors. Report of the British Nutrition Foundation Task Force,* Oxford, Blackwell Publishing Ltd.

BRITISH NUTRITION FOUNDATION. 2009. *Heart disease- Dietary factors and CVD risk* [Online]. Available: http://www.nutrition.org.uk/nutritionscience/disease/heart-disease?start=4 [Accessed 2011].

BRITTON, A., MILNE, B., BUTLER, T., SANCHEZ-GALVEZ, A., SHIPLEY, M., RUDD, A., WOLFE, C. D., BHALLA, A. & BRUNNER, E. J. 2012. Validating self-reported strokes in a longitudinal UK cohort study (Whitehall II): Extracting information from hospital medical records versus the Hospital Episode Statistics database. *BMC Med Res Methodol,* 12**,** 83.

BROWN, L., ROSNER, B., WILLETT, W. W. & SACKS, F. M. 1999. Cholesterol-lowering effects of dietary fiber: a meta-analysis. *American Journal of Clinical Nutrition,* 69**,** 30-42.

BRUNNER, E., STALLONE, D., JUNEJA, M., BINGHAM, S. & MARMOT, M. 2001. Dietary assessment in Whitehall II: comparison of 7 d diet diary and food-frequency questionnaire and validity against biomarkers. *British Journal of Nutrition,* 86**,** 405-14.

BURNS, E. M., RIGBY, E., MAMIDANNA, R., BOTTLE, A., AYLIN, P., ZIPRIN, P. & FAIZ, O. D. 2012. Systematic review of discharge coding accuracy. *Journal of Public Health,* 34**,** 138-48.

BUTTRISS, J. L. & STOKES, C. S. 2008. Dietary fibre and health: an overview. *British Nutrition Foundation: Nutrition Bulletin* 33**,** 186-200.

BUYKEN, A. E., FLOOD, V., EMPSON, M., ROCHTCHINA, E., BARCLAY, A. W., BRAND-MILLER, J. & MITCHELL, P. 2010. Carbohydrate nutrition and inflammatory disease mortality in older adults. *American Journal of Clinical Nutrition,* 92**,** 634-43.

CADE, J., THOMPSON, R., BURLEY, V. & WARM, D. 2002. Development, validation and utilisation of food-frequency questionnaires - a review. *Public Health Nutrition,* 5**,** 567-87.

CADE, J. E., BURLEY, V. J. & GREENWOOD, D. C. 2004a. The UK Women's Cohort Study: comparison of vegetarians, fish-eaters and meat-eaters. *Public Health Nutrition,* 7**,** 871-8.

CADE, J. E., BURLEY, V. J. & GREENWOOD, D. C. 2007. Dietary fibre and risk of breast cancer in the UK Women's Cohort Study. *International Journal of Epidemiology,* 36**,** 431-8.

CADE, J. E., BURLEY, V. J., WARM, D. L., THOMPSON, R. L. & MARGETTS, B. M. 2004b. Food-frequency questionnaires: a review of their design, validation and utilisation. *Nutr Res Rev,* 17**,** 5-22.

CALVERT, C., CADE, J., BARRETT, J. H. & WOODHOUSE, A. 1997. Using cross-check questions to address the problem of mis-reporting of specific food groups on Food Frequency Questionnaires. United Kingdom Women's Cohort Study Steering Group. *European Journal of Clinical Nutrition,* 51**,** 708-12.

CAPEWELL, S., ALLENDER, S., CRITCHLEY, J., LLOYD-WILLIAMS, F., O'FLAHERTY, M., RAYNER, M. & SCARBROUGH, P. 2008. Cardio and Vascular Coalition. Modelling the UK burden of cardiovascular disease to 2020. [Online] Available at www.cardiovascularcoalition.org.uk [Accessed November 2013]. British Heart foundation.

CHAO, A., YANG, H. C. & YIP, P. S. F. 2003. The use of capture-recapture methodology in epidemiological surveillance. *In:* LU, Y. & FANG, J. Q. (eds.) *Advanced Medical Statistics.* Singapore: World Scientific Publishing Co. Pte. Ltd.

CHUANG, S.-C., NORAT, T., MURPHY, N., OLSEN, A., TJØNNELAND, A., OVERVAD, K., BOUTRON-RUAULT, M. C., PERQUIER, F., DARTOIS, L., KAAKS, R., TEUCHER, B., BERGMANN, M. M., BOEING, H., TRICHOPOULOU, A., LAGIOU, P., TRICHOPOULOS, D., GRIONI, S., SACERDOTE, C., PANICO, S., PALLI, D., TUMINO, R., PEETERS, P. H. M., BUENO-DE-MESQUITA, B., ROS, M. M., BRUSTAD, M., ÅSLI, L. A., SKEIE, G., QUIRÓS, J. R., GONZÁLEZ, C. A., SÁNCHEZ, M.-J., NAVARRO, C., AICUA, E. A., DORRONSORO, M., DRAKE, I., SONESTEDT, E., JOHANSSON, I., HALLMANS, G., KEY, T., CROWE, F., KHAW, K.-T., WAREHAM, N., FERRARI, P., SLIMANI, N., ROMIEU, I., GALLO, V., RIBOLI, E. & VINEIS, P. 2012. Fiber intake and total and cause-specific mortality in the European Prospective Investigation into Cancer and Nutrition cohort. *American Journal of Clinical Nutrition,* 96**,** 164-174.

COLDITZ, G. A., MARTIN, P., STAMPFER, M. J., WILLETT, W. C., SAMPSON, L., ROSNER, B., HENNEKENS, C. H. & SPEIZER, F. E. 1986. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. *American Journal of Epidemiology,* 123**,** 894-900.

COMA 1991. *Committee on Medical Aspects of Food Policy, Dietary Reference Values for Food Energy and Nutrients for the United Kingdom,* London, HMSO

COMA 1994. *Committee on Medical Aspects of Food Policy, Nutritional aspects of cardiovascular disease,* London, HMSO.

COULTATE, T. P. 2009. Chapter 3 Polysaccharides. *In:* COULTATE, T. P. (ed.) *Food The Chemistry of its components.* Cambridge: The Royal Society of Chemistry.

COVIELLO, V. 2001. Ordinary case-cohort design and analysis. *Stata Technical Bulletin,* 59**,** 12-18.

COX, D. R. & OAKES, D. 1984. *Analysis of Survival Data,* London, Chapman & Hall/CRC.

CRAMER, D. 2003. Log-linear analysis. *In:* BRYMAN, A. (ed.) *Advanced Quantitative Data Analysis.* Berkshire: Open University Press.

CRAWLEY, H. 1993. *Food Portion Sizes 2nd Edition,* London: HMSO.

CRD 2008. Centre for Reviews and Dissemination. Systematic Reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination, University of York.

CROWE, F. L., KEY, T. J., APPLEBY, P. N., OVERVAD, K., SCHMIDT, E. B., EGEBERG, R., TJONNELAND, A., KAAKS, R., TEUCHER, B., BOEING, H., WEIKERT, C., TRICHOPOULOU, A., OURANOS, V., VALANOU, E., MASALA, G., SIERI, S., PANICO, S., TUMINO, R., MATULLO, G., BUENO-DE-MESQUITA, H. B., BOER, J. M., BEULENS, J. W., VAN DER SCHOUW, Y. T., QUIROS, J. R., BUCKLAND, G., SANCHEZ, M. J., DORRONSORO, M., HUERTA, J. M., MORENO-IRIBAS, C., HEDBLAD, B., JANSSON, J. H., WENNBERG, P., KHAW, K. T., WAREHAM, N., FERRARI, P., ILLNER, A. K., CHUANG, S. C., NORAT, T., DANESH, J. & RIBOLI, E. 2012. Dietary fibre intake and ischaemic heart disease

mortality: the European Prospective Investigation into Cancer and Nutrition-Heart study. *European Journal of Clinical Nutrition,* 66**,** 950-956.

CUMMINGS, J. & MANN, J. 2007. Carbohydrates. *In:* MANN, J. & TRUSWELL, A. S. (eds.) *Essentials of Human Nutrition. Third edition.* Oxford: Oxford university press.

DAHM, C. C., KEOGH, R. H., SPENCER, E. A., GREENWOOD, D. C., KEY, T. J., FENTIMAN, I. S., SHIPLEY, M. J., BRUNNER, E. J., CADE, J. E., BURLEY, V. J., MISHRA, G., STEPHEN, A. M., KUH, D., WHITE, I. R., LUBEN, R., LENTJES, M. A., KHAW, K. T. & RODWELL BINGHAM, S. A. 2010. Dietary fiber and colorectal cancer risk: a nested case-control study using food diaries. *J Natl Cancer Inst,* 102**,** 614-26.

DAUCHET, L., AMOUYEL, P. & DALLONGEVILLE, J. 2005. Fruit and vegetable consumption and risk of stroke: a meta-analysis of cohort studies. *Neurology,* 65**,** 1193-7.

DAVEY, G. K., SPENCER, E. A., APPLEBY, P. N., ALLEN, N. E., KNOX, K. H. & KEY, T. J. 2003. EPIC-Oxford: lifestyle characteristics and nutrient intakes in a cohort of 33 883 meat-eaters and 31 546 non meat-eaters in the UK. *Public Health Nutrition,* 6**,** 259-69.

DE ROOS, B., MAVROMMATIS, Y. & BROUWER, I. A. 2009. Long-chain n-3 polyunsaturated fatty acids: new insights into mechanisms relating to inflammation and coronary heart disease. *Br J Pharmacol,* 158**,** 413-28.

DEEKS, J. J., ALTMAN, D. G. & BRADBURN, M. J. 2008. Statistical methods for examining heterogeniety and combining results from several studies in meta-analysis. *In:* EGGER, M., SMITH, G. D. & ALTMAN, D. G. (eds.) *Systematic Reviews in Health Care: Meta-analysis in Context.* Second ed. London: BMJ Publishing Group.

DEFRA 2012. Department for Environment, Food & Rural Affairs. The Food Information Regulations 2013. Guide to Compliance. [Online] Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/82663/consult-fic-guidance-20121116.pdf [Accessed Jan 2014].

DREHER, M. L. 2001. Dietary fiber overview. *In:* CHO, S. S. & DREHER, M. L. (eds.) *Handbook of Deitary Fiber.* New York: Marcel Dekker, Inc.

DROGAN, D., HOFFMANN, K., SCHULZ, M., BERGMANN, M. M., BOEING, H. & WEIKERT, C. 2007. A food pattern predicting prospective weight change is associated with risk of fatal but not with nonfatal cardiovascular disease. *Journal of Nutrition,* 137**,** 1961-1967.

DU, H., VAN DER, A. D., BOSHUIZEN, H. C., FOROUHI, N. G., WAREHAM, N. J., HALKJAER, J., TJONNELAND, A., OVERVAD, K., JAKOBSEN, M. U., BOEING, H., BUIJSSE, B., MASALA, G., PALLI, D., SORENSEN, T. I., SARIS, W. H. & FESKENS, E. J. 2010. Dietary fiber and subsequent changes in body weight and waist circumference in European men and women. *American Journal of Clinical Nutrition,* 91**,** 329-36.

EFSA 2010. European Food Safety Authority: Scientific Opionion on Dietary Reference Values for Carbohydrates and Dietary Fibre. *European Food Safety Authority Journal,* 8**,** 1462.

EGGER, M., SMITH, G. D. & O'ROURKE, K. 2008a. Rationale, potentials, and promise of systematic reviews. *In:* EGGER, M., SMITH, G. D. & ALTMAN, D. G. (eds.) *Systematic Reviews in Health Care: Meta-analysis in Context.* Second ed. London: BMJ Publishing Group.

EGGER, M., SMITH, G. D. & SCHNEIDER, M. 2008b. Systematic reviews of observational studies. *In:* EGGER, M., SMITH, G. D. & ALTMAN, D. G. (eds.) *Systematic Reviews in Health Care: Meta-analysis in Context.* Second ed. London: BMJ Publishing Group.

ELOVAINIO, M., FERRIE, J. E., SINGH-MANOUX, A., SHIPLEY, M., BATTY, G. D., HEAD, J., HAMER, M., JOKELA, M., VIRTANEN, M., BRUNNER, E., MARMOT, M. G. & KIVIMAKI, M. 2011. Socioeconomic differences in cardiometabolic factors: social causation or health-related selection? Evidence from the Whitehall II Cohort Study, 1991-2004. *Am J Epidemiol,* 174**,** 779-89.

ENGLYST, H. N. & CUMMINGS, J. H. 1988. Improved method for measurement of dietary fiber as non-starch polysaccharides in plant foods. *J Assoc Off Anal Chem,* 71**,** 808-14.

ENGLYST, K. N., LIU, S. & ENGLYST, H. N. 2007. Nutritional characterization and measurement of dietary carbohydrates. *European Journal of Clinical Nutrition,* 61**,** S19-S39.

ESHAK, E. S., ISO, H., DATE, C., KIKUCHI, S., WATANABE, Y., WADA, Y., WAKAI, K., TAMAKOSHI, A. & GROUP, J. S. 2010. Dietary fiber intake is associated with reduced risk of mortality from cardiovascular disease among Japanese men and women. *Journal of Nutrition,* 140**,** 1445-53.

EUROFIR 2010. European Food Information Resource. EuroFIR eSearch Prototype [Online] Available from http://esearch.eurofir.org/ [Accessed 11/05/2011].

FEHILY, A. M., YARNELL, J. W., SWEETNAM, P. M. & ELWOOD, P. C. 1993. Diet and incident ischaemic heart disease: the Caerphilly Study. *British Journal of Nutrition,* 69**,** 303-314.

FLIGHT, I. & CLIFTON, P. 2006. Cereal grains and legumes in the prevention of coronary heart disease and stroke: a review of the literature. *European Journal of Clinical Nutrition,* 60**,** 1145-59.

FREEDMAN, L. S., POTISCHMAN, N., KIPNIS, V., MIDTHUNE, D., SCHATZKIN, A., THOMPSON, F. E., TROIANO, R. P., PRENTICE, R., PATTERSON, R., CARROLL, R. & SUBAR, A. F. 2006. A comparison of two dietary instruments for evaluating the fat-breast cancer relationship. *International Journal of Epidemiology,* 35**,** 1011-21.

FREEDMAN, L. S., SCHATZKIN, A., MIDTHUNE, D. & KIPNIS, V. 2011. Dealing with dietary measurement error in nutritional cohort studies. *Journal of the National Cancer Institute,* 103**,** 1086-92.

FRIZZELL, J. P. 2005. Acute stroke: pathophysiology, diagnosis, and treatment. *AACN Clin Issues,* 16**,** 421-40.

FSA 1999. Food Standards Agency. Food Labelling & Standards Division 'A'. Guidance notes on nutrition labelling. [Online] Available at: http://www.food.gov.uk/multimedia/pdfs/nutlabel2.pdf [Accessed Jan 2014].

GALE, C. P., CATTLE, B. A., WOOLSTON, A., BAXTER, P. D., WEST, T. H., SIMMS, A. D., BLAXILL, J., GREENWOOD, D. C., FOX, K. A. & WEST, R. M. 2012. Resolving inequalities in care? Reduced mortality in the elderly after acute coronary syndromes. The Myocardial Ischaemia National Audit Project 2003-2010. *European Heart Journal,* 33**,** 630-9.

GEORGE, S. J. & LYON, C. 2010. Pathogenesis of Atherosclerosis. *In:* JOHNSON, J. & GEORGE, S. J. (eds.) *Atherosclerosis: molecular and cellular mechanisms.* Wiley-VCH.

GLASS, G. V. 1976. Primary, secondary and meta-analysis of research. *Educational Researcher,* 10**,** 3-8.

GOLDSTEIN, L. B., BUSHNELL, C. D., ADAMS, R. J., APPEL, L. J., BRAUN, L. T., CHATURVEDI, S., CREAGER, M. A., CULEBRAS, A., ECKEL, R. H., HART, R. G., HINCHEY, J. A., HOWARD, V. J., JAUCH, E. C., LEVINE, S. R., MESCHIA, J. F., MOORE, W. S., NIXON, J. V. & PEARSON, T. A. 2011. Guidelines for the primary prevention of stroke: a guideline for healthcare

professionals from the American Heart Association/American Stroke Association. *Stroke,* 42**,** 517-84.

GRECH, E. D. 2003. Pathophysiology and investigation of coronary artery disease. *BMJ,* 326**,** 1027-30.

GREENLAND, S. 1980. The effect of misclassification in the presence of covariates. *American Journal of Epidemiology,* 112**,** 564-9.

GREENLAND, S. & LONGNECKER, M. P. 1992. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *American Journal of Epidemiology,* 135**,** 1301-1309.

GREENLAND, S., PEARL, J. & ROBINS, J. M. 1999. Causal diagrams for epidemiologic research. *Epidemiology,* 10**,** 37-48.

GREENWOOD, D. C., GILTHORPE, M. S., GOLDING, C. & CADE, J. E. 2003. Stability over time of dietary patterns in the UK Women's Cohort Study. *Proceedings of the Nutrition Society,* 62, 89A.

HAMLING, J., LEE, P., WEITKUNAT, R. & AMBUHL, M. 2008. Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category. *Statistics in Medicine,* 27**,** 954-970.

HANSELL, A., BOTTLE, A., SHURLOCK, L. & AYLIN, P. 2001. Accessing and using hospital activity data. *Journal of Public Health Medicine,* 23**,** 51-6.

HAUNER, H., BECHTHOLD, A., BOEING, H., BRONSTRUP, A., BUYKEN, A., LESCHIK-BONNET, E., LINSEISEN, J., SCHULZE, M., STROHM, D. & WOLFRAM, G. 2012. Evidence-based guideline of the German Nutrition Society: carbohydrate intake and prevention of nutrition-related diseases. *Annals of Nutrition and Metabolism,* 60 Suppl 1**,** 1-58.

HE, F. J., NOWSON, C. A. & MACGREGOR, G. A. 2006. Fruit and vegetable consumption and stroke: meta-analysis of cohort studies. *Lancet,* 367**,** 320-6.

HE, M., VAN DAM, R. M., RIMM, E., HU, F. B. & QI, L. 2010. Whole-grain, cereal fiber, bran, and germ intake and the risks of all-cause and cardiovascular disease-specific mortality among women with type 2 diabetes mellitus. *Circulation,* 121**,** 2162-8.

HENDERSON, L., GREGORY, J., IRVING, K. & SWAN, G. 2003. The National Diet & Nutrition Survey: adults aged 19 to 64 years. Volume 2: Energy, protein, carbohydrate, fat and alcohol intake. HMSO London.

HEROUX, M., JANSSEN, I., LAM, M., LEE, D. C., HEBERT, J. R., SUI, X. & BLAIR, S. N. 2010. Dietary patterns and the risk of mortality: Impact of cardiorespiratory fitness. *International Journal of Epidemiology,* 39**,** 197-209.

HERRETT, E., SHAH, A. D., BOGGON, R., DENAXAS, S., SMEETH, L., VAN STAA, T., TIMMIS, A. & HEMINGWAY, H. 2013. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. *British Medical Journal,* 346**,** f2350.

HERRETT, E., SMEETH, L., WALKER, L. & WESTON, C. 2010. The Myocardial Ischaemia National Audit Project (MINAP). *Heart,* 96**,** 1264-7.

HES. 2013. *The HES Processing Cycle and HES Data Quality. [Online] Available from: http://www.hscic.gov.uk/media/1366/The-HES-processing-cycle-and-HES-data-quality/pdf/HES_data_quality_for_web_FINAL.pdf [Accessed 2013]* [Online].

HIGGINS, J. P. & THOMPSON, S. G. 2002. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine,* 21.

HILL, A. B. 1965. The environment and disease: Association of causation? *Proceedings of the Royal Society of Medicine,* 58**,** 295-300.

HIPPISLEY-COX, J., COUPLAND, C., VINOGRADOVA, Y., ROBSON, J., MINHAS, R., SHEIKH, A. & BRINDLE, P. 2008. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ,* 336**,** 1475-82.

HLEBOWICZ, J., PERSSON, M., GULLBERG, B., SONESTEDT, E., WALLSTROM, P., DRAKE, I., NILSSON, J., HEDBLAD, B. & WIRFALT, E. 2011. Food patterns, inflammation markers and incidence of cardiovascular disease: the Malmo Diet and Cancer study. *Journal of Internal Medicine,* 270**,** 365-76.

HOLLAND, B., WELCH, A. A., UNWIN, I. D., BUSS, D. H., PAUL, A. A. & SOUTHGATE, D. A. T. 1991. *McCance & Widdowson's The Composition of Foods*, London: Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food.

HOOK, E. B. & REGAL, R. R. 1995. Capture-recapture methods in epidemiology: methods and limitations. *Epidemiol Rev,* 17**,** 243-64.

IWGDMF 1995a. International Working Group for Disease Monitoring and Forecasting. Capture-recapture and multiple-record systems estimation I: History and theoretical development. *American Journal of Epidemiology,* 142**,** 1047-58.

IWGDMF 1995b. International Working Group for Disease Monitoring and Forecasting. Capture-recapture and multiple-record systems estimation II: Applications in human diseases. *American Journal of Epidemiology,* 142**,** 1059-68.

JAMES, S. L., MUIR, J. G., CURTIS, S. L. & GIBSON, P. R. 2003. Dietary fibre: a roughage guide. *Internal Medicine Journal,* 33**,** 291-6.

JOHANSSON, E. V., NILSSON, A. C., OSTMAN, E. M. & BJORCK, I. M. 2013. Effects of indigestible carbohydrates in barley on glucose metabolism, appetite and voluntary food intake over 16 h in healthy adults. *Nutr J,* 12**,** 46.

JUNI, P., ALTMAN, D. G. & EGGER, M. 2008. Assessing the quality of randomised controlled trials. *In:* EGGER, M., SMITH, G. D. & ALTMAN, D. G. (eds.) *Systematic Reviews in Health Care: Meta-analysis in Context.* Second ed. London: BMJ Publishing Group.

KAUSHIK, S., WANG, J. J., WONG, T. Y., FLOOD, V., BARCLAY, A., BRAND-MILLER, J. & MITCHELL, P. 2009. Glycemic index, retinal vascular caliber, and stroke mortality. *Stroke,* 40**,** 206-212.

KEY, T. J., APPLEBY, P. N., CAIRNS, B. J., LUBEN, R., DAHM, C. C., AKBARALY, T., BRUNNER, E. J., BURLEY, V., CADE, J. E., GREENWOOD, D. C., STEPHEN, A. M., MISHRA, G., KUH, D., KEOGH, R. H., WHITE, I. R., BHANIANI, A., BORGULYA, G., MULLIGAN, A. A. & KHAW, K. T. 2011. Dietary fat and breast cancer: comparison of results from food diaries and food-frequency questionnaires in the UK Dietary Cohort Consortium. *American Journal of Clinical Nutrition,* 94**,** 1043-52.

KIPNIS, V. & FREEDMAN, L. S. 2008. Impact of exposure measurement error in nutritional epidemiology. *Journal of the National Cancer Institute,* 100**,** 1658-9.

KIPNIS, V., SUBAR, A. F., MIDTHUNE, D., FREEDMAN, L. S., BALLARD-BARBASH, R., TROIANO, R. P., BINGHAM, S., SCHOELLER, D. A., SCHATZKIN, A. & CARROLL, R. J. 2003. Structure of dietary measurement error: results of the OPEN biomarker study. *American Journal of Epidemiology,* 158**,** 14-21.

KNEKT, P., REUNANEN, A., JARVINEN, R., SEPPANEN, R., HELIOVAARA, M. & AROMAA, A. 1994. Antioxidant vitamin intake and coronary mortality in a longitudinal population study. *American Journal of Epidemiology,* 139**,** 1180-1189.

KOETH, R. A., WANG, Z., LEVISON, B. S., BUFFA, J. A., ORG, E., SHEEHY, B. T., BRITT, E. B., FU, X., WU, Y., LI, L., SMITH, J. D., DIDONATO, J. A., CHEN, J., LI, H., WU, G. D., LEWIS, J. D., WARRIER, M., BROWN, J. M., KRAUSS, R. M., TANG, W. H., BUSHMAN, F. D., LUSIS, A. J. & HAZEN, S. L. 2013. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med,* 19**,** 576-85.

KOKUBO, Y., ISO, H., SAITO, I., YAMAGISHI, K., ISHIHARA, J., INOUE, M. & TSUGANE, S. 2011. Dietary fiber intake and risk of cardiovascular disease in the Japanese population: The Japan Public Health Center-based study cohort. *European Journal of Clinical Nutrition,* 65**,** 1233-1241.

LAAKSONEN, D. E., NYYSSONEN, K., NISKANEN, L., RISSANEN, T. H. & SALONEN, J. T. 2005. Prediction of cardiovascular mortality in middle-aged men by dietary and serum linoleic and polyunsaturated fatty acids. *Archives of Internal Medicine,* 165**,** 193-199.

LAM, C. S. & LITTLE, W. C. 2012. Sex and cardiovascular risk: are women advantaged or men disadvantaged? *Circulation,* 126**,** 913-5.

LANDBERG, R. 2012. Dieraty fiber and mortality: convincing observations that call for mechanistic investigations. *American Journal of Clinical Nutrition,* 2**,** 3-4.

LARSSON, S. C., MANNISTO, S., VIRTANEN, M. J., KONTTO, J., ALBANES, D. & VIRTAMO, J. 2009. Dietary fiber and fiber-rich food intake in relation to risk of stroke in male smokers. *European Journal of Clinical Nutrition,* 63**,** 1016-1024.

LATTIMER, J. M. & HAUB, M. D. 2010. Effects of dietary fiber and its components on metabolic health. *Nutrients,* 2**,** 1266-89.

LEAN, M. E. J. 2006. *Fox and Cameron's Food Science, Nutrition & Health 7th Edition,* London, Edward Arnold Ltd.

LIU, S., BURING, J. E., SESSO, H. D., RIMM, E. B., WILLETT, W. C. & MANSON, J. E. 2002a. A prospective study of dietary fiber intake and risk of cardiovascular disease among women. *Journal of the American College of Cardiology,* 39**,** 49-56.

LIU, S., MANSON, J. E., BURING, J. E., STAMPFER, M. J., WILLETT, W. C. & RIDKER, P. M. 2002b. Relation between a diet with a high glycemic load and plasma concentrations of high-sensitivity C-reactive protein in middle-aged women. *American Journal of Clinical Nutrition,* 75**,** 492-8.

LIU, S., WILLETT, W. C., STAMPFER, M. J., HU, F. B., FRANZ, M., SAMPSON, L., HENNEKENS, C. H. & MANSON, J. E. 2000. A prospective study of dietary glycemic load, carbohydrate intake, and risk of coronary heart disease in US women. *American Journal of Clinical Nutrition,* 71**,** 1455-61.

LOGUE, J., MURRAY, H. M., WELSH, P., SHEPHERD, J., PACKARD, C., MACFARLANE, P., COBBE, S., FORD, I. & SATTAR, N. 2011. Obesity is associated with fatal coronary heart disease independently of traditional risk factors and deprivation. *Heart,* 97**,** 564-8.

LOUCKS, E. B., LYNCH, J. W., PILOTE, L., FUHRER, R., ALMEIDA, N. D., RICHARD, H., AGHA, G., MURABITO, J. M. & BENJAMIN, E. J. 2009. Life-course socioeconomic position and incidence of coronary heart disease: the Framingham Offspring Study. *Am J Epidemiol,* 169**,** 829-36.

LUCAS, R. M. & MCMICHAEL, A. J. 2005. Association or causation: evaluating links between "environment and disease". *Bull World Health Organ,* 83**,** 792-5.

LUDWIG, D. S., PEREIRA, M. A., KROENKE, C. H., HILNER, J. E., VAN HORN, L., SLATTERY, M. L. & JACOBS JR, D. R. 1999. Dietary fiber, weight gain, and cardiovascular disease risk factors in young adults. *Journal of the American Medical Association,* 282**,** 1539-1546.

LUNN, J. & BUTTRISS, J. L. 2007. Carbohydrates and dietary fibre. *British Nutrition Foundation Nutrition Bulletin,* 32**,** 21-64.

LYONS-WALL, P. 2007. Food analysis, food composition tables and databases. *In:* MANN, J. & TRUSWELL, A. S. (eds.) *Essentials of Human Nutrition. Third edition.* Oxford University Press.

MACKAY, J. & MENSAH, G. (eds.) 2004. *World Health Organisation The Atlas of Heart Disease and Stroke. [Online] Available at http://www.who.int/cardiovascular_diseases/en/cvd_atlas_15_burden_stroke.pdf. [Accessed August 2012],* Geneva: World Health Organisation.

MATTHEWS, K. A., CRAWFORD, S. L., CHAE, C. U., EVERSON-ROSE, S. A., SOWERS, M. F., STERNFELD, B. & SUTTON-TYRRELL, K. 2009. Are changes in cardiovascular disease risk factors in midlife women due to chronological aging or to the menopausal transition? *J Am Coll Cardiol,* 54**,** 2366-73.

MCFADDEN, E., LUBEN, R., WAREHAM, N., BINGHAM, S. & KHAW, K. T. 2008. Occupational social class, risk factors and cardiovascular disease incidence in men and women: a prospective study in the European Prospective Investigation of Cancer and Nutrition in Norfolk (EPIC-Norfolk) cohort. *European Journal of Epidemiology,* 23**,** 449-58.

MCMACKIN, E., DEAN, M., WOODSIDE, J. V. & MCKINLEY, M. C. 2013. Whole grains and health: attitudes to whole grains against a prevailing background of increased marketing and promotion. *Public Health Nutr,* 16**,** 743-51.

MENTE, A., DE KONING, L., SHANNON, H. S., ANAND, S. S., MENTE, A., DE KONING, L., SHANNON, H. S. & ANAND, S. S. 2009. A systematic review of the evidence supporting a causal link between dietary factors and coronary heart disease. *Archives of Internal Medicine,* 169**,** 659-69.

MICHELS, K. B. 2001. A renaissance for measurement error. *International Journal of Epidemiology,* 30**,** 421-422.

MINAP 2010. MINAP Dataset Version 9.1 December. 2010. [Online] Available from http://www.ucl.ac.uk/nicor/audits/minap/dataset.

MOSCA, L., BANKA, C. L., BENJAMIN, E. J., BERRA, K., BUSHNELL, C., DOLOR, R. J., GANIATS, T. G., GOMES, A. S., GORNIK, H. L., GRACIA, C., GULATI, M., HAAN, C. K., JUDELSON, D. R., KEENAN, N., KELEPOURIS, E., MICHOS, E. D., NEWBY, L. K., OPARIL, S., OUYANG, P., OZ, M. C., PETITTI, D., PINN, V. W., REDBERG, R. F., SCOTT, R., SHERIF, K., SMITH, S. C., JR., SOPKO, G., STEINHORN, R. H., STONE, N. J., TAUBERT, K. A., TODD, B. A., URBINA, E. & WENGER, N. K. 2007. Evidence-based guidelines for cardiovascular disease prevention in women: 2007 update. *J Am Coll Cardiol,* 49**,** 1230-50.

MOSCA, L., BENJAMIN, E. J., BERRA, K., BEZANSON, J. L., DOLOR, R. J., LLOYD-JONES, D. M., NEWBY, L. K., PINA, I. L., ROGER, V. L., SHAW, L. J., ZHAO, D., BECKIE, T. M., BUSHNELL, C., D'ARMIENTO, J., KRIS-ETHERTON, P. M., FANG, J., GANIATS, T. G., GOMES, A. S., GRACIA, C. R., HAAN, C. K., JACKSON, E. A., JUDELSON, D. R., KELEPOURIS, E., LAVIE, C. J., MOORE, A., NUSSMEIER, N. A., OFILI, E., OPARIL, S., OUYANG, P., PINN, V. W., SHERIF, K., SMITH, S. C., JR., SOPKO, G., CHANDRA-STROBOS, N., URBINA, E. M.,

VACCARINO, V. & WENGER, N. K. 2011. Effectiveness-based guidelines for the prevention of cardiovascular disease in women--2011 update: a guideline from the American Heart Association. *J Am Coll Cardiol,* 57**,** 1404-23.

MOZAFFARIAN, D., KUMANYIKA, S. K., LEMAITRE, R. N., OLSON, J. L., BURKE, G. L. & SISCOVICK, D. S. 2003. Cereal, fruit, and vegetable fiber intake and the risk of cardiovascular disease in elderly individuals. *Journal of the American Medical Association,* 289**,** 1659-1666.

MRC 2010. Medical Research Council. Strategic Framework for Health Informatics in Support of Research. [Online] Available from http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC006669 [Accessed May 2013].

NDNS 1994. The Dietary and Nutritional Survey of British Adults-Further analysis. London: HMSO.

NDNS 2011. National Diet and Nutrition Survey: Headline results from Years 1 and 2 (combined) of the rolling programme 2008-9 - 2009-10 (Tables) [Online] Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216485/dh_128556.pdf [Accessed Jan 2014].

NICHOLS, M., TOWNSEND, N., LUENGO-FERNANDEZ, R., LEAL, J., GRAY, A., SCARBOROUGH, P. & RAYNER, M. 2012. European Cardiovascular Disease Statistics. European Heart Network, Brussels, European Society of Cardiology, Sophia Antipolis.

NICKLAS, T. A., O'NEIL, C. E., LISKA, D. J., ALMEIDA, N. G. & FULGONI, V. L. 2011. Modeling dietary fibre intakes in US adults: Implications for public policy. *Food and Nutrition Sciences,* 2**,** 925-931.

NIGB. 2011. *National Information Governance Board for Health and Social Care: Section 251. [Online] Available from http://www.nigb.nhs.uk/s251 [Accessed 2011]* [Online].

NIH 2013. National Institutes of Health. National Heart, Lung and Blood Institute. Risk Assessment Tool for Estimating Your 10-year Risk of Having a Heart Attack. [Online]. Available from: http://cvdrisk.nhlbi.nih.gov/calculator.asp [Accessed November 2013].

NILSSON, A., JOHANSSON, E., EKSTROM, L. & BJORCK, I. 2013. Effects of a brown beans evening meal on metabolic risk markers and appetite regulating hormones at a subsequent standardized breakfast: a randomized cross-over study. *PLoS ONE,* 8**,** e59985.

NORDESTGAARD, B. G., PALMER, T. M., BENN, M., ZACHO, J., TYBJAERG-HANSEN, A., DAVEY SMITH, G. & TIMPSON, N. J. 2012. The effect of elevated body mass index on ischemic heart disease risk: causal estimates from a Mendelian randomisation approach. *PLoS Med,* 9**,** e1001212.

NOVELLA, S., DANTAS, A. P., SEGARRA, G., MEDINA, P. & HERMENEGILDO, C. 2012. Vascular Aging in Women: is Estrogen the Fountain of Youth? *Front Physiol,* 3**,** 165.

O'DONNELL, M. J., XAVIER, D., LIU, L., ZHANG, H., CHIN, S. L., RAO-MELACINI, P., RANGARAJAN, S., ISLAM, S., PAIS, P., MCQUEEN, M. J., MONDO, C., DAMASCENO, A., LOPEZ-JARAMILLO, P., HANKEY, G. J., DANS, A. L., YUSOFF, K., TRUELSEN, T., DIENER, H. C., SACCO, R. L., RYGLEWICZ, D., CZLONKOWSKA, A., WEIMAR, C., WANG, X. & YUSUF, S. 2010. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet,* 376**,** 112-23.

OBA, S., NAGATA, C., NAKAMURA, K., FUJII, K., KAWACHI, T., TAKATSUKA, N. & SHIMIZU, H. 2010. Dietary glycemic index, glycemic load, and intake of carbohydrate and rice in

relation to risk of mortality from stroke and its subtypes in Japanese men and women. *Metabolism: Clinical & Experimental,* 59**,** 1574-82.

OH, K., HU, F. B., CHO, E., REXRODE, K. M., STAMPFER, M. J., MANSON, J. E., LIU, S. & WILLETT, W. C. 2005. Carbohydrate intake, glycemic index, glycemic load, and dietary fiber in relation to risk of stroke in women. *American Journal of Epidemiology,* 161**,** 161-169.

OKURA, Y., URBAN, L. H., MAHONEY, D. W., JACOBSEN, S. J. & RODEHEFFER, R. J. 2004. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *Journal of Clinical Epidemiology,* 57**,** 1096-103.

ONS 2013. Office for National Statistics. Information Paper. Quality and methodology information. Mortality Statistics in England and Wales. [Online]. Available from qmimortalityjuly13finalforpub_tcm77-248785 (1).pdf [Accessed December 2013].

ORSINI, N., BELLOCCO, R. & GREENLAND, S. 2006. Generalized least squares for trend estimation of summarized dose-response data. *Stata Journal,* 6**,** 40-57.

ORSINI, N. & GREENLAND, S. 2011. A procedure to tabulate and plot results after flexible modeling of a quantitative covariate. *Stata Journal,* 11**,** 1-29.

PARK, Y., SUBAR, A. F., HOLLENBECK, A. & SCHATZKIN, A. 2011. Dietary fiber intake and mortality in the NIH-AARP diet and health study. *Archives of Internal Medicine,* 171**,** 1061-1068.

PEARSON-STUTTARD, J., BAJEKAL, M., SCHOLES, S., O'FLAHERTY, M., HAWKINS, N. M., RAINE, R. & CAPEWELL, S. 2012. Recent UK trends in the unequal burden of coronary heart disease. *Heart,* 98**,** 1573-82.

PEREIRA, M. A., O'REILLY, E., AUGUSTSSON, K., FRASER, G. E., GOLDBOURT, U., HEITMANN, B. L., HALLMANS, G., KNEKT, P., LIU, S., PIETINEN, P., SPIEGELMAN, D., STEVENS, J., VIRTAMO, J., WILLETT, W. C. & ASCHERIO, A. 2004. Dietary fiber and risk of coronary heart disease: a pooled analysis of cohort studies. *Archives of Internal Medicine,* 164**,** 370-6.

PERK, J., DE BACKER, G., GOHLKE, H., GRAHAM, I., REINER, Z., VERSCHUREN, W. M., ALBUS, C., BENLIAN, P., BOYSEN, G., CIFKOVA, R., DEATON, C., EBRAHIM, S., FISHER, M., GERMANO, G., HOBBS, R., HOES, A., KARADENIZ, S., MEZZANI, A., PRESCOTT, E., RYDEN, L., SCHERER, M., SYVANNE, M., SCHOLTE OP REIMER, W. J., VRINTS, C., WOOD, D., ZAMORANO, J. L. & ZANNAD, F. 2012. European guidelines on cardiovascular disease prevention in clinical practice (version 2012): The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts). *Atherosclerosis,* 223**,** 1-68.

PIETINEN, P., RIMM, E. B., KORHONEN, P., HARTMAN, A. M., WILLETT, W. C., ALBANES, D. & VIRTAMO, J. 1996. Intake of dietary fiber and risk of coronary heart disease in a cohort of Finnish men. The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study. *Circulation,* 94**,** 2720-2727.

PRENTICE, R. L. 1986. A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika,* 73**,** 1-11.

PROSKY, L., ASP, N. G., FURDA, I., DEVRIES, J. W., SCHWEIZER, T. F. & HARLAND, B. F. 1985. Determination of total dietary fiber in foods and food products: collaborative study. *Journal of the Association of Official Analytical Chemists,* 68**,** 677-9.

PROSKY, L., ASP, N. G., SCHWEIZER, T. F., DEVRIES, J. W. & FURDA, I. 1988. Determination of insoluble, soluble, and total dietary fiber in foods and food products: interlaboratory study. *Journal of the Association of Official Analytical Chemists,* 71**,** 1017-23.

REGAL, R. R. & HOOK, E. B. 1984. Goodness-of-fit based confidence intervals for estimates of the size of a closed population. *Statistics in Medicine,* 3**,** 287-91.

REIMER, R. A., PELLETIER, X., CARABIN, I. G., LYON, M., GAHLER, R., PARNELL, J. A. & WOOD, S. 2010. Increased plasma PYY levels following supplementation with the functional fiber PolyGlycopleX in healthy adults. *Eur J Clin Nutr,* 64**,** 1186-91.

REINTJES, R., THELEN, M., REICHE, R. & CSOHAN, A. 2007. Benchmarking national surveillance systems: a new tool for the comparison of communicable disease surveillance and control in Europe. *Eur J Public Health,* 17**,** 375-80.

RIBOLI, E. & KAAKS, R. 1997. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *International Journal of Epidemiology,* 26 Suppl 1**,** S6-14.

RIMM, E. B., ASCHERIO, A., GIOVANNUCCI, E., SPIEGELMAN, D., STAMPFER, M. J. & WILLETT, W. C. 1996. Vegetable, fruit, and cereal fiber intake and risk of coronary heart disease among men. *Journal of the American Medical Association,* 275**,** 447-451.

RIMM, E. B., WILLIAMS, P., FOSHER, K., CRIQUI, M. & STAMPFER, M. J. 1999. Moderate alcohol intake and lower risk of coronary heart disease: meta-analysis of effects on lipids and haemostatic factors. *BMJ,* 319**,** 1523-8.

ROGER, V. L., GO, A. S., LLOYD-JONES, D. M., BENJAMIN, E. J., BERRY, J. D., BORDEN, W. B., BRAVATA, D. M., DAI, S., FORD, E. S., FOX, C. S., FULLERTON, H. J., GILLESPIE, C., HAILPERN, S. M., HEIT, J. A., HOWARD, V. J., KISSELA, B. M., KITTNER, S. J., LACKLAND, D. T., LICHTMAN, J. H., LISABETH, L. D., MAKUC, D. M., MARCUS, G. M., MARELLI, A., MATCHAR, D. B., MOY, C. S., MOZAFFARIAN, D., MUSSOLINO, M. E., NICHOL, G., PAYNTER, N. P., SOLIMAN, E. Z., SORLIE, P. D., SOTOODEHNIA, N., TURAN, T. N., VIRANI, S. S., WONG, N. D., WOO, D. & TURNER, M. B. 2012. Heart disease and stroke statistics--2012 update: a report from the American Heart Association. *Circulation,* 125**,** e2-e220.

SACN 2008. Scientific Advisory Committee on Nutrition Statement on dietary fibre. [Online] Available at http://www.sacn.gov.uk/pdfs/final_sacn_position_statement_for_website_dietary_fibre.pdf [Accessed 2011].

SELF, S. G. & PRENTICE, R. L. 1988. Asymptotic distribution theory and efficiency results for case-cohort studies. *Annals of Statistics,* 16**,** 64-81.

SHEN, J., JOHNSON, V. M., SULLIVAN, L. M., JACQUES, P. F., MAGNANI, J. W., LUBITZ, S. A., PANDEY, S., LEVY, D., VASAN, R. S., QUATROMONI, P. A., JUNYENT, M., ORDOVAS, J. M. & BENJAMIN, E. J. 2011. Dietary factors and incident atrial fibrillation: the Framingham Heart Study. *American Journal of Clinical Nutrition,* 93**,** 261-6.

SLAVIN, J. 2003. Why whole grains are protective: biological mechanisms. *Proceedings of the Nutrition Society,* 62**,** 129-34.

SLAVIN, J. 2004. Whole grains and human health. *Nutrition Research Reviews,* 17**,** 99-110.

SLAVIN, J. L. 2005. Dietary fiber and body weight. *Nutrition,* 21**,** 411-8.

SMITH-WARNER, S. A., SPIEGELMAN, D., ADAMI, H. O., BEESON, W. L., VAN DEN BRANDT, P. A., FOLSOM, A. R., FRASER, G. E., FREUDENHEIM, J. L., GOLDBOHM, R. A., GRAHAM, S.,

KUSHI, L. H., MILLER, A. B., ROHAN, T. E., SPEIZER, F. E., TONIOLO, P., WILLETT, W. C., WOLK, A., ZELENIUCH-JACQUOTTE, A. & HUNTER, D. J. 2001. Types of dietary fat and breast cancer: a pooled analysis of cohort studies. *Int J Cancer,* 92**,** 767-74.

SOUTHGATE, D. A. 1969. Determination of carbohydrates in foods. II. Unavailable carbohydrates. *J Sci Food Agric,* 20**,** 331-5.

SPENCE, M., CADE, J. E., BURLEY, V. J. & GREENWOOD, D. C. 2002. Ability of the UK Women's Cohort food frequency questionnaire to rank dietary intakes: a preliminary validation study (Abstract) *Proceedings of the Nutrition Society,* 61**,** 117A.

STAMPFER, M. J., HU, F. B., MANSON, J. E., RIMM, E. B. & WILLET, W. C. 2000. Primary Prevention of Coronary Heart Disease in Women through Diet and Lifestyle. *New England Journal of Medicine,* 343**,** 16-22.

STATACORP 2009. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP.

STATACORP 2011. Stata Statistical Software: Release 12. College Station, TX. StataCorp LP.

STRANGES, S. & GUALLAR, E. 2012. Cardiovascular disease prevention in women: a rapidly evolving scenario. *Nutr Metab Cardiovasc Dis,* 22**,** 1013-8.

STREPPEL, M. T., ARENDS, L. R., VAN 'T VEER, P., GROBBEE, D. E. & GELEIJNSE, J. M. 2005. Dietary fiber and blood pressure: a meta-analysis of randomized placebo-controlled trials. *Archives of Internal Medicine,* 165**,** 150-6.

STREPPEL, M. T., OCKE, M. C., BOSHUIZEN, H. C., KOK, F. J. & KROMHOUT, D. 2008. Dietary fiber intake in relation to coronary heart disease and all-cause mortality over 40 y: the Zutphen Study. *American Journal of Clinical Nutrition,* 88**,** 1119-1125.

TALATI, R., BAKER, W. L., PABILONIA, M. S., WHITE, C. M. & COLEMAN, C. I. 2009. The effects of barley-derived soluble fiber on serum lipids. *Ann Fam Med,* 7**,** 157-63.

THREAPLETON, D. E., BURLEY, V. J., GREENWOOD, D. C., ALDWAIRJI, M. & CADE, J. E. 2012a. Dietary fibre intake and risk of fatal coronary heart disease in a cohort of British women. *Proceedings of the Nutrition Society,* 71 (OCE3) E224 [Abstract].

THREAPLETON, D. E., BURLEY, V. J., GREENWOOD, D. C., ALDWAIRJI, M. & CADE, J. E. 2013a. Dietary fibre and risk of haemorrhagic or ischemic stroke in the UK Women's Cohort Study. *Eur J Epidemiol,* 28 (S1) S25, O-047. [Abstract].

THREAPLETON, D. E., BURLEY, V. J., GREENWOOD, D. C. & CADE, J. E. 2012b. Dietary fibre intake and fatal stroke risk in the UK Women's Cohort Study. *J Epidemiol Community Health,* 66 (S1) OP27. [Abstract].

THREAPLETON, D. E., GREENWOOD, D. C., BURLEY, V. J., ALDWAIRJI, M. & CADE, J. E. 2013b. Dietary fibre and cardiovascular disease mortality in the UK Women's Cohort Study. *European Journal of Epidemiology,* 28**,** 335-46.

THREAPLETON, D. E., GREENWOOD, D. C., BURLEY, V. J. & CADE, J. E. 2013c. Cardiovascular disease incidence and dietary fibre intake in the UK Women's Cohort Study. *J Epidemiol Community Health,* 67 (S1) OP26. [Abstract].

THREAPLETON, D. E., GREENWOOD, D. C., EVANS, C. E., CLEGHORN, C. L., NYKJAER, C., WOODHEAD, C., CADE, J. E., GALE, C. P. & BURLEY, V. J. 2012c. Dietary fibre intake and cardiovascular disease: A systematic review and meta-analysis of prospective studies. *Proceedings of the Nutrition Society,* 71 (OCE3) E223 [Abstract].

THREAPLETON, D. E., GREENWOOD, D. C., EVANS, C. E., CLEGHORN, C. L., NYKJAER, C., WOODHEAD, C., CADE, J. E., GALE, C. P. & BURLEY, V. J. 2012d. Dietary fibre intake and

risk of stroke: A systematic review and meta-analysis of prospective studies. *Proceedings of the Nutrition Society,* 71 (OCE3)**,** E248 [Abstract].

THREAPLETON, D. E., GREENWOOD, D. C., EVANS, C. E., CLEGHORN, C. L., NYKJAER, C., WOODHEAD, C., CADE, J. E., GALE, C. P. & BURLEY, V. J. 2013d. Dietary fiber intake and risk of first stroke: a systematic review and meta-analysis. *Stroke,* 44**,** 1360-8.

THREAPLETON, D. E., GREENWOOD, D. C., EVANS, C. E., CLEGHORN, C. L., NYKJAER, C., WOODHEAD, C., CADE, J. E., GALE, C. P. & BURLEY, V. J. 2013e. Dietary fibre intake and cardiovascular disease: A systematic review and meta-analysis. *BMJ,* 347**,** f6879.

TODD, S., WOODWARD, M., TUNSTALL-PEDOE, H. & BOLTON-SMITH, C. 1999. Dietary antioxidant vitamins and fiber in the etiology of cardiovascular disease and all-cause mortality: results from the Scottish Heart Health Study. *American Journal of Epidemiology,* 150**,** 1073-1080.

TOWNSEND, N., WICKRAMASINGHE, K., BHATNAGAR, P., SMOLINA, K., NICHOLS, M., LEAL, J., LUENGO-FERNANDEZ, R. & RAYNER, M. 2012. Coronary heart disease statistics. A conpendium of health statistics. British Heart Foundation Health Promotion Research Group. Department of Public Health, University of Oxford. [Online]. Available at http://www.bhf.org.uk/publications/view-publication.aspx?ps=1002097 [Accessed 2013].

TROWELL, H. 1972. Ischemic heart disease and dietary fiber. *American Journal of Clinical Nutrition,* 25**,** 926-32.

TRUSWELL, A. S. 2002. Cereal grains and coronary heart disease. *European Journal of Clinical Nutrition,* 56**,** 1-14.

UNAL, B., CRITCHLEY, J. A. & CAPEWELL, S. 2004. Explaining the decline in coronary heart disease mortality in England and Wales between 1981 and 2000. *Circulation,* 109**,** 1101-7.

USDA 2010. Unites States Department of Agriculture, Agricultural Research Service. USDA National Nutrient Database for Standard Reference, Release 23. Nutrient Data Laboratory Home Page [online]. [Accessed 11 May 2011]. Available from: http://www.ars.usda.gov/ba/bhnrc/ndl.

USDA/HHS 2010. United States Department of Agriculture & United States Department of Health and Human Services. Chapter 4, foods and nutrients to increase. *Dietary Guidelines for Americans.* 7th ed. Washington, DC: U.S. Government Printing Ofice.

VAIDYA, D., BECKER, D. M., BITTNER, V., MATHIAS, R. A. & OUYANG, P. 2011. Ageing, menopause, and ischaemic heart disease mortality in England, Wales, and the United States: modelling study of national mortality data. *British Medical Journal,* 343**,** d5170.

VAN HORN, L., MCCOIN, M., KRIS-ETHERTON, P. M., BURKE, F., CARSON, J. A., CHAMPAGNE, C. M., KARMALLY, W., SIKAND, G., VAN HORN, L., MCCOIN, M., KRIS-ETHERTON, P. M., BURKE, F., CARSON, J. A. S., CHAMPAGNE, C. M., KARMALLY, W. & SIKAND, G. 2008. The evidence for dietary prevention and treatment of cardiovascular disease. *Journal of the American Dietetic Association,* 108**,** 287-331.

VAN HORN, L., TIAN, L., NEUHOUSER, M. L., HOWARD, B. V., EATON, C. B., SNETSELAAR, L., MATTHAN, N. R. & LICHTENSTEIN, A. H. 2012. Dietary Patterns Are Associated with Disease Risk among Participants in the Women's Health Initiative Observational Study. *Journal of Nutrition,* 142**,** 284-91.

VARRASO, R., KABRHEL, C., GOLDHABER, S. Z., RIMM, E. B. & CAMARGO, C. A., JR. 2012. Prospective study of diet and venous thromboembolism in US women and men. *American Journal of Epidemiology,* 175**,** 114-26.

WACHOLDER, S., MCLAUGHLIN, J. K., SILVERMAN, D. T. & MANDEL, J. S. 1992a. Selection of controls in case-control studies. I. Principles. *Am J Epidemiol,* 135**,** 1019-28.

WACHOLDER, S., SILVERMAN, D. T., MCLAUGHLIN, J. K. & MANDEL, J. S. 1992b. Selection of controls in case-control studies. II. Types of controls. *Am J Epidemiol,* 135**,** 1029-41.

WACHOLDER, S., SILVERMAN, D. T., MCLAUGHLIN, J. K. & MANDEL, J. S. 1992c. Selection of controls in case-control studies. III. Design options. *Am J Epidemiol,* 135**,** 1042-50.

WALKER, A. E., ROBINS, M. & WEINFELD, F. D. 1981. The National Survey of Stroke. Clinical findings. *Stroke* 12 (2 pt 2 suppl 1)**,** 13-44.

WALLSTROM, P., SONESTEDT, E., HLEBOWICZ, J., ERICSON, U., DRAKE, I., PERSSON, M., GULLBERG, B., HEDBLAD, B. & WIRFALT, E. 2012. Dietary fiber and saturated fat intake associations with cardiovascular disease differ by sex in the Malmo diet and cancer cohort: A prospective study. *PLoS ONE,* 7.

WANDERS, A. J., VAN DEN BORNE, J. J., DE GRAAF, C., HULSHOF, T., JONATHAN, M. C., KRISTENSEN, M., MARS, M., SCHOLS, H. A. & FESKENS, E. J. 2011. Effects of dietary fibre on subjective appetite, energy intake and body weight: a systematic review of randomized controlled trials. *Obes Rev,* 12**,** 724-39.

WANG, J.-D. 2002. *Basic prinples and practical applications in epidemiological research,* Singapore, World Scientific Publishing Co. Pte. Ltd.

WARD, H. A., KEOGH, R., LENTJES, M., LUBEN, R. N., WAREHAM, N. J. & KHAW, K. T. 2012. Fibre intake in relation to serum total cholesterol levels and CHD risk: A comparison of dietary assessment methods. *European Journal of Clinical Nutrition,* 66**,** 296-304.

WCRF 2007. World Cancer Research Fund. Second Expert Report. Food, Nutrition, Physical Activity and the Prevention of Cancer: a Global Perspective. Systematic Literature Review Specification Manual. [Online]. Available at: http://www.dietandcancerreport.org/cancer_resource_center/downloads/SLR_Manual.pdf [Accessed July 2013].

WHELTON, S. P., HYRE, A. D., PEDERSEN, B., YI, Y., WHELTON, P. K. & HE, J. 2005. Effect of dietary fiber intake on blood pressure: a meta-analysis of randomized, controlled clinical trials. *Journal of Hypertension,* 23**,** 475-81.

WHITE, I. R. 2009. Multivariate random-effects meta-analysis. *Stata Journal,* 9**,** 40-56.

WHO 2008. World Health Organisation: The 10 leading causes of death by broad income group. [Online] Available from: http://www.who.int/mediacentre/factsheets/fs310/en/ [Accessed August 2012].

WHO 2013. World Health Organisation Regional Office for Europe Definition of Cardiovascular Diseases. [Online] Available from: http://www.euro.who.int/en/health-topics/noncommunicable-diseases/cardiovascular-diseases/cardiovascular-diseases2/definition-of-cardiovascular-diseases [Accessed November 2013].

WILLETT, W. 2013a. Correction for the Effects of Measurement Error. *In:* HOFMAN, A., MARMOT, M., SAMET, J. & SAVITZ, D. Z. (eds.) *Nutritional Epidemiology.* Third ed. Oxford: Oxford University Press.

WILLETT, W. 2013b. Implications of total energy Intake for Epidemiologic Analyses. *In:* HOFMAN, A., MARMOT, M., SAMET, J. & SAVITZ, D. Z. (eds.) *Nutritional Epidemiology.* Third ed. Oxford: Oxford University Press.

WILLETT, W. 2013c. Nature of Variation in Diet. *In:* HOFMAN, A., MARMOT, M., SAMET, J. & SAVITZ, D. Z. (eds.) *Nutritional Epidemiology.* Third ed. Oxford Oxford University Press.

WILLETT, W. 2013d. Overview of Nutritional Epidemiology. *In:* HOFMAN, A., MARMOT, M., SAMET, J. & SAVITZ, D. Z. (eds.) *Nutritional Epidemiology.* Third ed. Oxford: Oxford University Press.

WILLETT, W. & LENART, E. 2013. Reproducibility and validity of Food Frequency Questionnaires. *In:* HOFMAN, A., MARMOT, M., SAMET, J. & SAVITZ, D. Z. (eds.) *Nutritional Epidemiology.* Third ed. Oxford: Oxford University Press.

WILLETT, W. W. 1998a. Diet and Coronary Heart Disease. *In:* WILLETT, W. W. (ed.) *Nutritional Epidemiology.* Second ed. Oxford: Oxford University Press.

WILLETT, W. W. 1998b. Issues in analysis and presentation of dietary data. *In:* WILLETT, W. W. (ed.) *Nutritional Epidemiology. Second Edition.* Oxford: Oxford University Press.

WILLETT, W. W. & STAMPFER, M. J. 1998. Implications of total energy intake for epidemiologic analyses. *In:* WILLETT, W. W. (ed.) *Nutritional Epidemiology.* Second ed. Oxford: Oxford University Press.

WOLK, A., MANSON, J. E., STAMPFER, M. J., COLDITZ, G. A., HU, F. B., SPEIZER, F. E., HENNEKENS, C. H. & WILLETT, W. C. 1999. Long-term intake of dietary fiber and decreased risk of coronary heart disease among women. *Journal of the American Medical Association,* 281**,** 1998-2004.

WORRALL-CARTER, L., SKI, C., SCRUTH, E., CAMPBELL, M. & PAGE, K. 2011. Systematic review of cardiovascular disease in women: assessing the risk. *Nursing and Health Science,* 13**,** 529-35.

WRIGHT, F. L., GREEN, J., CANOY, D., CAIRNS, B. J., BALKWILL, A. & BERAL, V. 2012. Vascular disease in women: comparison of diagnoses in hospital episode statistics and general practice records in England. *BMC Med Res Methodol,* 12**,** 161.

YE, E. Q., CHACKO, S. A., CHOU, E. L., KUGIZAKI, M. & LIU, S. 2012. Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease, and weight gain. *Journal of Nutrition,* 142**,** 1304-13.

Appendices

Appendix I: Medline search strategy

1. exp cohort studies/
2. case control study/
3. cohort$.tw.
4. epidemiologic methods/
5. or/1-4
6. (animals not (humans and animals)).sh.
7. 5 not 6
8. oligosaccharide$.tw.
9. (resistant adj3 starch).tw.
10. cellulose/
11. lignin/
12. methylcellulose/
13. carboxymethylcellulose/
14. inulin/
15. alginates/
16. exp oligosaccharides/
17. mannans/
18. pectins/
19. plant gums/
20. gum arabic/
21. tragacanth/
22. karaya gum/
23. dietary fibre/
24. fibre$.tw.
25. fibre$.tw.
26. "guar gum".tw.
27. psyllium/
28. psyllium$.tw.
29. "beta glucan$".tw.
30. beta-glucans/
31. or/8-30
32. exp cardiovascular diseases/
33. stroke.ab,ti.
34. "acute coronary syndrome".tw.

35. stemi.tw.

36. nstemi.tw.

37. (transient isch$emic adj3 (accident or incident)).tw.

38. exp coronary diseases/

39. exp heart diseases/

40. exp heart diseases/

41. (CHD or CVD).tw.

42. (myocardial adj3 infarction).tw.

43. exp myocardial infarction/

44. exp myocardial ischemia/

45. or/32-44

46. 7 and 31 and 45

47. limit 46 to english language

48. limit 47 to yr="2009 -current"

49. limit 48 to (addresses or bibliography or biography or case reports of clinical conference or comment or congresses or consensus development conference or dictionary or directory or editorial or interview or letter)

50. 48 not 49

51. limit 50 to (cats or cattle or chick embryo or dogs or goats or guinea pigs or hamsters or horses or mice or rabbits or rats or sheep or swine)

52. 50 not 51

Appendix II: Structured guidelines for article screening

Article Relevancy Criteria

MARK AS: 'article not relevant' / Reject if you can determine:

- Study published before 1990
- Study is not published in English
- Participants outside age range 5-70 years
- Study includes animals only
- The reference is not an original research article (e.g. news, letter, review)
- The study is not a cohort or an RCT (e.g. case study, cross-sectional study)
- The study does not relate to Carbohydrate intake at all (e.g. Meat, Soy etc)
- All participants have a health condition, are pregnant or have an eating disorder (e.g. Polycystic Ovary Syndrome/Cancer Patients/ Type1Diabetes/ Type2Diaberes/ Hypertension/ CVD/ Angina etc)
- The study does not relate carbohydrate intake to a clinical outcome
- The study relates to exercise and dietary components cannot be separated from the exercise.
- The study is clearly not relevant to our scope (e.g. cancer treatment studies or surgical operations)
- It does not cover satiety outcomes and intervention duration is 1 day or less

MARK AS: 'population not relevant' / Reject if you can determine:

- Participants are Oriental, African or Asian
- Participants are exclusively a native subgroup in an otherwise included country (e.g. native sub-groups in USA, Australia, New Zealand etc)

MARK AS: 'potentially relevant' / Allow if you cannot reject on the above criteria:

- Anything which appears to be relevant or where insufficient information is available to make a decision that it is 'article not relevant' or 'population not relevant'
- Studies which appear to be relevant even if the duration is too short to be formally included at a later stage.

Appendix III: Structured flow chart for article inclusion

Inclusion/Exclusion Form

Dietary Carbohydrates and Cardio-Metabolic Health and Disease

Assessor name..... **Date:** / /2010

Citation Details

First Author	
Ref Manager ID	
Publication Year	
Journal Details	

Status of Study (circle one):

Excluded	Included	Pending	Status/Code Updated in RefMan? ✔
Code.....	Code.....		☐

Population: Potentially non relevant? Give Detail .....

.....

.....

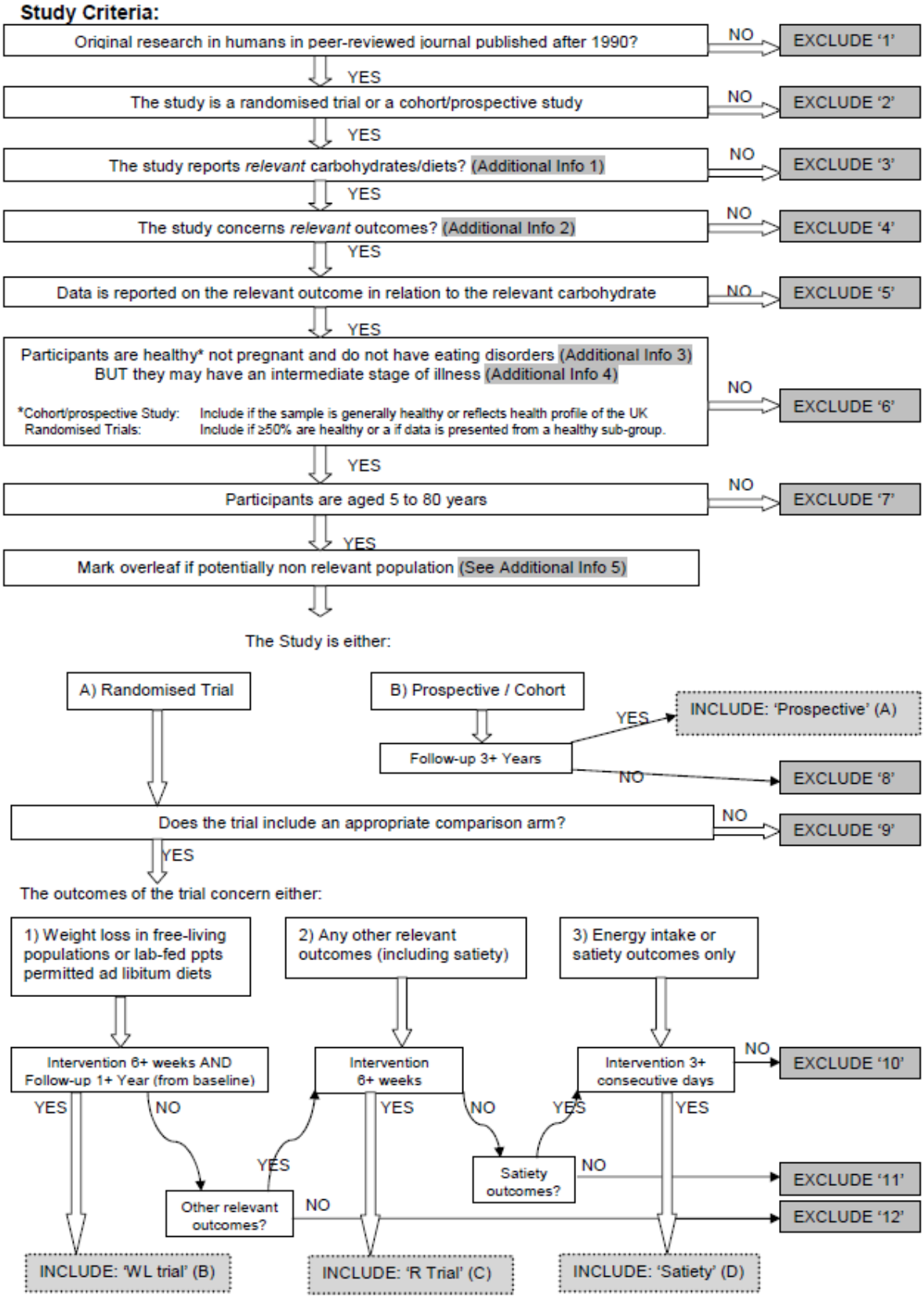
If included, mark this study as PA, PB, PC or PD to denote the population difference

If included:

		RefMan Updated ✔	Date Updated
Determine study type (see flow chart)			
Decide Study name			

If included as a trial, circle as appropriate

Was the allocation sequence adequately generated?	Bias	No bias	Unclear
Was allocation adequately concealed?	Bias	No bias	Unclear
Were participants blinded to treatment status?	Bias	No bias	Unclear
Were assessors blinded to treatment status?	Bias	No bias	Unclear
Were incomplete outcome data adequately addressed?	Bias	No bias	Unclear
Was the study free of suggestion of selective outcome reporting?	Bias	No bias	Unclear
Was the study free of other problems that could cause bias?	Bias	No bias	Unclear


Study Criteria:
Original research in humans in peer-reviewed journal published after 1990?
NO
EXCLUDE '1'
YES
The study is a randomised trial or a cohort/prospective study
NO
EXCLUDE '2'
YES
The study reports relevant carbohydrates/diets? (Additional Info 1)
NO
EXCLUDE '3'
YES
The study concerns relevant outcomes? (Additional Info 2)
NO
EXCLUDE '4'
YES
Data is reported on the relevant outcome in relation to the relevant carbohydrate
NO
EXCLUDE '5'
YES
Participants are healthy* not pregnant and do not have eating disorders (Additional Info 3)
BUT they may have an intermediate stage of illness (Additional Info 4)
*Cohort/prospective Study: Include if the sample is generally healthy or reflects health profile of the UK
Randomised Trials: Include if ≥50% are healthy or a if data is presented from a healthy sub-group.
NO
EXCLUDE '6'
YES
Participants are aged 5 to 80 years
NO
EXCLUDE '7'
YES
Mark overleaf if potentially non relevant population (See Additional Info 5)
The Study is either:
A) Randomised Trial
B) Prospective / Cohort
YES
INCLUDE: 'Prospective' (A)
Follow-up 3+ Years
NO
EXCLUDE '8'
Does the trial include an appropriate comparison arm?
NO
EXCLUDE '9'
YES
The outcomes of the trial concern either:
1) Weight loss in free-living populations or lab-fed ppts permitted ad libitum diets
2) Any other relevant outcomes (including satiety)
3) Energy intake or satiety outcomes only
Intervention 6+ weeks AND Follow-up 1+ Year (from baseline)
YES
NO
Intervention 6+ weeks
YES
NO
Intervention 3+ consecutive days
NO
EXCLUDE '10'
YES
YES
YES
NO
Satiety outcomes?
NO
EXCLUDE '11'
Other relevant outcomes?
NO
EXCLUDE '12'
INCLUDE: 'WL trial' (B)
INCLUDE: 'R Trial' (C)
INCLUDE: 'Satiety' (D)

Appendix IV: Individual study results for studies included in Chapter 2 systematic literature review and meta-analyses

Table IV.i: Results from cohort studies identified in the systematic review: Total dietary fibre and CVD, CHD and stroke events

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Sub-group Detail	Exposure Contrast and units	RR (CI) / Mean exposure (SD)	p	p trend	Adjustments
*(Akbaraly et al., 2011) Whitehall II	141/7319	17.7y	Total fibre intake (Englyst)	Total CVD [fatal]		Per decile of fibre intake based on recommendation of 24g/day	$ 0.87 (0.71, 1.05)		0.15	Sex, age, ethnicity, occupation, marital status, smoking, energy intake, physical activity, BMI, prevalent CVD, diabetes, hypertension, dyslipidaemia, metabolic syndrome, inflammatory markers
(Appleby et al., 1999) Oxford Vegetarian Study	(525) /11140	13.3 y	Dietary Fibre (Englyst)	Ischaemic heart disease [fatal]		Q3 vs Q1	2.25 (0.92, 5.53)		NS	Age, SES/class, sex, smoking
(Ascherio et al., 1998) HPFS	(328) /51529	8 y	Dietary Fibre (AOAC)	Total stroke [fatal + non-fatal]		(28.9) vs (12.4) g/d	‡ 0.70 (0.48, 1.0)		0.028	Age, alcohol, BMI, energy intake, smoking hypercholesterolaemia, hypertension, occupation, parental MI, physical activity,
(Bazzano et al., 2003) NHANES I	(1198) /14407	19 y (4)	Dietary Fibre (Assessment by Englyst and Southgate)	Total CVD [fatal]		10 g/ 1735 kcal/d	0.96 (0.9, 1.03)	0.29		Age, alcohol, BMI, total cholesterol intake, smoking, education, ethnicity, diabetes, physical activity, saturated fatty acid intake, sex, systolic blood pressure
	(3762)			Total CVD [fatal + non-fatal]			$ 0.93 (0.89, 0.97)	<0.001		
	(668)		Fibre Density	CHD events [fatal]			0.91 (0.83, 1.0)	0.06		
	(1843)			CHD events [fatal + non-fatal]			# 0.92 (0.86, 0.98)	0.01		
	(233)			Total stroke [fatal]			1.02 (0.85, 1.24)	0.8		
	(928)			Total stroke [fatal + non-fatal]			‡ 0.94 (0.87, 1.02)	0.12		
*(Buyken et al., 2010) Blue	(109)/1490	13y	Dietary fibre (AOAC)	Total CVD[fatal]	Women	36.2 vs 19.7 g/d	$ 0.88 (0.53, 1.46)	0.67		Age, energy intake, glyacemic index residuals, alcohol, smoking,

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Sub-group Detail	Exposure Contrast and units	RR (CI) / Mean exposure (SD)	p	p trend	Adjustments
Mountains Eye Study										diabetes
	(151)/1245				Men	36.4 vs 18.4 g/d	$ 0.84 (0.53, 1.34)	0.55		Age, energy, glycaemic index residuals, total fat, underweight, smoking, use of corticosteroids
*(Chuang et al., 2012) EPIC-Heart	2489/ 322153	12.7y	Dietary fibre assessed as AOAC in most countries and values were then calibrated across Europe	Total circulatory disease [Fatal]	Women	Per 10g/d	$ 0.88 (0.81, 0.97)		<0.001	Stratified by recruitment age, sex and centre. Adjusted for education, smoking, alcohol, BMI, physical activity, total energy intake
	2115/ 130564				Men		$ 0.90 (0.84, 0.97)		0.032	
*(Crowe et al., 2012) EPIC-Heart	2381/3063 31	11.5y	Dietary fibre assessed as AOAC in most countries and values were then calibrated across Europe	IHD mortality		Per 10g/d	# 0.85 (0.73, 0.99)		0.031	Age, alcohol, BMI, physical activity, marital status, education, employment, hypertension, hyperlipidaemia, angina pectoris, diabetes, PUFA:SFA, energy intake
	1596				Men		0.89 (0.76, 1.05)		0.156	
	785				Women		0.74 (0.57, 0.95)		0.017	
	909				Age<60 at recruitment		0.81 (0.67, 0.97)		0.022	
	1472				Age >=60 at recruitment		0.89 (0.75, 1.05)		0.177	
	747				Never smoker		0.80 (0.63, 1.01)		0.055	
	834				Former smoker		0.89 (0.72, 1.09)		0.254	
	156				Current smoker <10/d		0.84 (0.51, 1.37)		0.475	
	272				Current smoker 10-19/d		0.85 (0.61, 1.20)		0.362	
	372				Current smoker >=20/d		0.86 (0.65, 1.15)		0.320	
*(Eshak et al., 2010) Japanese Collaborative cohort study	(1063)/ 110792	13.4y	Dietary fibre (similar to AOAC values)	Total CVD[fatal]	Men	>12.6 vs. <7.8 g/d	$ 0.83 (0.63, 1.09)		0.054	Age, BMI, hypertension, diabetes, alcohol, smoking, education, exercise, walking, stress, sleep, fish, saturated fatty acid intake, n3 fatty acid, sodium intake, folate, vitamin Energy intake
	(1017)			Total CVD[fatal]	Women	>12.7 vs. <8.5 g/d	$ 0.82 (0.57, 0.97)		0.044	
	(231)			CHD[fatal]	Men	>12.6 vs. <7.8 g/d	# 0.81 (0.61, 1.09)		0.022	
	(191)			CHD[fatal]	Women	>12.7 vs. <8.5 g/d	# 0.80 (0.57, 0.97)		0.014	
	(499)			Stroke[fatal]	Men	>12.6 vs. <7.8 g/d	‡1.09 (0.75, 1.58)		0.555	
	(484)			Stroke[fatal]	Women	>12.7 vs. <8.5 g/d	‡1.05 (0.73, 1.51)		0.775	

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Sub-group Detail	Exposure Contrast and units	RR (CI) / Mean exposure (SD)	p	p trend	Adjustments
	(333)			Other CVD[fatal]	Men	>12.6 vs. <7.8 g/d	0.78 (0.54, 1.13)		0.313	
	(342)			Other CVD[fatal]	Women	>12.7 vs. <8.5 g/d	1.06 (0.74, 1.51)		0.212	
*(Kokubo et al., 2011) Japan public health centre-based cohort	1984 /40046	10.4	Energy-adjusted intake of total fibre (similar to AOAC)	Total CVD [fatal + non-fatal]	Men	19.9 vs. 6.0 g/d	$ 0.94 (0.74, 1.20)		0.649	Age, sex, smoking, alcohol, BMI, history of diabetes, medication for hypertension or hyper-cholesterolaemia, exercise, fruit, vegetables, fish, sodium, isoflavone, energy, public health centre
	1253 /46341				Women		$ 0.65 (0.48, 0.87)		0.002	
	1499 /40046			All stroke [fatal + non-fatal]	Men		‡1.00 (0.76, 1.32)		0.976	
	1054 /46341				Women		‡0.64 (0.46, 0.88)		0.005	
	910 /40046			Cerebral infarction [fatal + non-fatal]	Men		0.94 (0.66, 1.34)		0.540	
	518 /46341				Women		0.73 (0.55, 0.97)		0.029	
	456 /40046			Intracerebral haemorrhage [fatal + non-fatal]	Men		1.08 (0.66, 1.78)		0.588	
	310 /46341				Women		0.53 (0.28, 0.97)		0.100	
	133 /40046			Subarachnoid haemorrhage [fatal + non-fatal]	Men		1.02 (0.45, 2.54)		0.672	
	226 /46341				Women		0.72 (0.37, 1.43)		0.419	
	485 /40046			CHD [fatal + non-fatal]	Men		# 0.76 (0.47, 1.25)		0.327	
	199 /46341				Women		# 0.68 (0.32, 1.42)		0.149	
	712 /unknown		Total fibre intake (similar to AOAC)	Total CVD [fatal + non-fatal]	Male non-smokers	Q5 vs. Q1	0.59 (0.38, 0.90)		0.045	Age, sex, smoking, alcohol, BMI, history of diabetes or hypertension, medication for hyper-cholesterolaemia, exercise, fruit, vegetables, fish, sodium, isoflavone, energy, public health centre
	1424 /unknown				Male smokers		1.05 (0.79, 1.40)		0.862	
	1152 /unknown				Female non-smokers		0.61 (0.45, 0.83)		0.001	
	218 /unknown				Female smokers		0.58 (0.24, 1.39)		0.158	
(Larsson et al., 2009) ATBC Study	(196) /29133	13.6y	Dietary Fibre (Englyst)	Stroke, haemorrhage Subarachnoid [fatal + non-fatal]		(35.8) vs (16.1) g/d	‡0.86 (0.47, 1.59)		0.49	Age, alcohol, BMI, total cholesterol intake, diastolic and systolic blood pressure, energy intake, folate, HDL-C, CHD, diabetes, physical

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Sub-group Detail	Exposure Contrast and units	RR (CI) / Mean exposure (SD)	p	p trend	Adjustments
	(383) /29133			Stroke, haemorrhage Intracerebral [fatal + non-fatal]			‡0.97 (0.61, 1.54)		0.63	activity, magnesium Intake, smoking, group allocation
	(2702) /29133			Stroke, ischaemic [fatal + non-fatal]			‡1.01 (0.85, 1.19)		0.83	
(Liu et al., 2002a) The Women's Health Study	(570) /39876	6 y	Dietary Fibre (AOAC)	Total CVD [fatal + non-fatal]		26.3 vs. 12.5 g/d	$ 0.79 (0.58, 1.09)		0.17	Age, alcohol, BMI, energy intake, familial MI, fat intake, folate, diabetes, Hypercholesterol-aemia, hypertension, physical activity protein intake, smoking, group allocation, supplements, postmenopausal HRT
	(177) /39876	6 y	Dietary Fibre (AOAC)	MI [fatal + non-fatal]		(26.3) vs (12.5) g/d	# 0.68 (0.39, 1.22)		0.13	
(Mozaffarian et al., 2003) Cardiovascular Health Study	(811) /5201	8.6 y	Dietary Fibre (AOAC)	CHD events [fatal + non-fatal]		Q5 vs Q1 g/d	# 0.84 (0.66, 1.07)		0.23	Age, alcohol, cereal fibre, education, fibre from fruit, fibre from vegetables, diabetes, physical activity, sex, smoking
(Oh et al., 2005) Nurses' Health Study	(1020) /121700	18 y	Dietary Fibre (AOAC)	Total stroke [fatal + non-fatal]		(21) vs (10) g/d	‡0.83 (0.66, 1.04)		0.07	Age, alcohol, aspirin, BMI, carbohydrate intake, energy intake, familial: diabetes, hypertriglyceride-aemia, hypertension or MI, menopausal status, physical activity, smoking, vitamin intake, postmenopausal HRT
	(279)			Stroke, haemorrhagic [fatal + non-fatal]			0.84 (0.54, 1.3)		0.34	
	(515)			Stroke, ischaemic [fatal + non-fatal]			0.78 (0.56, 1.09)		0.09	
*(Park et al., 2011) NIH-AARP Diet and Health Study	5248 /388122	9y	Dietary Fibre (AOAC)	Total CVD [fatal]	Male	10 g/d	$ 0.88 (0.86, 0.91)		<0.001	Age, race, education, marital status, health status, BMI, physical activity, smoking, alcohol, red meat, fruit, vegetables, total energy
	2417 /388122				Female	10 g/d	$ 0.76 (0.69, 0.84)		<0.001	
	1134 /388122				Male never smokers	Q5 vs. Q1	0.95 (0.74, 1.21)		0.81	
	729 /388122				Female never smokers	Q5 vs. Q1	0.69 (0.50, 0.95)		0.02	
(Pietinen et al., 1996) ATBC Study	(1399) /29133	6.1 y	Dietary Fibre (Englyst)	Fatal CHD, MI		(34.8) vs (16.1) g/d	# 0.87 (0.73, 1.04)		0.8	Age, alcohol, beta-carotene, BMI, diastolic and systolic blood pressure, education, saturated fatty acid, energy intake, physical
	(635)			CHD events [fatal]			0.73 (0.56, 0.95)		0.004	

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Sub-group Detail	Exposure Contrast and units	RR (CI) / Mean exposure (SD)	p	p trend	Adjustments
										activity, smoking, group allocation, vitamin C and E
(Rimm et al., 1996) HPFS	(232) /51529	6 y	Dietary Fibre (AOAC)	CHD events [fatal]		(28.9) vs (12.4) g/d	# 0.45 (0.28, 0.72)		<0.001	Age, alcohol, BMI, saturated fatty acid intake, familial MI, smoking, vitamin E, hypercholesterolaemia, occupation, physical activity, hypertension
	(734)			MI [fatal + non-fatal]			0.59 (0.46, 0.76)		<0.001	
	(511)			MI [non-fatal]			# 0.65 (0.49, 0.88)		0.02	
(Streppel et al., 2008) Zutphen Elderly Study	(348) / 1373	40 y (0.2)	Dietary Fibre (Energy adjusted, recent intake)	CHD events [fatal]		10 g/d	# 0.83 (0.70, 0.98)			Fatty acid intake (Trans, cis-PUFA and saturated), BMI, smoking, energy intake, fish, prescribed diet, SES/Class, alcohol
			Dietary Fibre (Intake during middle age. Energy adjusted, long term intake)	CHD events [fatal]		10 g/d	0.87 (0.71, 1.07)			
(Todd et al., 1999) Scottish Heart Health Study	(296) /11629	9 y (0.1)	Fibre density (Inclusive of NSP resistant starch and lignin)	CHD events [fatal + non-fatal]	Men	Q4 vs Q1	0.64 (0.45, 0.9)			Age, alcohol, BMI, total cholesterol intake, energy intake, fibrinogen, HDL-C, diabetes, personality score, physical activity, smoking, systolic blood pressure, blood triglycerides
	(97)				Women	Q4 vs Q1	0.56 (0.29, 1.08)			
*(Wallstrom et al., 2012) Malmo Diet and Cancer Cohort	(1,077)/ 8038	13.2y	Fibre density (non-starch polysaccharide)	Fatal and non-fatal ischaemic coronary and cerebrovascular events	Men	11.4 vs. 5.8 g/1000kcal	$ 0.85 (0.70, 1.04)		0.30	Age, diet assessment method version, energy intake, season, BMI, Smoking, education, alcohol, SBP, antihypertensive medication, antihyperlipidaemic medication, leasure physical activity, quintiles of energy-adjusted dietary fibre [Note adjustment for fibre was listed in the publication although actual adjustment for this is unlikely in models dealing with dietary fibre intake]
	(687)/ 12,535	13.6y			Women	12.9 vs. 6.5 g/1000kcal	$ 0.76 (0.59, 0.97)		0.022	
	(680)/ 8038	13.2y		MI and death from ischaemic heart disease	Men	11.4 vs. 5.8 g/1000kcal	# 0.97 (0.75, 1.25)		0.85	
	(330)/ 12,402	13.6y			Women	12.9 vs. 6.5 g/1000kcal	# 0.78 (0.55, 1.11)		0.067	
	(397)/ 8,038	13.2y		Ischaemic stroke	Men	11.4 vs. 5.8 g/1000kcal	‡0.69 (0.49, 0.96)		0.050	
	(346)/ 12,402	13.6y			Women	12.9 vs. 6.5 g/1000kcal	‡0.73 (0.52, 1.04)		0.18	
*(Ward et al., 2012) EPIC-Norfolk	(1294) / 4347	11y	Dietary fibre g/d (Englyst) assessed with FFQ	Fatal and non-fatal CHD	Men	Risk per 6g/d increase	0.98 (0.89, 1.08)		0.68	Age, BMI, physical activity, smoking, family history of MI, class, diabetes, antihypertensive medication use, lipid-lowering
	(712) / 2728				Women		0.90 (0.80, 1.01)		0.072	

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Sub-group Detail	Exposure Contrast and units	RR (CI) / Mean exposure (SD)	p	p trend	Adjustments
	(1294) / 4347		Dietary fibre g/d (Englyst) assessed with food diary		Men		0.86 (0.79, 0.95)		0.001	medication use, aspirin use, energy from total fat, energy from non-fat, alcohol, saturated fat intake, serum cholesterol
	(712) / 2728				Women		0.86 (0.74, 0.99)		0.036	
(Wolk et al., 1999) Nurses' Health Study	(429) /121700	10 y (20)	Dietary Fibre (Long-term intake over 6 years. (AOAC))	Non-fatal MI		22.9 vs 11.5 g/d	0.57 (0.42, 0.77)		<0.001	Age, alcohol, aspirin, BMI, carbohydrate intake, saturated fatty acid intake, energy intake, hypertension, menopausal status, parental MI, period of exposure, physical activity, smoking, postmenopausal HRT, vitamin intake
	(162)			CHD events [fatal]			0.41 (0.23, 0.7)		0.002	
	(591)			Non-fatal MI, fatal CHD			# 0.77 (0.57, 1.04)		0.07	

*Identified during update search; # Result was used in total fibre and CHD meta-analysis; ǂ Result was used in total fibre and stroke meta-analysis; $ Result was used in total fibre and CVD meta-analysis

Table IV.ii: Results from cohort studies identified in the systematic review: Soluble and insoluble fibre and CVD, CHD and stroke events

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Subgroup	Contrast (mean)	RR (CI)	p	p trend	Adjustments
(Bazzano et al., 2003) NHANES I	(1198) /14407	19 y (4)	Soluble fibre Density ((Multiple assessment methods) of assessment, including Englyst and Southgate)	Total CVD [fatal]		5g/1735 kcal	0.98 (0.93, 1.04)	0.48		Age, alcohol, BMI, total cholesterol intake, smoking, education, ethnicity, diabetes, physical activity, saturated fat intake, sex, systolic blood pressure
	(3762)			Total CVD [fatal + non-fatal]			0.94 (0.9, 0.99)	0.01		
	(928)			Total stroke [fatal]			1.03 (0.83, 1.28)	0.78		
	(928)			Total stroke [fatal + non-fatal]			‡0.95 (0.88, 1.03)	0.18		
	(668)			CHD events [fatal]			0.91 (0.83, 0.99)	0.03		
	(1843)			CHD events [fatal + non-fatal]			#0.92 (0.87, 0.97)	0.004		
*(Eshak et al., 2010) Japanese Collaborative cohort study	(1063)/ 110792	13.4y	Insoluble fibre (similar to AOAC)	Total CVD [fatal]	Men	>9.2 vs. <5.9 g/d	0.82 (0.65, 0.98)	0.042		Age, BMI, hypertension, diabetes, alcohol, smoking, education, exercise, walking, stress, sleep, fish, saturated fatty acid intake, n3 fatty acid, sodium, folate, vitamin E
	(1017)				Women	>9.1 vs. <6.2 g/d	0.69 (0.53, 0.91)	0.017		
	(231)			CHD [fatal]	Men	>9.2 vs. <5.9 g/d	#0.48 (0.27, 0.84)	<0.001		
	(191)				Women	>9.1 vs. <6.2 g/d	#0.49 (0.27, 0.86)	0.004		
	(499)			Stroke [fatal]	Men	>9.2 vs. <5.9 g/d	0.96 (0.64, 1.45)	0.715		
	(484)				Women	>9.1 vs. <6.2 g/d	0.90 (0.63, 1.28)	0.128		
	(333)			Other CVD [fatal]	Men	>9.2 vs. <5.9 g/d	1.15 (0.78, 1.62)	0.798		
	(342)				Women	>9.1 vs. <6.2 g/d	0.83 (0.51, 1.33)	0.698		
	(1063)		Soluble fibre (similar to AOAC)	Total CVD [fatal]	Men	>2.3 vs. <1.3 g/d	0.81 (0.63, 1.04)	0.042		
	(1017)				Women	>2.4 vs. <1.5 g/d	0.83 (0.53, 1.02)	0.043		
	(231)			CHD [fatal]	Men	>2.3 vs. <1.3 g/d	#0.71 (0.41, 0.97)	0.043		
	(191)				Women	>2.4 vs. <1.5 g/d	#0.72 (0.43, 0.99)	0.035		
	(499)			Stroke [fatal]	Men	>2.3 vs. <1.3 g/d	‡0.90 (0.61, 1.31)	0.790		
	(484)				Women	>2.4 vs. <1.5 g/d	‡1.02 (0.73, 1.42)	0.643		
	(333)			Other CVD [fatal]	Men	>2.3 vs. <1.3 g/d	1.08 (0.75, 1.64)	0.573		
	(342)				Women	>2.4 vs. <1.5 g/d	0.96 (0.61, 1.50)	0.613		
*(Kokubo et al., 2011) Japan public	1253 /46341	10.4	Soluble fibre (similar to AOAC)	Total CVD [fatal + non-fatal]	Women	Q5 vs. Q1	0.74 (0.56, 0.97)		0.012	Age, sex, smoking, alcohol, BMI, history of diabetes, medication for hypertension or hyper-cholesterolaemia, exercise, fruit,
	1054			All stroke [fatal + non-fatal]			0.78 (0.58, 1.06)		0.031	

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Subgroup	Contrast (mean)	RR (CI)	p	p trend	Adjustments
health centre-based cohort	518			Cerebral infarction [fatal + non-fatal]			0.73 (0.47, 1.14)		0.051	vegetables, fish, sodium, isoflavone, energy, public health centre
	310			Intracerebral haemorrhage [fatal + non-fatal]			0.71 (0.40, 1.26)		0.183	
	199			CHD [fatal + non-fatal]			0.60 (0.29, 1.21)		0.252	
	1253		Insoluble fibre (similar to AOAC)	Total CVD [fatal + non-fatal]			0.64 (0.47, 0.85)		<0.001	
	1054			All stroke [fatal + non-fatal]			0.62 (0.45, 0.85)		0.001	
	518			Cerebral infarction [fatal + non-fatal]			0.62 (0.40, 0.98)		0.006	
	310			Intracerebral haemorrhage [fatal + non-fatal]			0.55 (0.30, 1.00)		0.070	
	199			CHD [fatal + non-fatal]			0.78 (0.48, 1.27)		0.396	
(Larsson et al., 2009) ATBC Study	(196) /29133	13.6 years	Soluble fibre (Englyst)	Stroke, haemorrhage, Subarachnoid [fatal + non-fatal]		(7.7) vs (3.8) g/d	‡0.95 (0.51, 1.79)		0.86	Age, alcohol, BMI, total cholesterol intake, diastolic blood pressure, energy intake, folate, HDL-C, CHD, diabetes, physical activity, magnesium intake, smoking, group allocation, systolic blood pressure
	(383)			Stroke, haemorrhage-Intracerebral [fatal + non-fatal]			‡0.99 (0.62, 1.59)		0.6	
	(2700)			Stroke, ischaemic [fatal + non-fatal]			‡0.86 (0.73, 1.02)		0.17	
	(196)		Insoluble fibre (Englyst)	Stroke, haemorrhage, Subarachnoid [fatal + non-fatal]		(28.3) vs (12.2) g/d	0.89 (0.49, 1.64)		0.58	
	(383)			Stroke, haemorrhage, Intracerebral [fatal + non-fatal]			0.88 (0.56, 1.39)		0.43	
	(2702)			Stroke, ischaemic [fatal + non-fatal]			1.03 (0.87, 1.21)		0.61	
(Liu et al., 2002a) The Women's Health Study	(570) /39876	6 y	Soluble fibre (AOAC)	Total CVD [fatal + non-fatal]		(8.6) vs (3.7) g/d	0.90 (0.68, 1.21)		0.5	Age, alcohol, BMI, energy intake, familial MI, fat intake, folate, diabetes, hypercholesterolaemia, hypertension, physical activity,
	(570)		Insoluble fibre (AOAC)	Total CVD [fatal + non-fatal]		(21.8) vs (9.5) g/d	0.78 (0.57, 1.06)		0.09	

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Subgroup	Contrast (mean)	RR (CI)	p	p trend	Adjustments
	(177)		Soluble fibre	MI [fatal + non-fatal]		(8.6) vs (3.7) g/d	#0.83 (0.47, 1.48)		0.4	protein, smoking, group allocation, supplements, postmenopausal HRT
			Insoluble fibre	MI [fatal + non-fatal]		(21.8) vs (9.5) g/d	#0.74 (0.42, 1.3)		0.12	
(Pietinen et al., 1996) ATBC Study	(1399) /29133	6.1 y	Insoluble fibre (based on Englyst)	Fatal CHD & non fatal MI		(27.7) vs (12.2) g/d	0.87 (0.73, 1.04)		0.13	Age, alcohol, beta-carotene, BMI, systolic and diastolic blood pressure, education, saturated fatty acid, energy intake, physical activity, smoking, group allocation, vitamin C and E
	(1399)		Insoluble non cellulosic polysaccharides (based on Englyst)	Fatal CHD & non fatal MI		(15.9) vs (6.8) g/d	0.86 (0.72, 1.03)		0.13	
	(635)		Insoluble fibre (based on Englyst)	CHD events [fatal]		(27.7) vs (12.2) g/d	#0.75 (0.58, 0.98)		0.01	
	(635)		Soluble fibre (based on Englyst)	CHD events [fatal]		(7.4) vs (3.7) g/d	#0.68 (0.5, 0.92)		0.003	
	(1399)		Soluble fibre (based on Englyst)	Fatal CHD & non fatal MI			0.83 (0.68, 1.01)		0.05	
(Rimm et al., 1996) HPFS	(740) 51529	6 y (6%loss)	Soluble fibre (Englyst)	MI + fatal coronary disease		10g/d	#1.07 (0.57, 2.02)			Saturated fat, vitamin E, age, BMI, physical activity, smoking, alcohol, hypertension, high cholesterol, familial MI, profession, insoluble fibre
			Insoluble fibre (Englyst)				#0.75 (0.59, 0.94)			Saturated fat, vitamin E, age, BMI, physical activity, smoking, alcohol, hypertension, high cholesterol, familial MI, profession, soluble fibre

*Identified during update search; # Result was used in the soluble or insoluble fibre and CHD meta-analysis; ‡ Result was used in the soluble fibre and stroke meta-analysis

Table IV.iii: Results from cohort studies identified in the systematic review: Fibre from cereal foods and CVD, CHD and stroke events

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Sub-group Detail	Contrast (mean)	RR (CI)/ Mean Exposure (SD)	p trend	Adjustments
*(Baer et al., 2011) Nurses' Health Study	(1026)/ 50112	18y	Cereal fibre (energy adjusted) (AOAC)	Total CVD [fatal]		Continuous risk estimate per 4g/d	0.82 (0.69, 0.97)		Competing risks model including: Age, BMI, weight change, height, smoking, physical activity, alcohol, nuts, PUFA, glycaemic load, dietary cholesterol, systolic blood pressure, medication, diabetes, parental MI, time since menopause
*(Bernstein et al., 2011) Nurses' Health Study	(2500)/ 72266	22y	Cereal fibre, cumulative average intake (AOAC)	Total CHD [fatal plus non-fatal]		Per 5g/d	#0.77 (0.69, 0.85)	<0.0001	Stratified on age, assessment period. Adjusted for saturated fat, monounsaturated fat, polyunsaturated fat, GI score, folate, protein, energy intake, alcohol trans-unsaturated fatty acids, BMI, smoking, menopausal status, parental eearly MI, multivitamin use, vitamin E supplement, weekly aspirin use, physical activity.
*(Buyken et al., 2010) Blue Mountains Eye Study	(109)/ 1490	13y	Cereal fibre (breakfast cereals, bread, pasta and rice) (AOAC)	Total CVD [fatal]	Women	10.9 vs 2.9 g/d	0.87 (0.55, 1.38)	0.54	Age, energy, GI residuals, alcohol, smoking, diabetes
	(151)/ 1245				Men	11.5 vs 3.0 g/d	1.04 (0.67, 1.61)	0.89	Age, energy, GI residuals, total fat, underweight, smoking, use of corticosteroids
*(Crowe et al., 2012) EPIC-Heart	2381/ 306331	11.5y	Cereal fibre assessed as AOAC in most countries and values were then calibrated across Europe	IHD mortality		Per 5g/d	#0.91 (0.82, 1.01)	0.084	Age, alcohol, BMI, physical activity, marital status, education, employment, hypertension, hyperlipidaemia, angina pectoris, diabetes, PUFA:SFA, energy intake, fibre from fruit, vegetables and other sources

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Sub-group Detail	Contrast (mean)	RR (CI)/ Mean Exposure (SD)	p trend	Adjustments
*(Eshak et al., 2010) Japanese Collaborative cohort study	(231)/ 110792	13.4y	Cereal fibre (similar to AOAC)	CHD [fatal]	Men	>2.1 vs. <1.4 g/d	#0.89 (0.65, 1.01)	0.060	Age, BMI, hypertension, diabetes, alcohol, smoking, education, exercise, walking, stress, sleep, saturated fatty acid, n3 fatty acid, sodium, folate, vitamin E, vegetable fibre, fruit fibre
	(191)				Women	>1.7 vs. <1.1 g/d	#0.76 (0.59, 0.97)	0.044	
(Kaushik et al., 2009) Blue Mountains Eye Study	(Cases not reported) /3654	13 y	Cereal fibre (Energy adjusted)	CHD events [fatal]		(3) vs (11) g/d	0.94 (0.73, 1.22)	0.65	Age, BMI, diastolic blood pressure, education ,MI, stroke, diabetes, self-rated health status, sex, smoking, systolic blood pressure, hypertension medication
	(95) /3654			Total stroke [fatal]		(3) vs (11) g/d	2.13 (1.19, 3.8)	0.02	
(Larsson et al., 2009) ATBC Study	(196) /29133	13.6 years	Cereal fibre (Englyst)	Stroke, haemorrhage Subarachnoid [fatal + non-fatal]		(27.5) vs (8.9) g/d	0.86 (0.5, 1.46)	0.6	Age, alcohol, BMI, total cholesterol intake, diastolic and systolic blood pressure, energy intake, folate, HDL-C, CHD, diabetes, physical activity, magnesium Intake, smoking, group allocation
	(383)			Stroke, haemorrhageIntracerebral [fatal + non-fatal]			0.94 (0.63, 1.42)	0.71	
	(2702)			Stroke, ischaemic [fatal + non-fatal]			1.06 (0.91, 1.23)	0.25	
(Liu et al., 2002a) The Women's Health Study	(177) /39876	6 y	Cereal fibre (AOAC)	MI [fatal + non-fatal]		(6.5) vs (3) g/d	#0.91 (0.56, 1.47)	0.74	Age, alcohol, BMI, energy intake, familial MI, fat intake, folate, diabetes, hypercholesterolaemia, hypertension, physical activity, protein intake, smoking, group allocation, supplements, postmenopausal HRT
	(570)			Total CVD [fatal + non-fatal]		(6.5) vs (3) g/d	1.11 (0.84, 1.46)	0.38	
(Mozaffarian et al., 2003) Cardiovascular Health Study	(811) /5201	8.6 y	Cereal fibre (AOAC)	CHD events [fatal + non-fatal]		>6.3 vs <1.7 g/d	#0.79 (0.62, 0.99)	0.02	Age, alcohol, education, fibre from fruit, fibre from Vegetables, diabetes, physical activity, sex, smoking
	(204)				Age 65-69y	80th vs 20th Centile	0.82 (0.67, 1.01)		
	(255)				Age 70-74y	80th vs 20th Centile	0.89 (0.75, 1.06)		
	(352)				Age >75y	80th vs 20th Centile	0.87 (0.73, 1.04)		

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Sub-group Detail	Contrast (mean)	RR (CI)/ Mean Exposure (SD)	p trend	Adjustments
	(434)				Women	80th vs 20th Centile	0.89 (0.74, 1.06)		
	(377)				Men	80th vs 20th Centile	0.83 (0.68, 1.02)		
	(575)				No T2DM	80th vs 20th Centile	0.85 (0.73, 1.0)		
(Oh et al., 2005) Nurses' Health Study	(1020) /121700	18y	Cereal fibre (AOAC)	Total stroke [fatal + non-fatal]		(5.7) vs (1.4) g/d	0.66 (0.52, 0.83)	0.001	Age, alcohol, aspirin, BMI, carbohydrate intake, energy intake, familial: diabetes, hypertriglycerideaemia, hypertension, MI, menopausal status, physical activity, smoking, postmenopausal HRT, vitamin intake
	(279)			Stroke, haemorrhagic [fatal + non-fatal]			0.51 (0.33, 0.78)	0.01	
	(515)			Stroke, ischaemic [fatal + non-fatal]			0.8 (0.57, 1.12)	0.23	
*(Park et al., 2011) NIH-AARP Diet and Health Study	5248 /388122	9y	Fibre from grains (AOAC)	Total CVD [fatal]	Male	Q5 vs. Q1	0.77 (0.71, 0.85)	<0.05	Age, race, education, marital status, health status, BMI, physical activity, smoking, alcohol, red meat, fruit, vegetables, total energy
	2417 /388122				Female	Q5 vs. Q1	0.72 (0.63, 0.82)	<0.05	
(Pietinen et al., 1996) ATBC Study	(635) /29133	6.1 y	Cereal fibre (Englyst)	CHD events [fatal]		(26.3) vs (8.8) g/d	#0.74 (0.57, 0.96)	0.01	Age, alcohol, beta-carotene, BMI, diastolic and systolic blood pressure, education, saturated fatty acid intake, energy intake, physical activity, smoking, group allocation, vitamin C and E
	(1399)			Fatal CHD, non-fatal MI		(26.3) vs (8.8) g/d	0.91 (0.77, 1.09)	0.18	
(Rimm et al., 1996) HPFS	(740) /51529	6 y	Cereal fibre (AOAC)	MI [fatal + non-fatal]		Continuous risk estimate 10g/d	#0.71 (0.55, 0.91)		Age, alcohol, BMI, saturated fatty acid intake, familial MI, hyper-cholesterolaemia, occupation, physical activity, hypertension, smoking, vitamin E
(Streppel et al., 2008) Zutphen Elderly Study	(348) /1373	40 y (0.2)	Cereal fibre (Energy adjusted, fibre contained within bread and other cereal products - recent intake)	CHD events [fatal]		Continuous risk estimate 10 g/d	#0.84 (0.64, 1.1)		Trans fatty acid intake, alcohol, BMI, smoking, Cis-PUFA, energy intake, fish, prescribed diet, saturated fatty acid, SES/class
			Cereal fibre (Intake in	CHD events [fatal]		Continuous risk	0.86 (0.64, 1.15)		

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Sub-group Detail	Contrast (mean)	RR (CI)/ Mean Exposure (SD)	p trend	Adjustments
			middle age)			estimate 10 g/d			
*(Ward et al., 2012) EPIC-Norfolk	(1294) / 4347	11y	Cereal fibre (Englyst) assessed with 7-day food diaries	Fatal and non + fatal CHD events	Men	Risk per 2g/day increase	0.97 (0.93, 1.01)	0.092	Age, BMI, physical activity, smoking, family history of MI, class, diabetes, antihypertensive medication, lipid-lowering medication, aspirin use, total fat energy, energy from non-fat, saturated fat intake, alcohol, plasma ascorbic acid
	(712)/ 2728				Women		0.97 (0.90, 1.04)	0.37	
(Wolk et al., 1999) Nurses' Health Study	(591) /121700	10 y (20)	Cereal fibre (Long-term intake over 6 years. AOAC)	Non-fatal MI, fatal CHD		Continuous risk estimate 5 g/d	0.63 (0.49, 0.81)	<0.001	Age, alcohol, aspirin, beta-carotene, BMI, carbohydrate intake, saturated fatty acid intake, energy intake, folate, hypertension, magnesium intake, menopausal status, parental MI, period of exposure, physical activity, smoking, post-menopausal HRT, vitamin B6 and C
	(289)				Age <60		0.63 (0.44, 0.90)	0.01	
	(302)				Age >60		0.76 (0.57, 0.99)	0.05	
	(319)				Never or former smoker		0.59 (0.43, 0.79)	<0.001	
	(272)				Smokers		0.87 (0.63, 1.2)	0.39	
	(249)				BMI <25		0.58 (0.4, 0.82)	0.003	
	(278)				BMI >25		0.85 (0.62, 1.17)	0.31	
	(177)				Lowest tertile of SFA		0.62 (0.44, 0.88)	0.007	
	(194)				Middle tertile of SFA		0.79 (0.54, 1.15)	0.21	
	(220)				Highest tertile of SFA		0.68 (0.43, 1.07)	0.1	
	(189)				Lowest tertile of TFA		0.69 (0.49, 0.97)	0.03	
	(211)				Middle tertile of TFA		0.77 (0.53, 1.12)	0.18	
	(191)				Highest tertile of TFA		0.57 (0.35, 0.92)	0.02	

*Identified during update search; # Result was used in the cereal fibre and CHD meta-analysis

Table IV.iv: Results from cohort studies identified in the systematic review: Fruit fibre and CVD, CHD and stroke events

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Subgroup Detail	Contrast (mean)	RR (CI) / Mean Exposure (SD)	p	p trend	Adjustments
*(Buyken et al., 2010) Blue Mountains Eye Study	(109)/ 1490	13y	Fruit fibre (AOAC)	Total CVD [fatal]	Women	11.7 vs 2.8 g/d	1.03 (0.61, 1.75)	0.84		Age, energy, GI residuals, alcohol, smoking, diabetes
	(151)/ 1245				Men	11.1 vs 2.4 g/d	0.61 (0.38, 0.99)	0.05		Age, energy, GI residuals, total fat, underweight, smoking, use of corticosteroids
*(Crowe et al., 2012) EPIC-Heart	2381/ 306331	11.5y	Fruit fibre assessed as AOAC in most countries and values were then calibrated across Europe	IHD mortality		Per 2.5g/d	#0.94 (0.88, 1.01)		0.090	Age, alcohol, BMI, physical activity, marital status, education, employment, hypertension, hyperlipidaemia, angina pectoris, diabetes, PUFA:SFA, energy intake, fibre from cereals, vegetables and other sources
*(Eshak et al., 2010) Japanese Collaborative cohort study	(231)/ 110792	13.4y	Fruit fibre (similar to AOAC)	CHD [fatal]	Men	>1.7 vs. <0.4 g/d	#0.55(0.32, 0.96)	0.032		Age, BMI, hypertension, DMT2, alcohol, smoking, education, exercise, walking, stress, sleep, SFA, n3 fatty acid, Sodium intake, folate, Vitamin E, vegetable fibre, cereal fibre
	(191)				Women	>2.2 vs. <0.7 g/d	#0.42 (0.33, 0.81)	0.014		
(Larsson et al., 2009) ATBC Study	(196) /29133	13.6 years	Fruit fibre (Englyst)	Stroke, haemorrhage, Subarachnoid [fatal + non-fatal]		(6.2) vs (0.7) g/d	1.28 (0.8, 2.06)		0.14	Age, Alcohol, BMI, total cholesterol intake, Blood pressure, energy intake, Folate, HDL-C, CHD, diabetes, physical activity, magnesium Intake, Smoking, Group allocation
	(383)			Stroke, haemorrhage, Intracerebral [fatal + non-fatal]			0.88 (0.61, 1.26)		0.44	
	(2702)			Stroke, ischaemic [fatal + non-fatal]			0.91 (0.8, 1.04)		0.83	
(Liu et al., 2002a) The Women's Health Study	(177) /39876	6 y	Fruit fibre (AOAC)	MI [fatal + non-fatal]		(6) vs (2.5) g/d	#1.11 (0.62, 1.96)		0.63	Age, Alcohol, BMI, energy intake, family history of MI, fat intake, Folate, Diabetes Hypercholest-erolaemia, hypertension, protein intake, physical activity, Smoking, Group allocation, Supplements, Postmenopausal HRT
	(570)	6 y		Total CVD [fatal + non-fatal]		(6) vs (2.5) g/d	0.82 (0.61, 1.09)		0.09	

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Subgroup Detail	Contrast (mean)	RR (CI) / Mean Exposure (SD)	p	p trend	Adjustments
(Mozaffarian et al., 2003) Cardiovascular Health Study	(811) /5201	8.6 y	Fruit fibre (AOAC)	CHD events[fatal + non-fatal]		>7.5 vs <2.8 g/d	#0.99 (0.78, 1.25)		0.98	Age, Alcohol, Cereal Fibre, Education, Fibre from Vegetables, diabetes, physical activity, sex, Smoking
	(204)				Age 65-69y	80^{th} vs 20^{th} Centile (C)	0.92 (0.75, 1.13)			
	(255)				Age 70-74y	80^{th} vs 20^{th} C	1.04 (0.89, 1.23)			
	(352)				Age >75y	80^{th} vs 20^{th} C	1 (0.85, 1.19)			
	(434)				Women	80^{th} vs 20^{th} C	1.04 (0.87, 1.24)			
	(377)				Men	80^{th} vs 20^{th} C	0.94 (0.77, 1.16)			
	(575)				No diabetics	80^{th} vs 20^{th} C	0.99 (0.85, 1.16)			
(Oh et al., 2005) Nurses' Health Study	(1020) /121700	18	Fruit fibre (AOAC)	Total stroke [fatal + non-fatal]		(7.3) vs (1.3) g/d	0.87 (0.7, 1.09)		0.28	Age, Alcohol, Aspirin, BMI, carbohydrate intake, energy intake, diabetes, family history of hypertriglycerideaemia/ hypertension/ MI, Menopause Status, physical activity, Smoking, Postmenopausal HRT, Vitamin intake
	(279)			Stroke, haemorrhagic [fatal + non-fatal]			0.86 (0.57, 1.29)		0.64	
	(515)			Stroke, ischaemic [fatal + non-fatal]			0.87 (0.63, 1.21)		0.22	
*(Park et al., 2011) NIH-AARP Diet and Health Study	5248 / 388122	9y	Fibre from fruits (AOAC)	Total CVD [fatal]	Male	Q5 vs. Q1	1.03 (0.93, 1.13)			Age, race, education, marital status, health status, BMI, physical activity, smoking, alcohol, red meat, fruit, vegetables, total energy intake
	2417 / 388122				Female	Q5 vs. Q1	1.06 (0.93, 1.22)			
(Pietinen et al., 1996) ATBC Study	(635) /29133	6.1 y	Fruit fibre (Englyst)	CHD events [fatal]		(5.3) vs (0.7) g/d	#1.16 (0.8, 1.67)		0.77	Age, Alcohol, Beta-carotene, BMI, blood pressure, education, saturated fat intake, energy intake, physical activity, Smoking, Group allocation, Vitamin C, Vitamin E
	(1399)			Fatal CHD, non-fatal MI		(5.3) vs (0.7) g/d	0.99 (0.78, 1.27)		0.57	
(Rimm et al., 1996) HPFS	(740) /51529	6 y	Fruit fibre (AOAC)	MI [fatal + non-fatal]		Continuous risk estimate 10g/d	#0.79 (0.6, 1.05)			Age, Alcohol, BMI, saturated fat intake, family history of MI, Smoking, Vitamin E Hypercholesterolaemia, Occupation, physical activity, hypertension

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Subgroup Detail	Contrast (mean)	RR (CI) / Mean Exposure (SD)	p	p trend	Adjustments
(Streppel et al., 2008) Zutphen Elderly Study	(348)/ 1373		Fruit fibre (Energy adjusted - *recent* intake, AOAC)	CHD events [fatal]		Continuous risk estimate 10g/d	#1.13 (0.75, 1.7)			Trans fatty acid, Cis-PUFA and saturated fat intake, Alcohol, BMI, Smoking, energy intake, Fish intake, Prescribed diet, , SES/Class
	(348)/ 1373	40y	*long term* intake in middle age				1.01 (0.43, 2.36)			
*(Ward et al., 2012) EPIC-Norfolk	(1294) / 4347	11y	Fruit fibre (Englyst) assessed with 7-day food diaries	Fatal and non + fatal CHD events	Men	Risk per 2g/day increase	0.93 (0.85, 1.02)		0.10	Age, BMI, physical activity, smoking, family history of MI, class, diabetes, antihypertensive medication, lipid-lowering medication, aspirin use, total fat energy, energy from non-fat, saturated fat intake, alcohol, plasma ascorbic acid
	(712)/ 2728				Women		0.90 (0.80, 1.01)		0.063	
(Wolk et al., 1999) Nurses' Health Study	(591) / 121700	10 y (20)	Fruit fibre (Long-term intake over 6 years. AOAC)	Non-fatal MI, fatal CHD		Continuous risk estimate 5g/d	#0.93 (0.74, 1.16)	0.51		Age, Alcohol, Aspirin, Beta-carotene, BMI, carbohydrate intake, saturated fatty acid intake, energy intake, Fibre, Folate, hypertension, Magnesium Intake, Menopause Status, Parental MI, Period of exposure, physical activity, Smoking, Postmenopausal HRT, Vitamin B6 intake, Vitamin C

*Identified during update search; # Result was used in the fruit fibre and CHD meta-analysis

Table IV.v: Results from cohort studies identified in the systematic review: Fibre contained within vegetables and CVD, CHD and stroke events

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Subgroup Detail	Contrast (mean)	RR (CI) /Mean Exposure (SD)	p	p trend	Adjustments
*(Buyken et al., 2010) Blue Mountains Eye Study	109/ 1490	13y	Vegetable fibre (AOAC)	Total CVD [fatal]	Women	13.5 vs 6.5 g/d	1.01 (0.64, 1.60)	0.88		Age, energy, GI residuals, alcohol, smoking, diabetes
	151/ 1245				Men	13.7 vs 6.0 g/d	0.88 (0.60, 1.30)	0.6		Age, energy, GI residuals, total fat, underweight, smoking, use of corticosteroids
*(Crowe et al., 2012) EPIC-Heart	2381/ 306331	11.5y	Vegetable fibre assessed as AOAC in most countries and values were then calibrated across Europe	IHD mortality		Per 2.5g/d	#0.90 (0.76, 1.07)		0.255	Age, alcohol, BMI, physical activity, marital status, education, employment, hypertension, hyperlipidaemia, angina pectoris, diabetes, PUFA:SFA, energy intake, fibre from fruit, cereals and other sources
*(Eshak et al., 2010) Japanese Collaborative cohort study	(231)/11 0792	13.4y	Vegetable fibre (similar to AOAC)	CHD [fatal]	Men	>4.5 vs. <2.8 g/d	#0.90(0.54, 1.51)	0.666		Age, BMI, hypertension, diabetes, alcohol, smoking, education, exercise, walking, stress, sleep, saturated fat and n3 fatty acid intake, sodium intake, folate, Vitamin E, cereal fibre, fruit fibre
	(191)				Women	>5.6 vs. <3.1 g/d	#0.97 (0.58, 1.62)	0.917		
(Larsson et al., 2009) ATBC Study	(196) /29133	13.6 years	Vegetable fibre (Englyst)	Stroke, haemorrhage Subarachnoid [fatal + non-fatal]		(7.1) vs (2.9) g/d	0.63 (0.43, 1.07)		0.06	Age, Alcohol, BMI, total cholesterol intake, blood pressure, energy intake, Folate, HDL-C, CHD, diabetes, physical activity, magnesium Intake, Smoking, Group allocation
	(383)			Stroke, haemorrhage Intracerebral [fatal + non-fatal]			0.81 (0.57, 1.14)		0.62	
	(2702)			Stroke, ischaemic [fatal + non-fatal]			0.86 (0.76, 0.99)		0.001	
(Liu et al., 2002a) The Women's Health Study	(177) /39876	6 y	Vegetable fibre (AOAC)	MI [fatal + non-fatal]		(8) vs (5.9) g/d	#0.89 (0.52, 1.53)		0.87	Age, Alcohol, BMI, energy intake, family history of MI, fat intake, Folate, diabetes, Hypercholesterolaemia, hypertension, physical activity, protein intake, Smoking, Group allocation, Supplements, Postmenopausal HRT
	(570)	6 y		Total CVD [fatal + non-fatal]		(8) vs (5.9) g/d	0.96 (0.72, 1.28)		0.78	

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Subgroup Detail	Contrast (mean)	RR (CI) /Mean Exposure (SD)	p	p trend	Adjustments
(Mozaffarian et al., 2003) Cardiovascular Health Study	(811) /5201	8.6 y	Vegetable fibre (AOAC)	CHD events [fatal + non-fatal]		>9.2 vs <4.2 g/d	#1.08 (0.86, 1.36)		0.95	Age, Alcohol, Cereal Fibre, Education, Fibre from fruit, diabetes, Smoking, physical activity, Sex
	(204)				Age 65-69y	80th vs 20th Centile	0.9 (0.76, 1.04)			
	(255)				Age 70-74y	80th vs 20th Centile	1.04 (0.9, 1.19)			
	(352)				Age >75y	80th vs 20th Centile	1.09 (0.95, 1.25)			
	(434)				Women	80th vs 20th Centile	1.08 (0.92, 1.27)			
	(377)				Men	80th vs 20th Centile	0.96 (0.82, 1.14)			
	(575)				No diabetics	80th vs 20th Centile	1.05 (0.92, 1.2)			
(Oh et al., 2005) Nurses' Health Study	(1020) /121700	18	Vegetable fibre (AOAC)	Total stroke [fatal + non-fatal]		(8.5) vs (2.9) g/d	0.92 (0.74, 1.14)		0.14	Age, Alcohol, Aspirin, BMI, carbohydrate intake, energy intake, family history of diabetes/ hypertriglycerideaemia/ hypertension/MI, Menopause, physical activity, Smoking, Postmenopausal HRT, Vitamin intake
	(279)			Haemorrhagic stroke [fatal + non-fatal]			0.76 (0.51, 1.13)		0.18	
	(515)			Ischaemic stroke [fatal + non-fatal]			1.01 (0.74, 1.38)		0.48	
*(Park et al., 2011) NIH-AARP Diet and Health Study	5248 /388122	9y	Fibre from vegetables (AOAC)	Total CVD [fatal]	Male	Q5 vs. Q1	0.96 (0.88, 1.05)			Age, race, education, marital status, health status, BMI, physical activity, smoking, alcohol, red meat, fruit, vegetables, total energy
	2417 /388122				Female	Q5 vs. Q1	0.96 (0.84, 1.10)			
(Pietinen et al., 1996) ATBC Study	(635) /29133	6.1 y	Vegetable fibre (based on Englyst)	CHD events [fatal]		(7.1) vs (2.9) g/d	#0.88 (0.66, 1.19)		0.08	Age, Alcohol, Beta-carotene, BMI, education, saturated-fatty acid intake, energy intake, physical activity, Smoking, Group allocation, blood pressure, Vitamin C, Vitamin E
	(1399)	6.1 y		Fatal CHD, non-fatal MI		(7.1) vs (2.9) g/d	0.94 (0.77, 1.14)		0.15	
(Rimm et al., 1996) HPFS	(740) /51529	6 y	Vegetable fibre (AOAC)	MI [fatal + non-fatal]		Continuous risk estimate 10 g/d	#0.78 (0.61, 1.0)			Age, Alcohol, BMI, saturated fatty-acid intake, mahily history of MI, Smoking, Vitamin E Hypercholesterolaemia, Occupation, physical activity, hypertension
(Streppel et al., 2008) Zutphen Elderly Study	(348) /1373	40 y (0.2)	Vegetable fibre (Energy adjusted **recent** intake, AOAC)	CHD events [fatal]		Continuous risk estimate 10 g/d	#0.88 (0.48, 1.65)			Trans-fat intake, Alcohol, BMI, Smoking, Cis-PUFA intake, energy intake, Fish, Prescribed diet, Saturated fat intake, SES/Class
	(348)	40 y (0.2)	Vegetable fibre (Energy adjusted **long term** intake in middle age)	CHD events [fatal]		Continuous risk estimate 10 g/d	1 (0.36, 2.77)			

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Subgroup Detail	Contrast (mean)	RR (CI) /Mean Exposure (SD)	p	p trend	Adjustments
*(Ward et al., 2012) EPIC-Norfolk	(1294) / 4347	11y	Vegetable fibre (Englyst) assessed with 7-day food diaries	Fatal and non + fatal CHD events	Men	Risk per 2g/day increase	0.97 (0.88, 1.05)		0.43	Age, BMI, physical activity, smoking, family history of MI, class, diabetes, antihypertensive medication, lipid-lowering medication, aspirin use, total fat energy, energy from non-fat, saturated fat intake, alcohol, plasma ascorbic acid
	(712)/ 2728				Women		0.96 (0.84, 1.09)		0.50	
(Wolk et al., 1999) Nurses' Health Study	(591) / 121700	10 y (20)	Vegetable fibre (Long-term intake over 6 years. AOAC)	Non-fatal MI, fatal CHD		Continuous risk estimate 5g/d	#1.06 (0.84, 1.32)	0.63		Age, Alcohol, Aspirin, Beta-carotene, BMI, carbohydrate intake, saturated fat intake, energy intake, Fibre, Folate, hypertension, magnesium intake, Menopause Status, Parental MI, Period of exposure, physical activity, Smoking, Postmenopausal HRT, Vitamin B6 intake, Vitamin C

*Identified during update search; # Result was used in the vegetable fibre and CHD meta-analysis

Table IV.vi: Results from cohort studies identified in the systematic review: Fibre from other sources or types and CVD and CHD events

Result ID/ Reference/ Cohort Name	(Cases)/ Total	Follow Up (% loss)	Exposure	Outcome/ Assessment Details	Sub-group details	Contrast (mean)	RR (CI)	p-trend	Adjustments
*(Crowe et al., 2012) EPIC-Heart	2381/ 306331	11.5y	'Other fibre' (Non cereal, fruit or vegetable based fibre. Assessed as AOAC in most countries and values were then calibrated across Europe)	IHD mortality		Per 5g/d	1.01 (0.87, 1.17)	0.890	Age, alcohol, BMI, physical activity, marital status, education, employment, hypertension, hyperlipidaemia, angina pectoris, diabetes, PUFA:SFA, energy intake, fibre from fruit, vegetables and cereals
*(Park et al., 2011) NIH-AARP Diet and Health Study	5248 /388122	9y	Fibre from beans (AOAC)	Total CVD [fatal]	Men	Q5 vs. Q1	0.93 (0.85, 1.01)		Age, race, education, marital status, health status, BMI, physical activity, smoking, alcohol, red meat, fruit, vegetables, total energy
	2417 /388122				Women	Q5 vs. Q1	0.83 (0.74, 0.95)	<0.05	
(Pietinen et al., 1996) ATBC Study	(635) /29133	6.1 y	Cellulose	CHD events [fatal]		(6.3) vs (3.1) g/d	0.72 (0.54, 0.97)	0.006	Age, alcohol, beta-carotene, BMI, diastolic and systolic blood pressure, education, saturated fatty acid, energy intake, physical activity, smoking, group allocation, vitamin C and E
	(1399) /29133		Cellulose	Fatal CHD, non-fatal MI		(6.3) vs (3.1) g/d	0.90 (0.75, 1.10)	0.07	
	(1399) /29133		Lignin	Fatal CHD, non-fatal MI		(5.8) vs (2.1) g/d	0.89 (0.75, 1.06)	0.21	
	(635) /29133		Lignin	CHD events [fatal]		(5.8) vs (2.1) g/d	0.75 (0.58, 0.97)	0.002	
(Streppel et al., 2008) Zutphen Elderly Study	(348) /1373	40 y (0.2)	Legume fibre (Energy adjusted *long term* intake in middle age, AOAC)	CHD events [fatal]		Continuous risk estimate 10 g/d	0.52 (0.25, 1.09)		TFA, Alcohol, BMI, Cis-PUFA, energy intake, Fish, Smoking, Prescribed diet, saturated fat, SES/Class
			Legume fibre (Energy adjusted *recent* intake)	CHD events [fatal]			0.64 (0.34, 1.2)		
			Potato fibre (Energy adjusted *recent* intake)	CHD events[fatal]			0.71 (0.48, 1.06)		
			Potato fibre (Energy adjusted *long term* intake in middle age)	CHD events[fatal]			0.94 (0.62, 1.45)		

Appendix V: FFQ items contributing to sources of fibre

Items from the baseline FFQ which have been chosen to contribute to estimates of fibre from a range of food sources. Separate FFQ items are indicated with '|' (some lines in the FFQ listed multiple foods e.g. Currants, raisins and sultanas).

Fibre from fruit includes the following FFQ items:

Apples | Avocado| Banana| Grapes| Kiwi| Mango| Citrus| Papaya| Pears| Pineapple| Apricots| Melon| Nectarines| Peaches| Plums| Raspberries| Currants red and white| Rhubarb| Strawberries| Dates| Figs| Prunes| Dried fruit| Currants, raisins and sultanas.

Fibre from nuts & seeds includes the following FFQ items:

Peanuts and pistachio| Cashews and almonds| Pecans and walnuts| Sunflower and sesame seeds.

Fibre from total cereal foods includes the following FFQ items:

White bread & rolls| Brown bread & rolls| Wholemeal bread & rolls| Chapattis, Nan, Paratha| Papadums| Tortillas| Pitta bread| Crispbread e.g. Ryvita| Cream crackers, cheese biscuits| Porridge, readybrek| Sugar coated cereals e.g. sugar puffs| Non-sugar coated cereals e.g. cornflakes, rice krispies| Muesli| All bran, bran flakes| Weetabix, shredded wheat| White pasta e.g. spaghetti, green pasta, red pasta, noodles| Wholemeal pasta, brown spaghetti| White rice| Brown rice| Wild rice| Barley| Bulgar wheat| Wheat germ| Cous-cous| Cereal bars & flapjack| Plain biscuits e.g. marie, nice, digestive| Chocolate biscuits| Sandwich or cream biscuit| Fruitcake| Sponge cake| Buns, pastries e.g. croissants doughnuts, tray bakes| Scones, pancakes, muffins, crumpets| Fruit pies, tarts crumbles| Sponge puddings.

Fibre from breakfast cereals includes the following FFQ items:

Porridge, readybrek| Sugar coated cereals e.g. sugar puffs| Non-sugar coated cereals e.g. cornflakes, rice krispies| Muesli| All bran, bran flakes| Weetabix, shredded wheat.

Fibre from vegetables includes the following FFQ items:

Bean sprouts| Beetroot| Broccoli, spring greens, kale| Brussels sprouts| Cabbage| Cauliflower| Celery| Coleslaw (low calorie coleslaw)| Courgettes, marrow, squash| Cucumber| Garlic| Green beans, runner beans| Leeks, lettuce, mushrooms| Aubergine, okra| Parsnips| Peas, mushy peas, mange-tout| Peppers| Swede| Sweet corn| Tomatoes (raw, canned, sauce)| Turnip| Watercress, mustard & cress.

Fibre from pulses includes the following FFQ items:

Lentils and dals| Chick peas, chanas| Hummus| Baked beans| Mung beans & red kidney beans| Black eyed beans| Butter beans and broad beans.

Appendix VI: Research Ethics Committee approval

NHS
Health Research Authority

NRES Committee Yorkshire & The Humber - Leeds East

Yorkshire and Humber REC Office
First Floor, Millside
Mill Pond Lane
Meanwood
Leeds
LS6 4RA

Tel: 0113 3050108

Professor Janet Cade
Nutritional Epidemiology Group
School of Food and Science Nutrition
University of Leeds
LS2 9JT

1 December 2011

Dear Prof Cade

Study title: **The UK Women's Diet and Lifestyle Cohort Study**

The above amendment was reviewed on 1 December 2011 by the Sub-Committee in correspondence.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Notice of Substantial Amendment (non-CTIMPS)	1	17 November 2011
Original Protocol sent to ethics committees	1	2 March 1983
Example consent from one committee	1	14 August 1995
NIGB Section 251 approval letter	1	4 August 2011
Copy of participant Invitation Letter (Baseline)	1	1995
Copy of participant Invitation Letter (Follow-up)	1	1999
Participant Details Form	1	1995

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

Yours sincerely

Alan Ebbutt
Vice Chair
E-mail: jade.thorpe@nhs.net

Enclosures: List of names and professions of members who took part in the review

Copy to: *Rachel de Souza*
Diane Threapleton

NRES Committee Yorkshire & The Humber - Leeds East

Attendance at Sub-Committee of the REC meeting on 01 December 2011

Committee Members:

Name	*Profession*	*Present*	*Notes*
Mrs Victoria Ajayi	Assistant Contracts Manager	Yes	
Prof Alan Ebbutt	Statistician	Yes	

Appendix VII: National Research Ethics Committee Section 251 approval

NIGB

Ethics and Confidentiality Committee

NHS Connecting for Health,
Floor 7,
New Kings Beam House,
22 Upper Ground,
London,
SE1 9BW.

Tel: (020) 7633 7052
Email: eccapplications@nhs.net

Professor Janet Cade
University of Leeds
Nutritional Epidemiology Group
School of Food Science and Nutrition
LS2 9JT

04 August 2011

Dear Professor Cade

ECC 6-05(e)/2011 Dietary fibre and cardiovascular disease in the UK Women's Cohort Study

Thank you for your application for support under section 251 of the NHS Act 2006 and the Health Service (Control of Patient Information) Regulations 2002 ('section 251 support') to process patient identifiable information without consent. This application was considered at the Ethics and Confidentiality Committee meeting on the 28 July.

Context

This application from the University of Leeds detailed a study to explore dietary fibre intake and fibre from different food sources in relation to the risk of cardiovascular disease (CVD). Section 251 support was requested to allow the disclosure of pseudonymised Hospital Episode Statistics (HES) data for patients with specified ICD10 codes. Initial consent for the study was obtained in 1995 and specific reference had not been made to either the HES database or the NHS Information Centre as a source of data. This application and consent form had previously been considered by the NHS Information Centre's Data Access Advisory Group who advised that the consent form was not sufficient to allow the release of HES data and that section 251 support would be required for the disclosure of data.

Outcome

Consent form

The Committee noted that the initial consent for this study had been obtained in 1995 and at the time it was not a requirement to complete consent forms. The Committee agreed that they were sympathetic to these types of historical studies and understood requests for section 251 to enable a secure legal basis for the disclosure of this data.

Practicability of gaining further explicit consent

Members discussed whether further explicit consent could be obtained by the applicant and agreed that given the low level of sensitivity of the data requested, the original consent taken and the historical nature of the cohort, it would be disproportionate to require the applicant to re-consent all participants. However members agreed that if, in future, there was a requirement to approach participants again consideration should be given to taking consent for

access to national databases. At this time the NHS Information Centre should be approached to determine suitable wording.

In considering the questionnaire which contained the consent provisions, whilst views were raised that this could be considered as adequate consent for the disclosure of HES data, as a whole the Committee understood the hesitation to allow the data to be released without section 251

Conclusion

As re-consenting the cohort was deemed unfeasible in the circumstances, and given that all participants had initially consented to the original study which had detailed follow up of the cohort using medical records, the Committee agreed that a recommendation of support could be provisionally given. This recommendation was subject to the following specific and standard conditions of approval:

Specific conditions of approval

1. The provision of a REC favourable opinion to access HES data or confirmation that the original REC approval covers this amendment.

2. Confirmation of satisfactory security arrangements (approved 12/07/2011)

Following confirmation of a favourable REC opinion final approval under section 251 can be issued; our register of approved applications on our website will be updated shortly to reflect this approval.

Please do not hesitate to contact me if you have any queries following this letter and I would be grateful if you could quote the above reference number in all future correspondence.

Yours sincerely

Claire Edgeworth
Deputy Approvals Manager

Appendix VIII: Definition for final diagnosis field in the MINAP dataset

Final diagnosis	Definition and notes
ST segment elevation MI	There will normally be a history consistent with the diagnosis. The diagnosis requires the presence of cardiographic changes of ST elevation consistent with infarction of =>2mm in contiguous chest leads and/or ST elevation of =>1 mm ST elevation in 2 or more standard leads. (New LBBB is included; although ST elevation is usually apparent in the presence of LBBB). There must be enzyme or troponin elevation. Where CK is used the peak value should exceed twice the upper limit of the reference range. Where troponin assay is used the locally accepted cut off value should be used. (See Threatened MI) This group includes all patients with STEMI regardless of whether typical changes were evident on the admission ECG or developed subsequently.
Threatened infarction	After early reperfusion treatment there may be rapid resolution of existing ST elevation associated with a CK rise less than twice the upper limit of normal or a small troponin release. If only troponin has been measured and is elevated; it is a local decision whether this is recorded as 'Definite infarction' or 'Threatened infarction'.
ACS troponin +ve	ACS troponin positive includes all those patients previously defined as nSTEMI. There must be symptoms consistent with cardiac ischaemia and there will normally be cardiographic changes consistent with this diagnosis. Troponon elevation above locally determined reference level is mandatory.
ACS troponin –ve	Use where there are symptoms consistent with cardiac ischaemia without troponin release. There must be dynamic ECG changes consistent with fluctuating ischaemia. Synonym unstable angina.
Chest pain cause uncertain	Use in any patient admitted with chest pain not accompanied by significant cardiographic change, without any enzyme / troponin release, and where no other clear diagnosis emerges. It is likely that at admission there was a high index of clinical suspicion that the pain was cardiac, but this remains unconfirmed.
Other diagnosis	Use where a patient is admitted with clinical suspicion of cardiac pain and where any diagnosis other than cardiac ischaemia is confirmed.
ACS troponin not stated (unconfirmed MI)	This diagnosis must only be applied to patients who die in hospital before biochemical confirmation of infarction can be confirmed.

Source: (MINAP, 2010)