The association between dietary fibre intakes and incidence of type 2 diabetes mellitus

Maryam A. O. D. Aldwairji

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Food Science and Nutrition

December 2013

Intellectual property and Publication Statements

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others. Part of chapter 6 is based on work from jointly authored publications as follow:

- Dietary fibre intake and risk of fatal coronary heart disease in a cohort of British women. Threapleton, D. E., Burley, V. J., Greenwood, D. C., Aldwairji, M., & Cade, J. E. (2012). *Proceedings of the Nutrition Society, 71* (OCE3), E224.
- Dietary fibre and cardiovascular disease mortality in the UK Women's Cohort Study. Threapleton, D. E., Greenwood, D. C., Burley, V. J., Aldwairji, M., & Cade, J. E. (2013). *European Journal of Epidemiology,* 28, 335-346.

In both publications, AOAC-fibre intake of the UKWCS obtained from added AOAC-fibre values in the baseline FFQ dataset that is directly attributed to my work and covered in chapter 6 and 7. Diane Threapleton was the paper lead author. Victoria Burley, Diane Threapleton, Darren Greenwood, Janet Cade conceived the study idea. Statistical analysis was undertaken by Diane Threapleton. Diane Threapleton wrote the manuscript. All authors reviewed the manuscript and contributed to manuscript revisions.

Acknowledgments

Thank you God for giving me the strength to complete my thesis, despite the hardships I encountered, still I manage to complete my thesis. It would not have been possible to write this doctoral thesis without the help and support of the people around me. So this is the best time to say thank you to all people were around me during my PhD journey.

Special thanks to my wonderful girl, Loulwa for her love and patient and for giving me the courage to make it. She is my best friend, my lovely daughter and everything in this life, thank you for sharing all challenges with me.

I would like to thank Dr Victoria Burley and Dr Caroline Orfila for their enthusiasm, their encouragement, and their guidance in all the time of research and writing of this thesis.

I have been fortunate to be blessed with loving and supporting family

- Thanks to my parents (Loulwa and Abdullah)
- Thanks to my sisters (Ghadeer, Ashwaq and Raba'a) and my brother (Othman)

I would also like to thank all of my friends for encouraging me to continue on. Wafa, thanks for listening and letting me get my anxieties out. Fatima, Khulood and Khawla, thanks for being my family in UK.

I would like to thanks my colleagues from the Nutrition Epidemiology group in University of Leeds, Diane Threapleton and Neil Hancock for their help in managing the dataset.

I would also like to thanks the principle investigators of the UKWCS for allowing me to use the dataset for my thesis and also many thanks to all women who participate in the UKWCS.

Last, but by no means least, I would like to convey thanks to the Kuwait Government particularly, Ministry of Health for providing the financial means.

Abstract

The incidence of diabetes is increasing alarmingly indicating a need for preventive strategies. Prospective evidence is inconclusive regarding the protective effect of high fibre intake on the risk of Type 2 diabetes mellitus (T2DM). Differences in fibre measuring methodologies may have contributed some of these inconsistencies. For the first time, this thesis employed both laboratory and epidemiological approaches to explore different aspects of dietary fibre.

Firstly, 14 commonly consumed legumes were analysed for fibre content using the Association of Official Analytical Chemists (AOAC) method, and the values were significantly higher than published non-starch polysaccharide (NSP) values, with a mean AOAC-fibre: NSP ratio of 1:1.43. The UK Women's Cohort Study (UKWCS) data was used to compare AOAC-fibre intakes with NSP intakes. Good agreement (Kappa = 0.9) was observed between intakes of AOAC-fibre and NSP with a resulting AOAC-fibre: NSP ratio being generated of 1:1.43.

Following this, the links between risk of incident T2DM and intakes of AOAC-fibre and NSP were investigated. There was no evidence of dose-response relationships between T2DM risk and total fibre intake (AOAC-fibre and NSP) or from key fibre sources, except an age-adjusted lower risk of T2DM with every 5g/day increment in cereal fibre (OR=0.86; 95%CI: 0.75, 0.99).

As an important fibre provider; the relationship between the risk of T2DM and legume consumption was also explored. Women in the highest dried legumes intake category experienced significantly lower odds of incident T2DM (OR=0.85; 95%:0.52, 0.89, p =0.03) compared to women in the lowest category. Agreement of fibre intake obtained using different dietary assessments approaches suggested fair agreement of diary derived NSP vs. FFQ derived NSP and poor agreement for vegetable fibre.

Overall, results do not support an increase in fibre intake to prevent diabetes in the studied population, although there may be a benefit of increased dried legume intake.

List of publications and presentations

Publications

1. Aldwairji, M., Burley, V., & Orfila, C. (2011). Are current fibre recommendations using NSP adequate? A comparison of Englyst and Association of Official Analytical Chemists fibre values. *Proceedings of the Nutrition Society, 70* (OCE4).This is covered in chapter 5.
2. Aldwairji, M., Orfila, C., & Burley, V. J. (2012). Degree of agreement between AOAC-based dietary fibre intake and Englyst-based dietary fibre intake in the UK Women's Cohort Study (UKWCS). *Proceedings of the Nutrition Society, 71* (OCE2). This is covered in chapter 7.
3. Threapleton, D. E., Burley, V. J., Greenwood, D. C., Aldwairji, M., & Cade, J. E. (2012). Dietary fibre intake and risk of fatal coronary heart disease in a cohort of British women. *Proceedings of the Nutrition Society, 71* (OCE3), E224. My contribution in this abstract: obtained AOAC-fibre intake of the UKWCS by added AOAC-fibre values for food items listed in the baseline FFQ. This is covered in chapter 6.
4. Aldwairji., M, Orfila, C. Orfila & Burley, V.J. (2013). Dietary fibre intake and risk of type 2 diabetes in British women. *Proceedings of the Nutrition Society, 72* (OCE4), E204. This is covered in chapter 8.
5. Aldwairji., M, Orfila, C. Orfila & Burley, V.J. (2013). Legume intake and risk of type 2 diabetes in British women. *Proceedings of the Nutrition Society (2013), 72* (OCE4), E275. This is covered in Chapter 9.
6. 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, 335-346. My contribution in this paper: obtained AOAC-fibre intake of the UKWCS by added AOAC-fibre values for food items listed in the baseline FFQ. This data is covered in chapter 6).

Conference presentation

Towards a predictive NSP: TDF ratio for cooked legumes: a comparison between non-starch polysaccharides and total dietary fibre. M. Aldwairji, V. Burley, C. Orfila, School of Food Science and Nutrition, University of Leeds, UK. 5th International dietary fibre conference in Rome, Italy, 7-9 May 2012.

In review

Analysis of dietary fibre of boiled and canned legumes commonly consumed in the United Kingdom. Maryam A. Aldwairji, Victoria J. Burley and Caroline Orfila. This paper submitted to the Journal of Food Composition and Analysis. This is covered in chapter 5.

In preparation

Intake and sources of dietary fibre and non-starch polysaccharides in the UK Women's Cohort Study. Does method of dietary fibre analysis matter? Maryam A. Aldwairji, Caroline Orfila, Janet E. Cade and Victoria J. Burley.

List of abbreviations

AOAC	Association of Official Analytical Chemist
AOAC-fibre	Total dietary fibre measured by AOAC analytical method only
AACC	American Association of Cereal Chemists
CV	Coefficient of variation
DP	Degree of polymerization
EuroFIR	European food information resources
FFQ	Food frequency questionnaire
FSA	Food Standards Agency
FOS	Fructo-oligosaccharides
Hb_{A1C}	Glycosylated haemoglobin
HDL	High density lipoprotein
IDF	Insoluble dietary fibre
LDL	Low density lipoprotein
MAFF	Ministry of Agriculture, Fisheries and Food
DANTE	Diet and Nutrition Tool for Evaluation
NDNS	National Diet and Nutrition Survey
NDO	Non-digestible oligosaccharides
NSP	Non starch polysaccharides
OGTT	Oral glucose tolerance test
RCT	Randomized controlled trial
RS	Resistant starch
SACN	Scientific Advisory Committee of Nutrition
SDF	Soluble dietary fibre
TDF	Total dietary fibre measured by any analytical method
T2DM	Type 2 diabetes mellitus
UKWCS	UK Women's Cohort Study
USDA	National Nutrient database for Standards Reference in USA
WHO /FAO	World Health Organization/ Food and Agriculture Organization

Table of contents

List of Tables

List of Figures

Chapter 1: Aims and Objectives

The aim of this thesis is to examine the association between dietary fibre intake and incidence of type 2 diabetes mellitus (T2DM). This aim will be addressed through the following objectives:

Chapter 2:

- To extend knowledge about the definitions of dietary fibre, and discuss the main analytical methods.
- To provide background on the structure of dietary fibre and physiochemical properties of the main sources of dietary fibre and key physiological and health benefits. Also to describe dietary fibre intakes particularly in the UK, and dietary fibre recommendations across Europe as well as in other parts of the world.

Chapter 3:

- To provide background on diabetes definition, aetiology, case ascertainment in prospective studies and trend of incidence, prevalence in UK.
- To describes prospective evidence on dietary and lifestyle practices in relationship to T2DM.

Chapter 4:

- To carry out a comprehensive review using a systematic approach aimed to assess whether or not the prospective evidence supports the presence of a relation between total dietary fibre (TDF) consumption and T2DM incidence by comparing extreme intake groups.
- To explore the association between fibre sources and fibre types (soluble and insoluble) on the risk of the developing T2DM from cohort studies.

Chapter 5:

- To adapt AOAC methodology to measure TDF of legumes and to compare the measured dietary fibre values obtained by AOAC (AOAC-fibre) with the published NSP values.
- To assess the effect of cooking methods (boiling and canning) on the TDF content of commonly consumed legumes.
- To predict the relationship between AOAC-fibre and NSP by generating AOAC-fibre: NSP ratio for the legume group.

Chapter 6

- To describe the UKWCS in terms of its initiation and follow-up.
- To describe the methodology of adding AOAC-fibre to baseline FFQ of the UKWCS.

Chapter 7

Within UKWCS:

- To describe women in the highest AOAC-fibre quintile in the UKWCS by determining food, dietary and other lifestyle predictors.
- To identify the characteristics of women in the highest AOAC-fibre quintile.
- To compare between AOAC-fibre and NSP intake values using Kappa statistics.
- To determine whether subjects are classified consistently using both methods.

Chapter 8

- To examine the association between risk of T2DM and intakes of AOAC-fibre and NSP in the UKWCS using logistic regression method.
- To incorporate the findings from the UKWCS in the meta-analysis generated in chapter 4 for TDF intake and risk of T2DM.
- To assess, through meta-analysis, the magnitude of the relation between TDF intake and the risk of T2DM.
- To examine the association between risk of T2DM and fibre source and fibre types in the UKWCS.

Chapter 9

- To examine specifically whether high legume intake lowers the risk of T2DM among the UKWCS using logistic regression methods.
- To examine the relationship between types of legume consumed and the risk of T2DM in the UKWCS.

Chapter 10

- To determine the degree of agreement between NSP obtained from food diaries and NSP obtained from baseline FFQ in the UKWCS
- To determine the degree of agreement between NSP sources obtained from food diary and NSP obtained from baseline FFQ in the UKWCS.
- To assess the repeatability of dietary fibre intakes obtained from baseline FFQ and repeated FFQ in the UKWCS.

Chapter 2: Literature review of dietary fibre

2.1 Introduction

The aim of this chapter is to extend knowledge about the definitions of dietary fibre, and discuss the most commonly used analytical methods. The structure of dietary fibre and physiochemical properties of the main sources of dietary fibre as well as key physiological and health benefits will be discussed. Also, this chapter describes dietary fibre intakes particularly in the UK, and dietary fibre recommendations across Europe as well as in other parts of the world.

2.2 Dietary fibre definition

Dietary fibre is a term first used 58 years ago by Hipsley (1953) to describe non-digestible plant cell wall components. The term was later used by Trowell to mean *"skeletal remains of plant cells that are resistant to hydrolysis by the enzymes of man"*

(Trowell, 1972, p.926)

This definition was limited to the cellular walls of plants. A few years later, Trowell (1976) expanded upon the definition of dietary fibre to include all digestion-resistant polysaccharides (plant storage polysaccharides) such as gums, modified cellulose, mucilages, oligosaccharides, and pectin, in addition to cellulose, hemicelluloses, lignin, and associated minor substances such as waxes, cutin, and suberin. The on-going debate about the definition of dietary fibre and the dietary fibre analytical methods are challenges facing nutritional epidemiologists who evaluate the effect of dietary fibre intake on the risk of nutrition related chronic diseases (Deharveng *et al.,* 1999). The definition of dietary fibre has been updated several times. Current dietary fibre definitions can be broadly based on chemical structure or/and physiological function (Slavin, 2003a).

In recent years, the argument has mainly centred on physiological aspects of dietary fibre and inclusion of different dietary fibre components within the definition (Camire *et al.*, 2001). The main sources of scientific advice which define dietary fibre are the American Association of Cereal Chemists, the European Commission, Department of Health in the UK, the Food and Nutrition Board, and Codex Alimentarius Commission (Table 2.1).

Table2.1 Dietary fibre definitions from different organizations

Organization	dietary fibre definition
The American Association of Cereal Chemists (AACC) (Camire *et al.*, 2001, p.112)	Defines dietary fibre as *'the remnants of the edible part of plants or analogous carbohydrates that are resistant to digestion and absorption in the human small intestine, with complete or partial fermentation in the large intestine. Dietary fibre includes polysaccharides, oligosaccharides, lignin, and associated plant substances. Dietary fibres promote beneficial physiological effects, including laxation, and/or blood cholesterol attenuation, and/or blood glucose attenuation'*
The Commission of the European Communities *(European Commission, 2008, p.L285)*	stated *'"fibre" means carbohydrate polymers with three or more monomeric units, which are neither digested nor absorbed in the human small intestine and belong to the following categories: edible carbohydrate polymers naturally occurring in the food as consumed; edible carbohydrate polymers which have been obtained from food raw material by physical, enzymatic or chemical means and which have a beneficial physiological effect demonstrated by generally accepted scientific evidence; edible synthetic carbohydrate polymers which have a beneficial physiological effect demonstrated by generally accepted scientific evidence.'*
Department of Health in the UK(1991)	*"intrinsic plant cell wall polysaccharides"*
Food and Nutrition Board(Food and Nutrition Board, 2001, p.3)	*"1. Dietary Fiber consists of nondigestible carbohydrates and lignin that are intrinsic and intact in plants.* *2. Added Fiber consists of isolated, nondigestible carbohydrates that have beneficial physiological effects in humans.* *3. Total Fiber is the sum of Dietary Fiber and Added Fiber"*
WHO/FAO (Codex Alimentarius Commission, 2008, p.46)	*"Dietary fibre means carbohydrate polymers*[1] *with 10 or more monomeric units*[2]*, which are not hydrolysed by the endogenous enzymes in the small intestine of humans, and thus belong to the following categories: edible carbohydrate polymers naturally occurring in food as it is consumed; carbohydrate polymers obtained from raw food material by physical, enzymatic, or chemical means, and which have been shown to have a physiological effect of being beneficial to health as demonstrated by generally accepted scientific evidence from competent authorities; and synthetic carbohydrate polymers that have been shown to have a physiological effect of being beneficial to health as demonstrated by generally accepted scientific evidence from competent authorities.*

[1] When derived from a plant origin, dietary fibre may include fractions of lignin and/or other compounds when associated with polysaccharides in the plant cell walls, and if these compounds are quantified by the AOAC gravimetric analytical method for dietary fibre analysis. Fractions of lignin and the other compounds (proteic fractions, phenolic compounds, waxes, saponins, phytates, cutin, phytosterols, etc.) intimately "associated" with plant polysaccharides are often extracted with the polysaccharides in the AOAC 991.43 method. These substances are included in the definition of fibre insofar as they are actually associated with the poly- or oligo-saccharidic fraction of fibre. However, when extracted or even re-introduced into a food containing non-digestible polysaccharides, they cannot be defined as dietary fibre. When combined with polysaccharides, these associated substances may provide additional beneficial effects (pending the adoption of the Section on Methods of Analysis and Sampling).

[2] The decision on whether to include carbohydrates from three to nine monomeric units should be left to national authorities".

The UK remained unconvinced of the EU definition as the Department of Health in the UK (1991) defined dietary fibre based on its chemical identity as non-starch polysaccharides (NSP) that referred to non-alpha-glucans substances as measured by the Englyst method (Englyst *et al.*, 1983, Englyst and Cummings, 1988). Englyst, a dietary fibre leader, with other scientists, reported a definition for dietary fibre in a comprehensive review as follows:

"Dietary fibre should be considered as a public health term supporting dietary guidelines to consume a plant-rich diet..... 'The definition "intrinsic plant cell wall polysaccharides" provides the only consistent link with the scientific evidence on which these guidelines are based."

(Englyst *et al.*, 2007, p.S36)

A year later a broader definition was suggested by the Scientific Advisory Committee of Nutrition (SACN) (2008) which was based on evidence that concluded that NSP and soluble fibres from oats, psyllium, pectin and guar gum were considered to be components of dietary fibre because they were proven to have a beneficial effects on health. Strong evidence was demonstrated regarding the beneficial effect of NSP on large intestine function and the beneficial effect of the soluble component of NSP on the blood lipid profile (Scientific Advisory Committee of Nutrition, 2008). A recent FAO/WHO scientific report (Mann *et al.*, 2007) agreed with the UK definition which defined dietary fibre as *"intrinsic plant cell wall polysaccharides*, p.S132*"* (Department of Health, 1991). The advantage of the chemical classification is that it offers useful information for labelling purposes (Scientific Advisory Committee of Nutrition, 2008). While, the nutritional importance of dietary fibre can be achieved when both chemical and food matrices are considered. The committee on the Scientific Evaluation of Dietary Reference Intakes in Food and Nutrition Board, Institute of Medicine of the United State, proposed dietary fibre definition as in Table 2.1.

After several years of debate and discussion, global agreement on dietary fibre definition was finally reached by WHO/FAO Codex Alimentarius Commission (Table 2.1) (Betteridge, 2009).

Global agreement on dietary fibre definition will provide clear boundaries on what to examine in research and will facilitate a consistent, well defined message for food labelling and health claims, especially if the analytical method is standardized. From the above dietary fibre definitions, the main arguments were focused on

whether short chain polysaccharides (DP 3-9)[3] should be included within the dietary fibre definition, however this decision was left for national governments in the final global definition by WHO/FAO (Codex Alimentarius Commission, 2008).

Furthermore, the concern about low molecular weight carbohydrates (3 – 9 monomer units) was reported by Betteridge (2009) mainly because this group of carbohydrates are not necessarily consumed from fruit, vegetables and wholegrain cereals and can be found in liquid products. Therefore, it would not consistently help in promoting plant-rich food as the unique source of fibre especially as production of non-digestible oligosaccharides (NDO) is increasing and they are added as functional foods in different food products (Mussatto and Mancilha, 2007). However, others agreed on including carbohydrate polymers containing 3 – 9 sugar units such as fructo-oligosaccharides (FOS), polydextrose and resistant dextrins (European Food Safety Authority, 2010) in the definition because of evidence which supports physiological properties of such non-digestible materials such as stool bulking and softening, and colonic fermentation (Tungland and Meyer, 2002).

A recent report on the dietary fibre definition from the Ninth Vahouny Fiber Symposium demonstrate the reasons for including carbohydrate polymers between 3 – 9 units as there is a lack of evidence that distinguish between the physiological effects of oligosaccharides and physiological effects of polysaccharides (Howlett *et al.*, 2010). Also, it has been emphasized that the cut-off point of 10 units cannot be determined through current analytical methods.

The definition proposed by Codex Alimentarius Commission (2008) was explored further by Mann and Cummings (2009) and they supported the three points in the footnotes. The first point is the definition of intrinsic plant carbohydrates where positive health effects were obtained from high quality studies that examined diets rich in fibre (fruit, vegetables and wholegrain) on health outcomes. The last two points defined the extrinsic sources (either synthesized or extracted types of carbohydrates) which can only be accepted as fibre if scientific evidence demonstrates their beneficial effects.

With the new definition, as NDO and resistant starch were part of total dietary fibre, the demands of a method which includes all dietary components resulted in an updating of analytical methods. The integrated dietary fibre method which

[3] Carbohydrates with 3-9 DP include Non-α-glucans oligosaccharides such as galato and fructo-oligosacchrides, inulin, polydextrose, raffinose and stachyose. (Cummings and Stephen, 2007)

determines the complete range of dietary fibre components was established by McCleary (McCleary, 2007) and a collaborative study recently by McCleary with other scientists (McCleary *et al.*, 2010) from 16 laboratories showed high repeatability and reproducibility of the new method with previous AOAC (991.43) method. This procedure aimed to standardize dietary fibre labelling across Europe, since the EU definition included resistant materials. However this method is quite new and more studies should be carried out in a wide range of food and food products before considering its use in nutrient databases.

In the latest Food and Nutrition Board definition, the dietary fibre category mainly included plant non-digestible carbohydrates which are largely intact with a plant matrix and integrated in to the plant cell wall, whereas "added fibre" referred to chemically or enzymatically isolated or extracted compounds and manufactured materials. This categorization also aimed to add animal origin polysaccharides in the "added fibre" category only when these are recognized as beneficial to human health (Food and Nutrition Board, 2001).

2.3 Dietary fibre components and structure

Researchers classify dietary fibre based on several factors such as chemical structure and physical properties that result in variable physiological effects in human. The following categorization is intended to provide information on cell wall derived dietary fibre components in the form of NSP and other non-NSP components. A wide range of dietary fibre components were included in the definition (Codex Alimentarius Commission, 2008) summarized in Appendix A.

2.3.1Cell wall derived components

Plant cell walls are the principle dietary fibre component derived from all plant foods (Selvendran, 1984). Plant cell walls contain two main compartments. First, cellulose microfibrils account for 30-40% of NSP (McDougall *et al.*, 1996). Secondly, the matrix of carbohydrates also called non-cellulosic polysaccharides that are embedded in cellulose microfibrils, which are made of hemicelluloses and pectin that varied widely in different food groups. Table 2.1 illustrates the different NSP components in the two main compartments with their main food sources.

2.3.1.1 Cellulose

Cellulose is a linear polymer of glucose units linked by β 1-4 glycosidic bonds which form crystalline microfibrils. Materials derived from cellulose are insoluble,

non-digestible and hardly fermented in the large intestine (Buttriss and Stokes, 2008).

2.3.1.2 Non-cellulosic polysaccharides

Non-cellulosic polysaccharides (the carbohydrate matrix) cover a wide range of substances (British Nutrition Foundation, 1990). As shown in Table 2.2, pectins are neutral polysaccharides rich in galacturonic acids, arabinose and galactose. Arabinogalactans are mainly found in the cell walls of most fruits and vegetables, which are solubilised by solutions that contain chelating agents such as Ca^{++} and Mg^{++} (British Nutrition Foundation, 1990). The second component of the carbohydrate matrix are the hemicelluloses that are solubilised by strong alkaline solutions, include β-glucans, xylo-glucans, arabinoxylans and glucurono-xylans (British Nutrition Foundation, 1990). More information on the non-cellulosic polysaccharides will be illustrated in the dietary fibre sources in a later section (2.4). Glucuronoxylans and xyloglucans are insoluble in water and are part of the hemicelluloses fraction. Both can be found in fruits and vegetables.

Within pectic and hemicellulosic components, the water soluble components include arabinogalactans that are mostly found in fruits and vegetables, mixed linked (1-3 and 1-4) β-glucans that are found in many cereals such as barley and oat, and arabinoxylans that are found in cereals.

Table 2.2 Non-starch polysaccharides components in plant foods

Major group	Components	Structure	Food sources
Cellulose		long chain β-glucan	All food groups
NCP[1]	Pectic substances	Galacturonans	Fruit and vegetables including legumes
		Arabinogalactans	
	HC[2]	Arabinoxylans	Cereals
		Glucuronoarabino-xylans	Cereals
		Glucurono-xylans	Fruits, vegetables
		Xylo-glucans	Fruits, vegetables
		β-glucans	Cereals
Gums and mucilages		Wide range of hetero-polysaccharides	Seeds and fruits

[1]NCP = non-cellulosic polysaccharides [2] HC = Hemicellulose adapted from British Nutrition Foundation (1990)

2.3.2Non-cell wall derived components

2.3.2.1 Mucilages and gums

Mucilages and gums are NSP that are mainly added to processed food. Gums are obtained from plant exudates such as gum arabic or from seeds such as guar and locust beans or seaweed extract such as agar. Mucilages such as psyllium, act as a gelling agent and have been found to reduce blood total cholesterol and low density lipoprotein (LDL) levels of hypercholesterolemic adults in clinical trials (Olson *et*

al., 1997). Gums are thickening agent used by food industry (Food and Nutrition Board, 2005). Both, gums and mucilages have a high water-binding capacity to form viscous solutions, and they are found in the outer layer of seeds. They are extensively used in food industries as gel formers and thickeners and have bulk laxative characteristics (Asp, 1987).Variation in physiological consequences appears to vary depending on subtle differences in physiochemical properties, in particular, rheological properties (Eastwood and Morris, 1992).

2.3.2.2 Lignin

Lignin is considered to be integrated within parts of the cell wall (Cosgrove, 2005). Differences in cell wall components have been observed with the maturity of the plant (British Nutrition Foundation, 1990). With complete maturity of cell walls, these walls become rigid because of the presence of the woody tissue (lignin) in the secondary wall (Smith, 2001). Lignin is a non-carbohydrate component formed by condensation of aromatic alcohol that links tightly to cellulose and hemicelluloses and infiltrates the polysaccharides matrix of the cell walls (British Nutrition Foundation, 1990). Therefore, lignin provides mechanical support (DeMan, 1999). Lignin is not included within NSP values in UK food composition tables (Food Standards Agency, 2002), although it is provided in a separate column of the tables. It is included in dietary fibre values derived by other analytical methods (Prosky *et al.*, 1985, Lee *et al.*, 1992), and this will be discussed later on in the dietary fibre analytical methods section 2.5.

2.3.2.3 Indigestible carbohydrates

Other components which have recently been included in the dietary fibre definition are indigestible carbohydrates such as resistant oligosaccharides, resistant starch and synthetic carbohydrates. The non-digestible oligosaccharides (NDO) also referred to as resistant oligosaccharides, occur naturally as 2-20 monosaccharide units (British Nutrition Foundation, 1990). This type of carbohydrate is not digested or absorbed in the small intestine (Buttriss and Stokes, 2008). It is identified that the cut-off point between oligosaccharides and polysaccharides is 10 sugar units (Cummings and Stephen, 2007). However, this is an arbitrary set point (Roberfroid and Salvin, 2001). Resistant oligosaccharides can be made enzymatically or chemically from simple carbohydrates such as mono or disaccharides or by enzymatic hydrolysis of polysaccharides (Buttriss and Stokes, 2008). Resistant oligosaccharide in the form of inulin can be obtained naturally from Jerusalem

artichoke and chicory root. Also, fructo-oligosaccharides can be synthesized from sucrose and galacto-oligosaccharides, or can be synthesized from lactose. It has been reported that inulin-type fructans were recovered in 86-88% of the ingested dose in the ileostomy study (Knudsen and Hessov, 1995) which indicates the high indigestibility of the substance.

Scientific evidence showed potential health benefits of non digestible oligosaccharides (NDO) such as promoting the growth of intestinal microflora (Mussatto and Mancilha, 2007). Other properties of NDO for example, inulin use to replace fat and fructo-oligosaccharides have used as sweetening sugars (Roberfroid and Salvin, 2001).

Starch that is resistant to digestion and absorption in small intestine is also included in several dietary fibre definitions. A human study by Stephen (1991) reported that 2-20% of ingested starch escapes the small intestinal absorption step when aspirated samples from the terminal ileum were analysed which reflects that a substantial proportion of starch eaten daily may reach the large intestine. Another experimental evidence by Cummings and Englyst (1991) suggested 10% of starch consumed was resistant starch. Resistant starch (RS) has been recognized to be resistant to digestive enzymes and is considered as one of dietary fibre component according to the recent definition (Codex Alimentarius Commission, 2008). There are 4 types of RS which are presented in Table 2.3.

A review by Sajilata *et al.* (2006) reported that colonic health is mediated by the production of short chain fatty acids from resistant starch fermentation in the large intestine. Improvement of blood glucose and insulin responses as well as improvement in blood lipid profile were also reported as a physiological effect of RS consumption (Nugent, 2005).

Table 2.3 Types of resistant starch, their common food sources and processing methods that minimize the resistance

Type of RS	RS1	RS2	RS3	RS4
Description	Physically protected	Ungelatinized resistant starch granules	Retrograded starch	Chemically modified starches
Common food source	Wholegrain and seeds, legumes	Raw potatoes, green bananas, some legumes, high amylose corn	Cooked and cooled potatoes, bread, cornflakes	Foods in which modified starches been used (for example, breads, cakes)
Resistance minimized by	Milling, chewing	Processing and cooking	Processing conditions	Less affected to digestibility in vitro

Adapted from Nugent (2005)

2.3.2.4 Other minor components

Glycoproteins are a proteins found in the cell wall structure which form up to 10% of the immature cell wall, but decrease as the wall matures (British Nutrition Foundation, 1990). Glycoproteins were found to be linked to neutral polysaccharides which are also present in the matrix (Selvendran, 1984). Cell wall proteins are included in the dietary fibre definition because they are less digestible than other proteins (British Nutrition Foundation, 1990). Other minor food components of plant cell wall structures are non-carbohydrate compounds including cutin, suberin, and plant waxes which are lipid compounds (Southgate, 2001).These minor components that are closely associated with cell walls are on the exterior surface of the food.

Cutin is a hydrophobic material that forms a waxy layer located in the external surfaces of some fruits and leaves. Cutin resists hydrolysis in the intestine and can be recovered in faeces (Southgate, 2001). Suberin is an indigestible component composed of insoluble aromatic and polyester components attached to the cell walls of plants that highly hydrophobic (Kolattukudy, 1981). It can be found in roots and tubular vegetables.

Inorganic constituents such as calcium, potassium and magnesium salts, phosphate and silica were found to be involved in plant cell wall structure and have an important integral role in the cell wall (British Nutrition Foundation, 1990).
Another minor components that have been reported previously to be part of the cell wall are polyphenolic compounds, a plant material that can also be found bound to dietary fibre (Bravo *et al.*, 1994).

2.4 Main dietary fibre sources in the human diet

2.4.1Cereal fibre

Cereals and cereal products are the main sources of dietary fibre intake in the UK (Buttriss and Stokes, 2008). Wheat, rice, maize, oat, rye, and barley are the main foods in the cereal group. Cereal grain is composed of three compartments; germ, starchy endosperm and peripheral layers (Slavin, 2004). Bran cell walls (most outer layers in cereal grains) mainly contain cellulose, arabinoxylans, lignin and phenolic acids substances. Cellulose, hemicelluloses and lignin form a strong matrix of cell wall which is highly insoluble. NSP content of cereals and cereals products varies widely and ranges between 0.12 – 41.5% (Englyst *et al.*, 1983). Content depends on

the processing methods and the extent to which the cereal has been refined (Southgate, 1978).

The cereal cell wall contains mainly arabinoxylans and β-glucans. Variation in the proportion of these polysaccharides in the cell wall was reported previously (Selvendran, 1984) and each types of polysaccharides has been recognised to have different properties. For example, arabinoxylans are mostly soluble in water but have an insoluble part linked to phenolic acids that may related to resistant and partial fermentability in large intestine (British Nutrition Foundation, 1990). While β-glucan is highly viscous and suggested to have benefit health effects (Wood, 2007). β-glucan content and solubility varies between different cereal foods. The wheat endosperm cell wall is rich in arabinoxylans (88%) but poor in β-glucans (3%): however, barley contains the opposite proportion of cell wall polysaccharides (19% and 80% respectively). The aleurone layer of barley is rich in β-glucans and arabinoxylan with a small amount of cellulose and phenolics (Selvendran, 1984). In oat and barley, as the richest sources of soluble fibre, three-quarter of β-glucan in oat is solubilised in hot water whereas only 10% of β-glucan in rye is soluble (Wood, 2010). The β-glucans content of cereal grains are important, as they have been associated with key health benefits in human. Meta-analysis of 11 RCT showed intake of barley and β-glucan isolated from barley improve blood lipid profile significantly (AbuMweis *et al.*, 2010).

2.4.2Fruit and vegetable fibre

Generally, the amount of total NSP in fruit and vegetables is higher than in cereal and cereal products as g/100g (Englyst and Englyst, 2005). Soluble NSP accounts for 32% of total dietary fibre in vegetables and 38% in fruits (Anderson and Bridges, 1988). However, the proportion of soluble and insoluble fractions varies between different types of fruit and vegetables. For example, insoluble fibre is particularly high in berries, avocado, apple, and others listed by Englyst (1988), and this reflects the higher cellulose in these fruits, while other type of fruit contain more soluble fibre such as fresh apricot, fig and mango. Vegetables such as cauliflowers, tomatoes and onions contain significant high amount of insoluble fibres (O'Shea *et al.*, 2012).

Fruit and vegetables fibres are rich in pectin, and these are recognised to have important physiological effects in the human body such as decreasing gastric emptying, shortening transit time in the small intestine and lowering serum

cholesterol level (Chawla and Patil, 2010). Englyst (2005) demonstrated that cereals are the major contributor for total NSP as eaten but, fruit and vegetables have higher proportions of NSP in dry matter, and this contributes to the low energy density characteristics of these food groups.

2.4.3 Legume fibre

Legumes may be eaten either as fresh pod of seed fruit from plants of the Fabiaceae (or Leguminosae) family. Englyst with his colleagues (1988) determined the NSP content in 173 foods items from different food groups and found that legumes contain very high NSP compared to leafy vegetables (range between 4.9 to 17% of fresh weight). Legume seeds contain two dietary fibre components, the inner fibre which is cell wall of cotyledon and outer fibre referred to as coat seed or hull. The main variation between inner and outer fibres is in the proportion of cellulose and non-cellulosic polysaccharides. The outer layer is rich in cellulose, whilst, the inner fibre layer of legumes mostly contains pectin, and a small amount of cellulose. For example, a review by Guillon and Champ (2002) reported that dietary fibre composition in peas and lupin differ in the outer layer and inner layer. The inner layer of mainly contains 55% pectic substances and 9% cellulose, while outer layer contains 35-57% cellulose and a low amount of hemicelluloses and pectin. The inner layer of lupin characterized by high water holding capacity which gives a physiologically beneficial effect in terms of faecal bulking property (McCleary and Prosky, 2001). This helps in the prevention of constipation (Tiwari and Cummins, 2011) and has the potential to lower blood cholesterol level (McCleary and Prosky, 2001).

It was reported that several studies showed the proportion of insoluble dietary fibre is higher than soluble dietary fibre in different legumes such as chickpeas, beans and peas (Oomah *et al.*, 2011). However, this does not rule out that legumes are a good source of soluble fibre (Englyst *et al.*, 1988). Processing methods related to legume fibre were demonstrated in a previous review, methods such as milling, grinding, soaking, cooking, canning, and boiling were found to affect dietary fibre content and composition. For example, milling results in removal of the hull, which then results in a decrease in total dietary fibre particularly the insoluble fibre; an increase in soaking time affects the water-holding capacity of coat seeds, and canning reduces soluble fibre content in common beans (Tiwari and Cummins, 2011).

2.4.4 Fibre from nuts and seeds

Nuts are also rich in dietary fibre. Analysis of eight types of nuts and seeds showed that amount of NSP range from 2.43% for pecans to 5.56% for peanuts (Li *et al.*, 1997). Ten types of nuts were analysed for NSP content range between 3.5% to 13.7% and it was found that insoluble fibre is the predominant component which was mainly cellulose (Englyst *et al.*, 1988). Almonds, peanuts, peanuts butter and English walnuts were analysed chemically and found that cellulose and hemicelluloses contributed to 31% and 35% respectively while pectin accounted for 18% of total dietary fibre measured. Soluble dietary fibre was only 3% of the total fibre content which was mainly hemicelluloses (Marlett, 1992).

Heterogeneity in dietary fibre components can be appreciated from the above fibre sources. In addition to analytical method used to determine dietary fibre, the variety and portion of plant food analysed were also possible explanations of dietary fibre differences within dietary fibre source.

2.5 Dietary fibre analysis

A number of different analytical methods have been established to measure dietary fibre contents and components in different type of foods. Analytical methods were categorized into enzymatic chemical methods and enzymatic gravimetric methods, based on the approach and aim achieved with each method (Englyst *et al.*, 2007). The two most used analytical methods are the Englyst method and Association of Official Analytical Chemists (AOAC) method. Another earlier analytical method, the crude fibre method (Furda, 2001) was carried out previously and because of error related to underestimation of dietary fibre, it is not used nowadays to measure total dietary fibre. In earlier days, the Southgate method, developed in the late 1950s by Southgate (1969), was used to measure unavailable carbohydrates. However, because more accurate and precise methods were established, this method became obsolete. None the less, dietary fibre values derived from the Southgate method are still found in the sixth edition of the UK food composition tables (Food Standards Agency, 2002).

2.5.1Southgate method

The Southgate method (1969) is basically an enzymatic chemical method where digestive enzymes are used for starch hydrolysis, then acid hydrolysis was used to yield components of hemicelluloses, cellulose, and lignin which are

measured chemically. The disadvantage of this colorimetric method is the non-specific colour reactions of analyzed reducing sugars (Lee and Prosky, 1992) as well as multiple steps which requires longer time. This method has been largely replaced by the Englyst method.

2.5.2Englyst method

The Englyst method identifies the different monomers that make up total NSP. In essence, the method involves measurement of sugar monomers following starch hydrolysis using digestive enzymes (Englyst *et al.*, 1988). Limitations in the Englyst method were reported in a previous study (Wolters *et al.*, 1992). The acid hydrolysis step which aims to breakdown the bonds between monosaccharide's can cause some loss of NSP monosaccharides (Wolters *et al.*, 1992). Whole wheat and dried apple fibre measured directly by Englyst method showed lower NSP amount than the sum of soluble and insoluble NSP that may be due to presence of glucose because of oligosaccharides contaminated in the soluble fibre residue or cellulose loss from the direct measure (Wolters *et al.*, 1992).

2.5.3Enzymatic gravimetric (AOAC) methods

Having a published definition to work with, there was an effort by many researchers to develop a method for dietary fibre analysis to meet the definition (DeVries, 2004). A general consensus had been reached on the official method of analysis referred to as the AOAC method 985.29 for measuring total dietary fibre in foods. This method was established by scientists led by Prosky (1985), and aimed to measure dietary fibre as defined by Trowell (1972). A further modification in the methodology was developed and collaboratively studied, aiming to improve the method and to measure dietary fibre components based on their solubility (soluble and insoluble dietary fibre) in foods and food products (Prosky *et al.*, 1992).

The AOAC method measures the sum of non-digestible carbohydrates and lignin using enzymatic digestion to eliminate non-fibre components and quantify the residue by weighing (McCleary and Prosky, 2001).

AOAC (985.29) was the first version of the enzymatic gravimetric method adopted by Prosky and his colleagues (1985) that measures total dietary fibre in foods and foods products and later on was modified by using a different buffer (4-morpholine-ethanesulfonic acid-TRIS buffer). This was labelled as the AOAC (991.43) method (Lee *et al.*, 1992) which was used to measure soluble and

insoluble fibre in addition to total dietary fibre in foods. The above methods are used to determine total dietary fibre in food and food products for labelling purposes and nutrients databases (DeVries, 2004). Table 2.4 presents the main steps in the three common methods described above.

A collaborative study (Prosky *et al.*, 1985) represented 5 countries where 9 collaborators tested mainly cereal and cereal products using the AOAC (991.43) method. In this study, coefficient of variation (CV%), which gave an estimate of precision by evaluating the ability of an AOAC method to be reproduced in different laboratories, showed, CV% ranged between 1.56% - 9.80%. Later on Prosky and his colleagues (1992) reported insoluble and soluble dietary fibre CV% for 22 foods from different sources (carrots, kidney beans, barley, figs and others) analyzed in 39 collaborators and found that reproducibility of insoluble fibre of almost half of analyzed foods was less than 10%. However, the soluble fibre reproducibility was much higher at 10-20% which may be due to filtration problems or because of low fibre content in the tested samples. Again, because of the high soluble fibre CV% reproducibility results found in the previous study, another collaborative study (Prosky *et al.*, 1994) carried out in ten laboratories, testing the same samples. The averaged CV% was 14.1% for soluble fibre, 8.0% for insoluble fibre and 4.9% for total dietary fibre after overcoming filtration problems.

AOAC methods 985.29 and 991.43 have the same principle. The main steps of AOAC methods include sample digestion, precipitation with 4 volumes of alcohol, filtration then drying and weighing the total dietary fibre residue. To obtain the dietary fibre value, the sample will first be digested by enzymes then the digested sample will be filtered, dried and weighed to yield the insoluble dietary fibre value. The soluble fibre values will then be obtained from the filtrate that is participated in 4 volumes of ethanol, which is then again filtered, dried and weighed to yield the soluble fraction value. All fibre residues are corrected for ash and protein (Lee and Prosky, 1992).

Modifications were performed in the AOAC method including the use of organic buffers (MES/TRIS) instead of phosphate buffers to limit the formation of co-precipitations with dietary fibre and omitting pH adjustment steps to reduce time spent in analyses (Lee and Prosky, 1992) as illustrated in Table 2.4. Lee and Prosky (1992) demonstrated that the overall precision of soluble, insoluble and total dietary fibre values was acceptable in all mentioned methods when they compared between

them and advised the generation of dietary fibre values for a database for food composition and food labelling purposes.

Table 2.4 Main steps in three dietary fibre analytical methods

	Enzymatic gravimetric methods		Enzymatic chemical methods
Method step	AOAC Prosky et al, 1988	AOAC modified Lee et al., 1992	Englyst method
Sample	1g	1g	50-300mg
Buffer	Na-phosphate pH 6	MES/TRIS pH 8.2	no buffer
Starch and protein hydrolysis	Termamyl[2] 100°C, 15-30min, Amyloglucosidase 60°C, 30 min Protease 60°C, 30 min	Termamyl[2] 100°C, 30 min Amyloglucosidase 60°C, 30 min Protease 60°C, 30 min	DMSO[1] (0.5h, boiling water bath) Termamyl[2], 10min, boiling water bath Pancreatin and pullulanse ,0.5h and 50°C + 10 min boiling
pH adjustment 1	To pH 7.5 before Protease	No pH adjustment	No pH adjustment
pH adjustment 2	To pH 4.0-4.6 before Amyloglucosidase	To pH 4.0-4.6 before Amyloglucosidase	No pH adjustment
Alcohol precipitation	280ml	225ml	No precipitation
Acid hydrolysis and sugars determination	no	no	Analysis of sugars: 12 M H_2SO_4 (1h, 35°C) 2 M H_2SO_4 (1h, 100°C) and sugar determine chemically
lignin	included	included	Not determined
Filtration aid	Celite 545	Celite 545	No
Protein and ash correction	yes	yes	No

[1] DMSO = Dimethylsulphoxide for starch solubilisation, [2]Termamyl is thermostable α-amylase from Novo Nordisk, Denmark. Table adapted from McCleary and Prosky (2001) p.81.

Errors measuring dietary fibre by the AOAC method have been reported in previous study (Mañas and Saura-Calixto, 1993). The precipitation step was found to be subject to two errors when measuring soluble dietary fibre. Firstly, incomplete precipitation of pectin was proven from analyzed citrus pectin. Secondly, co-precipitation of substances other than fibre such as organic acids (Englyst *et al.*, 1996) have resulted in overestimation of what should be considered as dietary fibre because of the presence of artefacts (Mañas and Saura-Calixto, 1993).

Incomplete precipitation of soluble fibre may affect the measured distribution of soluble and insoluble fractions (Mañas *et al.*, 1994). Another possible error was reported such as errors related to weighed sample (Mertens, 2003).

2.5.4The use of dietary fibre analytical method

AOAC 985.29 method was approved as the official method for determining total dietary fibre in all foods while AOAC 991.43 method is applicable to measure the total, soluble and insoluble dietary fibre in all foods (Joint FAO/WHO Food

Standards Programme, 2011). As demonstrated in Table 2.5, Both methods are used to determine dietary fibre contents in foods in the USA and in most of the European countries' nutrients databases (DeVries and Rader, 2005).

In the United Kingdom, the Englyst method is used to determine NSP in foods. Dietary fibre values reported in McCance and Widdowson's food composition was accepted by the Ministry of Agriculture, Fisheries, and Food (MAFF) and remained the recommended method for nutrition and food labelling until 1999. In the same year, the Food Standards Agency accepted the role of resistant starch and lignin as a dietary fibre and they adapted AOAC method to measure dietary fibre in foods (Food Standards Agency, 2002).

Table 2.5 Global acceptance of the AOAC official method

Country	Type of AOAC official Method used
USA	AOAC 985.29 and 991.43 method
Australia	AOAC 985.29 method
Canada	AOAC 985.29 method
Japan	AOAC 985.29 and 991.43 methods
Mexico	AOAC 985.29 method
UK	Englyst method and 6thedition of UK food composition reported some values by AOAC 985.29 method.
Other countires[1]	AOAC 985.29 methods

[1]Denmark, Finland, Italy, and Sweden; Table adapted from (DeVries and Rader, 2005)

2.5.5Differences in dietary fibre analytical methods

Regarding the measured dietary fibre constituents, the Englyst method measures non-starch polysaccharides, which fit the definition of plant-rich diet "*intrinsic plant cell wall polysaccharides*" (Englyst et al., 2007, p.S27) based on their chemical components rather than the sum of indigestible materials as in the AOAC method (Englyst *et al.*, 2007). The higher dietary fibre values are measured by the AOAC method as compared with the Englyst method relate to the presence of resistant starch, lignin and other indigestible materials measured by the AOAC method. Resistant starch is measured by the AOAC method; in particular, retrograde resistant starch (RS3) and inaccessible starch granules (RS2) (Lunn and Buttriss, 2007).

Table 2.6 Dietary fibre components measured by selected analytical methods

Method	Lignin	NSP	Resistant starch	Resistant Polysaccharides Oligosaccharides[1]	Other[2]
Englyst method	X	Yes	X	X	X
Southgate method	Yes	Yes	Some	X	X
Lee method (AOAC 991.43)	Yes	Yes	Some	X	Some
Prosky method (AOAC 985.29)	Yes	Yes	Some	X	some

[1]Such as polydextrose, resistant maltodextrin and oligosaccharides, [2]Non-carbohydrates defined as other non-lignin non-carbohydrates products such as tannins, cutin, and protienaceous products. Selected methods adapted from Lunn and Buttriss (2007)

Table 2.6 summaries dietary fibre constituents measured by selected analytical methods that are used to determine the dietary fibre in foods and food products which are then used in nutrients databases. All the methods in the Table 2.6 agree in measuring NSP as a major component in the dietary fibre definition. Lignin and resistant starch, mainly measured by enzymatic gravimetric methods, are not counted in the values obtained from the Englyst method. Resistant starch are resistant to human enzymes but mostly fermented in the large intestine (Sajilata *et al.*, 2006) partly measured by the Southgate (1969) and Prosky (1985) methods.

Recently, an integrated procedure for total dietary fibre measurement which includes total resistant starch and low molecular weight non-digestible oligosaccharides was established (McCleary, 2007). The latest AOAC Method 2009.01 was validated and adopted by the Codex Commission as the official method for measurement of total dietary fibre (McCleary *et al.*, 2010).

The variation in the dietary fibre values between analytical methods become a concern in comparative epidemiological studies and it has been reported that dietary fibre values are highly dependent on the analytical method used which should be considered (Deharveng *et al.*, 1999).

The food composition of nine European countries of the EPIC study suggested that dietary fibre method comparability depends on food groups (Deharveng *et al.*, 1999). This finding is in agreement with recent review findings (Slimani *et al.*, 2007) which showed dietary fibre values obtained from the AOAC method were comparable to values obtained by the Englyst method for fruit and vegetables (excluding potato). While fibre from cereal, legumes and potatoes contain significant amounts of fibre other than non-starch polysaccharides components that are measured by the AOAC method but not with the Englyst method, which needs to be considered. It was reported by DeVries (2004) that most of the updated food composition databases include fibre values obtained from the AOAC method. In

addition to analytical method, portion of food analyzed and variety of food samples tested were all other possible explanation of differences in dietary fibre values (Marlett, 1992).

2.6 Dietary fibre recommendations and intakes

2.6.1 Dietary fibre recommendations

Generally, the aims of dietary recommendations are to prevent overt deficiencies of essential nutrients. However for fibre, the approach was from evidence of bowel habits that linked to the lowest risk of bowel disease. The aim was to set a fibre intake associated with stool weight of greater than 100g/day. As stool weight of less than 100g/day is associated with increased risk of bowel diseases, which occurs when the NSP intake is less than 12g/day and intakes above 32g/day of NSP will not further increase stool weight. Furthermore, it is expected that an increase of NSP intake from 13 to 18g/day would give an increase in average stool weight of 25% (Department of Health, 1991). The Dietary Reference Value (DRV) for NSP measured by Englyst in the UK is 18g/day (individual range 12-24g/d), applicable for adults only, based on quantitative evidence related to bowel habit (Department of Health, 1991).

In other countries, the recommendations were established based on disease prevention. For healthy American and Canadian populations, dietary fibre recommendations were published by the Food Nutrition Board (2005) which was based on coronary heart disease protective intake levels obtained from prospective and clinical data. An adequate intake of dietary fibre for people aged between 19 -50 years old was set at 25g/day for women. European dietary fibre recommendations in the scientific opinion of the European Food Safety Authority (EFSA) (2010) recently reported the DRVs for dietary fibre. This was set at 25g/day to ensure normal laxation in adults. Additionally, this value was set considering lowering risk of chronic diseases such as coronary heart disease and diabetes.

Table 2.7 demonstrates the dietary fibre recommendations in different countries. It has been noticed that the amount of dietary fibre recommendations are highest for Americans and Canadians, while the lowest amount is found for the UK recommendation, which could be explained by the variation in dietary analytical methods used to measure dietary fibre. Lunn and Buttriss (2007) demonstrated the variation in the recommended amount of dietary fibre intake across countries may also be due to whether the recommended amount was based on physiological

function as in the UK recommendation or other health benefits as in the France, Germany and USA recommendations or both beneficial effects as in the Netherlands.

Table 2.7 Dietary fibre recommendations in different countries

Country	Dietary fibre recommendation
Worldwide (FAO/WHO)	>20g[1]; 25g[2]
Denmark	20-30g[2]
Finland/Sweden/Norway	25-35g[2]
France	25-30g[3]
Germany	30g[3]
Netherlands	30-40g[3]
Spain	30g[2]
UK	18g[1]
USA/Canada	25g (women 19-50years); 21g (women +50years)[2]
Australia/New Zealand	30g (men); 25g (women)[3]
Japan	20-30g[2]
South Africa	30-40g[2]

[1] based on Englyst method [2] based on AOAC method [3] analytical method not specified. Table Adapted from Lunn and Buttriss (2007)

2.6.2Dietary fibre intakes

Complexities are still present in determining dietary fibre consumption especially for researchers interested in international comparison. This is partly because of analytical methods used to measure dietary fibre content. In term of dietary fibre intake among different populations, a recent National Diet Nutrition Survey (NDNS) (Department of Health, 2012) reported that for NSP, the average intake for adults over 19 years old was 13.8g/day (the upper 2.5 percentile was 25.9g/day and the lower 2.5 percentile was 5.5g/day). Based on gender, average dietary fibre intake for men was 14.8g/d and for women equal to 12.8g/day.

The main dietary fibre source among the UK population is cereal and cereal products which represent 37% of total non-starch polysaccharides intake. Potato and vegetables as groups are the second biggest contributors of NSP intake (34%) for adults. This outcome was the same as the previous report in 2004 (Department of Health, 2011).

A European Food Standard Agency panel reported that the average daily intake of dietary fibre across European countries ranges between 18 -29.7 g for men and 15.7-23g for women. In addition, three quarters of people aged over 65 years

have dietary fibre intake ranging between 19 and 25g/day (European Food Safety Authority, 2010). Green (2001) summarized that the typical daily fibre intake of adults consuming a Western diet as 11.8-16.4g of total NSP where 50% was from cereals, 40% from vegetables and 10% from fruits. Estimated daily fibre components were also reported by Green (2001) as the insoluble NSP intake ranged between 6.5 – 7.0g/day and soluble NSP was 5.3-8.7g/day.

Total dietary fibre intake across European countries was obtained from the EPIC study (Cust *et al.,* 2009) and showed that mean dietary fibre intake adjusted for age, total energy, weight, height, season and days of recall was the highest among the health conscious UK population (26.8g/day) and lowest in Sweden (15.1g/day) and in the UK general population (17.4g/day). In general, fruit and vegetables groups were the major contributor to total dietary fibre intake. Thus evidence indicates that dietary fibre intake for the UK general population is still below the recommended value.

The third National Health and Nutrition Examination Survey (NHANES III) (1988–94) (Bialostosky *et al.,* 2002) which provides information on the health and nutritional status of the U.S population reported that the mean total dietary fibre intake which includes unavailable carbohydrates was 17.9g for people aged between 20 -59 years whereas for women the mean fibre intake was 14g/day and for men the mean fibre intake was 20g/day. Insoluble dietary fibre intake was 11.2g/day while soluble dietary fibre intake was 6g/day for those aged between 20-59 years old. The daily fibre intake among US population was still under the desirable level as the recommended amount is 25g/day for women and 38g/day for men. In general, the daily dietary fibre intake globally across different countries still does not reach the equivalent recommended intake which reflects the importance of a more comprehensive health promotion strategy that tackles dietary fibre intake.

2.7 Dietary fibre related effects: physiological and health effects

A wide range of substances are included in the dietary fibre term which results in a wide range in physiochemical properties. Guillon and Champ (2000) concluded in their review that properties such as viscosity of dietary fibre are key contributors to health related effects on glucose and lipid metabolism while water holding capacity and fermentation[4] are more related to colonic function. This indicates the

[4] The fermentation process that occurs in the large intestine includes digestion and absorption of carbohydrates that escape the small intestine which results in formation of gases and short chain fatty

usefulness of nutritionally related characteristics which help in understanding the potential mechanisms related to the beneficial effects of dietary fibre intake. The main areas of research investigated the dietary fibre beneficial effects were gastrointestinal health, glucose and insulin responses, risk factors of coronary heart disease, some cancers, and satiety were summarized in a review by Lunn and Buttriss (2007).

Englyst (2005) classified carbohydrates based on gastrointestinal handling into glycaemic and non-glycaemic carbohydrates. Non-glycaemic carbohydrates were defined as carbohydrates which enter the large bowel for fermentation. Indigestible carbohydrates are resistant to digestion and absorption partly due to the presence of specific glucosidic linkage that cannot be hydrolysed by human digestive enzymes. Other factors which may influence digestion and absorption were illustrated by Englyst and Englyst (2005). For example, meal macronutrient components, subject biological variation, food matrix and chemical structure all affect gastrointestinal digestion and absorption. The presence of NSP in the stomach and small intestine results in a slow rate of carbohydrate digestion and absorption which is related to the viscosity property of soluble fibre. When non-glycemic carbohydrates enter the large intestine, fermentation occurs but with varying degrees of fermentation which then leads to the production of short chain fatty acids (SCFA) as an important source of energy for the intestinal cells (Englyst and Englyst, 2005). The physiological effects of dietary fibre were determined based on three main factors, solubility, viscosity and fermentability as suggested by Li and Uppal (2010). Classifications of dietary fibre are illustrated in Table 2.8.

---

acids by gut microflora and substrates which will be absorbed and metabolized by the human body. (FAO, 1998).

Table 2.8 Classifications of dietary fibre based on their solubility, viscosity and fermentability properties.

Physiochemical properties	Sub-classes	Chemical structure and food sources
Solubility	Soluble	• Glucans (oats and barley) • Pentoses (rye) • Oligosaccharides (pulses, onion, Jerusalem artichoke, garlic)
	Insoluble	• Cellulose, hemicelluloses and lignin (cereal: wheat and rice) • Resistant starch (wholegrains and pulses)
Viscosity	Viscous	• Gum • Psyllium • Pectin • Cellulose (oat bran)
	Non-viscous	• Inulin • Resistant starch • Polydextrose • Cellulose (e.g. wheat and rice bran)
Fermentability	Partially fermented	• Cellulose (vegetables, sugar beets, brans) • Hemicelluloses (cereal) • Lignin • Cutin/suberin/ other plant waxes in plant fibres • Chitin and chitosan, collagen (fungi, yeasts) • Resistant starches (corn, potatoes, grains, legumes, bananas)
	Highly fermented	• Beta-glucans (oat, barley and rye) • Pectins (fruits, vegetables, legumes, sugar beets and potatoes) • Gums (legumes, seaweed extracts, plant extracts) • Inulin (chicory, Jerusalem artichoke, onions, wheat) • Oligosaccharides

Adapted from Li and Uppal (2010)

2.7.1Solubility

Dietary fibre is divided into soluble and insoluble dietary fibre where each group has its physiological characteristics. Insoluble fibre properties include shortened transit time, laxation by attracting water, softened stool and increased stool bulk which is more likely to be un-fermentable in the large intestine (Buttriss and Stokes, 2008). Whereas soluble fibre is characterized by forming a gel-like solution and is more likely to be fermented in the large intestine (Buttriss and Stokes, 2008). Cellulose, hemicelluloses and lignin were categorized into insoluble dietary fibre while non-cellulosic polysaccharides such as pectin, gums and mucilages were all under the soluble fibre category (Oakenfull, 2001). Most grain products such as wheat and rye are rich sources of insoluble fibres while soluble

fibre is mostly found in fruit and vegetables including legumes (Lunn and Buttriss, 2007).

Resistant starch and oligosaccharides were also allocated into the insoluble and soluble groups as they have physiological consequences (Lunn and Buttriss, 2007). Resistant starch was included in the insoluble component of non-digestible carbohydrates where other components such as cellulose, hemicelluloses and lignin were present while oligosaccharides were categorized under the soluble component which suggested that this is a useful approach especially with regard to the new dietary fibre definitions.

Solubility of dietary fibre is a key factor for its physiological effect. The type of linkage between sugar units in the chemical structure of dietary fibre is one of the main determinants of this physical property of dietary fibre (Oakenfull, 2001). For example, β-glucan is a soluble polysaccharide but cellulose is an insoluble polysaccharide. This because of the ordered structure of the polysaccharides chain with β 1-4 linkage which prevent cellulose from being solubilised while, the irregular type of polysaccharides structure in β-glucans allows solubilisation in water (Oakenfull, 2001). Another reason for being soluble is the presence of a charged group in the polysaccharides which prevents molecules grouping together in an ordered structure (like cellulose) in pectin.

The challenges in solubility property were demonstrated by some researchers. The extraction of oligosaccharides from polysaccharides is achieved by 80% ethanol however the separation depends on the chemical structure and degree of polymerization (DP). However, high branched polymers with more than ten DP were found to be solubilised in ethanol solutions such as arabinan in sugar beet fibre (McCleary and Prosky, 2001) which indicates that analytical separation is arbitrary (FAO, 1998). Also difficulties can be appreciated especially when categorization is not totally based on solubility as some insoluble fibre is fermentable (Lunn and Buttriss, 2007). It had been suggested earlier by WHO/FAO (1997) that carbohydrate categories based on solubility are less useful for nutritional research. However, based on the new definition by WHO/FAO, resistant oligosaccharides and resistant starch were part of total dietary fibre and the categorization was recently addressed by Lunn and Buttriss (2007).

2.7.2Fermentability

It has been demonstrated previously by Stephen (1991) that insoluble fibre has a great effect on stool weight because it is hardly fermented and is thus more effective as a bulking agent. Cumming and Stephen (2007) support a physiologically based carbohydrate classification as this is useful for evaluating potential health benefits as well as helping to determine foods linked to a healthy diet. It is important to consider other components which are associated with dietary fibre such as phenolic substances, waxes, phytates and cutin which are also indigestible and may have potentially beneficial effects (Lunn and Buttriss, 2007). An in vivo study reported that pectin is highly digestible by microbial flora, with a faecal recovery of 3% of the amount ingested (Bravo *et al.*, 1994).

2.7.3Viscosity

Viscosity characteristics refer to the physical interaction of the polysaccharides molecule in solution (Oakenfull, 2001). Viscosity characteristics of some polysaccharides have shown a beneficial role in the human body such as reduction in glucose absorption and lowering of plasma cholesterol level by the inhibition of cholesterol and bile acid absorption. Increasing viscosity, delays gastric emptying, interferes with food mixing process by reducing muscle contraction activity of the gastrointestinal tract and this slows the rate of glucose absorption (British Nutrition Foundation, 1990). This is a suggested mechanism of viscosity to explain the reduction of blood glucose response with certain fibre intake (Jenkins *et al.*, 1978). However, the long term effect of viscous fibre on chronic disease such as coronary heart disease uncertain (Jenkins *et al.*, 2000).

A water holding ability is found in both soluble and insoluble fibres. Soluble fibres form a gel-like compound while insoluble fibre acts as a sponge by entrapping water in the matrix (Cho and Dreher, 2001). It has been reported that laxation depends on water holding capacity and bacterial cell mass so the faecal bulking property of different types of fibres depends on the chemical, physical and bacterial mass in the large intestine (Cho and Dreher, 2001).

The most recent dietary fibre definition includes a wide range of materials with overlapping physiological properties resulting in variation in the classification of dietary fibre that makes it more complicated for researchers. As shown in Table 1.8, viscous fibres are mostly soluble and fermentable like pectin. On the other hand,

inulin is soluble but not viscous and well fermented in the large intestine. While, poorly fermentable fibres are generally insoluble fibres (Li and Uppal, 2010).

2.7.4Dietary fibre and health

The health benefits and physiological effects of dietary fibre have been reported by many researchers (Brown *et al.*, 1999 , Jenkins *et al.*, 2002). This has led to emphasise on the physiological and health benefits in the dietary fibre definitions (Codex Alimentarius Commission, 2008, AOAC, 2001, European Commission, 2008). When Trowell (1972) suggested a dietary fibre hypothesis which assumed an inverse relationship between dietary fibre consumption and the incidence of diseases, it stimulated research investigating the relationship between dietary fibre intake and risk of chronic illnesses.

Pooled data from 10 European and US cohort studies reported an inverse relationship between dietary fibre intake and the risk of coronary heart diseases with a 14% decrease risk (RR=0.86; 95%CI: 0.78, 0.96) with every 10g increase in energy adjusted fibre intake (Pereira *et al.*, 2004). Cancer-related beneficial effects have been evaluated by other prospective studies. A meta-analysis from 16 prospective studies reported a significant risk reduction of colonic cancer by 10% (RR=0.90; 95%CI: 0.86, 0.94) with every 10g daily increment of total dietary fibre (Aune *et al.*, 2011). The relationship between the risk of breast cancer and intake of dietary fibre was examined prospectively. The UK Women's Cohort Study reported a significant protective effect of high dietary fibre intake on the risk of breast cancer in pre-menopausal women (Cade *et al.*, 2007). Other beneficial effects of dietary fibre such as lowering blood pressure and the protective effect from weight gain were summarized in a recent European scientific opinion report (European Food Safety Authority, 2010)

Further evaluation was carried out by several studies to determine the effects of fibre fractions (soluble and insoluble fibres) on health outcomes and aimed to identify the potential fibre component that may explain the beneficial effects of high fibre intake on the risk of chronic disease. A large cohort study showed that a high intake of insoluble fibre did show significant beneficial effects against the risk of myocardial infarction among women (Liu *et al.*, 2002). A meta-analysis pooled result from 6 prospective studies and found that with every 10g increment in soluble fibre, the risk was significantly reduced by 28% for all coronary events and by 54% for coronary death whilst the effect size of insoluble fibre was smaller. Thus, with

every 10g increment in insoluble fibre, the risk was reduced significantly by 10% for all coronary events and by 20% for coronary death (Pereira *et al.*, 2004).

In terms of the risk factors for cardiovascular disease, the effects of soluble fibre intake in the form of pectin, oat bran, guar gum and psyllium were pooled from 67 controlled trials (Brown *et al.*, 1999). With every 2-10g/day increase in soluble fibre, a small but significant decrease in total blood cholesterol level by 0.054mmol/l and lower LDL-cholesterol level by 0.057mmol/l was reported. However, the effect on blood HDL cholesterol and triglyceride levels were not significant.

2.8 Conclusion

The on-going argument about defining dietary fibre centres mainly on the non-NSP compounds. Growing evidence of the beneficial effect of indigestible compounds drives towards including such compounds in the definition but deviation from the great benefit of plant-based food to other sources is of great concern to many researchers. Dietary fibre classification is based on several factors such as chemical structure, solubility and fermentability which will provide information to understand the health beneficial effects of dietary fibre. It has to be appreciated that not all foods have similar plant cell wall composition: however, NSP was the basic component on which all definitions agree. Adding a wide range of other indigestible materials into the dietary fibre definition that have physiological properties similar to dietary fibres makes it complicated in terms of analytical methods, and in nutrition epidemiological research. Nevertheless, the updated analytical AOAC method by McCleary (2007) established to include all indigestible materials considered as fibre in the recent dietary fibre definition. Dietary fibre recommendations varied across countries due to variation in the intended aims and analytical methods whereas the intake of dietary fibre in most countries was still under the recommended amount. Dietary fibre intake based on AOAC fibre value in UK prospective study is lacking and comparison between dietary fibre intake estimated from NSP and AOAC values will provide a clear picture on ranking participants in observational studies. This part will be demonstrated in chapter 7.

Chapter 3: Literature review of type 2 diabetes mellitus

3.1 Introduction

Type 2 diabetes mellitus (T2DM) is one of the common public health problems that require more focus on the primary prevention strategies as this will have a great benefit on individual and population levels. Environmental and genetic factors play an important role in developing T2DM. Environmental factors mainly focus on lifestyle and dietary factors. Modifications in some health related behaviours were advocated by several health organizations, aiming to delay or prevent the development of diabetes. Some researchers were focused on the genetic factors in the development of T2DM which may help in identifying people who are at higher risk of developing diabetes which then can target the population for future interventional trials. Diagnosis and classifications of diabetes will be provided in this chapter, alongside incidence, trends and complications of T2DM among adults. In attempting to investigate the effect of dietary fibre intake on the risk of diabetes, it is important to first explore other diabetes related risk factors. The current chapter describes epidemiological evidences on dietary and lifestyle practices in relationship to T2DM. Specifically, dietary and non-dietary risk factors such as body fatness, physical activity, macronutrient intake, intakes of the main micronutrients and main food groups that are related to T2DM as well as the common dietary patterns examined in relation to risk of diabetes will be included in this literature review.

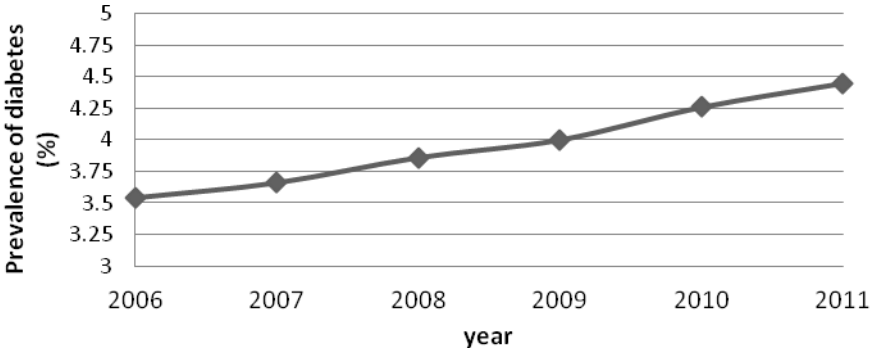
3.2 Trends in incidence, prevalence, mortality and morbidity

Globally, evidence shows an increase in the prevalence and incidence of diabetes among adults (World Health Organization, 2003). In the UK, a slow rate of increase in prevalence as well as incidence of diabetes has been observed between year 1996-2005 (González *et al.*, 2009). It was estimated in a 2012 report that 1 in 20 people in the UK have diabetes (Diabetes UK, 2012). The total number of people with diabetes in the UK Diabetes 2012 report was 2.9 million (Diabetes UK, 2012) and this number is estimated to increase by 2025 to five million. A previous UK report showed that the average prevalence of diabetes increased from 3.86% in 2008 to 4.45% in 2011 (Figure 3.1) and found that more men than women have diagnosed diabetes, as the prevalence of diabetes was 4.3% in men and 3.4% in women.

It has been concluded that the overall prevalence of diabetes increased by 54% and incidence increased by 63% in the UK between 1996 and 2005, an increase mainly attributed toT2DM, as the incidence of type 1 remained constant over the

same period time (González *et al.*, 2009). Suggested reasons for the number of T2DM include an increase in the number of overweight and obese people as well as the aging factor.

Figure 3.1 Trend of prevalence of T2DM in UK



Data obtained from (Department of Health, 2001).

Rates of morbidity and mortality are still high in people with diabetes. The WHO Multinational Study of Vascular Diseases in Diabetes reported that people with T2DM are at particularly high risk for developing cardiovascular diseases; 52% of deaths were attributed to cardiovascular diseases (Morrish *et al.*, 2001). The risk of ischemic stroke was higher by 60% for diabetic people in comparison to the general population (HR=1.56; 95%CI: 1.19, 2.05) which lead to a varied degree of disability and substantial financial cost (Sarwar *et al.*, 2010). In the same study where 102 prospective studies were included in the analyses, diabetic women had a hazard ratio of coronary heart disease significantly higher than diabetic men (women HR=2.59; 95%CI: 2.29, 2.93; men HR= 1.89; 95%CI: 1.73, 2.06). Ischemic stroke was also significantly higher among women with diabetes aged between 40 – 59 years in comparison to men with diabetes (women HR=2.83; 95%CI: 2.45, 3.40 and men HR=2.16; 95%CI: 1.84, 2.52). No explanation for the higher risk in women than men was reported.

Renal diseases account for 11% of deaths among T2DM, while women have a higher percentage of renal disease death than men (14% and 8% respectively) (Morrish *et al.*, 2001). Retinopathy is one of the main microvascular complications of diabetes; 60% of T2DM patients have been diagnosed with different stages of retinopathy (Scanlon, 2008). It has been summarized that diabetes is the most prevalent cause of blindness among people of working age in the UK (Diabetes UK, 2012).

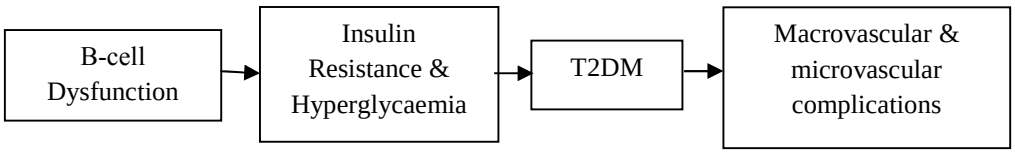
As mentioned before, diabetes complications have a tremendous impact on quality of life, financial costs and life expectancy. Diabetes accounts for 10% of hospital inpatient resources (Diabetes UK 2012b) and cost was estimated by Hex *et al.* (2012) to be about £23.7 billion. On average, ten years reduction in life expectancy was reported for patients with T2DM.

As T2DM is one of the main public health challenges with great health and cost burdens facing the whole world and in particular the UK, an increased awareness of the risk factors of type 2 is required to address this challenge. Thus, the next section will focus on the non-dietary and dietary related risk factors of T2DM.

3.3 Diagnosis and classification of type 2 diabetes mellitus

The term diabetes mellitus has been described by the World Health Organization (1999, p.2) as *"a metabolic disorder of multiple etiology, characterized by chronic hyperglycaemia with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action, or both"* . In general, β-cells in the pancreas are responsible for insulin secretion which regulates blood glucose. Dysfunction or/and destruction of pancreatic cells results in chronic hyperglycaemia that is linked to short term and long term organ damage resulting in failure and dysfunction (Figure 3.2).

Figure 3.2 Figure Pathogenesis process of type 2 diabetes mellitus



From the definition, as a result of insulin deficiency or/ and resistance, abnormalities in carbohydrate, protein and fat metabolism occur. Diabetes was categorized based on aetiology into three main groups, T2DM, type 1 diabetes mellitus (T1DM) and gestational diabetes. In addition to, other minor types of diabetes listed in Table 3.1.

Table 3.1 Classification of diabetes mellitus (World Health Organization, 1999)

1	Genetic defects of beta–cell function (maturity–onset diabetes of the young (MODY) and genetic defects in insulin action.
2	Diseases of the exocrine pancreas (Acquired processes include pancreatitis, trauma, infection, pancreatic carcinoma, and pancreatectomy).
3	Endocrinopathies (Acromegaly, Cushing's Syndrome).
4	Drug– or chemical–induced diabetes (glucocorticoids).
5	Infections (congenital rubella, Coxsackie B, cytomegalovirus).
6	Specific forms of immune–mediated diabetes mellitus (Anti–insulin receptor antibodies due to systemic lupus erythematosus and other autoimmune diseases).
7	Genetic syndromes sometimes associated with diabetes (Down's syndrome).

In terms of public health impact, T2DM is the commonest type of diabetes where evidence found a close link between T2DM and diet. Therefore, the current literature review will focus on T2DM.

Type 2 diabetes mellitus is also known as non-insulin dependent diabetes mellitus (NIDDM) which accounts for 90% of diabetes in the UK (Diabetes UK, 2012). Diagnostic values of fasting blood glucose and 2 hours post oral glucose load ingestion have been updated four times over the past 40 years as shown by WHO report (2006). Table 3.2 presents the changes in blood glucose cut-off values as reported by WHO/FAO (World Health Organization, 1999, World Health Organization, 2006).

Table 3.2 History of diagnostic criteria for diabetes mellitus

year of Publication	Fasting blood glucose		2-h blood glucose[1]
1965	Not specified		$\geq$7.2 mmol/l
1980	$\geq$8.0 mmol/l	and/or	$\geq$11.0 mmol/l
1985	$\geq$7.8 mmol/l	or	$\geq$11.1 mmol/l
1999/2006	$\geq$7.0 mmol/l	or	$\geq$11.1 mmol/l

[1]After ingestion of 75g oral glucose load

The decision to lower the values of fasting plasma glucose in the criteria for diagnosis of diabetes was based on the association of $\geq$7.0mmol/l of fasting plasma glucose level with a significantly higher risk of microvascular complications, mainly retinopathy (McCane *et al.*, 1994). The risk of coronary heart disease death rate is significantly doubled with a fasting plasma glucose level above 6.9mmol/l in comparison to a fasting blood glucose of less than 5.8mmol/l even though 2-h blood glucose is normal (RR=2.63; 95%CI: 1·72, 4·03) (Charles *et al.*, 1996).

Diabetes mellitus is a disease that develops over time and progress from norm-glycaemia; through pre-diabetes stage where higher blood glucose levels can be determine which then result into clear picture of diabetes. Several pre-diabetes

stages were identified by blood glucose levels such as impaired glucose tolerance and impaired fasting glycaemia (World Health Organization, 2006).

Epidemiological studies use values from measured fasting blood glucose concentrations and or 2-hours post 75g glucose load for diagnosis of diabetes. However to avoid the difficulties with fasting status and because of a strong correlation between 2-hour values and fasting blood glucose, 2-hour values are still used. High cost of OGTT had led to the use of fasting blood glucose in epidemiological research (World Health Organization, 1999). Diabetes UK recommended that all healthcare professionals adopt the new criteria in Table 3.3 for diabetes diagnosis from 1 June 2000.

Table 3.3 Criteria for diagnosing diabetes mellitus (World Health Organization, 2006)

1. Diabetes symptoms (i.e. polyuria[5], polydipsia[6] and unexplained weight loss) **plus** a random venous plasma glucose concentration $\geq$ 11.1 mmol/l.

   or

2. Fasting plasma glucose concentration $\geq$ 7.0 mmol/l (whole blood $\geq$6.1mmol/l).

   or

3. Two hour plasma glucose concentration $\geq$ 11.1 mmol/l two hours after 75g glucose in an oral glucose tolerance test (OGTT).

Polyuria and polydypsia are common symptoms that occur with hyperglycaemia. Asymptomatic patients may remain undiagnosed for many years (Beck *et al.*, 1995). However, diagnosis of individuals with asymptomatic hyperglycaemia can be made after individual trigger by infections, stress and acute illness. Other patients will be diagnosed when macrovascular complications such as coronary heart disease appear (World Health Organization, 1999).

An updated report on the diagnosis of T2DM was compiled by WHO (2011) which suggested that glycated haemoglobin form (Hb_{A1C}), which explains extent of blood glucose control in the past 12 weeks (National Institute of Health and Care Excellence, 2012). Hb_{A1C} can be used as a diagnostic test for T2DM with cut-off point of 6.5%. However, less than 6.5% does not exclude a diagnosis of diabetes because there is still not enough evidence available to have a formal recommendation for Hb_{A1C} below this point (World Health Organization, 2006). In addition, NICE guidelines (National Institute of Health and Care Excellence, 2012) reported that an individual with Hb_{A1C} of 6.0-6.5% should be categorized as being at

[5] Polyuria: excessive secretion of urine

[6] Poydipsia: increasing fluid intake due to increase in thirst (National Institutes of Health)

high risk of diabetes. The utility of Hb_{A1C} level in predicting incidence of diabetes has been investigated in a prospective study recently (Choi *et al.*, 2011). Women with Hb_{A1C} of more than 5.6% had a three times increased risk of T2DM over the 6-year follow-up (RR= 3.06; 95%CI: 2.46, 3.81; $p<0.01$) in comparison to less than 5.6% in the fully adjusted model. The advantage of Hb_{A1C} is that it can be performed anytime and does not need fasting status. However, it is not readily available worldwide in comparison to glucose assays (World Health Organization, 2011).

3.3.1Case ascertainment in T2DM in prospective studies

Identification of new cases of diabetes usually is a challenging aspect in large prospective studies. Several sources were reported on how T2DM incidents were collected. General practitioners, drug prescription and hospital records were some of the resources. Self-reporting diabetes mellitus, and/or current dietary treatment for diabetes in questionnaires is commonly used to identify new onset of diabetes in large prospective studies. Confirmation of diabetes diagnosis by primary care physician or medical records was reported in some prospective studies before considering subject as a case (Hopping *et al.*, 2010, Stevens *et al.*, 2002, Schulze *et al.*, 2004a). A few cohort studies reported using fasting blood glucose as another approach to identify new cases with or without the self-reports (Weng *et al.*, 2012, Barclay *et al.*, 2007). Only recently, one Swedish prospective study used Hb_{A1C} registry records to diagnosis some of the participants with new onset of diabetes (Hindy *et al.*, 2012). As mentioned before, this an expensive method for large sample sizes. Often, the validity of the subsample of cases only is checked, which leaves potential misclassification within the non-cases. Because of the lack of centralized databases for diabetes in UK, it is quite difficult to identify new onset cases, and this has a great impact on epidemiological research.

There is no national registry of diabetes in the UK that similar to the unique centralized system across the UK for cancer (Office of National Statistics; http://www.ons.gov.uk). However, there are two databases where medical information on diabetes and other diseases can be obtained for research purposes. The General Practice Research Database (GPRD) is an earlier database which provides medical data from 400 primary care practices in the UK for medical research purposes (Walley and Mantgani, 1997). Second database is the Health Improvement Network (THIN) database that contains medical information entered by primary care physicians under the terms of the UK's National Health Service

(www.epic.gov.uk) (González *et al.*, 2009). These databases may cover large proportion of UK population however, they do not cover the whole UK. Other variables that may be of interest such as ethnicity and/or socio-economic status are not available at patient level data.

3.3.2Development of type 2 diabetes and insulin resistance measurement

Progression from normal blood glucose to impaired glucose tolerance due to insulin resistance which then ends with development of diabetes is well-established pathway for T2DM. Insulin resistance measurement can help in exploring the pathophysiological in the development of diabetes.

Several insulin assessment methods were reported in a recent review, and these are listed in Table 3.4 (Singh and Saxena, 2010). Quantitative assessment of insulin sensitivity helps in evaluating the pathogenesis and the causes of several diseases. Insulinogenic index and HOMA were both used in some of epidemiological studies to determine insulin resistance, however the hyperinsulinemic euglycemic glucose clamp is the gold standard. Different studies use different methods in measuring insulin making it difficult to look across studies.

Table 3.4 Insulin resistance measurements used for researches

Method	Description
Hyperinsulinemic euglycemic glucose clamp	The gold standard method for determining insulin resistance however it is difficult to apply for large epidemiological studies
Fasting insulin	Most practical method to measure IR
Glucose/insulin ratio (G/I ratio)	Index of insulin resistance
Insulinogenic index (IGI)	Index of insulin secretion after OGTT
HOMA	Reflects the relationship between glucose and insulin by measuring insulin and glucose concentration
Fasting insulin resistance index	Measure fasting glucose and insulin and apply values in equation which estimate IR

Adapted from (Singh and Saxena, 2010)

3.4 Metabolic syndrome

A review by Eckel *et al.* (2005) reported that metabolic syndrome is a metabolic disorder which has been diagnosed with different criteria by several organizations. However insulin resistance and fatty acids dysfunction were the main aspects in pathophysiology process. Metabolic syndrome found to be a predictor for T2DM. The suggested underlying mechanism was increasing the release of fatty acids derived from adipose tissue which accompanied with insulin resistance. Insulin resistance refers to the inability of body organs such as liver and muscles to

respond to insulin, resulting in an inability to increase glucose uptake by these organs and inability to decrease hepatic gluconeogensis (Chaplin, 2005). Impairment of insulin sensitivity with hypertension and dyslipidemia were all present in metabolic syndrome.

3.5 Non-dietary related T2DM risk factors

Some of the following risk factors were considered to be non-modifiable risk factors as they cannot be changed by behavioural modifications. These are family history of diabetes, ethnicity, age and gender. However, other factors which also important in the development of T2DM are smoking and physical activity that can be modified.

*3.5.1 Famil*y *history and genetic factors*

One evidence-based review suggested that people with genetic predisposing factors were more likely to progress to T2DM (National Institute of Health and Care Excellence, 2012). A study on twins showed that genes play a great role in the development of T2DM, where chance of T2DM in the second twin reached 83% (Japan Diabetes Society, 1988). A positive family history of diabetes could reflect a genetic liability of an individual. Genes that have been associated with T2DM were demonstrated in previous review (Adeghate *et al.*, 2006) . For example, IAPP gene defect (islet amyloid polypeptide), which is expressed in β-cell that lead to impairment of the pancreatic tissue function and glucose intolerance. Second example is the insulin receptors and insulin-regulated glucose transporter (GLUT4) genes defect, which results in abnormal expression of receptors in insulin tissue and decreased number of insulin transporters leading to insulin resistance (Bell, 1991). Evidence suggests the important role of environmental factors modifications despite the presence of genetic susceptibility in the individual can prevent the progression of the disease (Lindström *et al.*, 2003).

Data from five prospective studies reported that the risk of diabetes increased with a positive family history of diabetes among first degree relatives and reported about two to six times increased risk of T2DM among participants with a positive family history in comparison to those with a negative family history (Harrison *et al.*, 2003). Some of those studies in that review found that participants with a positive parental history of diabetes engaged in some healthy behaviours and were more likely to be screened for diabetes than control people, which may partly explain the higher number of cases with a positive family history other than for genetic reasons.

It also may point towards the strong effect of genetic factors on developing diabetes. A recent UK Diabetes statistical report showed that parental history of diabetes has a strong link with the development of T2DM. There was a 15 % risk of T2DM when a single parent has a history of diabetes, while if both parents have a history of diabetes then the risk of diabetes is much greater (75%). In addition, the highest probability of diabetes was reported among identical twin (90%) (Diabetes UK, 2009). This indicates the strong influence of genetic factors on the development of diabetes.

3.5.2 Age and gender

A well-established positive association between age and the risk of diabetes has been reported in a previous study (Wild *et al.*, 2004). The UK Office for National Statistics recently reported that numbers of older people continue to increase in the UK. A particularly fast increase in number was noticed among population segment aged over 85 years. In addition the median age of the UK population increased from 1985 to 2010 (from 35.4 years to 39.7years; respectively) (Office for National Statistics, 2012). It was reported that one in twenty people have diabetes among those aged above 65 years while one in five people have diabetes among those people aged above 85 years (Department of Health, 2001). During the last decade the onset of diabetes has moved to younger age groups, which suggests a change in environmental risk factors (Rosenbloom *et al.*, 1999).

Worldwide, the increase in prevalence of diabetes with age was found to be similar for both genders, however the prevalence was much higher among women aged above 65 years than men (Wild *et al.*, 2004). In addition, the risks of death and morbidity were higher among diabetic women in comparison to diabetic men (Sarwar *et al.*, 2010). This suggested the effect of hormonal changes which may link to higher prevalence of T2DM via substantial weight gain in women after the menopause. Evidence suggests changes in body composition include increased body fat mass and decreased lean body mass in postmenopausal women in comparison to premenopausal women possibly linked to impairment in insulin sensitivity (Szmuilowicz *et al.*, 2009). In addition, evidence showed association between hormonal therapy use and improvement of insulin sensitivity (Lobo, 2008).

3.5.3 Ethnicity

Globally, the prevalence of T2DM differs among populations of different ethnic origin (Adeghate *et al.*, 2006). The Department of Health (2001) reported

that in the UK, T2DM is six fold higher in the South Asian population and three fold higher among African and African-Caribbean population in comparison to the white population. Ethnicity was considered in recent National Institute for Health and Care Excellence guidelines (National Institute of Health and Care Excellence, 2012) to identify populations at risk of developing T2DM as particular ethnic populations are affected by T2DM at a younger age than white European population.

3.5.4Smoking

Recent data from prospective studies suggest a positive association between smoking and diabetes (Willi *et al.*, 2007). Heavy smokers (>20 cigarettes/day) have a 61% increased risk of diabetes compared with light smokers (<20 cigarettes/day) and the risk among active smokers is 44% higher in comparison to non-smokers (RR= 1.44; 95%CI: 1.31, 1.58). Risk remain elevated, even after stopping smoking, ex-smokers have a 23% higher risk of diabetes in comparison to never smokers (RR=1.23; 95%CI: 1.14, 1.33). The explanation of the higher risk of diabetes after a short term smoking cessation is still unclear (Yeh *et al.*, 2010). There are limited data on long term smoking cessation effects on the development of T2DM (Tonstad, 2009). Tobacco smoke contains many toxins that may directly damage pancreatic tissue (Rimm *et al.*, 1993). Additionally, other behavioural factors such as physical inactivity accompanying smoking were suggested to be the underlining reason of association between diabetes and smoking (Will *et al.*, 2001).

3.5.5Physical activity

In general, sufficient evidence indicates that physical activity has a protective effect against T2DM (World Health Organization, 2003, National Institute of Health and Care Excellence, 2012). A meta–analysis including 10 prospective studies from the USA, Japan, the UK and other European countries reported an overall significant risk reduction of T2DM by 31% (RR= 0.69; 95% CI: 0.58, 0.83) among those who regularly exercised with moderate intensity in comparison to those leading a sedentary life. Also, the risk of diabetes is reduced by 30% (95%CI: 16%, 42%) among those who reported regular walking (>2.5hours/week) in comparison to those with no or minimal walking (Jeon *et al.*, 2007).

Prospective studies have evaluated potential sex differences in relation to the effect of physical activity and risk of diabetes (Meisinger *et al.*, 2005, Jeon *et al.*, 2007). An inverse relationship between physical activity (>2 h per week vs. no activity) and the risk of diabetes was significant in women who participated in the

MONICA/ KORA Augsburg Cohort study (women: 0.24; 95%CI: 0.06, 0.98 vs. men:0.83; 95%CI: 0.50, 1.36) however the risk of reduction was varied, there were strong gender differences (Meisinger *et al.*, 2005). Differences in intensity of physical activity and a threshold protective effect may partly explain the difference in gender. On the other hand, no significance difference between males and females (p=0.17) was reported in recent pooled data from 10 cohort studies (Jeon *et al.*, 2007).

Researchers has explored whether exercise is particularly beneficial in obese rather than lean women in limited number of cohort studies. Among U.S nurses who reported vigorous exercise, the risk of diabetes was reduced significantly by 21% in women with a BMI of more than 27kg/m^2 and 27% in women with a BMI of less than 27kg/m^2 in comparison to those who did not report regular exercise (Manson *et al.*, 1991). In another cohort study, women with a BMI of less than 30kg/m^2 who exercised at least once a week showed a substantial risk reduction of 76% (HR=0.24; 95%CI: 0.09, 0.65) in comparison to inactive women. However the beneficial effect of one hour/week of moderate to high level of physical activity among obese women was not observed (Meisinger *et al.*, 2005).

There are a number of potential mechanisms discussed in the previous studies which may explain the beneficial effect of physical activity in reducing the risk of T2DM. Improvements in insulin sensitivity and glycemic control among non-diabetic and diabetic populations who regularly exercise was reported in the Insulin Resistance Atherosclerosis Study (Mayer-Davis *et al.*, 1998). An increase in vigorous exercise (metabolic equivalent level $\geq$6) as well as non-vigorous exercise (metabolic equivalent level <6) both showed similar increases in insulin sensitivity after adjusting for BMI and WHR. For every one unit decrease in BMI, a 3.2% increase in insulin sensitivity was reported. It was suggested that body fatness may partly mediate the effect of physical activity on insulin sensitivity thus reducing the risk of diabetes.

Many trials have examined the effectiveness of physical activity on delay or prevention of the development of diabetes among participants at greater risk, such as the obese, those with impaired glucose tolerance or a history of gestational diabetes. A recently published Cochrane review including eight trials, found that exercise was vs. Standard was not beneficial (Pooled estimate =0.69; 95%CI: 0.29, 1.65) and diet vs. exercise was also not beneficial (pooled estimate = 0.69; 95%CI: 0.37, 1.29), but

when exercise combined with diet reduced the risk of diabetes by 37% (RR 0.63; 95%CI: 0.49, 0.79) compared with standard recommendations (Orozco *et al.*, 2008).

Many trials have investigated the frequency of physical activity and intensity of physical activity in relation to development of T2DM, as summarized by NICE guidelines (National Institute of Health and Care Excellence, 2012). Increase level of physical activity of at least 150 minutes of moderate intensity per week is needed to reduce the risk of T2DM. Participants who increased their level of physical activity reduced their risk of diabetes by 51% independently of weight loss. Thus evidence is consistent on the beneficial effect of physical activity on the development of diabetes.

3.6 Diet related T2DM risk factors

Many factors are attributed to the development of T2DM. It is important to consider the environmental factors that have an important role in diabetes development. Modifiable diabetic risk factors are mainly dietary-related factors.

3.6.1Obesity and risk of type 2 diabetes mellitus

3.6.1.1 BMI versus other measurement of adiposity

Epidemiological studies use body mass index (BMI), waist circumference (WC), and waist hip ratio (WHR) to reflect body fatness among studied populations. All are useful assessment tools for both sexes and for all ages of adults (Vazquez *et al.*, 2007). *"Abnormal or excessive fat accumulation that may impair health"* is the WHO definition of overweight status and obesity (World Health Organization, 2013). Body mass index is a simple index of weight-for-height that is widely used to classify adults into underweight, normal, overweight and obesity (World Health Organization, 2003), however it is argued that the association between body fatness % and BMI is not considered to be strong especially with BMI between 20-25kg/m^2 and such association is affected by age (Meeuwsen *et al.*, 2010). Another study found changes in BMI in adults strongly predicts changes in fat mass rather than lean mass and suggested it could be used for monitoring purposes in elderly black and white women (Arngrímsson *et al.*, 2009).

A meta-analysis compared between BMI which reflects general obesity and WC and WHR which reflect central obesity in relation to risk of diabetes (Vazquez *et al.*, 2007). All three parameters have similar associations with incident diabetes from pooled data from 32 cohort studies, for every one standard deviation increment of obesity markers, for BMI, the relative risk =1.87(95%CI: 1.67, 2.10), for WC

relative risk =1.87 (95%CI: 1.58, 2.20) and for WHR relative risk =1.88(95%CI: 1.61, 2.19). An earlier cohort of US middle-aged women followed for 14 years reported that body mass index (BMI) was the dominant predictor of T2DM and BMI between 23-25kg/m^2 was associated with four times higher risk for T2DM compared to women with BMI less than 22kg/m^2 (Colditz *et al.*, 1995). It is probable therefore that the BMI range of 18.5 – 24.9kg/m^2 is not appropriate for all population. From all of this, categorizing people into groups based on BMI is false setting of boundaries. The risk of T2DM is likely to be linear with the increase in body fat parameters. Therefore, when people are categorized into groups, the potential of losing linearity is high.

In a meta-analysis of 18 prospective studies reported by Abdullah *et al.* (2010) that the risk of diabetes is 7 times higher among obese participants (RR=7.19, 95% CI: 5.74, 9.00) and almost 3 times higher among overweight participants (RR=2.99, 95% CI: 2.42, 3.72) in comparison to normal weight. They also found some evidence of gender difference in relation to obesity (Abdullah *et al.*, 2010).

On the other hand, WC is also a measure of excess central fat and considered as a stronger predictor of diabetes than BMI among older women in the British Regional Heart Study and the British Women's Heart and Health study (Wannamethee *et al.*, 2010). In addition, comparison between highest vs. lowest quartile, the adjusted risk of diabetes was higher for WC (RR=12.18; 95% CI 4.83, 30.74) than BMI (RR= 4.10; 95% CI 2.16, 7.79). No clear explanation was reported in this study which requires further research.

A recent study among a white European population with mean BMI= 25(4)kg/m^2 showed that for every 1 kg/m^2increase in BMI, the risk of T2DM increased by 8.4% (RR= 1.08; 95%CI: 1.03, 1.14; $p<0.01$) and with every 1 cm increased in WC, the risk of T2DM increased by 3.2% (RR= 1.03; 95%CI: 1.01, 1.05; $p<0.01$) (Bombelli *et al.*, 2011) after adjustment for potential confounders. The authors conclude that BMI and WC were independently predicting incidence of T2DM.

3.6.1.2 Weight changes and risk of type 2 diabetes mellitus

The European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study (Schienkiewitz *et al.*, 2006) assessed weight changes over lifetime in relation to the risk of T2DM among women. With every unit increment of BMI, the risk of T2DM was higher if the weight gain occurred in early life (25-40 years) than

later life (40-55 years) (24% and 11% respectively). The risk of T2DM was 4 times higher among women who gain weight (above 4kg/m2) in early life as compared to women who had stable weight in early life but gained weight in later life. It has been suggested that longer accumulative effect of excess weight at early adulthood is a critical, where parity and obesity may both play a role in development of diabetes.

Another cohort study (Oguma *et al.*, 2005) investigated weight changes over time in comparison to initial BMI in relation to the risk of T2DM. This cohort, limited to men with a mean age of 46 years, found that both initial BMI and weight change from university entry to baseline showed positive associations with the risk of T2DM (p for trend <0.01). Among men with BMI at university between 21 to 23kg/m^2, those with severe weight gain have 9 times higher risk of T2DM (RR=9.57; 95%CI:5.90, 15.53) in comparison to men with stable weight in the same group. While, among men with initial BMI of less 21kg/m^2, the risk of diabetes was 7 times higher among men with severe weight gain (RR=7.68; 95%CI: 4.72, 12.5) in comparison to those with stable weight in the same group. It was suggested that overweight men also are likely to have health problems and more likely to be screened for diabetes.

Overall, obesity is strongly linked to development of diabetes. BMI found to be independent risk factor of diabetes and weight changes in another strong useful predictor of diabetes.

3.6.1.3 Weight loss effect on risk of diabetes

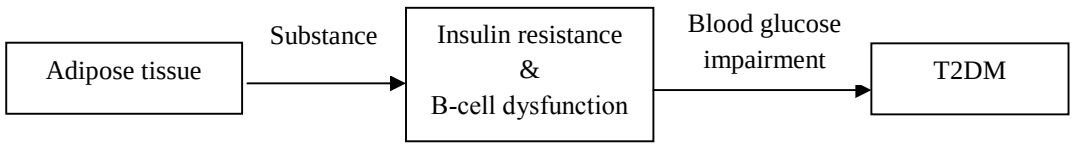
Effectiveness of weight loss by diet and physical activity on preventing risk of T2DM has been examined in trials. The Diabetes Prevention Program study that showed lifestyle intervention (weight loss, reduction in total fat and saturated fat intakes, increase in dietary fibre and an increase in physical activity) over an average of 2.8 years is superior to pharmaceutical approaches in reducing the incidence of diabetes (by 58% and 31%) in comparison to control group (Knowler *et al.*, 2002). This may suggest potential for reversing the epidemic of diabetes. Convincing evidence from randomized clinical trials showed weight loss of 5-7% from initial weight after lifestyle intervention (physical activity and diet) in overweight and obese people has reduced the risk of T2DM by 58% (Tuomilehto *et al.*, 2001) and even four years post intervention, the results showed a sustained risk reduction of T2DM by 36% (Lindstrom *et al.*, 2006).

It was reported in a human study, that energy restriction (less than 1,000kcal/day with 45% carbohydrates, 20% fat, 35% protein) resulted into weight loss (about 3 kg) and low fat mass after 3 weeks of intervention and this was found to be associated with lowered inflammatory markers that may lead to improvement of insulin sensitivity (Bastard *et al.*, 2000).

3.6.1.4 Underlying mechanism linking adiposity and type 2 diabetes mellitus

The possible underlying mechanisms linking obesity to T2DM were demonstrated in a review (Kahn *et al.*, 2006). The increases in the release of different substances which adversely affect insulin action have a crucial role in pathogenesis of T2DM. In simple way, it was found that adipose tissue increases the release of substances such as non-esterified fatty acids (NEFA) resulting insulin resistance and impairment of β-cell function (Figure 3.3). The increase in NEFA release modulates insulin sensitivity leading to defects in insulin release and resulting in dysregulation of glucose. It was found that when fasting hyperglycaemia manifested, pancreatic cell function has decreased by three quarters (Kahn *et al.*, 2006).

Figure.3.1 The relationship between adipose tissue, insulin resistance and T2DM



In conclusion, there is robust evidence of a strong relation between obesity and risk of diabetes and effectiveness of weight loss on reducing or delaying the risk of T2DM. From public health prospective, the above evidence shows that it is important to emphasize three main aspects of obesity and risk of T2DM. Prevention or delaying the progression to T2DM may possibly be achieved by avoiding excess weight gain during life; maintaining healthy weight and the loss excess weight.

3.6.2Macronutrient intakes and risk of type 2 diabetes mellitus

3.6.2.1 Carbohydrate intake and risk of type 2 diabetes mellitus

The majority of energy in human diets comes from carbohydrates (FAO, 1998).

3.6.2.1.1 Structure definition of carbohydrates

3.6.2.1.1.1 Total carbohydrates and type 2 diabetes mellitus

Several studies have looked at total carbohydrates and risk of T2DM. It was reported earlier (Colditz *et al.*, 1992) that total carbohydrates was not related to the risk of T2DM among female nurses aged 30-55 years living in U.S over a period of 6 years. A later prospective study examined the risk of T2DM among nurses between (1986-1992) from a more extended FFQ in 1986 (134 food items was 61 food items) and found no clear relationship between the risk of diabetes and total carbohydrate intake among these women (Salmeron *et al.*, 1997b). This finding did not differ among male health professionals (Salmeron *et al.*, 1997a). Similarly no association was reported among postmenopausal Iowa women and middle aged US women after adjustment for potential confounders (Meyer *et al.*, 2000, Schulze *et al.*, 2004a). Even with total carbohydrate intake obtained from 7 days diary, the association with risk of T2DM is not apparent (Ahmadi-Abhari *et al.*, 2013). Prospective studies which examined the percentage of energy from carbohydrates in relation to risk of T2DM reported no significant association after adjustment of potential confounders comparing extreme quintiles (Schulze *et al.*, 2004a, Schulze *et al.*, 2008).

In conclusion, there is no clear association between total carbohydrates intake and risk of T2DM, possibly because the term covers a wide range of different components which could be meaningless, except that dietary recommendation suggest carbohydrates intake as a percentage of energy. This leaves unclear answer for the recommended carbohydrates as energy percentage in relation to risk of T2DM.

3.6.2.1.1.2 Monosaccharides and disaccharides intakes with risk of type 2 diabetes mellitus

Total sugars intake was examined in relation to risk of T2DM in a few prospective studies (Barclay *et al.*, 2007, Hodge *et al.*, 2004). Among Australian participants, no relation was seen between risks of T2DM with every 100g/day of sugar consumed (Barclay *et al.*, 2007). While in Melbourne Collaborative Cohort Study, the same increment of total sugar (100g/day) intake was found to have significant risk reduction of T2DM (OR=0.61, 0.47– 0.79) (Hodge *et al.*, 2004). After 12 years of follow up in Finnish study, the risk of T2DM was not associated with intake of sugar estimated from dietary history interview (p for trend =0.1)

(Montonen *et al.*, 2007). In conclusion, the few studies that assessed intake of sugars and risk of T2DM showed inconsistent results and no conclusion can be drawn.

In terms of specific sugars, few cohort studies have reported the relationship between dietary glucose or fructose or sucrose intakes and risk of T2DM. Two prospective studies reported the risk of T2DM was significantly increased (30% and 68%) with high intake of glucose in comparison to lower intakes group (Meyer *et al.*, 2000, Montonen *et al.*, 2007) while, the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam did not show any association (Schulze *et al.*, 2008). For fructose intake, the risk of T2DM was significantly higher in the highest group in comparison to lowest group among Iowa women (Meyer *et al.*, 2000) and men and women participating in the Finnish Mobile Clinic Health Examination survey (Montonen *et al.*, 2007) but not among participants in EPIC-Potsdam (Schulze *et al.*, 2008). Among Iowan postmenopausal women, the highest sucrose consumers had lower risk of T2DM (RR=0.81; 95%CI: 0.67, 0.99) in comparison to the lowest sucrose intake group (Meyer *et al.*, 2000). However, other cohorts did not find any association between sucrose and T2DM (Montonen *et al.*, 2007, Schulze *et al.*, 2008).

From the above cohort studies, the role of glucose, fructose and sucrose intakes in the development of T2DM is therefore inconclusive. A recent systematic review (Sonestedt *et al.*, 2012) showed inconsistency in the results and "*a limited no conclusion grade"* was given by the authors.

3.6.2.1.1.3 Starch intake and risk of type 2 diabetes mellitus

Similar to simple sugars, dietary starch intake was investigated in relatively few cohort studies. Melbourne Collaborative Cohort Study reported that with every 100 g/day of starch, the risk of T2DM significantly increased (multivariate adjusted OR=1.47 (95%CI: 1.06, 2.05) and OR=1.52 (95%CI: 1.09, 2.11) with further adjustment for BMI and WHR (Hodge *et al.*, 2004). Other cohort studies did not find any significant association between starch intake and risk of T2DM (Schulze *et al.*, 2008, Meyer *et al.*, 2000). A limited number of epidemiological studies have investigated the intakes of starch with risk of diabetes and are inconsistent in their findings.

From the carbohydrate structure definition, it can be appreciated that from the few prospective studies that examined the relationship between risk of T2DM and simple sugars or starch did not reach into clear conclusion because of inconsistency

in their findings. The fibre intake and type 2 diabetes mellitus will be discussed in more detail in chapter 4 as this is one of the objective of the current thesis.

3.6.2.1.2 Physiological definition of carbohydrates and risk of type 2 diabetes mellitus

3.6.2.1.2.1 Glycaemic index and glycaemic load

The concept of Glycemic index (GI) was introduced by Jenkins and his colleagues (Jenkins *et al.*, 1981) to measure the change in blood glucose following ingestion of carbohydrate containing foods. Glycemic load (GL) was also introduced and used (Salmeron *et al.*, 1997a) as an indicator of the glucose response and insulin demand produced by a total amount of carbohydrates. Many cohort studies have evaluated the association between GI, GL and chronic disease risk. A meta-analysis summarized seven prospective studies, most of them involving American females (Barclay *et al.*, 2008). Pooled data from fully adjusted models showed an overall significant positive association between GI and risk of T2DM (RR=1.40; 95%CI: 1.23, 1.59) and a significant positive association between GL and risk of T2DM (RR= 1.27; 1.12, 1.45) when comparing highest with lowest groups. This showed the adverse effect of diet with high GI and/or high GL as the risk of T2DM increased by 40% and 27% respectively independent of cereal fibre intake. Another meta-analysis (Livesey *et al.*, 2013) found that healthy people who consumed diets with high GL seemed to have a 45% increased risk of T2DM with every 100g increment in GL.

3.6.2.1.2.2 Whole grain versus not: refinement

Whole grain may be defined as "*whole grain foods contain intact, flaked or broken grain kernels, coarsely ground kernels or flour that made from whole gains"*(Priebe *et al.*, 2008, p.3). Whole grains have become a focus of attention where earlier hypothesis by Burkitt (1975) on replacing whole grain with refined gains is related to the development of non-communicable diseases such as diabetes.

Many epidemiological studies reported the beneficial effect of whole grain intake on the prevention of T2DM. A recent review by Priebe *et al.* (2008) showed risk reduction of T2DM for high intake of wholegrain by 27% to 30% in comparison to low wholegrain intake. A recent meta-analysis (Ye *et al.*, 2012) reported that intake of 3-5 servings of wholegrain equivalent to 48-50g was associated with a lower risk of diabetes by 26% (RR=0.74, 95%CI: 0.69, 0.80) in comparison to never or rare consumers. In the same meta-analysis, data from randomized controlled trials

showed that wholegrain consumption for 4-16 weeks duration, reduced fasting blood glucose concentration among healthy participants and participants who had one or more major risk factors for T2DM or cardiovascular disease in comparison to control groups (mean difference -0.93mmol/l; 95%CI: -1.65, -0.21).

Possible explanation of the findings can be that refined cereals have a difference composition than whole grain due to processing methods that lead to removal of important nutrients from the cereal such as bran where phenolic compounds are rich (Slavin *et al.*, 1999).

In conclusion, evidence is consistent regarding the beneficial effect of wholegrain foods on the prevention of T2DM. It has been suggested that the health related beneficial effect of wholegrain cereals may be related to the presence of bioactive compounds such as minerals, vitamins, lignin and phenolic compounds in addition to fibre (Slavin, 2003b).

3.6.2.1.2.3 Liquid versus solid: Does it matter

Evidence showed a great increase in consumption of sugar and sweetened beverages (SSB) worldwide among different age groups (Malik *et al.*, 2010a). Beverages such as soft drinks, fruit drinks, energy and vitamins water drinks, are a source of added sugar, contain sucrose, high fructose syrup or fruit juice concentrates.

Meta-analysis of 9 cohort studies showed a significant high risk of T2DM by 26% (RR=1.26; 95%CI: 1.12, 1.41) with high intake of SSB in comparison to the lowest quintile. Further analysis showed that with every 1 severing of SSB (12-oz) per day, the risk of T2DM significantly increased by 25% (95%CI: 1.10, 1.42) (Malik *et al.*, 2010b). At the same time another review by Malik *et al.* (2010a) reported a significant positive association between weight gain and risk of T2DM in prospective studies. Hence, it was suggested that positive association between SSB intake and risk of T2DM may be partly attributed to high energy intake and obesity.

In Multi-Ethnic Study of Atherosclerosis study, when the relationship between diet soda and risk of diabetes assessed, the risk of T2DM was positively associated with diet soda even after BMI adjustment. However, the high intake of diet soda was not significantly associated with metabolic syndrome after BMI adjustment (Nettleton *et al.*, 2009). The pooled estimate from dose-response meta-analysis before and after excluding coefficients with energy and obesity adjustment showed a

raised effect size if BMI and energy were not adjusted in comparison to estimate adjusted for BMI and energy (RR=1.35 and 1.18 respectively) (Malik *et al.*, 2010b).

Cohort studies are liable for reverse causation, where false association between T2DM and intake of diet soda after BMI adjustment occurs because of possibility of having overweight subject who may end with cutting down calories by drinking beverages with artificial sweeteners which may result into a positive association between diet soda and T2DM. Artificial sweeteners contained SSB seem to be linked with obesity even if it contains zero calories (artificial sweeteners) as reported in recent EPIC-InterAct study (Romaguera *et al.,* 2013). It was suggested earlier by Hu and Malik (2010) that artificial sweeteners may tend to increase the desire for overconsumption of food. Earlier study by Schulze *et al.* (2004b) reported significant positive associations between drinking beverages and both weight gain as well as risk of T2DM among young and middle-aged women. Another possible explanation is that consumption of SSB tends to be grouped with other unhealthy dietary factors such as high fat and low fibre intakes, therefore overestimation of effect could possibly occurred from residual confounders.

From the carbohydrates section, the most evidence points toward the important of carbohydrate form. This due to the presence of strong evidence which suggest a positive association between sweet beverages and risk of T2DM while, evidence on the effect of carbohydrates from food on T2DM varies. This does not rule out the importance of other aspects as stated by UK Diabetes guidelines (Diabetes UK, 2011) that diets of low glycaemic index or low in glycaemic load and higher in dietary fibre and wholegrain reduce the risk of T2DM. Carbohydrate classifications based on chemical characteristics have the advantage of appropriate measurements of the nutrient. However, the complexity occurs when this needs to be translated into nutritional effects and health effects. This because of the wide variation in physiological properties found in each class of carbohydrates (Cummings and Stephen, 2007). Therefore pooled data on observational studies that evaluate carbohydrates' components and the risk of T2DM is quite challenging in term of terminology used to define the exposure. Dietary guideline recommend carbohydrates as percentage of energy, however this is not very clear regards risk of T2DM

3.6.2.2 Dietary fat intake

Many organizations agree that maintaining a low intake of total fat, saturated fat may be beneficial to reduce the risk of T2DM (National Institute of Health and Care Excellence, 2012, American Diabetes Association (ADA), 2008). This is mainly based on the strength of evidence[7] in a WHO report (World Health Organization, 2003) which showed that saturated fat intake probably[8] increases the risk of T2DM while total fat and trans-fatty acids may possibly[9] increase the risk of T2DM.

3.6.2.2.1 Total fat intake

Prospective studies of women have found the association between the risk of diabetes and total fat and saturated fat was not statistically significant after adjustment for age, BMI, WHR, physical activity, alcohol intake, and smoking (Meyer *et al.*, 2001, Salmerón *et al.*, 2001). The Health Professionals Follow-Up Study prospectively found significant positive associations where relative risk for extreme quintiles for total fat intake was 1.27 (95% CI: 1.04 ,1.55, P trend= 0.02) and relative risk of diabetes for saturated fat was 1.34 (95%CI: 1.09, 1.66, P trend=0.01) but when the risk was adjusted further for BMI, the association of total fat and saturated fat intakes with the risk of diabetes was not significant (Van Dam *et al.*, 2002).

3.6.2.2.2 Fat types intake

Many observational studies have reported an inverse association between polyunsaturated fatty acids (PUFA) with the development of T2DM and four studies in the systematic review found a significant positive association between trans-fatty acids and the risk of T2DM (Murakami *et al.*, 2005).

---

[7] Strength of evidence is based on modified criteria used by the World Cancer Research Fund (World Health Organization, 2003) as listed below:

[8] '***Probable evidence*** *is based on epidemiological studies showing fairly consistent associations between exposure and disease, but where there are perceived shortcomings in the available evidence or some evidence to the contrary, which precludes a more definite judgement. Shortcomings in the evidence may be any of the following: insufficient duration of trials (or studies); insufficient trials (or studies) available; inadequate sample sizes; incomplete follow-up. Laboratory evidence is usually supportive. The association should be biologically plausible'page.55.*

[9] '***Possible evidence*** *is based mainly on findings from case-control and cross-sectional studies. Insufficient randomized controlled trials, observational studies or non-randomized controlled trials are available. Evidence based on non-epidemiological studies, such as clinical and laboratory investigations, is supportive. More trials are required to support the tentative associations, which should also be biologically plausible'page.55.*

The ratio of polyunsaturated fatty acids to saturated fat has been used a marker of fat quality (Harding et al., 2004). In the European Prospective study, the diet of men and women aged 40–78 years was estimated by semi-quantitative FFQ. After 3-7 years of follow up, an increase in the ratio expressed as per standard deviation change (0.22) was associated with a reduced risk of diabetes by 12% after adjustment for age, sex, family history of diabetes, smoking, physical activity, total fat, protein, and alcohol (Harding *et al.*, 2004). The association become not statistically significant (OR = 0.91, 95% CI: 0.81, 1.03) after further adjustment of BMI and WHR.

3.6.2.2.3 Is the effect of fat mediated by obesity

Prospective studies suggested that obesity may mediate the effect of fat and saturated fat and the development of diabetes in the causal pathway (Meyer *et al.*, 2001, Van Dam *et al.*, 2002).

A cross-sectional study found that habitual dietary fat intake inversely and significantly associated with insulin sensitivity; however, no significant results were found after BMI adjustment. A significant inverse association was found only among obese participants (p=0.03) and not for non-obese participants (p=0.16) (Mayer-Davis *et al.*, 1997). This may support the assumption of obesity as an intermediate stage in the possible causal pathway that links quality of fat to the risk of diabetes. On the other hand, recent European interventional study on 417 participants with metabolic syndrome reported that isoenergetic diet with reduced saturated fat showed no effect on fasting insulin, insulin sensitivity and glucose concentration after 12 weeks (Tierney *et al.*, 2011). However, biological variation and short interventional period were some of the suggested explanation.

Overall, observational studies were relatively consistent in replacing saturated fat with PUFA to reduce the risk of diabetes. As demonstrated in a recent review, the underlying link between dietary fatty acids and incidence of diabetes is unclear (Risérus *et al.*, 2009). It was suggested in many of the studies that obesity may act as an intermediate step in the possible causal pathway between dietary fat intake and the risk of diabetes.

Review from experimental studies reported fat types have direct effect on insulin action and indirect effects mediated by obesity (Storlien *et al.*, 2000). Changes in cell membrane fatty acids composition resulted into changes in insulin action. If more saturated fatty acids found in cell membrane, then adverse effect on

insulin binding and action was reported, while if unsaturated fatty acids were the predominant in the cell membranes, this suggested being beneficial to insulin. Another potential mechanism is indirect effect of fat on insulin action through obesity where PUFA intake more likely to be utilized by human body for energy while saturated fat intake is more likely to be stored in the adipose tissues.

3.6.2.3 Protein intake

The prospective evidence based is minimal discussing the relationship between protein intake and risk of T2DM (Halton *et al.*, 2008, Sluijs *et al.*, 2010a). This is inappropriate to draw a conclusion on the effect on dietary protein intake and risk of T2DM.

3.6.2.4 Alcohol intake

Consumption of a moderate amount of alcohol is recommended by many nutritionally based organizations which aim to prevent T2DM (American Diabetes Association (ADA), 2008, National Institute of Health and Care Excellence, 2012). On the other hand, the WHO (World Health Organization, 2003) reported insufficient evidence[10] on the relationship between excess alcohol intake and increased risk of T2DM. An earlier cohort study (Stampfer *et al.*, 1988) reported that 85,051 female nurses aged 34 to 59 years who consumed more than 15g/day (more than 10 drinks/week) of alcohol had a significant risk reduction of 40% (RR=0.6; 95%CI: 0.3, 0.9) in comparison to non-consumers.

Whether the association between alcohol intake and risk of diabetes is modified by gender was examined in the Atherosclerosis Risk in Communities Study (Kao *et al.*, 2001). Middle aged men and women 12,261 followed between 3 to 6 years showed a significantly increased risk of diabetes among men who consumed more than 21 drinks per week in comparison to men who consumed less than one drink per week but not for women. Differences in beverage preferences could partly explain the differences in the results between men and women as this evidence showed that intake of wine was associated with healthier lifestyle which was more likely to be consumed by women. However, spirits and beer intake which were more likely to be consumed by men, were likely to have unhealthy lifestyle characteristics.

---

[10] **Insufficient evidence** referred to *"evidence based on findings of a few studies which are suggestive, but are insufficient to establish an association between exposure and disease. Limited or no evidence is available from randomized controlled trials. More well designed research is required to support the tentative associations. Page 55"*.

The most recent meta-analysis (Baliunas *et al.*, 2009) that included 20 cohort studies in the pooled analysis, supports the U-shaped relationship between alcohol consumption and the risk of diabetes. Compared with women who exhibited lifelong abstinence from alcohol, risk was reduced by 40% among those who consumed 24g/day of alcohol while a protective effect was not seen among those with 50g/day alcohol consumption.

The underlying mechanisms of protective effect of moderate alcohol consumption on diabetes was reported in a recent review (Rehm *et al.*, 2010). Ethanol was found to improve insulin sensitivity and lower plasma insulin level. Additionally, the anti-inflammatory effect of ethanol is another plausible mechanism. Still not very clear, as moderate amount of alcohol intake can be protective, higher intake found to be link to increased body weight.

Overall, the evidence on the beneficial effect of moderate alcohol consumption on the risk of T2DM supported by prospective studies and a concern on the margins between sufficient and limited consumption requires further research. There is still an issue regarding the extreme consumption of alcohol, whether this clusters with other unhealthy behaviours or is a high in itself.

3.6.3Micronutrient intakes and risk of type 2 diabetes mellitus

In 2003 the World Health Organization regarded the evidence that micronutrients might be implicated in development of T2DM to be inconclusive. However, since that time, further summaries of the epidemiological literature have suggested links with some micronutrient in particular, magnesium and vitamin D.

3.6.3.1 Magnesium intake and risk of type 2 diabetes

Recent evidence on magnesium intake and the risk of T2DM was reported in a meta-analysis (Larsson and Wolk, 2007). Pooled data from seven cohort studies where four studies included only women, found that overall diabetes risk was significantly reduced by 14% (RR=0.86; 95%CI: 0.77, 0.95) for a 100 mg/day increase in magnesium intake after adjustment for age, sex, BMI, physical activity and alcohol intake However heterogeneity between the studies was significant with I^2 = 72.3% (p=0.003). An inverse relationship between magnesium intake and the development of diabetes persisted even after adjustment of cereal fibre intake or/and wholegrain intake in three cohort studies (RR=0.81; 95%CI: 0.77, 0.86). Later on, six cohort studies included in another recent meta-analysis (Dong *et al.*, 2011) found a similar risk reduction of T2DM by 14% for every 100mg/day increase in

magnesium intake. Additionally, a significant inverse association was found only among overweight and obese participants but not in those with BMI< 25kg/m^2. Overall, consistent evidence on inverse association between risk of T2DM and magnesium intake was observed. Plausible mechanisms such as low intracellular magnesium level which was found in diabetes participants may interfere with enzymes activity in insulin mediated glucose uptake pathway and showed decrease cellular glucose utilization and increase insulin resistance (Barbagallo *et al.*, 2003). Observational studies showed low serum magnesium and intracellular magnesium in pre-diabetic-patients in comparison to healthy individuals (Lima *et al.*, 2009) as well as in T2DM participants (Resnick *et al.*, 1993).

3.6.3.2 Vitamin D and risk of type 2 diabetes mellitus

As vitamin D intake is of interest for many researchers, recent prospective studies have examined the role of vitamin D in the prevention of T2DM. The role of vitamin D on insulin action and insulin sensitivity was examined previously in relation to T2DM (Alvarez and Ashraf, 2009). However, the findings are inconsistent. In 2011, Mitri et al, summarised the results of 8 cohort studies and found that the risk reduction of T2DM was 13% among people with vitamin D intake above 500 international units (IU)/day in comparison to people with intake of less than 200IU/day. In addition, participants in the highest group of vitamin D status had a lower risk of diabetes (43%) in comparison to participants in the lowest vitamin D status group. However, Mitri and his colleagues (2011) concluded in the systematic literature review from randomized controlled trials, that vitamin D supplementation have no effect on participants with normal glucose tolerance however people with insulin resistance and vitamin D deficiency showed improvement of insulin resistance with daily vitamin D and calcium supplementation.

Therefore, protective effect reported in cohort studies was not supported by trials. The possibility that dietary vitamin D intake and vitamin D supplementation may have different effects on the risk of diabetes should be considered.

A meta-analysis (Afzal *et al.*, 2013) of 16 studies (10 cohorts and 6 nested case-control studies) examined whether low plasma 25-hydroxyvitamin D as an indicator of vitamin D status was associated with an increased risk of T2DM. The odds ratio of T2DM was 1.50 (95%CI: 1.33, 1.67) among participants in the lowest concentration of 25 hydroxyvitamin D in comparison to those with the highest

concentrations of 25 hydroxyvitamin D. The results from epidemiological studies support the beneficial effect of dietary vitamin D intake on the risk of diabetes. However vitamin D status is not just determined by diet. Research that examines the long term effect of dietary vitamin D or/and vitamin D supplements on populations free from diabetes as well as on diabetes markers are needed.

Overall, effects of micronutrient intakes on the risk of diabetes are still under investigation and high quality trials to determine the role of micronutrients in the prevention of T2DM are still needed (Wyness, 2009).

3.6.4The relationship between food groups and risk of type 2 diabetes mellitus

3.6.4.1 Fruit and vegetables

Data on fruits and vegetables intakes were pooled in a meta-analysis of five cohorts: the results did not show a significant association between five or more servings of fruits and vegetables with the risk of diabetes (RR=0.96; 95%CI: 0.79, 1.17) nor with three or more servings of fruit (RR=1.01; 95%CI: 0.88, 1.15) or three or more servings of vegetables intake (RR=0.97; 95%CI:0.86, 1.10) (Hamer and Chida, 2007). The authors reported that most of the cohort studies in the review were from the U.S and as obesity incidence is dramatically increasing that may overcome the intakes effects. In addition, obesity-related underreporting of usual dietary intake is considered as estimated bias in prospective studies which may also partly contribute to the findings. A recent meta-analysis (Cooper *et al.,* 2012) where two more cohort studies (European and Chinese) were included in the analyses found that the highest (4-11 portions per day) versus the lowest (0-2.5 portions per day) quintile of intakes of fruit and vegetables was weakly associated with a lower relative risk of diabetes (HR=0.90; 95%CI: 0.80, 1.01) independent of potential confounders. Also no relationship was found with the risk of diabetes when each food intake was examined separately. A significant inverse association was only noticed between green leafy vegetables intake and the risk of diabetes where the risk was reduced by 16% among those who consumed 9 portions of leafy vegetables per week (RR=0.84; 95%CI: 0.74, 0.94) in comparison to one portion per week.

A critical review (Boeing *et al.,* 2012) explained the lack of association between fruits and vegetables and the risk of diabetes may partly be explained by measurement error of the dietary assessment method used to estimate the intake. Also, it was assumed that increased intakes of fruits and vegetables may indirectly

decrease the development of T2DM by preventing body weight gain. Prospectively, fruits and vegetables intake still have no clear association with the risk of T2DM. However, leafy vegetables may have a potential benefit to the risk of diabetes.

3.6.4.2 Tea and coffee intakes and risk of type 2 diabetes mellitus

Consumption of tea and coffee in relation to the risk of diabetes has also been investigated in cohort studies. Pooled data from nine prospective studies that examined tea consumption and the risk of T2DM (Jing *et al.*, 2009), showed lack of association between less than 3 cups a day and never consumed. However, when pooled estimate from studies which included high tea consumption (more than four cups of tea) a reduction of the risk of diabetes by 20% (RR=0.80; 95%CI: 0.70, 0.93) in comparison to never consumers was observed. Also recent cohort study of African American women found no effect of tea consumption on the risk of T2DM (p trend=0.17) (Boggs *et al.*, 2010). Type of tea was not reported in majority of studies which may be relevant to risk of T2DM.

With regard to coffee consumption, pooled data from nine cohort studies reported a significant risk reduction of T2DM. Risk was reduced by 28% with consumption of 4-6 cups of coffee per day and further risk reduction by 35% with consumption of more than 6-7cups per day in comparison to participants who consumed less than 2 cups a day. Results were not altered by gender, BMI and regional adjustment (Van Dam and Hu, 2005). Another recent meta-analysis found similar results when data were pooled from 13 cohort studies (Muley *et al.*, 2012). It has been suggested that antioxidants in coffee may have the beneficial effect on insulin sensitivity. Also chlorogenic acid is one of the suggested coffee component that may lower hepatic glucose output and reduce glucose level (van Dam, 2006). Further investigation on the effect of coffee and tea intake on metabolic intermediate factors such as insulin action and glucose concentration are needed.

3.6.4.3 Legumes and risk of type 2 diabetes mellitus

Legumes are another important food group suggested to have a link to diabetes. The beneficial effect of non-oil seed pulses such as chickpeas, peas, lentils and beans on fasting blood glucose and insulin level in pooled analysis of interventional trials was seen (Sievenpiper *et al.*, 2009). Very few prospective studies have examined the intake of legumes and the risk of diabetes. In the Nurses' Health Study, no association was found between legumes intake and the risk of diabetes among women aged 38-63 years followed for 18 years (Bazzano *et al.*,

2008). Data on middle-aged Chinese women followed for an average of 4.6 years, yielded 64,227 participants who completed validated FFQ and reported a lower relative risk of T2DM (RR=0.62; 95%CI: 0.51, 0.74) among women in the highest consumption quintile in comparison to the lowest legumes quintile (p for trend <0.01). Women in the highest intake of legumes other than soybeans and peanuts had a reduced risk of T2DM by 24% (RR=0.74; 95%CI: 0.64, 0.90; p for trend<0.01) in comparison to the lowest quintile (Villegas *et al.,* 2008). Overall, interventional trials reported the beneficial effect of legumes intake among diabetic and non-diabetic participants whilst an unclear association has been found in the few prospective studies which have been carried out. Legume intake is low in Western populations (Schneider, 2002) which may explain the lack of epidemiological studies in this area. Further research is needed to determine the association between intake of legumes and the risk of diabetes prospectively. Detail on legumes consumption and risk of T2DM will be discussed in chapter 9.

3.6.5Dietary patterns

Dietary pattern analysis such as data driven and hypothetical driven approaches have been used in many prospective studies to examine the relationship of diet to the risk of diseases. Approach based on indices tended to be easily being translated into public health message but with respect to dataset driven method aren't. Nine dietary components were included in the Alternate Healthy Eating Index (AHEI) score to measure quality of diet where the highest score indicates that the dietary intake of participants met the American recommendations (i.e rich in fruits and vegetables and wholegrain with lower consumption of red and processed meat). The lowest score reflects the dietary intake of participants with the least healthy dietary intake (McCullough *et al.,* 2002). In a prospective cohort study including 80,029 women who were followed for 18 years, the risk of diabetes was significantly reduced by 36% among women in the highest quintile of AHEI (where AHEI was mainly rich in fruits and vegetables and wholegrain) in comparison with women in the lowest AHEI quintile (Fung *et al.,* 2007). On the other hand, an earlier study found the risk of diabetes was positively associated with an intake of meat and processed meat among U.S women followed for 14 years prospectively (Fung *et al.,* 2004). Women allocated to the highest Western dietary pattern quintile (which is characterized by high intakes of red meat and processed meat, sweet and desserts,

French fries, and refined gains) had a significantly increased risk of T2DM by 51% in comparison to the lowest quintile.

The Mediterranean diet has been a focus of much research interest. Data from ten large prospective studies pooled in a systematic review which included 190,000 participants free from diabetes followed for at least 2 years, reported 8,932 cases of diabetes (Esposito *et al.*, 2010a). They found a dietary pattern which is rich in fruits and vegetables, wholegrain, fish and poultry and is low in red meat, processed foods, sweetened beverages and starchy foods, reduces the risk of T2DM from 15% to 83%. Furthermore, people adherent to the Mediterranean diet which is mainly a plant based diet, were less likely to develop diabetes derived from prospective studies (Esposito *et al.*, 2010b).

A recent comprehensive meta-analysis of 15 prospective studies (Nettleton *et al.*, 2013) also supported the inverse association between diet score and metabolic markers of diabetes. The diet score categorized nine food groups' intakes into quartiles and assigned values for favourable (wholegrain, fruits, vegetables, fish and nuts/seeds) and unfavourable foods (red/processed meats, sweets, sugared beverages and fried potatoes) with the highest score being healthier diet. Significant inverse associations were seen between diet score and fasting blood glucose as well as diet score and fasting insulin concentration after adjustment for BMI, lifestyle and demographic factors.

Overall, observational studies consistently support the protective effect of diet rich in fruit and vegetables, whole grains, with less red meat and processed foods on the development of diabetes. However the total intake of fruits and vegetables as a suggested food group was found to have a weak inverse association with the risk of T2DM; evidence suggests the importance of the role of green leafy vegetables in the prevention of diabetes. Several dietary patterns have been suggested in epidemiological studies that are favourably associated with the prevention of T2DM. A diet rich in plant food was the main characteristic dietary pattern. On the other hand, a diet rich in red and processed meats may contribute to a high risk of T2DM.

3.7 Conclusion

Criteria of diabetes diagnosis have been updated over the past 30 years which need to be considered when comparison between epidemiological prospective studies that investigate the risk of diabetes and diet is carried out. As demonstrated

that T2DM is public health problem that have great health and economic burden which can be reduced by primary prevention. Evidence points to genetic and environmental factors that have promote development of diabetes. However, factors which effectively reduce the risk of T2DM were mainly dietary factors and few others such as physical activity and smoking.

Considering the importance of other factors such as age, gender and ethnicity which showed a strong association with T2DM is mainly to identify the target population at risk of T2DM in prevention plan strategies. Several epidemiological studies and trials reported the direct and indirect effect of dietary factors in development of T2DM. The association between diet, food intake and nutrients intakes with the risk of diabetes have been reported prospectively.

Evidence reported the potential benefit effect of diet which is a rich in wholegrain, but evidence really weak in supporting fruit and vegetable intakes. Prospective studies showed the potential beneficial effect of nutrient intake such as magnesium on risk of diabetes. It is difficult to separate the effect of carbohydrates on the risk of T2DM in research because if carbohydrates examined are based on structure then it will depend on whether it is liquid or solid and if assessing total carbohydrates, will then it depend whether it is low or high GI.

Literature on the recent studies particularly prospective studies that have examined the association between effect of dietary and non-dietary factors with the risk of T2DM helps in building the best models for examining the association between dietary fibre intake and risk of T2DM which will be discussed in chapter 8 and 9.

Chapter 4: Comprehensive review of dietary fibre intake and risk of type 2 diabetes mellitus

4.1 Introduction

As discussed in chapter 2, T2DM prevalence is growing nationally and internationally and is resulting in an enormous health and economic burden at both individual and population level. Both environmental and genetic factors' interact in the development of T2DM; however, lifestyle and obesity are suggested reversible factors that may play an important role in delaying the development of T2DM (Knowler *et al.*, 2002). Evidence suggests that some dietary nutrients and dietary fibre have a potential influence on the risk of T2DM; to play a role in primary prevention or delay in the development of T2DM (World Health Organization, 2003). Contradictory knowledge on the association between dietary fibre intake and risk of T2DM in prospective studies had led to further exploration of existing prospective evidence in this chapter.

A systematic review is a comprehensive review of the literature addressing a research question especially when many researchers focus on the same topic. Systematic review methodology aims to select all relevant studies reaching conclusions that answer similar questions of interest. After that, statistical analysis, also referred to as meta-analysis, may be used apply to statistically summarise the results of all selected studies. The aim of the present review is to summarise results from existing cohort studies published in the last 22 years which have evaluated the effect of total dietary fibre (TDF)[11] intake and risk of T2DM, and to present further evaluation on the selected prospective studies on the effect of the main food sources of fibre intakes, insoluble dietary fibre (IDF) and soluble dietary fibre (SDF) intakes on the development of T2DM. The current comprehensive review was undertaken using a systematic approach, and aimed to generate pooled estimate of risk by examining highest versus lowest dietary fibre intake groups in relation to risk of T2DM. Also Forest plots were draw from included cohort studies, aiming to illustrate the overall effect of intakes of TDF, fibre sources, and types of fibres on the risk of T2DM. In addition to, it would be possible to explore whether the effect of dietary fibre intake on the risk of diabetes may vary in subgroup analysis.

[11] In this chapter, TDF abbreviation referred to total dietary fibre measured by any analytical method. While in the following chapters, AOAC-fibre referred to total dietary fibre intake obtained from AOAC analytical method only.

4.2 Methodology

4.2.1 Study strategy

Original cohort studies that were published in English between Jan 1991 and March 2013 were collected mainly through two computer-based databases. Also articles were retrieved from examined references of published meta-analyses and reviews where additional studies were identified. The search started from 1st February 2011 until 1st February 2013. The search was carried out on MEDLINE(R) 1948 to January Week 4 2013 and on EMBASE using Ovid resources, EMBASE Classic + EMBASE 1947, to February, 2013. Studies were identified using a combination of key words related to the disease (T2DM) and dietary exposure (dietary fibre intake) as shown in appendix B. The search strategy included both exploded MeSH (Medical Subject Headings) terms and truncated free text, aiming to include all the relevant articles.

4.2.2 Study selection

According to the inclusion criteria in Table 4.1, identification of the prospective studies that report the estimated effect of dietary fibre intake on incidence of T2DM. The studies included were limited to English language studies. Studies that were excluded from the review were not prospectively evaluating the association between dietary fibre intake and the risk of diabetes, such as cross-sectional studies and interventional studies. Other exclusion criteria are listed in Table 4.2. The coding strategy used in the Endnote software program (a web-based reference organizer) aimed for data extraction according to established criteria. The primary search gave 1,554 articles that were checked to ensure that they met the search criteria; 884 articles were from Embase database and 670 articles were from Medline. Out of all the articles only 16 papers were relevant and others were excluded for different reasons as listed in Table 4.2. Figure 4.1 represents the results of the search strategy in the databases for the review with exclusion criteria. Articles were assessed based on title, abstract and the full manuscript. Quality of the studies was evaluated based on the number of studied participants, the follow up period and potential confounders that were included in the adjusted model. All the studies included in the results tables (4.5, 4.6 and 4.8) were published between1997 – 2013. Results were reported from the selected cohort studies aiming to evaluate the association of TDF, IDF and SDF intakes as well as fibre intakes from cereal,

vegetables, fruit, and legumes and the risk of T2DM from prospective evidence. Meta-analysis was carried out aim to summarize the existing evidence from eligible prospective studies and to pool estimate risk of T2DM by Forest plot graphs via assessing highest dietary fibre intake group versus lowest dietary fibre group.

Table 4.1 Inclusion criteria used to select relevant studies

1	Cohort and prospective studies
2	Articles investigating population/s in UK, Europe, USA, Canada, worldwide.
3	Studies that describe the TDF intake or/and types of fibres or/and soluble and insoluble fibre intake
4	Publication date between Jan 1990 and March 2013
5	Human studies
6	English language published articles
7	Original research article
8	Provide estimates of T2DM risk

Table 4.2 Coding of exclusion criteria for primary identified articles

code	Comments
0	Not relevant (reports or comments , further duplication)
1	Experimental, animal
2	Recommendations, guidelines
3	Management, diagnosis
4	Not adult diabetes (children, pregnancy)
5	Not prospective, cohort study (other study designs e.g. RCT, cross-sectional study)
6	Reviews
8	Potentially relevant – met inclusion criteria

4.2.3Data extraction

Data that were collected included the author's name, year of publication, country of origin, duration of follow-up, number of participants, mean or range of age, proportion of women if available, number of events, the amount of fibre consumption (g/d, g/1000kcal and g/MJ), methods for measurement of dietary fibre intake, analytical method of dietary fibre values, adjusted covariates, as well as estimated risks and 95% CIs of T2DM. Data extracted are in Tables 4.5, 4.6, 4.8.

4.2.4Statistical analysis

Meta-analysis was conducted if three or more studies were identified. The estimated effect of TDF, IDF and SDF intakes as well as for key fibre sources (cereal fibre, fruit fibre and vegetable fibre) intakes on the risk of T2DM was pooled from the selected cohort studies performed using the "metan" meta-analysis command in the Stata software statistical program (Corp-Stata, 2010). A Forest plot was created to summarize quantitative findings of and to illustrate the direction of the effect.

4.2.4.1 Total dietary fibre intake

Figure 4.1 demonstrates 16 papers were relevant, two papers included same study population (Nurses' Health Study) which reported by Colditz *et al.* (1992) and Salmeron *et al.* (1997b). Thus, the latest was included in the current review because of the larger sample size (914 cases vs. 702 cases). The second study that also excluded is because of the studied population were pre-diabetes participants rather than from general population (Lindstrom *et al.*, 2006). In total, ten cohorts reported the intake of dietary fibre as g/day and risk of T2DM obtained from comparing extreme categories, while the remaining four, either reported dietary fibre intake as g/MJ (Ericson *et al.*, 2013) or fibre density (Hopping *et al.*, 2010) or reported risk of T2DM per unit increment of dietary fibre intake (continuous variable) (Sluijs *et al.*, 2010b, Barclay *et al.*, 2007). Thus, the Forest plot using meta-analysis was carried out on the ten studies that provided 11 risk results as Stevens *et al.* (2002) reported risk of T2DM based on ethnicity (White and African American population) (Table 4.3). Further subgroup analyses based on follow-up period, origin of study and gender were carried out to help explain the potential sources of heterogeneity in the pooled estimate

Table 4.3 Number of studies identified and included in the meta-analysis

	TDF	Insoluble fibre	Soluble fibre	Cereal fibre	Vegetable fibre	Fruit fibre
Total publications	16	4	4	13	11	11
Cohort studies in meta-analysis	10	4	4	9	8	7
Excluded from meta-analysis	5	-	-	4	3	4
Risk results in forest plot	11	6	4	11	8	7

4.2.4.2 Types of fibre intakes

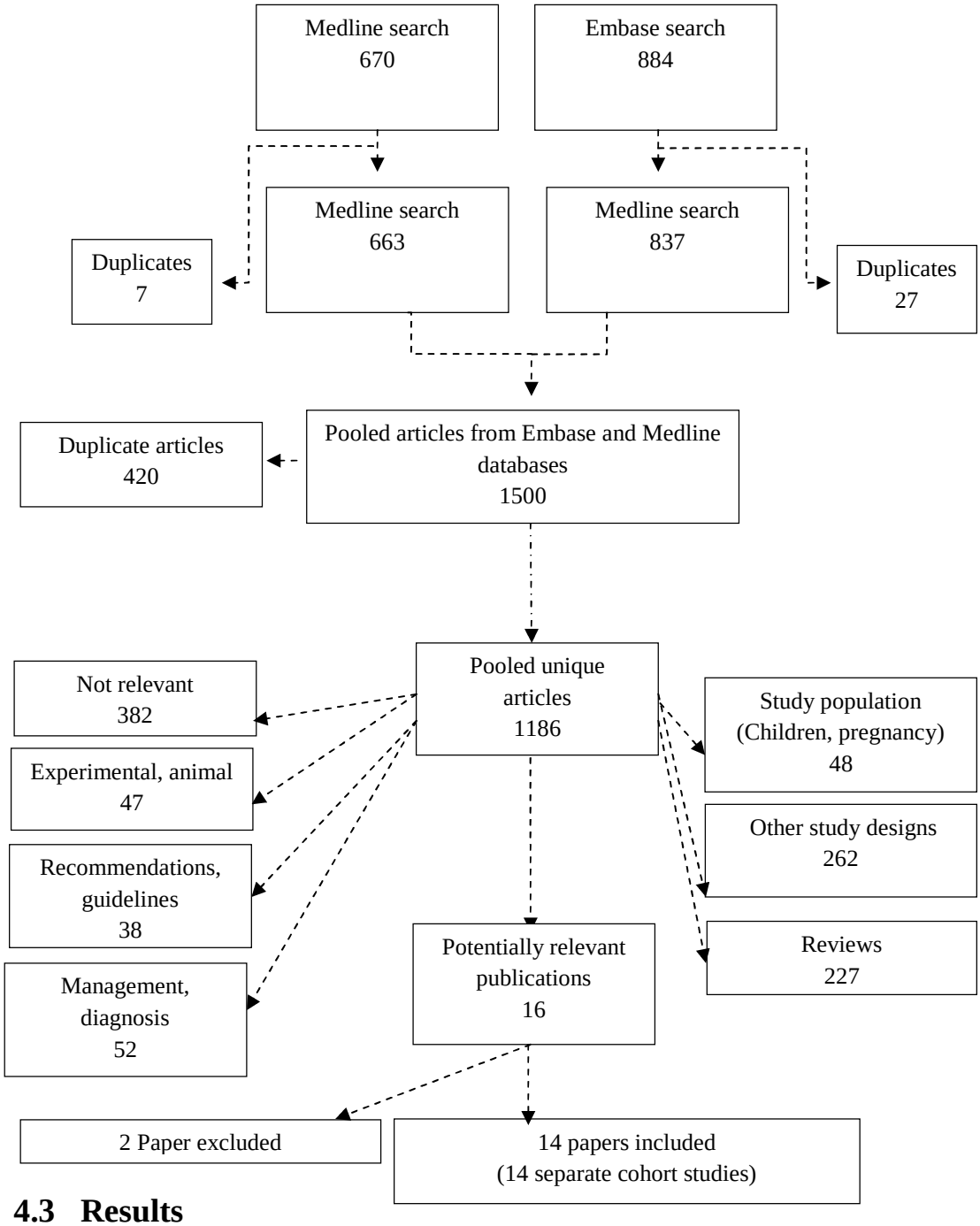
A pooled risk estimate was obtained for insoluble dietary fibre intake (IDF) from four cohort studies that provided 6 risk estimate results. Montonen *et al.* (2003) reported separate risk of T2DM with intakes of insoluble fibre subtypes (insoluble non-cellulosic polysaccharides, cellulose and lignin). Forest plots were created for pooled risk of T2DM and main fibre sources and for insoluble fibre intake (Figure 5.c in appendix C) and risk of T2DM from number of prospective studies as shows in Table 4.3. Soluble dietary fibre was reported in four studies and meta-analysis was performed and Forest plot was created in Appendix C, Figure 9.c.

4.2.4.3 Fibre sources intakes

Cereal fibre intake was reported in 13 cohort studies, eight of them were eligible for meta-analysis. Four studies were excluded from the meta-analysis because of either dietary fibre intake expressed as fibre density (Hopping *et al.*, 2010) or no reported risk result for dietary fibre categories (Barclay *et al.*, 2007, Hodge *et al.*, 2004, Stevens *et al.*, 2002). Another meta-analysis was carried out aimed for estimate the risk of T2DM with intake of cereal fibre intake as continuous variable (g/day). Three cohort studies were identified (Stevens *et al.*, 2002, Hodge *et al.*, 2004, Barclay *et al.*, 2007) for meta-analysis. For vegetable fibre intake, eight studies were eligible for pooled analysis. Three were either fibre intake expressed as fibre density or risk estimate for each unit increment in dietary fibre (Hopping *et al.*, 2010, Hodge *et al.*, 2004, Barclay *et al.*, 2007). For fruit fibre intake, seven studies were eligible for meta-analysis. Four were either fibre intake expressed as fibre density or risk estimate for each increment (Hopping *et al.*, 2010, Hodge *et al.*, 2004, Barclay *et al.*, 2007, Stevens *et al.*, 2002). Meta-analysis was not carried out for legumes fibre because one study met the criteria for meta-analysis (Meyer *et al.*, 2000).

Regards reported risk results, it was noticed that all cohorts studies reported the highest fibre intake effect on the risk of T2DM in comparison to lowest fibre intake group except in two studies (Weng *et al.*, 2012, Wannamethee *et al.*, 2009) which reported risk of T2DM aim to show the effect of low fibre intake on the risk of T2DM in comparison to high fibre intake. Therefore, reciprocal odd ratio that equal one divided by the odd ratio or risk ratio was calculated (Bewick *et al.*, 2004).

Figure 4.1 Flowchart demonstrating the results of the systematic review with exclusion criteria



4.3 Results

Figure 4.1 demonstrates that sixteen publications that equals 15 cohorts studies where study by Colditz *et al.* (1992) was excluded from the review because of another study by Salmeron *et al.* (1997b) examined the risk of T2DM on the same population using modified FFQ. Within the 14 cohorts, ten of them report risk estimate for the whole studied population while three cohorts provided separate risks based on sex (Hopping *et al.*, 2010, Ericson *et al.*, 2013) and ethnicity (Stevens *et al.*, 2002). The 14 cohort studies eligible for inclusion in this review are summarised

in Table 4.3. Generally, thirteen cohorts provided data on fibre source/s and four cohorts provided data on types of fibre. Additionally, one study reported cereal fibre intake but not the TDF in relation to the risk of diabetes (Krishnan *et al.*, 2007) which was included in the dietary fibre sources.

4.3.1Description of the evidence based publications

4.3.1.1 Exposure measures

The Food Frequency Questionnaire (FFQ) was found to be the most used dietary assessment tool; one cohorts used dietary records (Ericson *et al.*, 2013) or interview (Montonen *et al.*, 2003). The number of food items listed in the FFQ range from 49 to 148. Dietary fibre intake in ten cohort studies was reported as AOAC obtained values (Official Methods of Analytical Chemists, 1995). This was either reported in methodology or through cross checking food composition database. Three studies did not specify clearly the analytical method that determined the dietary fibre values in foods (Sluijs *et al.*, 2010b, Meyer *et al.*, 2000, Weng *et al.*, 2012). Only one study (Wannamethee *et al.*, 2009) reported the NSP intake which was derived by the Englyst method (Englyst *et al.*, 1982). Another study reported NSP components (lignin, cellulose, soluble and insoluble non-cellulosic polysaccharides) that usually measured by Englyst method (Montonen *et al.*, 2003). Detail on different types of dietary fibre analysis was demonstrated in the chapter 1.

4.3.1.2 Case ascertainment

The number of cases of T2DM ranged from 114 to 8,587 within each study. Percentages of diabetic cases out of total studied population were calculated and ranging from 0.8% (Schulze *et al.*, 2004a) to 12% (Stevens *et al.*, 2002). Supplementary questionnaires were most often used for self-reported incidents of diabetes. Confirmed self-reported diabetes by hospital registry department in some of the cohorts were recorded in the article (Sluijs *et al.*, 2010b, Ericson *et al.*, 2013, Weng *et al.*, 2012, Hodge *et al.*, 2004). Other studies reported measured fasting blood glucose was used to identify the incidence of diabetes with or without the self-reported cases (Barclay *et al.*, 2007, Hopping *et al.*, 2010, Weng *et al.*, 2012).

4.3.1.3 Origin of study/size of cohort

This evidence based was mainly dominated by developed world studies. Six studies were from USA, one from Finland, one from Germany, one from UK, two from Australia, one from Netherland, one from Sweden and one from Taiwan. The

number of subjects ranged from 1,604 (Weng *et al.*, 2012) to 91,249 (Schulze *et al.*, 2004a).

4.3.1.4 Duration of follow-up

The follow-up period ranged from 4 to 14 years. As shown in Table 4.5. One cohort was followed for four years (Hodge *et al.*, 2004). One study reported an average of 5 years of follow up duration. Three studies reported duration of six years (Salmeron *et al.*, 1997b, Salmeron *et al.*, 1997a, Meyer *et al.*, 2000), five studies reported follow up period range between 7 and 10 years and four studies reported follow up of more or equal 10 years.

4.3.1.5 Other aspects

All cohorts were run on adult participants. Most studies included both gender but some were specifically on women and few were specifically on men. Six cohort studies reported estimate risk among women participants (Ericson *et al.*, 2013, Hopping *et al.*, 2010, Schulze *et al.*, 2004a, Meyer *et al.*, 2000, Salmeron *et al.*, 1997b) and four cohorts reported estimate risk among men participants alone (Ericson *et al.*, 2013, Hopping *et al.*, 2010, Wannamethee *et al.*, 2009, Salmeron *et al.*, 1997a). While, eight studies could not be separated as estimated effect was reported for both genders.

4.3.2Characteristics of high fibre consumers in the cohort studies

The cohort studies in this review reported the characteristics of high fibre consumers (Weng *et al.*, 2012, Schulze *et al.*, 2007, Montonen *et al.*, 2003, Stevens *et al.*, 2002, Meyer *et al.*, 2000). High dietary fibre consumers are less likely to be men, smoke and consume alcohol in comparison to lower fibre consumers groups with a statistically significant trend. On the other hand, high fibre consumers were more likely to be older, educated, engaged in physical activity and have lower BMI and WC. Regarding the main dietary characteristics, participants in the high fibre group were found to have significantly higher energy intake, higher carbohydrate intake but lower protein and total fat intake in comparison to the lowest fibre intake group. In terms of fat constituents, participants in the high fibre category were more likely to consume polyunsaturated and monounsaturated fat and less saturated fat.

Table 4.4 Cohort studies linking intake of total dietary fibre (TDF) to the risk of T2DM

N	Author	No of participants	Age years	Follow-up years	DM cases (Dg)	Dietary assessment	Dietary fibre categories	Estimated risk (95%CI)	*P*-trend	Variables adjusted**
1	Ericson *et al.* (2013) Malmo Diet and Cancer Study (MDCS) (Sweden)	27,140 Women & men	45-74	12	1709 Self-report	7 days menu book, 168-itmes FFQ & 45min interview AOAC method	TDF (g/MJ) Women Q5 >2.9 Q1 <1.8 Men Q5 >2.5 Q1<1.6 **Q5 vs. Q1[4]**	Women M(1) 0.98(0.78,1.24) M(2) 0.97(0.77,1.22) Men M(1)0.85(0.68,1.05) M(2)0.84(0.68,1.04)	0.98 0.91 0.21 0.20	M(1) age, season[2], method version[3], EI M(1) age, method version, EI, education, PA, smoking, alcohol, BMI.
2	Weng *et al.* (2012) Cardiovascular Disease risk Factor Two-Township Study (CVDFACTS) (Taiwan)	1,604 58% Women	Mean DM (52.7± 10.5) Non DM (49.8±12.2)	4.6	141 FBG>7mmol/l or self-report	49-items FFQ Fibre analysis not specified	TDF Q1 <24.1 g/d Q5≥38.6g/d **Q1 vs. Q5[1]**	M(1) 1.91(1.11,3.28) M(2) 2.08(1.21,3.60) M(3) 2.04(1.17,3.53)	0.01 0.005 0.004	M(1) age, sex, age-sex interaction, EI, residential area M(2) M1 plus FH, BMI, central obesity. M(3) M2 plus education, smoking, alcohol, PA, HTN, high cholesterol, high TG, low HDL.
3	Sluijs *et al.* (2010b) EPIC-NL (Netherland)	37,844 74% Women	51yrs 21-70	10	915 Self-report and hospital records cases	79-item FFQ Fibre analysis not specified	Per 4.8g/d increment	M(1) 0.98(0.91,1.05) M(2) 0.81(0.92,1.06) M(3) 0.89(0.82,0.98)	<0.05	M(1) Age, sex M(2) 1 plus energy adjusted alcohol, PA, smoking, BMI, WC, SBP, Education, FH. M(3) 2 plus EI, energy adjusted intakes of GL, vitamin C, E, protein, SFA, PUFA,

** abbreviations used in column : EI= energy intake; PA= physical activity; FH= family history of diabetes; WC= waist circumference; BMI= body mass index; WHR=waist hip ratio; SFA= saturated fatty acid; PUFA=polyunsaturated fatty acids, GL=glycaemic load; SBP= systolic blood pressure; AOAC= Association of Official Analytical Chemist; (S)= significant; (NS)= not significant, NA=not available; [1] Q1 vs. Q5 estimated risk among lowest fibre group in comparison to highest fibre group; [2]Season (summer, winter, autumn and spring); [3]Method (2 categories of coding dietary data before and after September 1994), [4]Q5 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group; [5]Q4 vs. Q1 estimated risk among lowest fibre group in comparison to highest fibre group.

	Continue									
N	Author	No of participants	Age years	Follow-up years	DM cases (Dg)	Dietary assessment	TDF categories	Estimated risk (95%CI)	*P*-trend	Variables adjusted**
4	Hopping *et al.* (2010) Multi ethnic Cohort in Hawaii (USA)	75,512 52% Women 39% Caucasians 46% Japanese American 14% native Hawaiian	45-75	14	8,587 Self-report cases	FFQ AOAC method	TDF (g/4184 kJ.d) Men Q1 <7.4 Q5 ≥14.2 Women Q1<8.9 Q5≥16.2 **Q5 vs. Q1[4]**	All men 0.75(0.67,0.84) Caucasian 0.66(0.53,0.82) Japanese 0.84(0.72,0.97) Native Hawaiian 0.70(0.52,0.96) All women 0.95(0.85,1.06) Caucasian 0.80(0.62,1.02) Japanese 1.04(0.90,1.20) Native Hawaiian 0.85(0.66,1.10)	<0.01 <0.01 0.09 0.19 0.05 0.04 0.59 0.21	M (1) ethnicity, BMI, PA, education, and EI.
5	Wannamethee *et al.* (2009) British regional Heart Study (UK)	3,428 100% Men	60 – 79	7	162 Self-report cases	FFQ. Englyst-NSP	TDF Q1 ≤20g/d Q4 ≥ 31 g/d **Q4 vs. Q1[5]**	M(1) 0.63(0.42,0.96) M(2) 0.82(0.51,1.32) M(3) 0.83(0.52,1.13) M(4) 0.86(0.50,1.20) M(5) 0.95(0.59,1.53)	NA	M(1) age M(2) age, WC, smoking, PA, social class, alcohol, preexisting MI, stroke, use statin, EI. M(3) M1 plus IL-6 M(4) M2 plus t-PA. M(5) M3 plus GGT
6	Barclay *et al.* (2007) (Australia)	1,833 Women & men	+49	10	138 Self-report cases or FBG>126mg/dl	145-item FFQ AOAC method	Per 5g/day increment	M(1) 0.90(0.81,1.01) M(2) 0.90(0.79,1.02)		M(1) age, sex. M(2) 1 plus FH, smoking, TG, HDL, METs.

** Abbreviations used in column: EI= energy intake; PA= physical activity; FH= family history of diabetes; WC= waist circumference; BMI= body mass index; WHR=waist hip ratio; IL-6 = initerleukin-6; t-PA = tissue plasminogen activator; GGT = Gamma-glutamyl transferase; AOAC= Association of Official Analytical Chemist; NSP= non-starch polysaccharides; PA = physical activity; MET= Metabolic Equivalent Task; [1] Q1 vs. Q5 estimated risk among lowest fibre group in comparison to highest fibre group; [4]Q5 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group; [5]Q4 vs. Q1 estimated risk among lowest fibre group in comparison to highest fibre group.

Continue

N	Author	No of people (%) women	Age years	Follow-up years	DM cases (Dg)	Dietary fibre assessment	Dietary fibre intake categories	Estimated risk (95%CI)	*P*-trend	Variables adjusted**
7	Schulze *et al.* (2007) EPIC-Potsdam study (Germany)	25,067 61% Women	35 – 65 Women 40-65 Men	7	844 Self-reported cases	148-items FFQ AOAC method	Median TDF Q1 15.8 g/d Q5 27.9 g/d **Q5 vs. Q1**[4]	M(1)0.86(0.68,1.09) M(2)0.86(0.65,1.14)	0.11 0.19	M(1) age, sex, education, sport activity, cycling, occupational activity, smoking, Mg alcohol, EI, BMI, WC. M(2)1 plus PUFA:SFA & MUPA:SFA ratio, CHO.
8	(Schulze *et al.*, 2004a) Nurses' Health Study II (USA)	91,249 100% Women	24-44	8	741 Self-reported cases	133-item FFQ AOAC method	TDF Q1<14.2g/d Q5 >22 g/d **Q5 vs. Q1**[4]	M(1) 0.53(0.42, 0.67) M(2) 0.78(0.62,0.98) M(3) 1.00(0.75,1.34)	<0.01 0.008 0.80	M(1) age M(2) age, BMI M(3) M2 plus EI, alcohol, PA, FH, HTN Smoking, high blood cholesterol, postmenopausal hormonal use, O/C use, GL, Mg and caffeine intakes
9	Hodge *et al.* (2004) Melbourne Collaborative Cohort Study (Australia)	31,641 59% Women	27-75	4	365 Self-reported cases	121-items FFQ AOAC method	TDF= 20g/d. 87^{th} vs. 12.5^{th} percentiles	M(1) 0.93(0.73,1.18) M(2) 1.02 (0.81,1.30)	0.53 0.85	M(1) Age country of birth, PA, FH, weight change, EI. education, alcohol, FH, M (2) 1 plus BMI, WHR
10	Montonen *et al.* (2003) Finnish Mobile Clinic Health Survey (Finland)	4,316 47% Women	40-69	10	156 Self-reported cases	Diet history interview (>100 food items) Fibre analysis not specified	TDF Q1<19.2 g/d Q4 >33.2g/d **Q4 vs. Q1**[5]	M(1) 0.57(0.30,1.08) M(2) 0.51(0.26,1.00)	0.07 0.04	M(1) age, sex, geographic area, and EI. M(2) M 1 plus smoking, BMI, intake of fruit and berries, and vegetables

** Abbreviations used in column: EI= energy intake; PA= physical activity; FH= family history of diabetes; WC= waist circumference; BMI= body mass index; WHR=waist hip ratio; CHO=carbohydrates; SFA= saturated fatty acid; GL=glycemic load; AOAC= Association of Official Analytical Chemist; NSP= non-starch polysaccharides; PA = physical activity; MET= Metabolic Equivalent Task; [1] Q1 vs. Q5 estimated risk among lowest fibre group in comparison to highest fibre group; [4]Q5 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group; [5]Q4 vs. Q1 estimated risk among lowest fibre group in comparison to highest fibre group.

Continue

N	Author	No of people (%) women	Age years	Follow-up years	DM cases (Dg)	Dietary fibre assessment	Dietary fibre intake categories	Estimated risk (95%CI)	*P*-trend	Variables adjusted**
11	Stevens *et al.* (2002) (ARIC) study (USA)	12,251 Women & men 78% White 22% African American	45-64 (44.8)m (55.2)w	9	1,447 Self-report drug use or measured blood glucose FBG>7mmol/l or RBG>11.1mmol/l	66-items FFQ AOAC method	White Q1 11.2 Q5 27.5 African American Q1 10.2 Q5 26.1 **Q5 vs. Q1**[4]	White M(1) 0.99(0.98,1.01) M(2) 0.99(0.98,1.01) African American M(1) 0.97(0.92,1.03) M(2) 0.99(0.98,1.01)	*0.28* *0.91* *0.73* *0.84*	M(1) age, BMI, sex, field Centre M (2) M 1 plus education, smoking, PA.
12	Meyer *et al.* (2000) Iowa Women's Health Study (USA)	35,988 100% Women	55-69	6	1141 Self-reported cases	127-items FFQ AOAC method	Q1 ≤15.3 g/d Q5 >23.6g/d **Q5 vs.Q1**[4]	M(1) 0.78(0.64,0.96)	*0.005*	M (1) EI, age, BMI, WHR, education, smoking, alcohol, PA.
13	Salmeron *et al.* (1997b) Nurses' Health Study (USA)	65,173 100% Women	40-65	6	915 Self-reported cases	134-items FFQ AOAC method	Median Q1 11.8g/d Q5 24.1g/d **Q5 vs.Q1**[4]	M(1) 0.78(0.62,0.98)	*0.02*	M(1) age, BMI, alcohol, smoking, PA, FH.
14	Salmeron *et al.* (1997a) Heath Professionals Study (USA)	42,759 100% Men	40-75	6	523 Self-reported cases	131-items FFQ AOAC method	Median Q1 13.4g/d Q5 29.7g/d **Q5 vs.Q1**[4]	M(1) **0.**98(0.73,1.33)	*0.70*	M(1)Age, BMI, alcohol, smoking, PA, FH

** Abbreviations used in column: EI= energy intake; PA= physical activity; FH= family history of diabetes; WC= waist circumference; BMI= body mass index; WHR=waist hip ratio; CHO=carbohydrates; SFA= saturated fatty acid; GL=glycemic load; AOAC= Association of Official Analytical Chemist; NSP= non-starch polysaccharides; PA = physical activity; MET= Metabolic Equivalent Task; [1] Q1 vs. Q5 estimated risk among lowest fibre group in comparison to highest fibre group; [2]Season (summer, winter, autumn and spring); [3]Method (2 categories of coding dietary data before and after September 1994), [4]Q5 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group; [5]Q4 vs. Q1 estimated risk among lowest fibre group in comparison to highest fibre group.

Table 4.5 Cohort studies of insoluble dietary fibre (IDF) and soluble dietary fibre (SDF) intakes with the risk of T2DM

N	Author	No participants	Age (years)	follow -up (years)	DM cases	Dietary fibre assessment	SDF/IDF intake	Estimated risk results	*P*-trend	Variables adjusted**
1	Schulze *et al.* (2007) Germany	25,067 61% Women	35–65 Women 40-65 Men	7	844 self - report	148-FFQ AOAC method	SDF (g/d) Q1 5.3 Q5 9.6 IDF (g/d) Q1 10.3 Q5 18.4 **Q5 vs. Q1[4]**	SDF M(1) 0.79(0.63,0.98) M(2) 0.78(0.6,1.01) M(3) 0.83(0.57,1.22) IDF M(1) 0.83 (0.65,1.05) M(3) 0.82(0.61,1.08) M(3) 0.93(0.62,1.40)	 0.05 0.09 0.45 0.06 0.10 0.62	M(1) age, sex, education, PA, smoking, alcohol, TE, BMI, WC M(2)1 plus PUFA:SFA and MUPF:SFA ratios and CHO M(3)2 plus IDF or SDF intake.
2	Montonen *et al.* (2003) Finland	4316 47% Women	40-69	10	156 self - report	Dietary history interview (>100 food items) AOAC method	SDF (g/d) Q1 (0.53-4.5) Q4 (7.4-22.7) Inso-NCP (g/d) Q1 (1.1-8.7) Q4 (16.6- 69.3) Cellulose (g/d) Q1 (0.48-3.2) Q4 (5.4-15.2) Lignin (g/d) Q1 (0.48-2.3) Q4 (4.2-14.5) **Q4 vs. Q1[5]**	SDF M(1) 0.80(0.44,1.45) M(2) 0.57(0.29,1.12) Insoluble NCP M(1) 0.48(0.25,0.91) M(2) 0.47(0.25,0.91) Cellulose M(1) 0.84(0.47,1.50) M(2) 0.6(0.29,1.21) Lignin M(1) 0.68(0.36,1.28) M(2) 0.68(0.36,1.30)	0.57 0.21 0.02 0.03 0.60 0.19 0.15 0.16	M(1) age, BMI, sex, field centre M (2) M 1 + education, smoking, PA.

** Abbreviations used in column: EI= energy intake; PA= physical activity; FH= family history of diabetes; WC= waist circumference; BMI= body mass index; WHR=waist hip ratio; CHO=carbohydrates; SFA= saturated fatty acid; PUFA=polyunsaturated fatty acids, MUPA= monounsaturated fatty acids; AOAC= Association of Official Analytical Chemist; MET= Metabolic Equivalent Task; Inso-NCP = insoluble non-cellulosic polysaccharides; [4]Q5 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group; [5]Q4 vs. Q1 estimated risk among lowest fibre group in comparison to highest fibre group.

Continue

N	Author	No participants	Age (years)	follow - up (years)	DM cases (Dg)	Dietary fibre assessment	SDF/IDF intake	Estimated risk results	*P*-trend	Variables adjusted**
3	Meyer *et al.* (2000) USA	35,988 100% Women	55-69	6	1141 Self-report	127-items FFQ AOAC method	SDF (g/d): Q1 <4.8 Q5 >7.2 IDF(g/d): Q1 <11.4 Q5 >17.7 **Q5 vs. Q1**[4]	SDF 0.89(0.73,1.08) IDF 0.75(0.61,0.91)	0.23 0.001	M(1) Age, BMI, alcohol, smoking, PA, FH
4	Salmeron *et al.* (1997b) Nurses' Health Study USA	65,173 100% Women	40-65	6	915 Self-report	134-items FFQ AOAC method	amount not reported	SDF 1.07(0.86-1.33) IDF 0.77(0.61-0.95)	--	Energy adjusted and other factors were not reported.

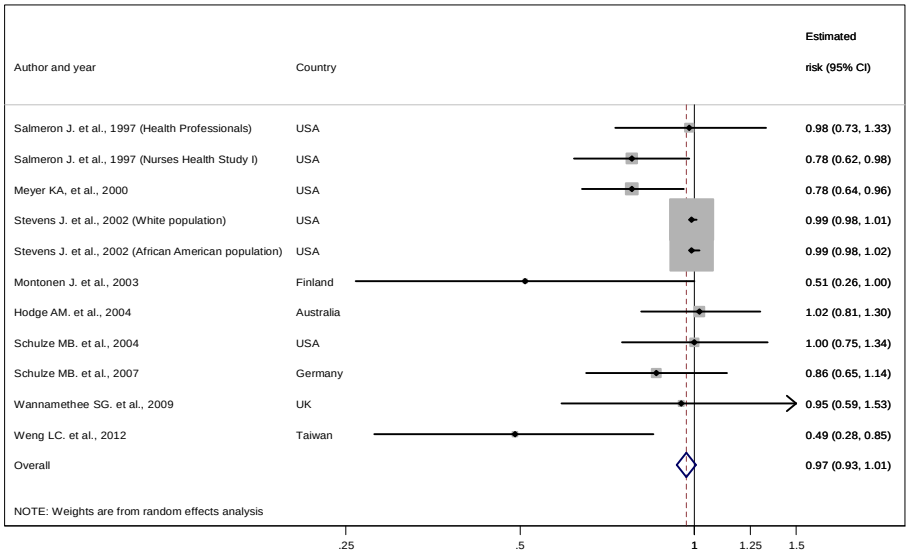
** Abbreviations used in column: EI= energy intake; PA= physical activity; FH= family history of diabetes; WC= waist circumference; BMI= body mass index; WHR=waist hip ratio; CHO=carbohydrates; SFA= saturated fatty acid; PUFA=polyunsaturated fatty acids, MUPA= monounsaturated fatty acids; AOAC= Association of Official Analytical Chemist; MET= Metabolic Equivalent Task; Inso-NCP = insoluble non-cellulosic polysaccharides; [4]Q5 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group.

4.3.3Total dietary fibre intake and risk of T2DM

Generally, six cohorts reported a significant inverse associations between TDF and risk of T2DM (Salmeron *et al.*, 1997b, Meyer *et al.*, 2000, Montonen *et al.*, 2003, Hopping *et al.*, 2010, Sluijs *et al.*, 2010b, Weng *et al.*, 2012). Eight cohorts reported no effect of TDF on the risk of T2DM. The overlap in the number of cohort studies relate to multi-ethnic cohort in Hawaii (Hopping *et al.*, 2010) study which showed significant effect in men but not in women.

Two of the 14 cohorts that reported a dose-response relationship and could not be included in the pooled estimate were inconsistent in their results. Barclay *et al.* (2007) showed for every 5g/day increment in dietary fibre, the risk of diabetes was significantly reduced by 10% but was not statistically significant. However, the second study reported that for every 4.8g/day increase in dietary fibre, the risk of diabetes reduced significantly by 11% (Sluijs *et al.*, 2010b) after adjustment for dietary and lifestyle factors (Details of the estimate risk and 95%CI provided in Table 4.5). A pooled estimate was obtained from 10 cohorts that gave 11 results where Stevens *et al.* (2002) study reported two estimate risk based on ethnicity. Figure 4.2 presents the estimated risks of T2DM of each prospective study for the highest dietary fibre intake group compared with the lowest dietary fibre intake group. The pooled results show highest fibre consumers have non-significant 3% lower risk than lowest fibre consumers (risk estimates=0.97; 95% CI: 0.92 to 1.01).

Figure 4.2 Estimated effect and 95% confidence intervals for dietary fibre intake (Highest vs. Lowest) and incidence of T2DM from all selected studies



4.3.3.1 Potential sources of heterogeneity

The lack of significant association between greater dietary fibre intake and risk of T2DM in compared with lowest fibre consumer showed heterogeneity test (I^2) equal to 49%. Higgins et al. (2003) describes the heterogeneity test as '*the percentage of total variation across studies that is due to heterogeneity rather than chance.'*. Moderate heterogeneity according to Higgins *et al.* (2003) was observed in the current analyses. Heterogeneity may reflect the diversity in studied population, gender; follow up period, origin of study, exposures and outcome measures. Therefore variation was explored by subgroup analyses. It have been suggested previously that study design and how outcome was measured were potential sources of heterogeneity (Higgins et al., 2003) however, the current meta-analyses included single study design which may rule out the possibility of design related variation. Furthermore, different population structures and variation in covariates adjusted in each study may also contribute to high heterogeneity in this analysis (Tu and Greenwood, 2012). In most of the studies, the cases were identified mainly by self-report with subsample verified by reference to medical records. This may have an element of variation related to undiagnosed cases as confirmation was carried out only on people with diabetes.

Also, heterogeneity may possibly be explained by study location, gender and follow up period (Higgins *et al.*, 2003). Therefore, studies were categorized into groups based on origin (USA, European and others), gender (men and women) and duration of follow-up (more than 7 years and equal or less than 7 years).

Table 4.6 Exploration of key sources of heterogeneity using subgroup analysis

Classification		No. studies	Pooled estimate(95%CI)	Results	I^2	P value[4]
Gender	Women	3	0.82(0.71, 0.92)	+[1]	10%	0.31
	Men	2	-	ND[3]	-	-
Origin	USA	6	0.98(0.95, 1.01)	-[2]	47%	0.09
	European	3	0.82(0.63, 1.07)	-[2]	16%	0.31
	Others	2	-	ND[2]	-	-
Follow-up duration	≤7 years	7	0.85(0.75, 0.96)	+[1]	24%	0.24
	>7 years	4	0.99(0.97, 1.01)	-[2]	20%	0.29

[1]Significant association; [2]non-significant association; [3]Not done (two or less cohort); [4]P for heterogeneity within each subgroup.

As illustrated in Table 4.4, pooled estimate effects based on gender using subgroup analysis showed that the risk of T2DM was reduced by 18% in high fibre consumers compared with low fibre consumers among women. Table 4.4 shows low heterogeneity within the subgroups, which suggest that estimates were consistent and evidence was homogenous. Two cohort studies reported risk of T2DM in men

and both showed lack of association between TDF intake and the risk of T2DM (Salmeron *et al.*, 1997a, Wannamethee *et al.*, 2009). Pooled estimate risk in women should be consider with caution as some eligible cohort studies (Stevens *et al.*, 2002, Hodge *et al.*, 2004, Montonen *et al.*, 2003) could not be included in the analysis due to lack of separate risk estimate among women (Appendix C Figure 10.c).

Additionally, two cohort studies were not included in pooled estimate (Hopping *et al.*, 2010, Ericson *et al.*, 2013) reported non-significant association with fibre intake expressed in g/1000kcal/day and g/MJ in comparison with low consumers among women. Further subgroup analysis of five studies (with six results) from USA populations showed people with high fibre intakes did not experience lower risk of diabetes as in Table 4.4. Also, pooled estimate from three European studies reported non-significant risk reduction among participants with high fibre consumption. Low heterogeneity was seen in European group. Forest plots for subgroup analyses are provided in appendix C (Figure 1c-4c). However, the three international studies were inconsistent in their findings (Hodge *et al.*, 2004, Weng *et al.*, 2012, Barclay *et al.*, 2007).

Four cohort studies with a longer follow up period, more than 7 years, found non-significant estimate risk, while seven studies with follow-up of less than 7 years had a pooled estimate of 0.85 (95%CI: 0.75, 0.96; I^2=24%).

Out of the 10 cohorts, seven studies estimated TDF intake from dietary fibre values measured by AOAC method and pooled risk did not differ from the overall risk estimate (risk=0.98; 95%CI: 0.96, 1.01; I^2 =21%). The remaining three studies either did not specify or reported other types of methods.

4.3.4Soluble and insoluble dietary fibre intakes and risk of T2DM

From the 16 papers, soluble and insoluble dietary fibre intakes were reported in only four studies (Table 4.6) (Salmeron *et al.*, 1997b, Meyer *et al.*, 2000, Montonen *et al.*, 2003, Schulze *et al.*, 2007). None of the studies documented any significant association between soluble fibre intake and the risk of T2DM, while three studies showed a significant reduction in diabetes risk with a high daily intake of insoluble fibre (Salmeron *et al.*, 1997b, Meyer *et al.*, 2000, Montonen *et al.*, 2003). Pooled estimate from 4 cohorts showed significant risk reduction in people with high insoluble fibre consumption (risk = 0.75; 95%CI: 0.66, 0.86) in compared to lowest intake group (Forest plot provided in Figure 5c in appendix C). From pooled meta-analysis, no significant association between SDF intake and risk of

T2DM compared extreme quintiles was observed (risk estimate=0.92; 95%CI: 0.77, 1.09; I^2=27% and p=0.24).

4.3.5Dietary fibre sources intakes and risk of T2DM

Table 4.7 demonstrates number of prospective evidence on fruit fibre and vegetable fibre intakes and legumes fibre with incidence of T2DM.

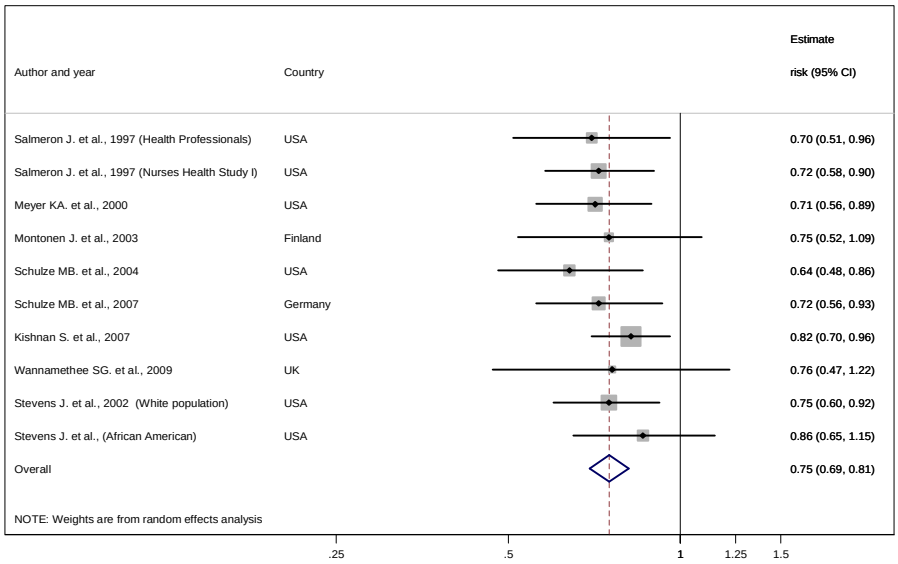
Table 4.7 Pooled estimate of key fibre sources and risk of T2DM

Fibre source	Total no of cohorts	Significant inverse association	no of studies in meta-analysis	Pooled estimate (95%CI; I^2)
Cereal fibre (High vs. low)	12	9	9	0.75(0.69, 0.81; 0.0%)
Cereal fibre (continuous)	3	1	3	0.96(0.93, 0.99; 0.0%)
vegetable fibre	11	3	7	1.02(0.89, 1.18; 41%)
Fruit fibre	11	1	6	0.98(0.83, 1.54; 51%)
Legume fibre	3	0	-	-

4.3.5.1 Cereal fibre intake

Risk of T2DM was significantly lower by 25% among participants in the highest cereal fibre intake in comparison to lowest intake (risk=0.75; 95%CI: 0.69, 0.81) as shown in Figure 4.3. Pooled estimate from other three cohorts showed significant 4% risk reduction of T2DM with increased intake of cereal fibre as shows in Table 4.7 (Figure 6.c in appendix C shows cereal fibre forest plot) (Stevens *et al.*, 2002, Hodge *et al.*, 2004, Barclay *et al.*, 2007). The remained cohort reported significant risk reduction of T2DM with high intake of cereal fibre among men and women (Hopping *et al.*, 2010). From Table 4.7, nine out of 12 cohorts reported significant effect of cereal fibre on the risk of T2DM.

Figure 4.3 Forest plot of estimated effect and 95%CI for highest vs. lowest cereal fibre intake and risk of T2DM



4.3.5.2 Other dietary fibre sources

Details of eligible cohort studies were provided in Table 4.8. The pooled risk estimate showed no association between risk of T2DM and people who consume high vegetable fibre compared with lower vegetable fibre consumers (Figure 7.c in appendix C). One of the eight studies reported significant results (Weng *et al.*, 2012). From the excluded studies, two reported significant inverse associations between the risk of T2DM and vegetable fibre intake (Hopping *et al.*, 2010, Barclay *et al.*, 2007). Risk reduction was 22% among men with high vegetable consumption but not among women and second cohort reported 24% risk reduction with every 5g increment in vegetable fibre among both gender.

Pooled estimate from seven cohorts showed no significant association between fruit fibre intake and risk of T2DM among high consumers in comparison to lowest consumers (Table 4.8 and appendix C provided forest plots in Figure 7.c and 8.c). Only one eligible cohort reported high risk of T2DM among lowest fruit fibre intake in comparison to highest group (Weng *et al.*, 2012) and none of the excluded cohorts showed an effect of high fruit fibre on the risk of T2DM (Hopping *et al.*, 2010, Barclay *et al.*, 2007, Hodge *et al.*, 2004, Stevens *et al.*, 2002). Moderate heterogeneity was seen in both meta-analysis, thus caution in interpretation is recommended (Higgins *et al.*, 2003).

Three cohort studies reported the associations between legume fibre intake and risk of T2DM, only two provided risk comparing highest versus lowest intake

thus meta-analysis was not conducted. None of these were found to be statistically significant (Hodge *et al.*, 2004, Meyer *et al.*, 2000, Stevens *et al.*, 2002).

Key findings from the systematic review in this chapter

- No association between risk of T2DM and fibre consumption, comparing highest and lowest consumers.
- Potential sources of moderate heterogeneity in pooled estimate of TDF may relate to gender, duration of follow up, country of origin and other possible sources not explored.
- Significant risk reduction of T2DM by 18% (95%CI: 0.71, 0.90) was observed among women with total high fibre intake in comparison to women in the lowest fibre intake group.
- People in the high intake of insoluble fibre group experienced lower risk of T2DM (risk estimate = 0.75; 95%CI: 0.66, 0.86) in compared to people with lower intake.
- People with greater cereal fibre intake have significantly lower risk of T2DM in comparison with lowest consumers.
- No associations were observed with the intakes of the remaining fibre sources and risk of T2DM comparing extreme groups.

Table 4.8 Cohort studies of dietary fibre sources intakes and risk of T2DM

	Author	No Participants	Age (years)	Follow-up (Years)	DM cases (Dg)	Dietary fibre assessment	Dietary fibre intake	Estimated risk	*P*-trend	Adjusted variables**
1	Weng *et al.* (2012) Taiwan	1,604 58% Women	Mean DM 52.7±10 Non DM 49.8±12	4.6	141 FBG ≥7mmol/l or self-report	64-items FFQ DF Not specified	Vegetable fibre Q1 <5.15g/d Q4 >12.9g/d Fruit fibre Q1 <5.6g/d Q4 >16.5g/d **Q1 vs. Q4[1]**	Vegetable fibre M(1)1.81(1.01,3.22) M(2)1.91(1.06,3.44) M(3) 2.23(1.22,4.08) Fruit fibre M(1)1.63 (0.95,2.81) M(2)1.74 (1.02,2.98) M(3)1.81 (1.05,3.13)	 0.02 0.01 <0.01 0.05 0.02 0.01	M(1) age, sex, age-sex interaction, EI residential area. M(2) M1 plus FH, BMI, central obesity. M(3) M2 plus education, smoking, alcohol, PA, HTN, high cholesterol, hypertriglyceridemia, low HDL-cholesterol
2	Hopping *et al.* (2010) USA	75,512 52% Women	45-75	14	8587 Self-report cases	FFQ AOAC[1] method	Men Grain fibre g/(4184kJ.d) Q1 <3.8 Q5 ≥9.6 Vegetable fibre g/(4184kJ.d) Q1 <4.4 Q5 ≥10.6 Fruit fibre g/(4184kJ.d) Q1 <1.6 Q5 ≥7.8 **Q5 vs. Q1[5]**	Grain fibre All 0.91(0.82,1.00) 0.81(0.67,0.99)[2] 0.98(0.87,1.11)[3] 0.83(0.64,1.07)[4] Vegetables fibre All 0.78(0.68, 0.88) 0.65(0.52,0.82)[2] 0.79(0.67,0.93)[3] 0.99(0.75,1.32)[4] Fruit fibre All 0.93(0.84,1.02) 0.88(0.71,1.08)[2] 0.97(0.85,1.11)[3] 0.93(0.72,1.19)[4]	 <0.01 0.02 0.30 0.15 <0.01 <0.01 0.01 0.89 0.17 0.34 0.75 0.46	M(1) ethnicity, BMI, PA, education, and EI.

** Abbreviations used in column: EI= energy intake; PA= physical activity; FH= family history of diabetes; HTN =hypertension; BMI= body mass index; acids; [1]AOAC= Association of Official Analytical Chemist; MET= Metabolic Equivalent Task; Inso-NCP = insoluble non-cellulosic polysaccharides; [1]Q1 vs. Q4 estimated risk among lowest fibre group in comparison to highest fibre group; [2]Caucasian, [3]Japanese American, [4]Native Hawaiian; [5]Q5 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group;.

Continue

N	Author	No Participants	Age (years)	Follow-up (Years)	DM cases (Dg)	Dietary fibre assessment	Dietary fibre intake	Estimated risk	*P*-trend	Adjusted variables**
Continue 2	Hopping *et al.* (2010) USA	75,512 52% Women	45-75	14	8587 Self-report cases	FFQ AOAC[1] method	Women Grain fibre g/(4184kJ.d) Q1 <4.2 Q5 ≥10.2 Vegetable fibre g/(4184kJ.d) Q1 <2.6 Q5 ≥10.4 Fruit fibre g/(4184kJ.d) Q1 4.2 Q5 ≥10.2 **Q5 vs. Q1[5]**	Grain fibre All 0.88(0.79,0.97) 0.79(0.62,1.01)[2] 0.91(0.79,1.04)[3] 0.87(0.69,1.09)[4] Vegetables fibre All 0.96(0.87-1.08) 0.94(0.73,1.22)[2] 1.00(0.86,1.16)[3] 0.92(0.73,1.16)[4] Fruit fibre All 0.0.95(0.85,1.06) 0.85(0.65,1.11)[2] 0.98(0.85,1.12)[3] 0.99(0.79,1.24)[4]	0.02 0.03 0.35 0.11 0.38 0.25 0.94 0.62 0.21 0.36 0.37 0.70	M(1) ethnicity, BMI, PA, education, and EI.
3	Wannamethee *et al.* (2009) UK	3428 100% Men	60 – 79	7	162 Self-report cases	FFQ. Englyst method	Cereal fibre < 6.9 vs. rest Vegetable fibre <11.3g/d vs. rest **Q5 vs. Q1[5]**	Cereal fibre M(1)1.61(1.16,2.23) M(2)1.43(1.00,2.06) M(3)1.32(0.91,1.91) Vegetable fibre M(1)1.64(1.18,2.29) M(2)1.40(0.98,1.98) M(3)1.28(0.89,1.82)	- - - - - -	M(1) age M(2) age, WC, smoking, PA, social class, alcohol, preexisting MI, stroke, use statin, EI. M(3) M2 plus IL-6, t-PA, and GGT
4	Krishnan *et al.* (2007) USA	40,078 100% Women	21-69	8	1938 Self-report cases	FFQ AOAC[1] method	Cereal fibre g/d: Q1 ≤2.3 Q5 ≥5.9 **Q5 vs. Q1[5]**	Cereal fibre M(1)0.67(0.58,0.77) M(2)0.81(0.69,0.96) M(3)0.82(0.70,0.96) BMI < 25: 0.41(0.24,0.72) BMI≥ 25: 0.88(0.75,1.04)	<0.01 0.04 0.01 0.003 0.11	M(1) age M(2) age, BMI,EI, FH, PA, smoking. M(3) 2 plus GI, protein, fat intake.

** Abbreviations used in column: EI= energy intake; PA= physical activity; FH= family history of diabetes; HTN =hypertension; BMI= body mass index; acids;; IL-6 = initerleukin-6; t-PA = tissue plasminogen activator; GGT = Gamma-glutamyl transferase; [1]AOAC= Association of Official Analytical Chemist; [2]Caucasian, [3]Japanese American, [4]Native Hawaiian; [5]Q5 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group.

Continue

N	Author	No Participants	Age (years)	Follow-up (Years)	DM cases (Dg)	Dietary fibre assessment	Dietary fibre intake	Estimated risk	*P*-trend	Adjusted variables**
5	Barclay *et al.* (2007) Australia	1833 Men & women	+49	10	138 self-report cases	FFQ AOAC method	Per 5g/d increment: Cereal fibre Vegetable fibre Fruit fibre	Cereal fibre M(1) 0.96(0.80,1.16) M(2)0.96(0.78,1.20) Vegetable fibre M(1) 0.72(0.57,0.93) M(2) 0.76(0.57,0.99) Fruit fibre M(1)0.95(0.79,1.13) M(2)0.94(0.78,1.15)	 0.69 0.74 0.01 0.04 0.53 0.56	M(1) age, sex. M(2) M1 plus FH, smoking, TG, HDL, PA.
6	Schulze *et al.* (2007) Germany	25,067 61% Women	35 – 65 Women 40-65 Men	7	844 Self-report cases	FFQ AOAC method	Cereal fibre g/d Q1 6.6 Q5 16.6 Vegetable fibre g/d Q1 0.7 Q5 3.4 Fruit fibre g/d Q1 0.2 Q5 4.7 **Q5 vs. Q1[2]**	Cereal fibre M(1)0.73(0.57,0.94) M(2) 0.72(0.56,0.93) Vegetable fibre M(1)0.93(0.75,1.17) M(2)0.93(0.74,1.17) Fruit fibre M(1)0.89(0.71,1.13) M(2)0.89(0.70,1.13)	 0.02 0.02 0.64 0.66 0.36 0.22	M(1) age, sex, education, PA, cycling, occupational activity, smoking, alcohol, EI, BMI, WC, Mg, PUFA:SFA, MUPA:SFA ratio, CHO. M(2) M1other fibres
7	Schulze *et al.* (2004a) USA	91,249 100% Women	24-44	8	741 Self-report cases	FFQ AOAC method	Cereal fibre g/d Q1 <3.8 Q5 >7.3 Vegetable fibre g/d Q1 <4.2 Q5 >8.6 Fruit fibre g/d Q1 <1.6 Q5 >4.8 **Q5 vs. Q1[2]**	Cereal fibre M(1)0.54(0.42,0.70) M(2)0.64(0.48,0.86) Vegetable fibre M(1)0.87(0.69,1.09) M(2)1.12(0.87,1.46) Fruit fibre M(1)0.70(0.56,0.88) M(2)0.79(0.6,1.02)	 <0.01 0.004 0.50 0.19 <0.01 0.04	M(1) age, BMI M(2) M1 plus EI, alcohol, PA, FH, Smoking, HTN, high blood cholesterol, postmenopausal hormonal use, oral contraceptive use, GL, Mg intake, caffeine intake, other fibre types

** Abbreviations used in column: EI= energy intake; PA= physical activity; FH= family history of diabetes; HTN =hypertension; BMI= body mass index; CHO=carbohydrates; SFA= saturated fatty acid; PUFA=polyunsaturated fatty acids, MUPA= monounsaturated fatty acids; HDL= high density lipoprotein; [1]AOAC= Association of Official Analytical Chemist; [2]Q5 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group.

Continue

N	Author	No Participants	Age (years)	Follow-up (Years)	DM cases (Dg)	Dietary fibre assessment	Dietary fibre intake	Estimated risk	*P*-trend	Adjusted variables**
8	Hodge *et al.* (2004) Australia	31,641 59% Women	27-75	4	365 Self-report cases	FFQ AOAC[1] method	Per 10g/d increment Cereal fibre Fruit fibre Per 5g/d increment Vegetable fibre Per 1g/d increment Potato fibre Legumes fibre	Cereal fibre M(1)0.97(0.79,1.20) M(2)1.08(0.88,1.32) Fruit fibre M(1)0.93(0.77,1.11) M(2)0.97(0.81,1.16) Vegetable fibre M(1) 1.01(0.87,1.18) M(2)1.00(0.86,1.17) Potato fibre M(1)1.04(0.92,1.17) M(2)1.03(0.91,1.16) Legumes fibre M(1)1.01(0.97,1.06) M(2)1.01(0.96,1.06)	 0.79 0.46 0.40 0.71 0.89 0.96 0.57 0.65 0.62 0.67	M(1) Age country of birth, PA, FH, weight change last 5 yrs, education, alcohol, FH, EI. M (2) M1 plus BMI, WHR
9	Montonen *et al.* (2003) Finland	4316 47% Women	40-69	10	156 Self-report cases	Dietary history and interview AOAC[1] method	Cereal fibre g/d Q1 (0.47-12) Q4 (24.5-111) Fruit fibre g/d Q1 (0-0.99) Q4 (3.4-36.8) Vegetable fibre g/d Q1 (0.11-3.7) Q4 (6.8-26.5) **Q4 vs. Q1[2]**	Cereal fibre M(1)0.36(0.18,0.70) M(2)0.39(0.20,0.77) Vegetable fibre M(1)1.33(0.81,2.18) M(2)1.19(0.46,3.04) Fruit fibre M(1)0.98(0.63,1.53) M(2)0.92(0.4,2.13)	 0.005 0.01 0.44 0.86 0.78 0.87	M(1) age, sex, geographic area, and EI. M(2) M 1 plus smoking, BMI, intake of fruit and berries, and vegetables

** Abbreviations used in column: EI= energy intake; PA= physical activity; FH= family history of diabetes; HTN =hypertension; BMI= body mass index; CHO=carbohydrates; SFA= saturated fatty acid; PUFA=polyunsaturated fatty acids, MUPA= monounsaturated fatty acids; HDL= high density lipoprotein; [1]AOAC= Association of Official Analytical Chemist; [2]Q4 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group.

Continue

N	Author	No Participants	Age (years)	Follow-up (Years)	DM cases (Dg)	Dietary fibre assessment	Dietary fibre intake	Estimated risk	*P*-trend	Adjusted variables**
10	Stevens *et al.* (2002) USA	12,251 70% Women 78% White 22% African American	45-64	9	1,447 Self-report cases	FFQ AOAC[1] method	Per 1g/d increment of cereal fibre, fruit fibre and legumes fibre	Cereal fibre M(1)0.94(0.91,0.97)[2] M(2)0.95(0.92,0.98)[2] M(1)0.97(0.92,1.03)[3] M(2)0.98(0.92,1.03)[3] Fruit fibre M(1)0.99(0.97,1.01)[2] M(2) 1.00(0.98,1.02)[2] M(1)1.00(0.98,1.03)[3] M(2)1.00(0.98,1.33)[3] Legume fibre M(1)1,00(0.95,1.05)[2] M(2)1.00(0.95,1.05)[2] M(1)0.95(0.87,1.04)[3] M(2)0.96(0.88,1.04)[3]	 <0.01 0.006 0.46 0.52 0.57 0.84 0.53 0.47 0.81 0.77 0.32 0.36	M(1) age, BMI, sex, field centre M (2) 1 plus education, smoking, PA.
11	Meyer *et al.* (2000) USA	35,988 100% Women	55-69	6	1141 Self-report cases	FFQ AOAC[1] method	Cereal fibre g/d Q1 <3.4 Q5 >7.5 Vegetable fibre g/d Q1 <5.75 Q5 >10.14 Legume fibre g/d Q1 <0.31 Q5 >1.21 Fruit fibre g/d Q1 <2.55 Q5 >7.02 **Q5 vs. Q1[4]**	Cereal fibre M(1)0.64(0.53,0.79) M(3)0.78(0.62,0.99) M(5)0.71(0.56,0.89) Vegetables fibre M(1)0.97(0.80,1.18) Legumes fibre M(1)1.10(0.91,1.33) Fruit fibre M(1)1.17(0.96,1.42)	 0.0001 0.025 0.0017 0.77 0.17 0.081	M(1) Age, EI,WHR, BMI education, smoking, alcohol, PA, M(3) M1 plus Mg and total grain M(5) M1 plus Mg and whole grain

** Abbreviations used in column: EI= energy intake; PA= physical activity; FH= family history of diabetes; HTN =hypertension; BMI= body mass index; CHO=carbohydrates; WHR= waist hip ratio; [1]AOAC= Association of Official Analytical Chemist; [2]White population; [3]African American population; [4]Q5 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group.

Continue

N	Author	No Participants	Age (years)	Follow-up (Years)	DM cases (Dg)	Dietary fibre assessment	Dietary fibre intake	Estimated risk	*P*-trend	Adjusted variables**
12	Salmeron *et al.* (1997b) USA	65,173 100% Women	40-65	6	915 Self-report	FFQ AOAC[1] method	Cereal fibre g/d Q1 2.0 Q5 7.5 Vegetable fibre g/d Q1 3.4 Q5 9.6 Fruit fibre g/d Q1 1.4 Q5 7.6 **Q5 vs. Q1[2]**	Cereal fibre 0.72(0.58,0.90) Vegetable fibre 1.17(0.93,1.46) Fruit fibre 0.87(0.70,1.08)	0.001 0.54 0.39	M (1) age, BMI, alcohol, FH, prior weight change, and time period.
13	Salmeron *et al.* (1997a) USA	42,759 100% Men	40-75	6	523 Self-report cases	FFQ AOAC[1] method	Cereal fibre g/d Q1 2.5 Q5 10.2 Fruit fibre g/d Q1 1.2 Q5 8.3 Vegetable fibre g/d Q1 3.5 Q5 11.3 (Q5 vs. Q1)	Cereal fibre 0.70(0.51,0.96) Vegetable fibre 1.12(0.84,1.49) Fruit fibre 1.01(0.76,1.36)	0.007 0.65 0.68	Adjusted for age, BMI, alcohol, smoking, FH

** Abbreviations used in column: EI= energy intake; PA= physical activity; FH= family history of diabetes; HTN =hypertension; BMI= body mass indexo; [1]AOAC= Association of Official Analytical Chemist; [2]Q5 vs. Q1 estimated risk among highest fibre group in comparison to lowest fibre group.

4.4 Discussion

This comprehensive review showed that among high fibre consumers, lifestyle and dietary factors in general were in the line with dietary guidelines (Food Standards Agency, 2007).

4.4.1 Total dietary fibre intake and risk of T2DM

Pooled estimate comparing highest versus lowest dietary fibre intake showed a lack of association between TDF and risk of T2DM with moderate heterogeneity that may possibly be explained by variation in gender and follow-up period and geographic location as demonstrated in the subgroup analysis. The current evidence suggests protective effect of TDF on risk of DM in women. Additionally, it showed significant inverse association between TDF and risk of T2DM among participants who were followed for less than 7 years duration but not from studies with a longer duration. Change in diet over time may have an influence on the risk of diabetes. On the other hand, exploring the degree of stability in dietary intake over a longer follow-up duration cannot be achieved by single dietary assessment which may possible partly contribute to heterogeneity.

From the total number of cohorts, half of the prospective studies reported non-significant effect of high intake of TDF on the risk of diabetes. One of the studies that not included in the pooled estimate showed gender variation in the dietary fibre and risk of T2DM and suggested that type of diet may differ between women and men, where women more likely to misreport their dietary intake (Hopping *et al.*, 2010). Even for the two studies that examine dose-response relationship, the results differ with almost similar increment in TDF (5g/day and 4.8g/day) (Barclay *et al.*, 2007, Sluijs *et al.*, 2010b). This reflects a considerable inconsistency between the studies and even within the study (Hopping *et al.*, 2010).

Second half of the cohorts that showed significant inverse association, high dietary fibre consumers in seven cohort studies showed a wide range of significant risk reductions of T2DM ranging from 11% to 49%. However, three out of them reported approximately a 25% significant risk reduction (Hopping *et al.*, 2010, Meyer *et al.*, 2000, Salmeron *et al.*, 1997b) in comparison to the lowest categories.

Another potential source of heterogeneity could be explained by the varied amount of daily dietary fibre consumed across the cohorts (Higgins *et al.*, 2003). In this review, the significant protective effects of dietary fibre on the risk of T2DM observed with daily dietary fibre intakes were more than 23.6g, 24.1g, and 33.2g in

comparison to less than 15.3g, 11.8g and 19.2g respectively (Meyer *et al.*, 2000, Salmeron *et al.*, 2001, Montonen *et al.*, 2003). On the other hand, non-significant risk reductions were reported with relatively similar amounts of daily fibre (more than 22g, 26g, 27g, 29.7g and 31g) (Schulze *et al.*, 2004a, Stevens *et al.*, 2002, Schulze *et al.*, 2007, Salmeron *et al.*, 1997a, Wannamethee *et al.*, 2009) in comparison to lowest fibre intake group (14.2g, 10.2g, 15.1g, 17.1g and 20g respectively).

Variation in characteristics of studied population also may contribute to diversity between and within cohort studies (Higgins *et al.*, 2003). Genetic background and ethnicity were associated with diabetes incidence as discussed in chapter 1. Significance difference in dietary fibre intake was reported across different ethnic background in England (South Asian, black African and white European) (Goff *et al.*, 2013). In a multi-ethnic cohort study (Hopping *et al.*, 2010) Caucasians showed an inverse association which was significant, but this relation was not statistically significant among the Japanese American group. Swedish study by Ericson *et al.* (2013) was further examined the risk of diabetes based on genetic background (Hindy *et al.*, 2012); a protective effect of high fibre intake on the risk of T2DM was reported among participants with CC genotype of TCF7L2 rs7903146 but not with other genotypes.

Another possible source of variation across cohort studies is the different sources of fibre which may have different effect of fibre on the development of diabetes. In vitro starch digestibility study demonstrated that incorporated psyllium or oat bran into ready to eat snack resulted in significant reduction of glucose response by 15% and 5.4% respectively compared to control groups after 120 minutes of consumption, while in vivo study showed that only psyllium snack reduced postprandial glucose concentration at 20 and 60 minutes in comparison with oat bran snack and control (Brennan *et al.*, 2012)

Hopping *et al.* (2010) suggested that variation based on ethnic background may be because wheat was the commonly consumed fibre-rich food among Caucasians in comparison to rice, which was the commonly consumed fibre rich food in the other group.

The majority of studies used the AOAC analytical method and one UK study reported the NSP. More detail on the dietary fibre analysis is in Chapter 2. However, it is hard to explore the fibre analysis method as a potential source of variation in the

subgroup analysis. Chapter 7 aims to examine whether high AOAC-fibre consumers similar to high NSP consumers in the UKWCS which can be useful to explore the impact of fibre analysis method in epidemiological aspect.

Also some of the heterogeneity may be attributed to variation in dietary assessment methods. The common dietary assessment tool used in prospective studies is the FFQ as most cost-effective way to estimate dietary intake of large studied populations (Cade *et al.*, 2004b). The FFQ is not without problems, as estimated dietary intake is subject to bias in term of estimated portion size, the influence of the number of food items included in the FFQ and misreporting by participants. Food item numbers of the FFQ in cohorts included in the current meta-analysis varied between 66 (Stevens *et al.*, 2002) and 162 (Wannamethee *et al.*, 2009). The Melbourne Collaborative Cohort Study (Hodge *et al.*, 2004) argued that non-significant findings may relate to the lack of details for some fibre-rich food such as breakfast cereal and type of pasta and rice which may resulted in underestimation of dietary fibre intake.

Another potential source of heterogeneity could be that the variation in the sources of cereal fibre may have a different impact. It was reported by Cummings and Stephen (2007) that specific grain types contributing to whole grain intake can vary between countries. Wheat was the main contributor of whole grain intake in the UK, while high percentage of oat contributed to whole grain intake in the USA. This suggests the impact of type of food that provided dietary fibre may be differ. However, in this meta-analysis, the pooled estimate on cereal fibre intake on the risk of diabetes was highly consistent and no heterogeneity was seen.

Methods of case ascertainment can also have an impact on risk estimate in prospective studies. Undiagnosed cases may mask the true association since people with asymptomatic diabetes are more likely to be undiagnosed (Salmeron *et al.*, 1997a, Stevens *et al.*, 2002). Most of confirmation on diabetes diagnosis in cohorts (Schulze *et al.*, 2004a) aimed to rule out non-cases in self-reported diabetes group but not to identify cases in diabetes free group.

Small sample size is another aspect that may partly explain the results in some studies (Wannamethee *et al.*, 2009, Barclay *et al.*, 2007).

However, lack of association between total dietary fibre intake and risk of T2DM was noticed even in cohort studies with a large sample size (741 and 844 diabetic cases) (Schulze *et al.*, 2007, Schulze *et al.*, 2004a). On the other hand,

prospective studies that included much smaller sample size reported significant inverse association between dietary fibre intake and risk of T2DM (156 and 141 diabetic cases) (Weng *et al.*, 2012, Montonen *et al.*, 2003).

Experimental evidence examined the association between fibre intake and diabetes status markers. Interventional study with high fibre intake from food (cereals and apples) over 3 month duration in healthy adults compared with control group with similar nutrient composition except for fibre showed significant decrease in fasting blood glucose level with high fibre group (Aller *et al.*, 2004). Another trial among overweight adult with intake of wholegrain that provide (30g/day of fibre) for 6 weeks in compared to refined grain group (provide 17g/day of fibre) did not show significant changes in the fasting blood glucose levels between the groups (Andersson *et al.*, 2007). Variation in studied population and duration of intervention may explain these results.

4.4.2Insoluble and soluble dietary fibre intakes and risk of T2DM

Further analysis in this review on the effect of types of dietary fibre on the risk of T2DM showed significant inverse association between the intake of insoluble dietary fibre but not with soluble fibre on the risk of T2DM comparing extreme fibre groups. The protective effect may be related to the higher amount consumed in comparison to SDF intake which may possibly be adequate to show a significant effect. In most of the studies, insoluble fibre consumption was double the amount of soluble fibre, which may indicate that the amount of soluble fibre intake may be insufficient to show significant protective results. Further research is needed on dietary fibre fractions in relation to chronic diseases such as diabetes. Overall, this demonstrates less clarity regards type of dietary fibre that effectively reduce the risk of T2DM thus, there is a need for further prospective studies to examine the effect of insoluble and soluble fibre intakes on the development of diabetes.

4.4.3Intakes of main fibre sources and risk of T2DM

Pooled cereal estimate showed significant risk reduction by 25% among people who consume high cereal fibre in comparison to lower consumers. Also pooled estimate from cohorts that reported dietary fibre intake as continuous showed significant risk reduction by 4% among highest cereal fibre intake in comparison to lowest cereal consumers. Some of the possible reasons for negative findings in the few identified studies were reported. For example, Atherosclerosis Risk in

Community (ARIC) study showed non-significant association among the African American population but not the white population which also may be because of small sample size as incidence of diabetes is higher in African-Americans (17.5% in comparison to whites (10.2%) (Stevens *et al.*, 2002). Also underestimation of dietary fibre intake due to lack of some fibre rich food in the FFQ such as breakfast cereal (Hodge *et al.*, 2004).

From the current review, the pooled estimate showed no relationship between risk of T2DM and intake of vegetable fibre. Also for the fruit fibre intake no significant link to risk of T2DM in pooled analysis. None of the few prospective studies found an effect of legumes fibre on the risk of T2DM. This may suggest that specific fibre types/sources are more or less potent in terms of influencing glycaemia and insulinemia. However, pectin consumption in interventional study showed significant reduction in blood glucose (Jenkins *et al.*, 1977). Zulphen study reported significant inverse association between habitual pectin intake and incremental area under the curve for glucose among middle aged participants (Feskens and Kromhout 1990). So there is conflict in these results.

Generally, current meta-analysis showed lack of significant associations between T2DM and main fibre sources (vegetables, fruit and legumes) except for cereal fibre. This could be explained by inadequate intake of fruit, vegetable and legumes fibres in comparison to cereal fibre intake.

Recent review suggested the presence of aleurone cell walls in cereal foods which are rich in phenolic acids and cross-linked with other bioactive substances but absent in fruit fibre (Lillioja *et al.*, 2013) possibly explain the beneficial effect of cereal fibre. It was also reported that cereal fibre could be a marker of other vital micronutrients which are present in wholegrain (Willett, 2012).

Another aspect which needs to be considered is the measurement error with estimating fruit and vegetable fibre intakes, especially as these types of food are usually consumed in composite dishes such as green salad and fruit salad which are hard for participants to estimate the actual amount consumed. More exploration on the measurement errors from FFQ by comparing two dietary assessment methods will be in chapter 10.

On the other hand, recent review by Palafox‐Carlos *et al.* (2011) demonstrated that dietary fibre from fruit and vegetables interferes with the

bioavailability of phenolic compounds and carotenoids by trapping micronutrients in the fibre matrix which decreases absorption of these bioactive compounds.

One of the issues that emerges from these findings, the British Regional Heart study (Wannamethee *et al.*, 2009) which reported an intake of less than 20g/day of NSP in comparison to higher than 20g/day increased the risk of diabetes by 47% among adult men after adjustment for potential confounders listed in Table 4.5. This is higher than the recommended NSP value (18g/d) among adult (Department of Health, 1991) however, this study was only on men and may not represent the general UK population, which restrict the generalization of their results in addition to overestimation bias should be considered with the use of FFQ dietary information.

4.4.4Strength and limitations

The main advantages of meta-analysis are to increase the chance of detecting a more reliable pooled estimated effect from existing relevant studies, to improve the estimation of the dietary fibre effect when it is based on more evidence and to allow investigation of the consistency of effect where studies characteristics' differ (Deeks *et al.*, 2008).

On the other hand, some disadvantages such as the presence of bias in individual studies may be combined in the meta-analysis, and diversity in the exposure and outcome measures of the studies should be considered when selecting primary studies otherwise the meta-analysis will be irrelevant and misleading (Ioannidis and Lau, 1999). Deeks *et al.* (2008) says random effect meta-analysis is the most appropriate approach for looking at cohort studies rather than fixed effect. This allows each of the studies to estimate a different effect size and thus to obtain strong assumptions.

A number of limitations need to be considered in the current chapter. First the gender subgroup analysis did not show significant heterogeneity within women studies. However, this was subject to bias as the presence of the small number of studies among women group may lower the statistical power for detecting variation within each group. Also some cohort studies were not included in pooled estimate because of unavailable separate estimated risk by gender which may have a potential effect on the pooled estimate. Studies from European countries were found to be more consistent than the USA. However, the subgroup analysis based on country of

origin showed lack of association between fibre intake and risk of T2DM in both European and USA groups.

Secondly, analysis was undertaken for only fully-adjusted models, but this may not rule out the potential effect of residual confounders. On the other hand, some studies adjusted for large numbers of confounders have the disadvantage of estimation bias due to over-adjustment (Schulze *et al.*, 2007, Hodge *et al.*, 2004).

Thirdly, the majority of cohort studies used the FFQ to assess the dietary fibre intake among studied populations and although several studies reported a reasonable validity and reproducibility of the FFQ used, measurement errors cannot be avoided (Cade *et al.*, 2004b). Finally, comparing highest to lowest categories has its own limitations. For example, there is variation in the number of categories as studies' populations may subdivide into 4 quartiles or 5 quintiles, depending on exposure distribution, and there is always a concern regard loss of information with categorization. This issue may have contributed to heterogeneity in the current pooled analyses (Tu and Greenwood, 2012). It was noticed that dietary fibre categories were overlapped across studies, where highest group in one study (Q5>23g/day) (Meyer *et al.*, 2000) could be lowest in another study (Q1<24.1g/day) (Weng *et al.*, 2012). Also variation within the categorical fibre groups in the cohort may not be so great to show significant lower risk (Schulze *et al.*, 2004a).

The differences across studies should be considered and this was addressed partly in this review by using random effect method. This review included prospective studies which have the advantage of eliminating the chance of dietary intake recall bias since the history of dietary intake is taken before the diagnosis of diabetes.

Further research in this field regarding the effect of different dietary fibre sources and types of dietary fibre on the development of diabetes would be of great help in optimizing potential health benefits. Also, more studies are needed to examine the amount of increment of dietary fibre intake in relation to the risk of diabetes giving more quantitative data in term of dose-response relationship to support dietary guidelines aiming for primary prevention of disease. Interventional trials focusing on dietary fibre intake could generate information on the optimal amount of fibre that may protect from T2DM.

4.4.5 Potential mechanisms of the dietary fibre effect on the development of diabetes

In general, delayed gastric emptying, slow rate of intestine digestion and absorption, and fibre fermentation, that results in short chain fatty acid production in the large intestine, were all demonstrated by Jenkins and his colleagues (1987) as physiological beneficial effects of dietary fibre. Early studies by Jenkins reported that intake of soluble fibres reduced the postprandial glucose response (Jenkins *et al.*, 1978) which may related to its viscosity properties. In the current review, insoluble fibre rather than soluble fibre has a clearer protective effect on the risk of T2DM. Although, the potential mechanisms involved are still not very clear. However, a review by Weickert and Pfeiffer (2008) demonstrated that studies on increased satiety via high fibre diets may explain the protective insoluble fibre effect that is found in prospective studies acting through body weight reduction or limit weight gain. Furthermore, improvement in insulin sensitivity with high fibre diets intake was also suggested as a potential mechanism (Weickert and Pfeiffer, 2008). Another possible mechanism that has been investigated in a previous study was the inflammatory markers which were reported as strong predictors of diabetes (Sattar *et al.*, 2008).

Evidence of the beneficial effect of magnesium intake on the risk of diabetes was demonstrated in two meta-analyses (Larsson and Wolk, 2007, Dong *et al.*, 2011). As significant beneficial effect of high cereal fibre intake on the risk of diabetes was demonstrated in a cohort study after adjustment for confounders including magnesium intake (Schulze *et al.*, 2007). However because of the high correlation between the intake of fibre and magnesium, it was hard to separate the effect of each so this was subject to estimation bias. A study showed that intracellular magnesium has a key role in insulin action where insufficient magnesium suggested an underlying reason for insulin resistance (Barbagallo *et al.*, 2003).

A number of mechanisms are suggested through which dietary fibre might prevent T2DM. Jenkins *et al.* (1978) showed significant reduction in blood glucose post fibre consumption which highly correlated with dietary fibre viscosity. But importantly Jenkins *et al.* (2002) studies of lente carbohydrates, all tended to suggest that water soluble fibres where most potent leading to a smaller glucose Area under the curve (AUC) not insoluble fibre. This believed to slow the rate of intestinal

starch digestion and absorption through delay gastric emptying that led to improve glucose response (Jenkins *et al.*, 1987). On the other hand, recall that some cohorts have found insoluble fibre or cereal fibre to reduce risk of T2DM, but soluble fibre less so (Salmeron *et al.*, 1997b, Meyer *et al.*, 2000, Montonen *et al.*, 2003).

In recent years, the attention on the bioactive compounds (polyphenols) that found attached to the plant cell walls was suggested to have the beneficial effect on human health (Saura-Calixto, 2010) where dietary fibre believed to be a carrier for these antioxidants substances.

4.5 Conclusion

Meta-analysis provides a useful summary of the extent of variations between the studies on the existing evidence by pooling risk that compare highest versus lowest, however the pattern in the relationship is less clear. Estimated risk was pooled from studies performed in different countries which allow investigation of the effect of dietary fibre based on a wide spectrum of dietary fibre intake.

Considerable variations between the studies may possibly be explained by difference in gender, origin of the study and follow-up duration which was explored in the meta-analysis. Method used to assess diet and publication bias were other potential reasons for heterogeneity that could not explore in the current meta-analysis. Low level of consistency of evidence on total dietary fibre intake and risk of T2DM may suggest further meta-analysis for dose-response relationship.

The findings discussed in this chapter support the hypothesis that diets with cereal fibre decrease the risk of T2DM. Total dietary fibre and the remaining fibre sources showed no significant association among high fibre consumers compared to lower consumers.

Current meta-analysis suggested a significant protective effect of insoluble rather than soluble fibre on the development of diabetes. No protective effect of high vegetables, fruit and legumes fibres intake on the risk of T2DM in compared to people with lower intakes. It has been suggested that the presence of other components such as magnesium associated with cereal fibre may partly explain the strong inverse association with the risk of T2DM. However, the dose-response relationship between intake of dietary fibre sources and the risk of T2DM needs exploration in future work.

Chapter 5: Dietary fibre analysis in commonly consumed legumes in UK: AOAC-fibre: NSP ratio for legumes group

5.1 Introduction

Legumes are a rich source of dietary fibre as well as providing energy from starch and protein content (Dilis and Trichopoulou, 2009). The beneficial therapeutic effects of legumes have been reported in a pooled analysis on participants with and without a history of diabetes which showed improvement in markers of glycaemic control with consumption of beans (Sievenpiper *et al.*, 2009). This result may be explained by the characteristic low glycaemic response generated by legumes (Jenkins, 1980). Some types of dietary fibre have been shown to slow rate of digestion and absorption which may be due to its viscosity characteristics (Jenkins *et al.*, 1987). It is this slow digestibility characteristic of legume starch is thought to be contribute to the low glucose response (Jenkins, 1980).

To date various methods have been developed and introduced to measure dietary fibre content in food and food products. The two analytical methods that are commonly used for dietary fibre analysis are the enzymatic gravimetric method and the enzymatic chemical method (Englyst *et al.*, 2007). The influence of dietary fibre analytical methods upon dietary fibre measurement have been extensively reviewed (Englyst *et al.*, 2007).

The Englyst method is based on chemical analysis of NSP which represent the intrinsic plant cell walls provided in a plant rich food. The enzymatic gravimetric method endorsed by the Association Of Analytical Chemists (AOAC) (Prosky *et al.*, 1985) is based on gravimetric measurement after removing enzymatic digestion of starch and protein. Details on the analytical methods were demonstrated in chapter 1.

From dietary fibre definitions included components other than NSP include, lignin, enzymatic resistant starch and other non-digestible material. This may more closely covered by AOAC method which used worldwide to provide fibre content for nutrient databases and food labelling purposes (DeVries, 2004). In the UK, the Englyst method was used to determine NSP in food composition tables and for labelling purposes and remained the recommended method for nutrition and food labelling until 1999. After that, the Food Standards Agency accepted the role of resistant starch and lignin as being part of dietary fibre and adapted the AOAC

method to measure dietary fibre in food for labelling purposes (Food Standards Agency, 2002)

The AOAC (991.43) method (Association of Official Analytical Chemist, 1995) was adopted for the current study because this method had been recommended by the WHO and the FAO as a method to determine total soluble and insoluble fibre in foods and food products and they advised that it be used for nutritional labeling in the UK (Food Standards Agency, 2002). The sixth edition of the UK food composition tables provide a list of AOAC measured AOAC-fibre[12] values for 47 food items however no values for legumes were listed, despite being one of the best sources of dietary fibre. Therefore the aim of this study was to determine the AOAC-fibre and insoluble dietary fibre (IDF) in selected legumes that are commonly consumed in the UK using the AOAC method with simple modifications. This study also set out with the aim of assessing the importance of effect of common cooking methods (boiling and canning) on the AOAC-fibre content of legumes.

Most UK epidemiological studies use NSP values, therefore a potential bias should be considered when comparing with other studies which estimate dietary intake from AOAC-fibre values. One way to tackle this issue is to use AOAC-fibre: NSP ratio. A mean ratio of NSP: AOAC-fibre of 1: 1.33 was generated for different types of food groups (Lunn and Buttriss, 2007). However, the legumes group was not well represented in this ratio (Peattie *et al.*, 1983). An early experiment by Reistad and Frolich (1984) reported a ratio of 1: 1.1 to 1: 1.4 for vegetables not including legumes. Therefore the next aim was to generate a ratio for legumes which may be useful to convert NSP values to AOAC-fibre values for some of the nutritional epidemiologists who investigate the dietary fibre intake in population characterized by high consumption of legumes such as ethnic minorities and vegetarians.

5.2 Materials and methods

5.2.1Materials

The samples for testing were selected based on commonly used legume products in the NDNS (Henderson *et al.*, 2002) and the UK Women's Cohort Study (UKWCS) (Cade *et al.*, 2004a) more detail on the UKWCS is in chapter 6. Briefly, the UKWCS is one of the largest population based prospective study in the UK

[12] AOAC-fibre referred to total dietary fibre obtained from AOAC analytical method only.

which has the advantage of including women with a wide range of dietary exposures including a high proportion of vegetarians or vegan.

A descriptive analysis was carried out to determine types of legumes consuming and usual cooking method used among the cohort. Data were obtained from the baseline FFQ for 35,372 participants showed that 88% of women reported consumed legumes. The most consumed legumes by mean (SD) among cohort women are baked beans, green beans, peas, lentils, mung and red kidney beans, butter beans and chickpeas (Table 5.1). In terms of cooking methods used for legumes; 64% of women used boiling and pressure cooking as their usual method for cooking. Half of legumes were eaten from cans compared to other forms of legumes (dried, frozen, and fresh). The National Diet and Nutrition Survey reported that green beans were eaten by 22% of all women while a higher percentage of all women consumed baked beans (41%) (Henderson *et al.*, 2002).

Table 5.1 Mean (SD) of legume intake (g/day) by cohort women

Food item	Dietary intake (g/d) Mean (SD)
Baked beans	23.8 (28.3)
Green beans	18.6 (20.2)
Green peas	11.8 (12.1)
Lentils	4.5 (8.3)
Mung beans and red kidney beans	3.3 (5.4)
Butter beans	3.1 (5.2)
Chickpeas	2.8 (6.2)

5.2.2Sample preparation

From the above findings, eight types of legumes were included in the dietary fibre analysis, namely yellow chickpeas (*Cicer arietinum L*), red kidney beans (*Phaseolus vulgaris*), red lentils and green and brown lentil (*Lens culinaris*), butter beans (*Phaseolus lunatus L*), green peas (*Pisum sativum*), and green beans (*Phaseolus vulgaris*), baked bean in tomato sauce (haricot or navy beans; *Phaseolus vulgaris*) and mung beans (*Vigna mungo*). For lentils, three subtypes were included in the analysis.

The sampling technique for the selected types of legumes was adopted from *McCance and Widdowson's The Composition of Foods* (2002) as in Figure 5.1 to allow future comparability with NSP values published in the UK composition foods tables (Food Standards Agency, 2002). The selected types of legumes were purchased in form of dried legumes (8 types) and canned legumes (6 types). A number of brands were purchased from various UK supermarkets for each type of

legumes. Table 5.2 shows all brands for each type of canned legumes (baked beans, chickpeas, red kidney beans, butter beans, green peas and green beans). Table 5.3 shows all brands for each type of dried legume (yellow chickpeas, red kidney beans, mung beans, red lentil, butter beans, green beans and peas) were pooled before analysis. Eight pooled samples (composite sample) of dried legumes were processed prior to analysis. Processing included soaking overnight in tap water (1:5 w/v) at room temperature followed by draining and then cooking in tap water at boiling temperature described in the UK food composition tables (description section as in Figure 5.1) in *McCance and Widdowson's The Composition of Foods* (2002). When cooking instructions were not available in the aforementioned book, packet instructions were followed as per normal domestic practice. Canned legumes were drained. Then, all composite samples were drained and homogenised prior to analysis. In total, fourteen pooled samples of legumes were for dried and frozen legumes, and canned legumes.

Figure 5.1 Example of some legumes description from vegetable section in *McCance and Widdowson's The Composition of Foods* **(2002)**

No.	Food	Description and main data sources
Beans and lentils *continued*		
755	**Mung beans**, whole, *dried, raw*	Literature sources
756	*dried, boiled in unsalted water*	As raw, soaked and boiled
757	**Red kidney beans**, *dried, raw*	Whole beans. Analytical and literature sources
758	*dried, boiled in unsalted water*	As raw, soaked and boiled
759	canned, *re-heated, drained*	LGC; 10 cans, 6 brands

.

Table 5.2 List of canned legumes purchased from local supermarkets

Canned legumes	N°	Brands	Mean NSP*g/100g	Code*
Baked beans in tomato sauce	1	Sainsbury's baked beans	3.7	13-044
	2	Heinz baked beans		
	3	Tesco light baked beans		
	4	ASDA Baked Beans in tomato sauce		
	5	Organic baked beans		
Yellow chickpeas	1	Sainsbury's chickpeas	4.1	13-078
	2	Tesco chickpeas		
	3	Waitrose chickpeas		
	4	Morrison chickpeas		
	5	Morrison organic chickpeas		
Red kidney beans	1	Tesco red kidney beans	6.2	13-111
	2	Waitrose red kidney beans		
	3	Tesco whole food red kidney beans		
	4	Morrison red kidney beans		
	5	Sainsbury's red kidney beans		
	6	Organic Tesco red kidney beans		
Butter beans	1	Morrison butter beans	4.6	13-72
	2	Essential Waitrose butter beans		
	3	Sainsbury's butter beans		
Green peas	1	Sainsbury's green peas in water	5.1	13-135
	2	Co-operative green peas		
	3	ASDA green peas		
	4	Daucy garden peas		
	5	Morrison green peas		
	6	Tesco garden peas		
Green beans	1	Bandwelle green beans in water	2.6	13-85
	2	Sainsbury's whole French green beans		
	3	Morrison cut green beans		
	4	Morrison whole green beans		
	5	Tesco whole green beans		
	6	Batchelor's cut green beans		

* (Food Standards Agency, 2002)

Table 5.3 List of dried legumes purchased from local supermarkets

Dried legumes	N°	Brands	Mean NSP* g/100g	Code*
Yellow Chickpeas	1	Sainsbury chickpeas dried		
	2	Tesco chickpeas dried		
	3	Waitrose chickpeas dried	4.3	13-077
	4	Chanadal chickpeas dried		
Red kidney beans	1	Morrison's whole food red kidney beans		
	2	Great scot red kidney beans	6.7	13-110
	3	Natco red kidney beans		
Mung beans	1	Moong whole heeva		
	2	Natco mung beans	3.0	13-097
	3	Tesco mung beans		
Red lentil	1	East End red lentil		
	2	Indus red lentil		
	3	Tesco red lentil	1.9	13-092
	4	Great Scot red lentil		
Butter beans	1	Whitworths butter beans		
	2	Whole food butter beans	5.2	13-071
	3	Great Scot butter beans		
Green brown lentil	1	East End Green lentil		
	2	Brown lentil Heera	3.8	13-090
	3	Waitrose green lentil		
Green peas frozen	1	Morrison green peas		
	2	Sainsbury's basic British garden peas		
	3	Bird's Eye field fresh garden peas		
	4	British garden peas by Sainsbury's	5.1	13-134
	5	Cooperative farm British garden peas		
Green beans Frozen	1	Tesco sliced green beans		
	2	Sainsbury's very fine whole green beans	4.1	13-084
	3	ASDA sliced green beans		

* (Food Standards Agency, 2002)

5.2.3Dietary fibre analysis: modified AOAC method

All necessary enzymes and chemicals were purchased from Sigma-Aldrich (Dorset, UK) unless otherwise stated. A fibre assay kit (K-TDFR 03/2009) obtained from Megazyme International Ireland Ltd was used. AOAC-fibre was determined in triplicate number of sample of 1.000g.

5.2.3.1 Total dietary fibre analysis

Samples were analysed for AOAC-fibre and IDF following a modified AOAC official method (991.43) (Association of Official Analytical Chemist, 1995). Figure 5.2 demonstrates all the steps of the AOAC method. The sample was suspended in MES/TRIS buffer. The starch gelatinization step[13] was performed by placing the sample in a boiling bath for 60 minutes as a first modification step. Evidence shows that the gelatinization temperature of legume starch ranges between 70ºC – 95ºC and

[13] Gelatinization is an irreversible process which occurs when starch placed in large amount of water and heated, the starch granules gradually swells and the crystalline structure of the starch will be disrupted and starch will be dissolve. (Bogracheva *et al.*, 1998)

in most legume starches, no measurable granule swelling occurs at temperatures below 60ºC (Hoover *et al.*, 2010). Therefore, a high temperature is suggested to break down the strong interactions between amylose chains within the granules. A review by Guillon and Champ (2002) stated that the temperature for gelatinization of a range of legume starch can reach up to 125ºC (temperature of gelatinization range 55 – 125ºC) compared with cereal starch (55 – 90 ºC).

From Figure 5.2, enzyme hydrolysis was performed by incubating the sample with heat stable α-amylase (for starch digestion) with manual shaking for 35 minutes, followed by incubation with protease (for protein digestion) for 30 minutes with shaking, followed by pH adjustment to 4.5 and incubation with amyloglucosidase for 30 minutes in a shaking water bath for further starch and maltodextrin hydrolysis. Table 5.4 illustrates the characteristics of the digestive enzymes used for AOAC-fibre analysis.

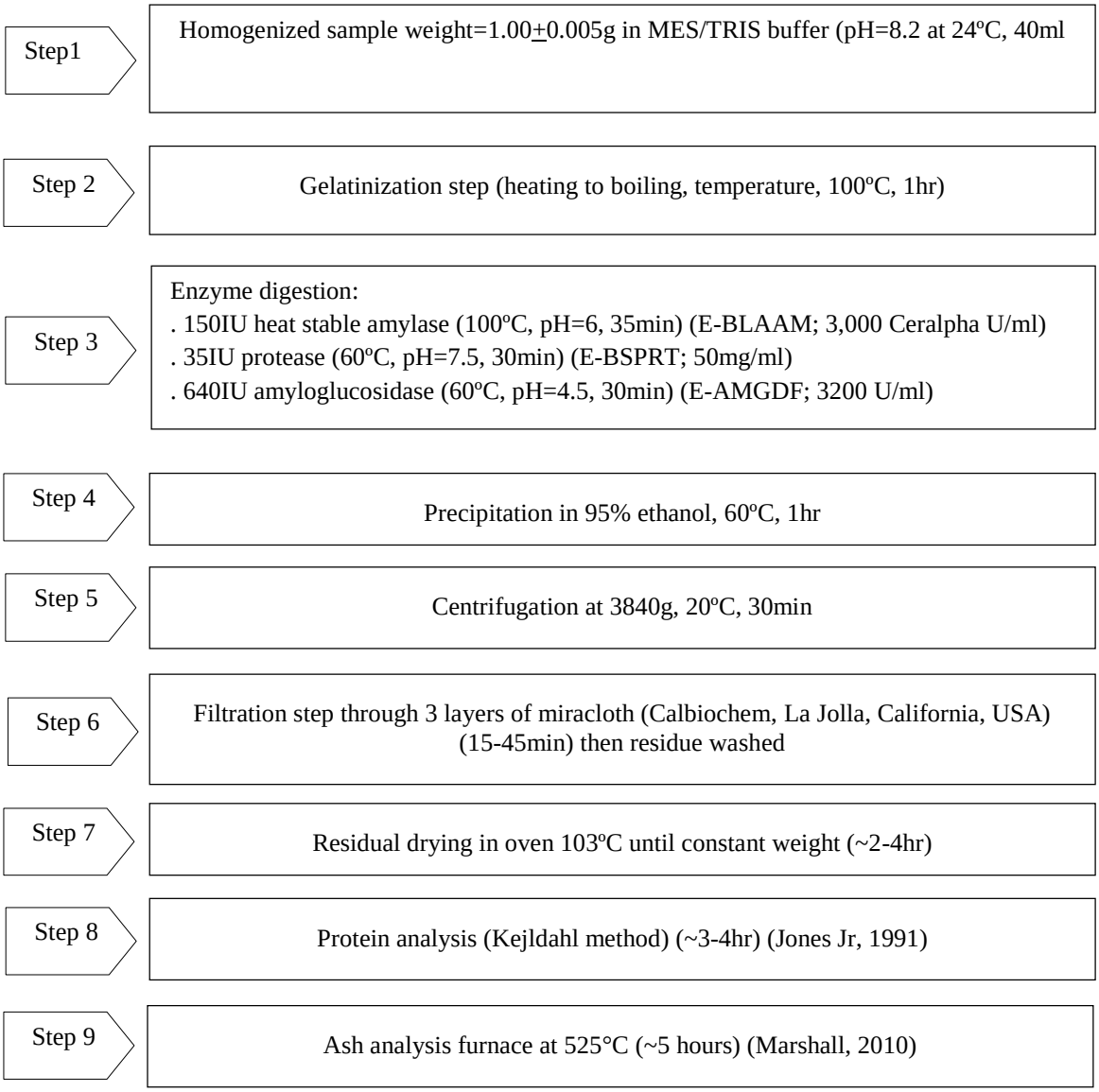
After step 3, the digested mixture was precipitated with four volumes of 95% ethanol that had been preheated to 60°C. The precipitated sample was centrifuged using a Beckman Coulter J2 Centrifuge using 250ml Beckman tubes at 3840 g for 30 minutes at 20°C. Centrifugation aimed to facilitate separation and reduce the filtration time (Oh and Grundleger, 1990). In step 6, the supernatant was removed from the centrifuged sample, and the residue filtered through 3 layers of Miracloth. Filtration with Miracloth was the second modification step that was carried out in the AOAC method. The residue was washed with ethanol, then acetone and dried in an oven at 103°C until constant weight was achieved in step 7. Step 8, residue was analysed for nitrogen content by the Kjeldahl method (Jones Jr, 1991). Appendix D illustrates the steps for Kjeldahl procedure. Nitrogen content was multiplied by a conversion factor of 6.25 to calculate protein content. Finally step 9, another residue was used for ash analysis by combustion in a furnace at 550ºC until a constant weight was achieved. AOAC-fibre values were recorded after subtracting protein and ash. Figure 5.2 presents the AOAC procedure with simple modification steps.

Table 5.4 Characteristics of enzymes used in AOAC method for dietary fibre analysis

Enzyme	Heat stable α-amylase E-BLAAM (3,000 Ceralpha U/ml)	Protease E-BSPRT (50mg/ml)	Amyloglucosidase E-AMGDF (3200 U/ml)
Source	Bacillus licheniformis	Bacillus licheniformis	Aspergillus niger
Substrate	*p*-nitrophenyle maltoheptaoside	Tyrosine	Soluble starch
Specific activity IU/volume	150 IU/50µl[1]	35 IU/100µl	640 IU/200µl[1]
Optimum pH	6 -6.5	7.0-7.5	4.0
PH stability	4.5–8.0	5.5-10	4.0 – 5.5
Optimum stability	75°C	60°C	70°C
Temperature stability	<80°C	≤60°C	<60°C

[1]One international unit (IU) of activity is defined as the amount of enzyme required to release one micro mol of glucose reducing sugar equivalent per minute under defined conditions of temperature and pH.

Figure 5.2 Modified AOAC method for AOAC-fibre measurement in selected legumes

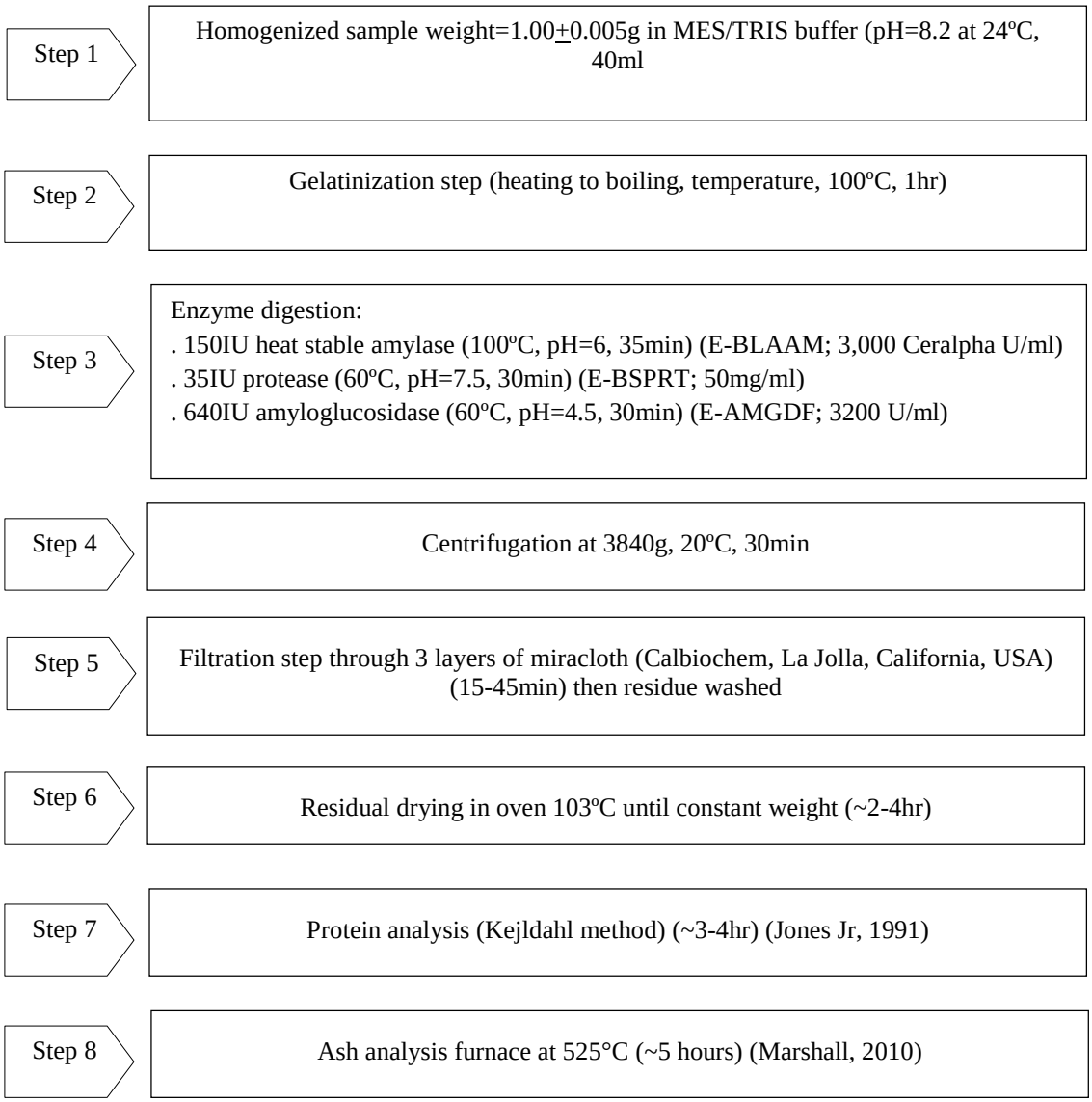


Modified from (Association of Official Analytical Chemist, 1995)

5.2.3.2 Insoluble dietary fibre analysis

The IDF from the same legumes samples was determined using the same kit. The procedure is similar apart from precipitation in ethanol. Triplicate samples of boiled and canned legumes were gelatinized and treated with the same digestive enzymes used to determine AOAC-fibre values. Figure 5.3 illustrates the steps for the modified AOAC method (991.43) (Association of Official Analytical Chemist, 1995). The SDF content was determined by the difference between AOAC-fibre and IDF values.

Figure 5.3 Modified AOAC method for IDF measurement in selected legumes.



Modified from (Association of Official Analytical Chemist, 1995)

5.2.3.3 Kjeldahl method and Ash analysis

Kjeldahl and ash methods were both part of the AOAC method used in this chapter. Values were used to correct the measured AOAC-fibre residues. Kejldahl method (appendix D) determines the total nitrogen in food. The fibre residue sample was hydrolysed by heating with concentrated acid (sulphuric acid). This reaction produces carbon dioxide, water and ammonium sulphate. Sodium hydroxide is then added to react with ammonium sulphate producing ammonia (gas). Distillation step carried out using boric acid to produce ammonium borate. This is then titrated against hydrochloric acid to determine the level of nitrogen present. Obtained nitrogen content was then converted into protein content by multiplying by 6.25 (conversion factor for food). Protein content was expressed as grams protein/100gram of food (%). The method of Kjeldahl was adopted from (Jones Jr, 1991). Ashing is a procedure to measure total mineral content in food. Dry fibre residues were analyzed to obtain ash content in the AOAC-fibre value. The sample was placed in silica crucible and in oven at temperature of 550°C (on average 4-5hours). Water is vaporized and organic compounds are burn in presence of oxygen to carbon dioxide of N_2. Minerals remained on the crucible are transferred into dessictor to cool then weigh the sample and crucible to obtain the weight of the ash. Finally ash was expressed in g/100g of food (Marshall, 2010).

5.3 Statistical analysis

Coefficient of variation (CV%) was calculated to compare the degree of analysis variation from one batch to another for each legume type. The differences between AOAC-fibre and IDF in canned and dried legumes were examined using Student's t-test and analysis of variance (ANOVA) as appropriate to analyse the effect of cooking method on total fibre values.

5.4 Results

5.4.1AOAC-fibre contents in boiled and canned legumes

Means and standard deviations of AOAC-fibre for canned and boiled legumes that are most commonly consumed in the UK were presented in Table 5.5. AOAC-fibre values were expressed as g per 100 g of legumes (wet weight as eaten). In total, 14 legume samples (8 boiled and 6 canned) were studied. The AOAC-fibre content in 8 boiled legumes ranged between 3.6 (0.1) g/100g in green beans to 11.2 (0.14) g/100g in red kidney bean, with an overall mean AOAC-fibre of 7.3 (0.34) g/100g.

The coefficient of variation for the boiled legumes ranged from 2.1% to 6.4%. The canned legumes showed a range of AOAC-fibre values from 2.7g/100g in canned green beans to 7.4g/100g for canned chickpeas, with a mean AOAC-fibre of 5.2g/100g. The CV% for canned legumes ranged between 1.4% to 5.7%. A report by AOAC Official Methods of Analytical Chemists (1995) showed a CV% of method 991.43 range between 0.89 – 6.26% for fibre rich food from different food groups. This reflects that measured total dietary fibre values were within an acceptable range. The AOAC-fibre values for boiled legumes were on average 47% higher than for the equivalent canned legume, and statistical analysis showed that boiled legume values were significantly higher than canned legumes by 1.95 g/100g ($p<0.01$). The greatest difference was found in red kidney beans, with AOAC-fibre values in canned samples (5.5 g/100 g) being half of the boiled equivalent (11.2 g/100 g). It seems that dietary fibre in legumes measured by AOAC method varied with different cooking method.

Table 5.5 Means and standard deviations of AOAC-fibre for cooked legumes (g/100g). Values are triplicate analyses from pooled samples.

Cooking method	Legume	AOAC-fibre g/100g Mean (SD)
Boiled legumes	Red kidney beans	11.2(0.14)
	Butter beans	8.4(0.35)
	Yellow chickpeas	9.2(0.46)
	Green beans	3.7(0.10)
	Green peas	5.9(0.16)
	Red lentil	9.2(0.21)
	Green brown lentil	5.2(0.11)
	Mung beans	4.4(0.07)
Canned legumes	Red kidney beans	5.5(0.44)
	Butter beans	4.5(0.14)
	Yellow chickpeas	7.4(0.34)
	Green beans	2.7(0.10)
	Green peas	5.2(0.13)
	Baked beans in tomato sauce	5.9(0.17)
Total legumes		6.3(2.43)

The insoluble protein and insoluble ash were measured in the AOAC-fibre residue and the mean insoluble protein was 0.19 (0.07) g/100g in the analysed legumes. Ash was found to be very low in amount in the dietary fibre residue with mean of 0.13 (0.1) g/100g. Indigestible protein and ash contributed slightly to the total dietary fibre residue (2% and 2.3% receptively).

5.4.2Insoluble and soluble dietary fibre contents in boiled and canned legumes

The means and standard deviations of measured insoluble and calculated soluble dietary fibre (SDF) for boiled and canned legumes are presented in Table 5.6

and 5.7. The results of this investigation show that IDF values in boiled legumes ranged from 2.6 (0.3) g/100g for green beans to 8.9 (0.67) g/100g for red kidney beans. Insoluble fibre for canned legumes ranged from 1.9 (0.36) g/100g green beans to 6.4(0.15) g/100g for yellow chickpeas. It can be noticed that insoluble fibre represents high proportion of AOAC-fibre (59% – 83% for boiled legumes and 70% – 87% for canned legumes) from all analyzed legumes.

One way analysis of variance showed that IDF values were significantly higher in boiled legumes by 1.7g/100g compared to their canned equivalents (F (1, 28) = 5.97, p= 0.02). The results of this research support previous study which reported that IDF in boiled soaked beans was higher than in canned beans with a difference of 1g/100g (Kutos *et al.*, 2003).

Table 5.6 Means and standard deviations of IDF and SDF of boiled legumes (g/100g).Values are triplicate analyses from pooled samples.

Boiled legumes	IDF g/100g Mean (SD)	SDF g/100g Mean (SD)	IDF%: SDF%
Red kidney beans	8.9(0.67)	2.3(0.70)	79:21
Butter beans	6.9(0.48)	1.5(0.68)	83:17
Yellow chickpeas	5.5(0.55)	3.7(0.67)	59:41
Green beans	2.6(0.30)	1.0(0.31)	73:27
Green peas	4.5(0.51)	1.4(0.61)	77:23
Red lentil	8.2(0.03)	1.1(0.23)	89:11
Green brown lentil	4.9(0.26)	0.4(0.14)	93:7
Mung beans	3.6(0.57)	0.8(0.56)	82:18
Total boiled legumes	5.6(1.76)	1.5(0.76)	79:21

Table 5.7 Means and standard deviations of IDF and SDF of canned legumes (g/100g). Values are triplicate analyses from pooled samples.

Canned legumes	IDF g/100g Mean (SD)	SDF g/100g Mean (SD)	IDF%: SDF%
Red kidney beans	3.8(0.73)	1.7(0.36)	70:30
Butter beans	3.5(0.28)	0.9(0.14)	78:22
Yellow chickpeas	6.4(0.15)	1.0(0.23)	87:13
Green beans	1.9(0.36)	0.8(0.30)	72:28
Green peas	4.3(0.22)	0.9(0.27)	82:18
Baked beans in tomato sauce	3.3(0.60)	2.6(0.43)	56:44
Total canned legumes	3.8(0.970	1.3(0.54)	74:26

In terms of proportion of dietary fractions, Figure 5.4 shows that the ratio between IDF and SDF did not differ with different cooking methods used (canned and boiled legumes) which reflect reduction in both fractions in canned and boiled legumes that did not greatly affect the proportions of fibre fraction in the measured total dietary fibre.

Figure 5.4 Percentage of IDF and SDF in boiled and canned legumes

5.4.3Comparison between measured AOAC-fibre, IDF values and available NSP values from UK food tables for selected cooked legumes

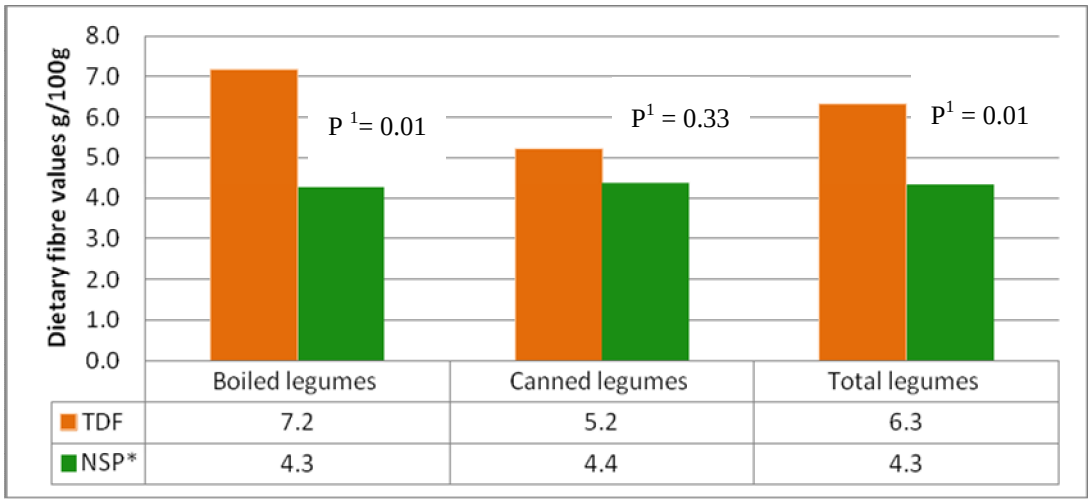
Table 5.8 presents the mean of total dietary fibre measured by modified AOAC method which was 6.3g/100g for all legumes, 5.2g/100g for canned legumes and 7.2g/100g for boiled legumes. Total legumes and boiled legumes values were significantly higher than mean NSP values (4.31g/100g and 4.23; respectively) from the UK food composition tables. However, there was no significant mean difference between AOAC-fibre and NSP content of canned legumes as shown in Figure 5.5.

Table 5.8 AOAC-fibre and NSP values for boiled and canned legumes. Values are triplicate analyses from pooled samples.

Cooking method	Legume	AOAC-fibre g/100g	NSP* g/100g	Mean difference (g/100g)
Boiled legumes	Red kidney beans	11.2(0.14)	6.7	4.5
	Butter beans	8.4(0.35)	5.2	3.2
	Yellow chickpeas	9.2(0.46)	4.3	4.9
	Green beans	3.6(0.05)	4.1	-0.5
	Green peas	5.9(0.16)	5.1	0.8
	Red lentil	9.2(0.21)	1.9	7.3
	Green brown lentil	5.2(0.11)	3.8	1.4
	Mung beans	4.4(0.07)	3.0	1.4
Boiled legumes mean ratio		1.68 (95%CI: 1.10, 2.25)		
Canned legumes	Red kidney beans	5.5(0.44)	6.2	-0.7
	Butter beans	4.5(0.14)	4.6	-0.1
	Yellow chickpeas	7.4(0.34)	4.1	3.3
	Green beans	2.7(0.07)	2.6	0.1
	Green peas	5.3(0.13)	5.1	0.2
	Baked beans in tomato sauce	5.9(0.17)	3.7	2.2
Canned legumes mean ratio		1.19 (95%CI: 0.78, 1.59)		
Total legumes		6.3	4.3	2.0
Total legumes mean ratio		1.43 (95%CI: 1.11, 1.81)		

*****NSP values are from McCance and Widdowson's The Composition of Foods (2002)

Figure 5.5 Mean AOAC-fibre values of the analysed cooked legumes compared to NSP values of equivalent legumes from UK food tables. Values are triplicate analyses from pooled samples.



[1]P value of <0.05 is significant.

AOAC-fibre values for all legumes were found to be on average 43% higher than published NSP values (Food Standards Agency, 2002). A previous analysis of ten food groups showed that AOAC-fibre fibre was higher than NSP by 20% for green vegetables to 77% for other vegetables (Englyst *et al.*, 1996). The results presented in Figure 5.5 also show that mean AOAC-fibre values of boiled legumes were higher than NSP values by 68% and the total dietary fibre value for canned legumes was also higher than NSP values by 19%. The difference may be partly explained by the presence of indigestible materials other than NSP which are mainly resistant starch and lignin (Wolters *et al.*, 1992). Among canned legumes, yellow chickpeas were found to be the richest source of fibre, and red kidney beans were the richest source of fibre among boiled legumes. Only three legume samples, boiled green beans, canned butter beans and canned kidney beans, showed slightly lower AOAC-fibre values compared to NSP. These observations were also found in some food items listed in the UK food composition table (Food Standards Agency, 2002) where unexpectedly, 5 of out of 47 food items had slightly lower AOAC-fibre values compared with NSP values within the range determined. This observation remained unexplained.

With regard to IDF values in legumes, a comparison between AOAC and Englyst method was carried out and showed that insoluble fibre measured by AOAC method was significantly higher than calculated IDF of NSP ($p<0.01$) for total, canned and boiled legumes (Table 5.9).

Table 5.9 Comparison between means of IDF for cooked legumes. Values of the IDF measured by AOAC method are triplicate analyses from pooled samples.

Legumes	IDF values		
	AOAC[1]	Englyst[2]	p values
Boiled legumes	g/100g	g/100g	
Red kidney beans	8.8	3.5	
Butter beans	6.9	3.0	
Yellow chickpeas	5.5	2.9	
Green beans	2.6	2.3	
Green peas	4.5	3.7	
Red lentil	8.2	1.3	
Green brown lentil	4.9	2.9	
Mung beans	3.6	2.3	
Mean of boiled legumes	5.6	2.7	<0.01
Canned legumes			
Red kidney beans	3.8	3.3	
Butter beans	3.5	2.6	
Yellow chickpeas	6.4	2.8	
Green beans	1.9	1.2	
Green peas	4.3	3.7	
Baked beans in tomato sauce	3.3	1.4	
Mean of canned legumes	3.8	2.5	0.04
Mean of total legumes	4.9	2.6	<0.01

[1] IDF measured by AOAC method [2] IDF of NSP value from (Holland *et al.*, 1991a)

On average an AOAC-fibre: NSP ratio of 1:1.43 was calculated for the whole legume group (n=14). For the whole group, the AOAC-fibre content of legumes was significantly correlated with NSP (r= 0.6, 95% CI: 0.10 to 0.87; p= 0.02). The ratio for the boiled legumes (1:1.68) was higher than for the canned legumes (1:1.19).

5.5 Discussion

Legumes are a rich source of dietary fibre and are considered an important part of a healthy diet (Leterme, 2002). The physiological and health benefits of dietary fibre have been reported consistently in many studies (Dilis and Trichopoulou, 2009). Some experimental studies have previously compared AOAC-fibre values and NSP values in few types of legumes (Englyst *et al.*, 1996, Mongeau and Brassard, 1989). In the sixth edition of the UK food composition tables (Food Standards Agency, 2002), there is lack of AOAC-fibre data for legumes. Therefore, this study was designed to measure total dietary fibre and IDF using the AOAC method with simple modifications in legumes consumed frequently in the UK. The current study explored the effect of usual cooking methods on dietary fibre content measured by AOAC method in legumes. The AOAC-fibre content in legumes ranged between 2.7g/100g and 11.2g/100g (w/w). Insoluble dietary fibre was found to be the major component in all legumes which represents at least 60% of total

dietary fibre. That is supported by many previous studies (Veena *et al.*, 1995, Kutos *et al.*, 2003, Li *et al.*, 2002).

It was observed that AOAC-fibre values for canned legumes were lower than for equivalent boiled legumes, suggesting an effect from food processing on AOAC-fibre. It was reported in previous studies that usual cooking method and autoclaving did not generate appreciable changes in the composition of non-cellulosic polysaccharides, cellulose and lignin (Apata, 2008). Another study reported little effect of cooking on the dietary fibre content of cooked vegetables when compared to raw vegetables without including legumes (Reistad and Frolich, 1984). The effect of cooking methods on dietary fibre content might be partly explained by the variation in stability of cell walls under different heating conditions (Chang and Morris, 1990).

The findings suggest that the fibre content of legumes is affected by processing methods. In particular, canned legumes were found to be significantly lower in fibre than boiled legumes. Previous study reported a higher total dietary fibre content measured by the AOAC method in boiled legumes (13.1 – 17.7g/100g) in comparison to the autoclaved legumes (7.1 – 9.29g/100g w/w) (Li and Cardozo, 1993). While, other studies showed that boiling and microwaving did not affect the NSP content of legumes (Marconi *et al.*, 2000). This suggests that canning affects compounds other than non-starch polysaccharides, most likely resistant starch. Enzyme-resistant starch is one of the components of dietary fibre that is included in the AOAC-fibre gravimetric measurement.

The unique structure of legume starch in comparison with cereal starch (Sandhu and Lim, 2008) as well as the fact that there is a high amount of resistant starch in cooked legumes are thought to be explanatory factors impeding its digestibility (Perera *et al.*, 2010). An in vivo study using rats showed that up to 35% of the starch from canned peas was undigested, a much higher percentage when compared to cereal starch (0.3% - 2.2%) (Marlett and Longacre, 1996). In addition, an in vitro study found that rapidly digestible starch accounts for 4.2 – 10.9% of total starch content, while resistant starch account for at least 50% of the total legume starch (Sandhu and Lim, 2008). Legume starch is relatively high in amylose (28-33%) which requires higher temperatures and longer heating times to gelatinise and shows higher propensity to retrogradation (Sandhu and Lim, 2008). That indicates that there is a positive correlation between amylose and resistant starch

content as reported in earlier review (Eerlingen and Delcour, 1995). An in vitro study showed that legume starch digestibility increased to 91% by heating at 121ºC (Rehman and Shah, 2005), suggesting that heating to high temperatures (e.g. canning) increases the availability of legume starch to amylase degradation, and therefore will reduce the amount of resistant starch residual in the fibre fractions. It was demonstrated in a previous study that exposure to high temperatures led to a breakdown of pectic substances (Anderson and Clydesdale, 1980). Preservation methods as in canning found to be attributed to increase softening and degradation of the cell wall in fruits/vegetables (McDougall *et al.*, 1996). This may partly explain the differences in NSP values between boiled and canned legumes.

The amount of IDF and SDF were significantly lower in canned legumes compared with boiled. Further research that focuses on determining the constituents of soluble and insoluble fibre in canned and boiled legumes may help in understanding the main components that contribute to the variability when a comparison between two cooking methods is needed. On the other hand, canning did not significantly change the proportion of IDF to SDF compared to boiled legumes.

The content of IDF was consistently around 60-80% of AOAC-fibre values, suggesting that canning affects both fibre subgroups. IDF is insoluble in buffer, and is thought to consist mainly of cellulosic and hemicellulosic cell wall polysaccharides, lignin, resistant starch (Saura-Calixto *et al.*, 2000) and polyphenolic compounds (Goñi *et al.*, 1996). It is likely that canning affects resistant starch, making it available for amylase digestion. Hemicellulosic polysaccharides may become soluble and recovered in the SDF fraction. Other components of IDF are likely to be unaffected. An in vivo study found that 98% of lignin and condensed tannins were recovered in faeces (Bravo *et al.*, 1994). It has been reported that insoluble polyphenols have a similar physiological effect to IDF in terms of low chemical degradation and high resistance to digestion (Bravo *et al.*, 1994). Canning may lead to the breakdown or solubilisation of pectic polysaccharides (Kutos *et al.*, 2003).

A ratio of 1:1.43 was obtained for the legume group, which is slightly higher than the published ratio of 1:1.33 for ten major food groups (Lunn & Buttriss, 2007). This ratio could be used to calculate AOAC-fibre values from NSP values, providing an opportunity to estimate AOAC-fibre intake and use the values to compare cohort studies in populations with high legume consumption. Moreover,

the ratio for boiled legumes (1:1.68) was dramatically higher than the ratio for canned legumes (1:1.19). Therefore, caution must be taken when applying the ratio without knowledge of the types of legume (boiled/canned) consumed. Characteristics of the studied population should be evaluated before considering the AOAC-fibre: NSP ratio. For example, boiled legume ratio may be more suitable for studies which focus on minority ethnic group in UK where boiled legumes are mostly consumed, compared to the rest of the UK general population which is more likely to consume canned legumes. More research on the AOAC-fibre: NSP ratio derived from a wide range of food items needs to be undertaken to understand the association between AOAC-fibre and NSP more clearly. Further research on generating ratio for a wider range of food items of legumes according to cooking method is also useful for epidemiological studies.

5.6 Conclusion

This is the first report of AOAC-fibre data for legumes commonly consumed in the UK. Insoluble dietary fibre is a major component in legumes. The result showed processing generated variation in AOAC measured fibre content in legumes. Cooking methods have a great effect on dietary fibre content in legumes and this result in variation in the generated ratio AOAC-fibre: NSP for canned and boiled legumes. Therefore, this need to be considered with the use of conversion factor in estimating dietary AOAC-fibre value from NSP value for future epidemiological studies. As legumes are not eaten raw, data on comparison between canned and boiled legumes are much more relevant than comparisons with raw food. Hence, such information is more relevant for nutritional research that investigates the impact of legume fibre intake on health outcomes. Dietary fibre analysis increases the researcher understanding of the complexity and heterogeneity of dietary fibre in legumes and possible error from analysis that may partly contribute to bias in estimated fibre intake in epidemiological research.

Chapter 6: UK Women's Cohort Study: Methodology

6.1 Introduction

The main objective of this chapter is to describe the UK Women's Cohort Study (UKWCS). A secondary objective is to describe how AOAC-fibre values were incorporated into the UKWCS dietary FFQ dataset.

6.2 The UK Women's Cohort Study

The UKWCS was funded by the World Cancer Research Fund (WCRF). Women in the UKWCS were identified from a half million responders in Wales, England and Scotland in the WCRF through a direct mail survey (Cade *et al.*, 2004a) . Recruitment was aimed to have a similar number of women with different diet where the selected vegetarian was matched for age to the next meat eater to ensure enough power for comparison between nutrients' intakes and diet with health outcomes (Greenwood *et al.*, 2000). Further recruitment was also achieved by asking eligible women to volunteer friends and relatives of a similar age group (Greenwood *et al.*, 2000). All women participated between 1995 and 1998 through a postal questionnaire.

From 61,000 women who were sent the baseline food frequency questionnaire (FFQ), 58% replied and were considered form the baseline of the UKWCS (n=35,372) (Cade *et al.*, 2007, Greenwood *et al.*, 2000). The women, who were aged between 35-69 years at the baseline, were asked about their diet and lifestyle. The original aim of the UKWCS was to investigate the relationship between diet and chronic disease (mainly cancer). Figure 6.1 demonstrates the flow of participants in the UKWCS. All original responders to the FFQ were sent a follow-up pack, 4.1 years, on average later (Cade *et al.*, 2004a). This was completed by 40% (14,244). Eighty seven percent of them (n=12,453) completed a 4 day semi-weighed food diary. At the time of this thesis, full dietary information was available electronically for only 1,883 women (15%) whose diaries were fully coded by the DANTE program and whose data were available for analysis. Figure 6.2 shows the time scale of women who completed different dietary assessment methods in the UKWCS. Subsample was carried out in another study by Greenwood *et al.* (2003) were FFQ repeated between 2000 to 2001 as shown in Figure 6.2 (Details discussed in section 6.6).

Figure 6.1Number of women who participated in the UKWCS at different levels

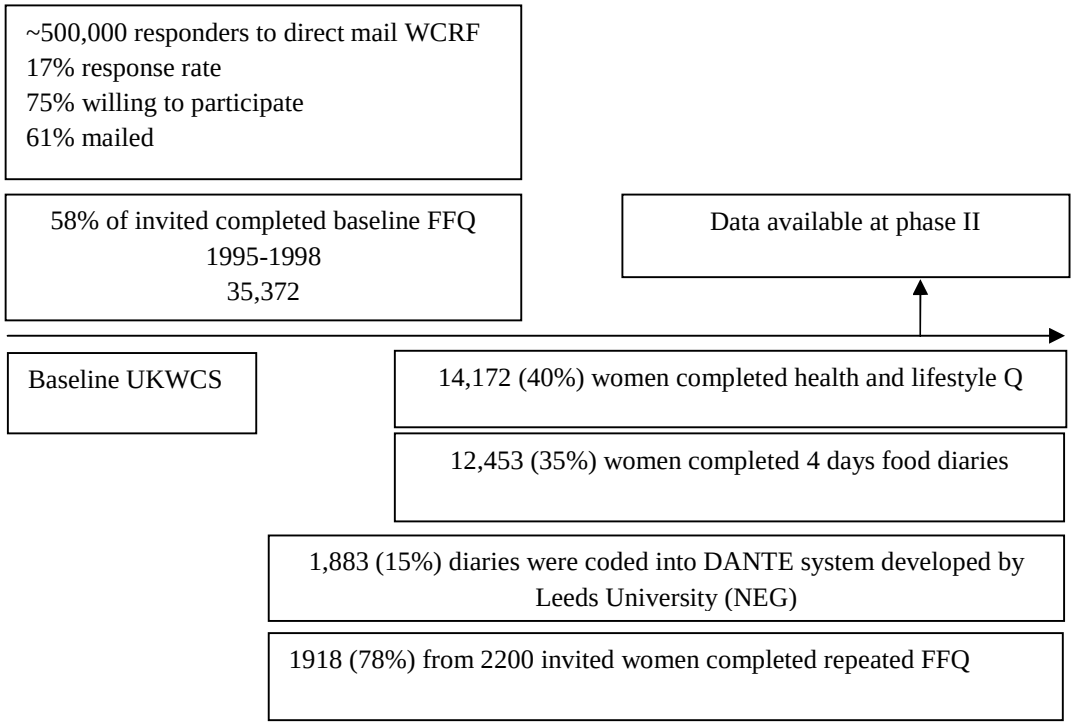
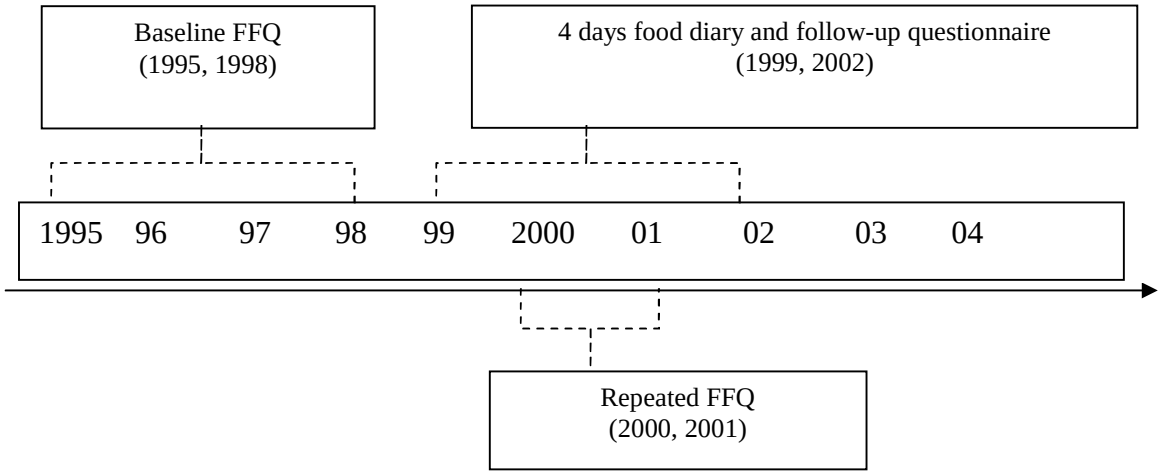


Figure 6.2 Time scale of the women who completed different dietary assessment methods in the UKWCS



6.2.1Outliers

Extreme dietary nutrient intake values in the dataset can greatly affect the findings especially when assessing continuous variables. These extreme intakes may not reflect the main body of data, and may indicate improper completion of the FFQ. This may be addressed with energy intake restriction (Willett, 2012). Outliers are unusual data that may be called error or noise in the dataset (Ben-Gal, 2005).

Outliers can occur due to measurement error or inaccurate data entry or can be due to natural variation within the studied population (Willett, 1998).

Statistical analyses were carried out after excluding women with unfeasible total energy intakes (more than 4,000 kcal/d) and less than 500 kcal/d) as suggested by Willett (1998) to address over and under-reporters. Adapting this approach, 918 women accounting for 2.6% of participants were excluded.

UKWCS data was obtained from women aged 35–69 years at baseline and were living in England, Wales and Scotland. In the mid 1990's, over 61,000 women were mailed and 58 % replied. At baseline, participants completed FFQ and also provided further dietary, lifestyle and health information after an average of 4 years.

6.3 Ethical consideration

Ethical approval was obtained from 174 local research ethics committees throughout the UK (Cade *et al.*, 2004a) (copies of two local ethical approval letters from two committees are in appendix M). All participants gave informed consent before they were included in the study. No new information was needed; therefore no additional ethical issues were related to this study. Participant identity was maintained confidentially and remained unidentified throughout the analyses.

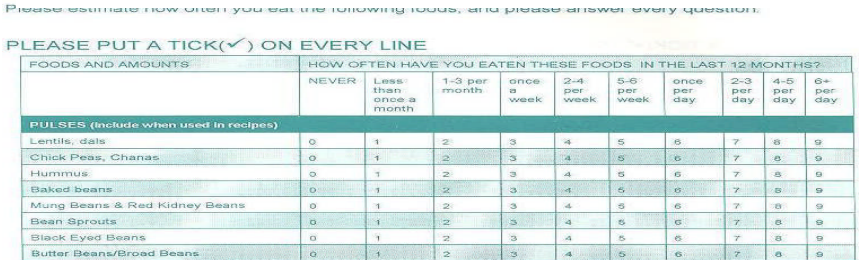
6.4 Food-frequency questionnaire

The self-administered Food Frequency Questionnaire (FFQ) (Appendix E) which had been adapted from the European Prospective Investigation into Cancer (EPIC) study (Cade *et al.*, 2007) was used at the baseline of the study to assess the participant's habitual diet over the past 12 months as well as health and lifestyle characteristics. Through inclusion of additional vegetable dishes, this FFQ had been modified to fit the studied population based on a pilot study using a subsample of vegetarian women (Cade *et al.*, 2004a). The pilot study also helped to generate portion sizes for use in nutrient intake calculation.

The self-administered food frequency questionnaire (FFQ) consists of frequency of consumption of 217 food items asked as "How often have you eaten those foods in the last 12 months?" and the answer is by ticking one box with the frequency ranging from "never" to "6 or more times per day". Figure 6.3 provides an example of the baseline question for some food items in the legumes section. A validation study was performed on a subgroup with 303 women and nutrient values from the FFQ were compared with nutrient values from 4 days' of food diary (Spence *et al.*, 2002). The results showed significant correlation between diary

intakes and baseline FFQ for selected micronutrients with a tendency to overestimate with FFQ compared to 4 days diaries. Details on food diary demonstrated in section 6.5.

Figure 6.3 A section of the baseline FFQ related to the intake of legumes



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
PULSES (include when used in recipes)										
Lentils, dals	0	1	2	3	4	5	6	7	8	9
Chick Peas, Chanas	0	1	2	3	4	5	6	7	8	9
Hummus	0	1	2	3	4	5	6	7	8	9
Baked beans	0	1	2	3	4	5	6	7	8	9
Mung Beans & Red Kidney Beans	0	1	2	3	4	5	6	7	8	9
Bean Sprouts	0	1	2	3	4	5	6	7	8	9
Black Eyed Beans	0	1	2	3	4	5	6	7	8	9
Butter Beans/Broad Beans	0	1	2	3	4	5	6	7	8	9

6.4.1 Generation of non-starch polysaccharides (NSP), and other dietary variables at the baseline

Non-starch polysaccharide intakes in the UKWCS dataset (Cade *et al.*, 2004a) were based on the UK Food Composition Tables (Holland *et al.*, 1992a, Holland *et al.*, 1988, Holland *et al.*, 1989, Holland *et al.*, 1991a, Holland *et al.*, 1992b, Holland *et al.*, 1993). Several types of possible food and cooking methods were used to represent each food item listed in the FFQ. The 217 food items listed in the baseline FFQ were derived from 545 individual food items. Number of items grouped together to represent one of the 217 FFQ items by calculating the estimated weighted proportion of food being eaten that contribute to the FFQ main items. For example, porridge was included in the breakfast cereal group. The NSP content of porridge (g/100g) was derived from proportions of NSP value (g/100g) of the four food items as shown in Table 6.1.

Table 6.1: Example of food item listed in the baseline FFQ with subtypes of food items obtained

	Food item	Code[1]	NSP value g/day	Proportion
10	Porridge		0.9	
10.1	Porridge with water	50-076	0.8	25
10.2	Porridge with whole milk	50-077	0.8	25
10.3	Porridge with water and milk	11-142	0.8	25
10.4	Readybrek	50-080	1.2	25

[1]Food code in the UK food composition tables

Values were assigned to each FFQ food item, and each participant's daily NSP intake from each food item was calculated by multiplying the frequency of food

consumption by an average portion size, and the product multiplied by grams of NSP fibre contained in 100g of the food item. The following equation in Figure 6.4 was used to calculate the NSP intake for the cohort participants. NSP with other nutrient intakes were created originally using SPSS, then computation from SPSS into a Microsoft Access program was developed by the University of Leeds - Nutrition Epidemiology Group to facilitate generation of dietary intakes from food groups. The estimated portions (Calvert *et al.*, 1997) were derived from an average of three sources: portion size derived from the pilot study of the weighed food diaries of vegetarians; from published 1993 and 1994 women's food portion sizes from national survey values (NDNS) and published values by Crawley (1993).

Figure 6.4 Equation used to calculate the NSP intake in the UKWCS database

Dietary NSP intake (g/day) per person	=	Estimated portion size X NSP content (g/100g) X Frequency of consumption per day

Other dietary variables examined were also from baseline FFQ dataset. For example, total energy intake which was expressed in kcal per day derived for all cohort women. Macronutrient intakes (carbohydrates, protein and total fat) were obtained from all food items listed in FFQ and expressed as percentage contribution to total energy intake. Vitamins such as vitamin C was assumed as a potential measures for overall diet quality (Drewnowski *et al.*, 1997) and was expressed in mg per day. Minerals such as calcium, iron and folate were also evaluated in relation to dietary fibre intake as evidence suggests these may also be a good reflection of diet quality (Aggarwal *et al.*, 2012). All macronutrient and micronutrients variables in the dataset were generated previously from baseline FFQ using nutrient values from McCance & Widdowson's The Composition of Foods (5^{th} edition) (Holland *et al.*, 1991b) for each food item listed in FFQ. Derived nutrient values were expressed in milligrams or micrograms per day and calculated similarly to equation in Figure 6.4. Total nutrient intake was derived from adding all food items consumed by the participant.

The use of dietary supplement was also reported in FFQ and women categorized into supplement users and non-users, through type of supplement taken was available only as a free text, so was not categorised for this study.

Percentages of self-reported vegetarians or vegans were calculated from the question “Would you describe yourself as a vegetarian? And “Would you describe yourself as vegan?” answered as “yes” or “no”. However, using food intakes data, Cade *et al.* (2004a) defined meat eaters as women who ate meat at least once a week, vegetarians were defined as women who ate meat or fish less than once a week and fish eaters were women who ate fish but not meat at least once a week, Finally, oily fish-eaters were women who ate meat less than once a week and oily fish 2–4 times per week. So vegetarians are described into two ways, by absence of meat in diet, and by their self-report status.

6.4.2Other measured variables

Information on socio-demographic, anthropometrics, lifestyle and dietary characteristics were also determined for the studied population to have a general overview on the characteristics of participants. Socioeconomic status was based on the National Statistics Socio-economic Classification (Rose and Pevalin, 2003) where women are categorized into professional and managerial occupations, intermediate occupations and routine and manual occupations. Education level comprised four categories, namely no educational records; O-level, A-level and degree achievement was also assessed. Marital status was categorized into married, divorced, widowed, single, and separated. The majority of the UKWCS participants (99%) are white, however 0.46% of women were from India and 0.59% from other ethnic group. Because of the very few numbers of other than white population, therefore ethnicity was not explored in subgroup analysis in chapter 8.

The physical activity variable was derived from a question in the baseline questionnaire: “In a typical week during the last 12 months, how many hours did you spend on the following activity? On housework, Do it yourself, gardening, walking, cycling, other physical exercise” which is expressed as metabolic equivalents tasks (METs*).* A MET is defined as *“the ratio of work metabolic rate to standard resting metabolic rate of 1.0 (4.184kJ) $kg^{-1}\cdot hr^{-1}$, 1 MET is considered a resting metabolic rate obtained during quiet sitting”*. Time that was spent in each activity mentioned in the questionnaire was multiplied by the MET intensity level listed in the published Compendium of Physical Activities (Ainsworth *et al.*, 1993). Physical activity was expressed as METs. This is a useful measure of physical activity to quantify the total amount of activity and to compare it across individuals. Women provided detailed information about smoking status, but for these analyses,

they were grouped into two categories as smoker and non-smoker. Body mass index (BMI) was calculated as weight (kg) divided by the square of height (m), and waist circumference (WC) were both self-reported in the baseline FFQ.

6.5 Four day food diary

Four day semi-weighed food diaries (n=12,453) were collected from cohort women between 1999 and 2002. The food diary booklet included instructions on how to complete the diary with one example. Women recorded food eaten in term of time of food or drink, description of food or drink consumed (including brand name where possible) and amount. In addition, recipes/description of food eaten away from home/any other comments and daily dose of any supplements used were all included in the diary. Some general questions about diet during the recorded days, for example "*Which type of bread did you eat most often in the four days that you filled in the diary?"* were also asked at the end of the booklet. A copy of the diary is provided in appendix K. All physical activities for one full day were recorded in the third day in the form of an activity diary. Diaries were coded mainly by nutrition students over the past 10 years. Nutrient intakes were obtained after completion of coding using a Microsoft Access-based food diary analysis program named the DANTE for 1,883 women (15%). This program was designed by the Nutrition Epidemiology Group at the University of Leeds. The program has the ability to manage the data and output (Cade *et al.*, 2006). The nutrient intakes were calculated as the average of over 3 or 4 days' recordings depending on completion.

6.6 Repeated Food frequency questionnaire

Second data collection was carried out to obtain dietary information by the same FFQ in a subsample from the baseline UKWCS as demonstrated in Figure 6.1and 6.2. After an average of 3.2 years from the baseline FFQ, 2,200 (6%) women were invited to complete a repeated FFQ, and 78% (n=1,918) did this (Greenwood *et al.*, 2003).

6.7 Baseline characteristics of the UKWCS

Table 6.1 provides the socio-demographic, lifestyle and dietary characteristics of 34,454 participants in the UKWCS expressed as mean (SD) for continuous variables and frequency (percentage) for categorical variables. The majority of participants are white (98.7%), the mean age was 52 years (9.3), two thirds of women were from the professional and managerial social classes, and most women were well educated, as 51% at least achieve the A-level, and 27% had a degree.

Three quarters of the women were married. Regarding the lifestyle characteristics, physical activity expressed as METs which reflect the energy cost of activities performed by the individual was 13.6 (12.9). Majority of women did not smoke at baseline or never smoked before. The mean (SD) body mass index (BMI) of the women was 24.4(4.3) kg/m^2, and mean waist circumference was 73cm (9.2). More than half (57.8%) of the women reported using dietary supplements. Data on medical histories are not shown in the Table 6.2 because of small number of women who reported the medical histories of chronic illnesses in the FFQ with 17.2% of participants complaining of high blood pressure, only 7.6% had a history of high cholesterol level, 7.5% reported a positive medical history of cancer and very few (2%) reported a history of diabetes mellitus.

In term of dietary habits, a relatively high carbohydrate, low fat diet was the average. Table 6.2, lifestyle and dietary characteristics are summarised, for example looking at the macronutrients, it clear that the women are following a healthy dietary pattern on average. However, there are wide variations in dietary patterns because of the high proportion of vegetarians in the UKWCS. On the other hand, there were women with unhealthy eating pattern in the UKWCS. However on average, the UKWCS participants were more health conscious in compared to women from the recent NDNS report (Department of Health, 2012). In total, 70% of women were meat eaters, while vegetarians accounted for 18%, 9% of women were in other fish eaters group and only 2.3% of the studied population were in the oily fish eaters. On the other hand, 28% of cohort participants were self-reported vegetarians or vegan.

Dietary fatty acids profile was generated from the FFQ as a percentage of energy. These are approximate because no information available on the fat used for cooking in addition to difficulties in reporting type of fat from ready food or meals. Therefore, these types of fat were not used further in analyses.

Table 6.2 Baseline characteristics for 34,454 women in the UKWCS

Baseline characteristics		Mean (SD) or %
No of participants		34,454
Socio-demographic factors		
Age (year)		52.3 (9.3)
White population		98.7%
Socio-economic Status (%)	Professional	63.3
	Intermediate	27.5
	Routine	9.2
Educational level (%)	No education	16.9
	O-level	31.1
	A-level	24.7
	Degree	27.3
Marital status (%)	Married	75.0
	Divorced	8.8
	Widowed	6.3
	Single	7.6
	Separated	2.2
Postmenopausal (%)		52.6
Lifestyle characteristics		
Physical activity (METs)		13.6(12.9)
Smoker at baseline (%)	No	89.3%
	Yes	10.7%
Anthropometric		
BMI (kg/m^2)		24.4(4.3)
WC (cm)		73.5(9.2)
Dietary characteristics		
Energy intake (kcal)		2202(607)
NSP intake (g/d)		25.6(9.8)
Carbohydrates (% E)		54.6(6.4)
Protein intake(%E)		15.9(5.9)
Fat intake (% E)		33.1(5.8)
Sat fat intake (% E)		11.5(3.2)
PUFA intake (%E)[1]		15.8(6.2)
MUFA intake (%E)[2]		26.9(9.9)
Alcohol (units/d)[3]		0.54(1.2)
Vitamin C (mg/d)		166.0(85.2)
Folate (µg/d)		388.6(115.8)
Iron (mg/d)		17.4(6.3)
Calcium (mg/d)		1117.8(354)
Vegetable portion (80g/d)		5.0(2.5)
Fruit portion (80g/ d)		4.9(3.4)
Common dietary patterns	Meat group	70.3
	Fish group	9.2
	Oily fish group	2.3
	Vegetarians	18.1
Self-reported vegetarians %		27.7
Dietary supplement users (%)		57.8

[1]PUFA is polyunsaturated fatty acids expressed as percentage contributed from total energy intake. [2]MUFA is monounsaturated fatty acids expressed as percentage contributed from total energy intake. [3]Alcohol intake express asmedian and IQR.

6.8 Adding AOAC-fibre values to the UKWCS dataset

This section provides the approach of adding AOAC-fibre values into the UKWCS as previously done for NSP values (g/day).

6.8.1 Protocol 1: Search strategy for AOAC-fibre values

Since there is no single source of information concerning the AOAC-fibre content of UK foods, AOAC-fibre values were identified based on a search of several reference sources in order of priority, such as the recent UK food composition tables (Food Standards Agency, 2002), British Nutrition Foundation (BNF) review (Lunn and Buttriss, 2007), EuroFIR databases (European Food Information Resource e-search, 2010), USDA database (U.S Department of Agriculture, 2010), and commercial food labels. Recipe calculations according to McCance and Widdowson (Food Standards Agency, 2002) were also undertaken for a small number of food items. Figure 6.5 provides a guide for steps considered when obtaining the AOAC-fibre values from different sources. This protocol aimed to link AOAC-fibre values for food items listed in the baseline FFQ from a number of equivalent foods available in the UK food composition tables as previously done for NSP values.

The first reference was the sixth edition of the McCance and Widdowson's, the composition of foods book (Food Standards Agency, 2002). AOAC-fibre values were found for a few foods items (n= 47) in this latest edition. If the AOAC-fibre value was not found for the wanted item then the search progressed to the next reference, a review by Lunn and Buttriss (2007). In the Lunn and Buttriss review, a table contains a list of foods (n=95) with dietary AOAC-fibre values. Third reference is the European Food Information Resource AISBL which is abbreviated as EuroFIR AISBL was used (European Food Information Resource e-search, 2010) to extract AOAC-fibre values. It is a European Network of Excellence on Food Composition Databank Systems (http://www.eurofir.net) which aims to provide nutrients data from different European countries (Church, 2006). The EuroFIR e-Search facility is a network that can be accessed online through a registration procedure. It allows its users to search more than 20 standardised and specialized food composition databases.

EuroFIR databases provide a wide range of European data linking foods and nutrients through standardized data descriptions and associated nutrient value information. Information on the type of analytical dietary fibre methods is available

for a large number of food items which facilitates the identification of the analytical method needed for each food item chosen. For the matched food item, exploration of the cooking methods, ingredients, and recipes in case of mixed dishes food item was carried out. In the database, the code number (MI307) in the EuroFIR e-search referred to the enzymatic gravimetric method (AOAC).

If the AOAC-fibre values were not found in the above reference, then the fourth reference, the National Nutrient Database for Standard Reference in the U.S.A (U.S Department of Agriculture, 2010) was used to obtain AOAC-fibre values. This nutrient databank has an online search function, which gives free access to information on the nutrient content of around 7,538 different foods. All food has AOAC-fibre values. Matched food items with similarity in the cooking method, ingredients and recipe were considered.

The AOAC method was adopted by the Food Standards Agency as an official UK method for dietary fibre analysis and to be used for food labelling purposes (The Institute of Food Science & Technology, 2007). Therefore, food labelling was the fifth reference chosen for the remaining unmatched food items listed in the baseline FFQ. AOAC-fibre values were derived from food labels on the food items that can be purchased in common supermarkets in UK such as ASDA, Sainsbury's, Tesco, Ocado, and Morrison's. An average AOAC-fibre value was obtained from different brands aiming to have more representative dietary AOAC-fibre values (g/100g). The recipe calculation according to McCance and Widdowson's food composition book was used for a few food items. Calculated dietary AOAC-fibre values were derived from recipes in UK food composition and using the equation below (Food Standards Agency, 2002). Appendix J provides a list of food items (22) for which AOAC-fibre values were calculated in this way

$$\text{Nutrient content of cooked dish per 100g} = \frac{\text{Total nutrient of raw ingredients}}{\text{Weight of cooked dish}} \times 100$$

After obtaining all AOAC-fibre values from different sources, each AOAC-fibre value was assigned to an FFQ food item expressed in g/100g. Overall, 545 AOAC-fibre values were obtained to generate 217 food items in the FFQ. The pie chart in Figure 6.6 shows that around half of AOAC-fibre values were derived from the USDA databank, 19% were derived from food labelling, 12% were derived from a national dietary fibre review, 11% were derived from EuroFIR and 6 % were

derived from UK food composition tables and calculated proportion. Table 6.3 illustrates the percentage of AOAC-fibre values from different resources based on the main fibre-rich food groups. AOAC-fibre for the cereal group was mainly derived from USDA and UK food composition tables (29% and 24%, respectively), while AOAC-fibre values for the vegetables group were mainly derived from the USDA and the national review (36%, 26%). AOAC-fibre values for the legume group were derived mainly from the USDA (84%), and finally, AOAC-fibre values for the fruit group and nuts and seeds group were mainly derived from the national review (60% for the fruit group, 37% for nuts and seeds group) and EuroFIR databases (24% for the fruit group, 20% for the nuts and seeds group).

Figure 6.5 Search strategy for identifying the AOAC-fibre values for food items listed in the baseline FFQ

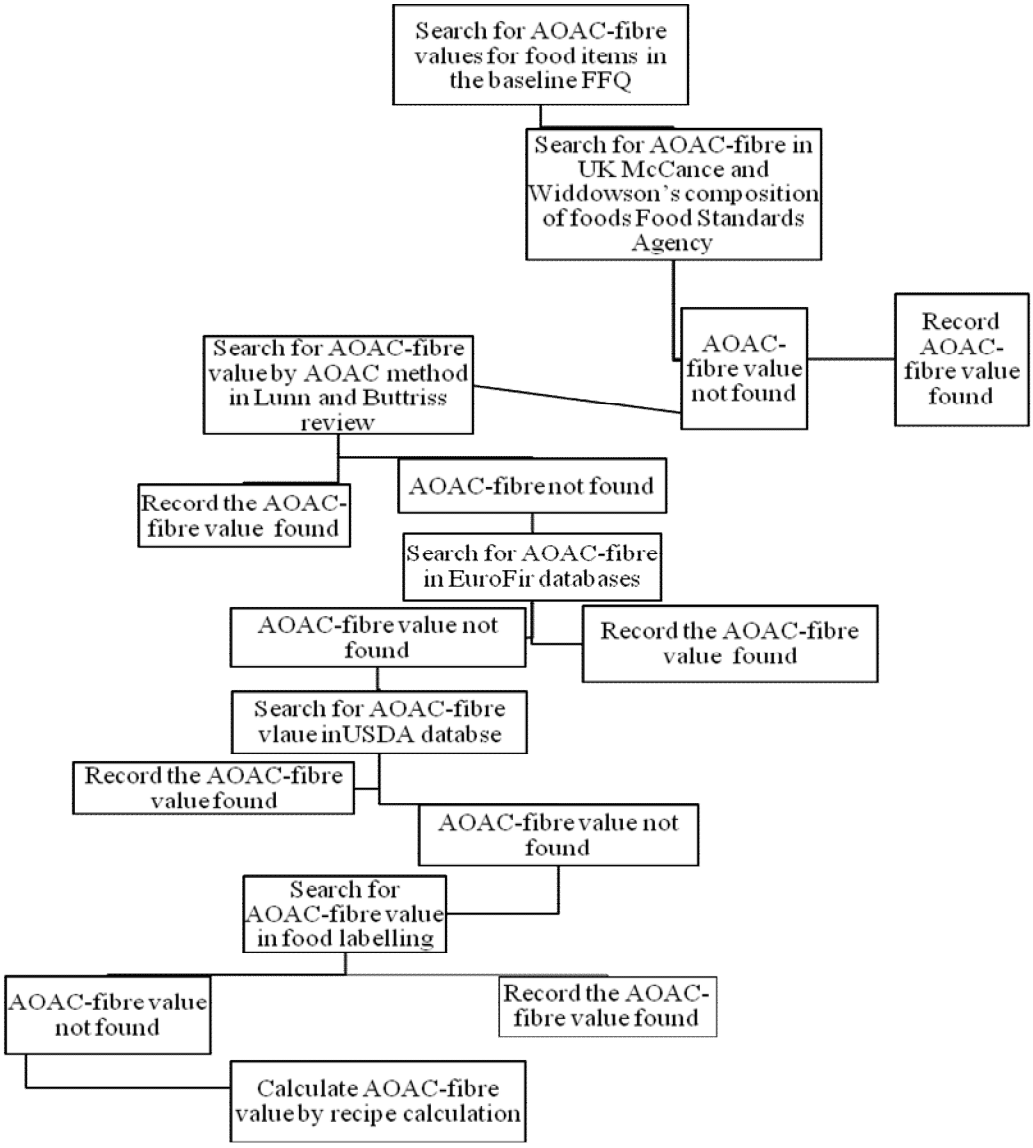


Figure 6.6 Distribution of information sources used to extract AOAC-fibre values

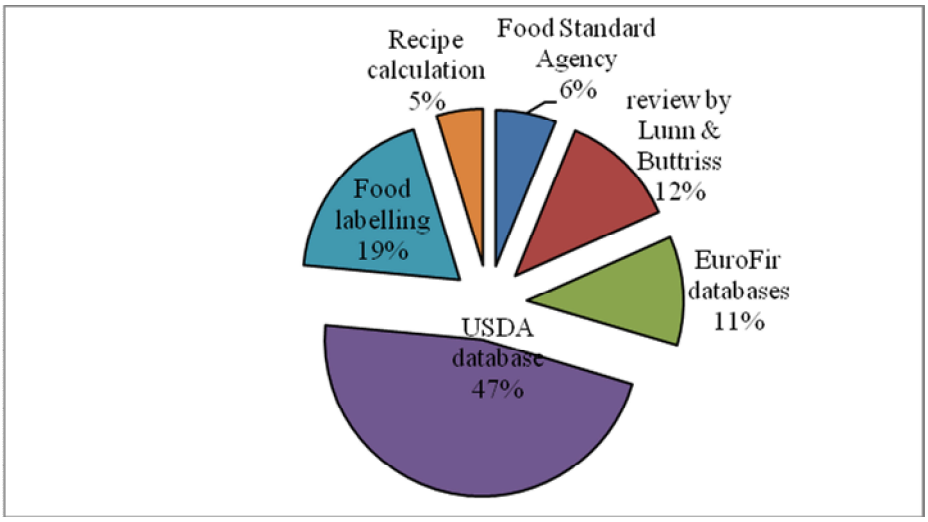


Table 6.3 Percentage of derived TDF values based on the top five food groups rich in fibre

Food groups	No. of food items	FSA 2002	National review	EuroFIR	USDA nutrient database	Food label	Recipe calculation
Cereal	35	24%	15%	15%	29%	13%	3%
Fruit	24	.	37%	24%	35%	4%	.
Vegetable	38	2%	28%	16%	36%	13%	5%
Legumes	9	.	.	5%	84%	5%	5%
Nuts and seeds	9	.	60%	20%	20%	.	.

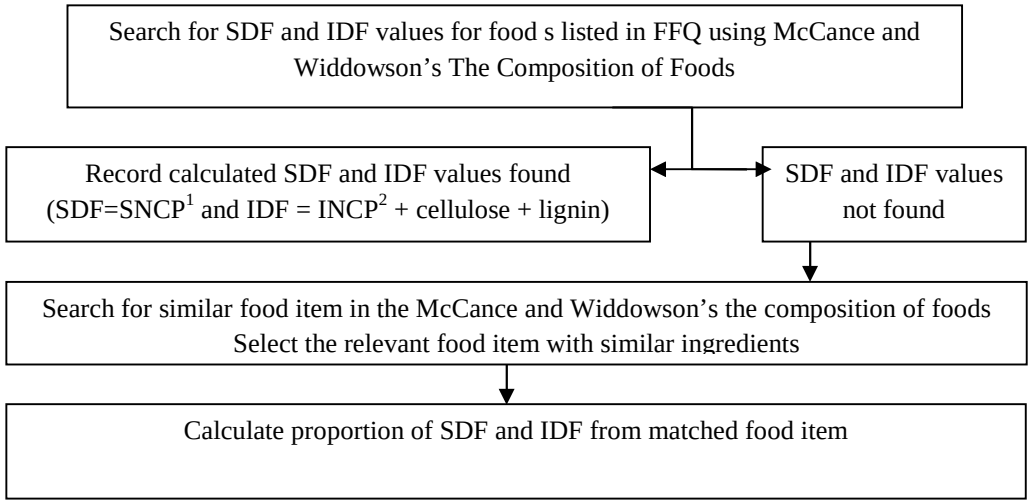
6.8.2Protocol 2: Search strategy for insoluble dietary fibre (IDF) and soluble dietary fibre (SDF) values

Similar to the search strategy used to obtain AOAC-fibre values, insoluble and soluble values were obtained from two resources. The majority (90%) of insoluble and soluble fibre values were identified in the UK food composition tables (Holland *et al.*, 1992a, Holland *et al.*, 1988, Holland *et al.*, 1989, Holland *et al.*, 1991a, Holland *et al.*, 1992b, Holland *et al.*, 1993). Values were assigned to each FFQ food item.

For the remaining 10% of food items in the baseline FFQ for which data were missing in UK food composition tables, an estimated proportion of soluble and insoluble fibre was made either by mapping to similar food items from available sources or estimating fraction proportions from food ingredients (of the food item) that rich in fibre then applied it to the available NSP value. Figure 6.7 presents the search strategy for insoluble and soluble fibre values. Similar to estimated AOAC-fibre intake within the UKWCS, daily dietary soluble and insoluble fibre intakes expressed as g/day were calculated for all cohort members by multiplying the

frequency of consumption of each FFQ item by the item portion size and the product was multiplied by the fibre content of each FFQ item.

Figure 6.7 Search strategy steps for identifying the soluble and insoluble fibre values for FFQ food items



[1]SNCP = soluble non-cellulosic polysaccharides
[2]INCP = insoluble non-cellulosic polysaccharides

6.9 Allocation of FFQ intakes to food groups

In the baseline FFQ, 217 food items were grouped into 18 food groups as presented in Table 6.4 mainly based on National Diet Nutrition Survey food groups categories (Department of Health, 2012). From the 18 food groups, the cereal group was subdivided into three main groups based on fibre content g per 100g of food (The Institute of Food Science & Technology, 2007). Cereal foods which contained 3-6g of dietary fibre per 100g were grouped into the 'cereals source of fibre' group, food items that contained more than 6g of dietary fibre per 100g were categorized into 'cereals high in fibre' group and the remaining cereal foods that contained less than 3 g/100g of food were put into 'other cereals' group.

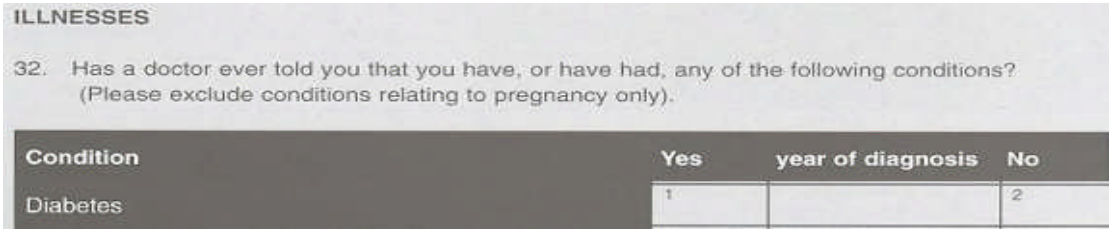
Table 6.4 List of food groups from 217 food items in the baseline FFQ

No	Food groups	The FFQ food items numbers	No of food items
1	Cereals “high in fibre”	3,5,8,13,14,15,68,70	8
2	Cereals “source of fibre”	2,4,6,9,10,11,22,69,173,204,206	11
3	“Other cereals”	1,7,12,21,23-26, 67,71,203,205,207,208,211	14
4	Potatoes	16-20	5
5	Fruit	53, 149-172 174	25
6	Vegetables	55, 90-96, 123-134, 136-148	33
7	Legumes	78-85, 135,142	10
8	Nuts	51, 56, 74-77,181-183	10
9	Milk and milk products	27-42 , 212-218	20
10	Meat and meat products	97-115	18
11	Fish and fish dishes	116-122	7
12	Non-alcoholic beverages	184-197	14
13	Alcoholic beverages	198-202	5
14	Eggs and eggs dishes	86-89	4
15	Fat spread	43-49	7
16	Savoury snacks	178-180	3
17	miscellaneous foods	57-66	10
18	Sugar and confectionary	50, 52-54, 175, 176, 177,219	8

6.10 Case ascertainment of T2DM as measure of outcome

Incident cases of T2DM were identified from a self-reported question completed at baseline and at phase II questionnaire (Figure 6.8). Phase II questionnaire was filled in after on average 4.2 (0.9) years from the baseline FFQ. (range 2-7 years). Question about year of diagnosis in phase II was used as a cross-check question to confirm incident cases in addition to data available from the baseline.

Figure 6.8 Questions to derive case ascertainment


ILLNESSES

32. Has a doctor ever told you that you have, or have had, any of the following conditions? (Please exclude conditions relating to pregnancy only).

Condition	Yes	year of diagnosis	No
Diabetes	1		2

From Table 6.5, out of 291 women who reported a diagnosis of diabetes in phase II, 152 women also reported diabetes at baseline. These women were excluded from the current study as well as 17 women who reported having diabetes at baseline and reported not having diabetes at phase II and 1,927 women who miss the question at baseline or phase II or both. 114 incident cases were therefore identified for the current study. Prevalence of diabetes from baseline is 1.8% (n=464) however, only 169 women who reported diabetes completed phase II questionnaire.

Table 6.5 Number of participants in the UKWCS who completed baseline and phase II.

		T2DM at phase II			
		Yes	No	Missing	Total
T2DM at baseline	Yes	152	17	14	183
	No	114	12,035	860	13,009
	missing	25	821	207	1,053
	Total	291	12,873	1,081	14,245

From 12,035, only 11,982 women were included as non-cases after excluding participants who reported unfeasible total energy intake (<500kcal/day and >5000 kcal/day) (n=53) Overall, the total number of women included in the analyses in chapter 8 and 9 was 12,096.

6.11 Conclusion

The UKWCS is one of the largest population based prospective studies in UK. This cohort is characterised by wide range of dietary pattern exposure that give the chance to examine diet and disease relationship. The baseline characteristics suggested that women were health conscious with high proportion of vegetarians.

Chapter 7: Comparison between dietary AOAC-fibre and NSP intakes in the UKWCS: Does method of dietary fibre analysis matter?

7.1 Introduction

Global nutrition guidelines are in general agreement that diets composed of foods rich in dietary fibre are beneficial with regard to the prevention of a wide range of nutrition-related chronic diseases (World Health Organization, 2003). However, a number of different analytical methods have been employed historically to estimate the proportion of food that is resistant to degradation by the human intestinal enzymes and that may therefore be classified as dietary fibre (DeVries and Rader, 2005). Currently, the two most commonly used analytical methods for measurement of dietary fibre for use in food composition tables and food labelling are the Englyst method (Englyst *et al.*, 1982), which estimates non-starch polysaccharide (NSP) and the methods of the Association of Official Analytical Chemists (AOAC) (Official Methods of Analytical Chemists, 1995) which provides an estimate of total dietary fibre that includes NSP, resistant starch and lignin. The latter method has been adopted by most European countries for the determination of total dietary fibre (TDF) for food composition tables, food labelling purposes (DeVries, 2004), and dietary fibre recommendations (Lunn and Buttriss, 2007).

In the UK, dietary fibre values reported in food composition tables and reference nutrient intakes set by the Committee on Medical Aspects of Food Policy are expressed as NSP (Food Standards Agency, 2002, Department of Health, 1991). More recently, resistant starch which by definition is not included in NSP, has been found to confer beneficial physiological effects (Nugent, 2005) and was added to the dietary fibre definition proposed by WHO/FAO (DeVries and Rader, 2005). As a result, the AOAC method has been recommended for the purposes of food labelling in the UK (Food Standards Agency, 2002). Buttriss and Stokes (2008) described how differences in methodology and analysis of dietary fibre have led to the confusing situation in the UK in which foods are labelled using AOAC fibre estimates, but dietary reference values, guidelines for consumption and food composition tables generally provide only NSP values. This is potentially confusing for the consumer, but also has impact for research concerning the impact of dietary fibre on health. Most surveys of UK participants report dietary fibre as NSP,

whereas other European and US studies use AOAC databases to determine consumption statistics. This makes cross-country comparisons of intake challenging, since it is not clear whether the highest NSP consumers in the UK would be similarly categorized as such using AOAC-fibre values. It is theoretically possible, that depending on the dietary pattern of an individual, a very high intake of foods rich in resistant starch may make a disproportionate contribution to AOAC-fibre intake, and may mean the highest AOAC-fibre consumer would not be similarly categorized as a high NSP consumer. Food groups which are particularly high in resistant starch include potatoes and potato products, bread, legumes and certain breakfast cereals such as cornflakes (Sajilata *et al.*, 2006).

A further challenge in nutrition epidemiology is the agreement of analytical methods for dietary fibre when pooling the results of studies that employed the different dietary fibre analysis methods. The relationship between AOAC-fibre and NSP values will determine whether conversion between the two analytical methods generates misclassification in studies using pooled data. A previous analytical approach (Englyst *et al.*, 1996) based on a small range of foods for ten major food groups (Peattie *et al.*, 1983, Food Standards Agency, 2002) suggested a mean ratio of AOAC-fibre to NSP of 1:1.33 (Lunn and Buttriss, 2007) may be employed to 'convert' NSP fibre values to AOAC fibre values. However, it is unclear whether this conversion approach may be equally valid when comparing dietary intakes within populations. The objectives for this chapter include the following.

- To examine the degree of agreement between AOAC-fibre and NSP intake in the UKWCS.
- To describe high fibre consumers as NSP or AOAC-fibre consumers in the UKWCS and to assess whether the characteristics of the high NSP consumers, are similar to high AOAC-fibre consumers.
- To determine the main food sources contributing to the NSP and AOAC-fibre intakes among cohort women
- To determine the dietary and non-dietary predictors for high fibre consumers.

7.2 Methodology

UKWCS data was used for this chapter as previously described in chapter 6.

7.3 Statistical analyses

A total of 34,454 participants were available for analysis after exclusion of women who reported consuming less than 500 kcal or more than 4,000 kcal per day.

Initially, graphical normality tests such as histograms, graph boxes and k-densities were carried out before using any hypothetical tests (Appendix I). Descriptive analyses were performed using mean (SD), median (IQR) of AOAC-fibre intake, NSP intake, and food groups that contributes to AOAC-fibre and NSP intakes. Dietary AOAC-fibre and NSP intakes were expressed as g/day and expressed in g per 1000 kcal per day which represents the proportion of NSP in the diet whilst holding total energy intake constant and so reduces variation related to energy intake. Statistical significance between the two measures was determined with P value < 0.05. Parametric and non-parametric tests were used as appropriate to examine the dietary NSP and AOAC-fibre intakes according to socioeconomic status, demographic, lifestyle and dietary variables.

Pearson's correlation between NSP and AOAC-fibre intakes was calculated to assess the strength of association between them. In order to determine the extent of agreement in the classification of the women into fibre intake quintiles, the subjects were divided into intake quintiles on the basis of both NSP and AOAC-fibre intakes. 1^{st} Quintile includes participants with the lowest fibre intakes and 5^{th} quintile includes participants with the highest fibre intakes. Demographic characteristics (age, ethnicity, education, and marital status), lifestyle characteristics (smoking status, alcohol intake, use of dietary supplements, and physical activity), anthropometric status (weight, BMI) for all participants were explored across the fibre quintiles. One-way analysis of variance and chi-squared tests were used to test for differences between the quintiles. The associations between dietary fibre and food groups were determined using Pearson correlation coefficients.

Both un-weighted and weighted Kappa analyses (Cook, 1998) were used to determine the extent of agreement of participant categorization into the same or adjacent quintiles respectively beyond the level expected by chance.

Food items listed in the baseline FFQ were grouped into 18 groups according to the National Diet Nutrition Survey food categories (Department of Health, 2012) as described in chapter 6. Further subgrouping was carried out for cereal group into three groups based on fibre content g per 100g (The Institute of Food Science & Technology, 2007) (appendix G provide the list of food items in each food group). Intake of NSP and AOAC-fibre from different sources were obtained and expressed in percentage. Fibre sources of insoluble and soluble fibre are presented in Table 7.4.

Analysis to explore dietary and non-dietary factors that predict high dietary AOAC-fibre and NSP consumption was carried out using multiple linear regression approach. Regression coefficients, beta (β), which denotes the increase in the dependent variable (intake of AOAC-fibre or NSP) for each unit increase in the explanatory variable (food groups or other women's characteristics), are presented with accompanying 95% confidence intervals. R^2 values denoting the variation in the outcome variable explained by the variable included in the regression model are also presented.

First multiple regression analysis was carried out to identify dietary determinants of high AOAC-fibre intake in cohort women. This help in identifying food groups consumed by women that result into a greater AOAC-fibre and NSP consumption. The second multiple regression analysis was carried out to predict lifestyle characteristics of participants with high AOAC-fibre and NSP consumption which may provide information about the target population (low fibre consumers) that may be the target for future health promotion and dietary intervention plans.

The above analyses were designed to explore whether there are similarities and differences in the characteristics of high fibre consumers when dietary fibre is estimated from different analytical methods. Mean (95%CI) NSP and AOAC-fibre intakes were obtained based on socioeconomic status, marital status, BMI, lifestyle and dietary factors. All statistical analysis was undertaken using STATA statistical software version 12 (Corp-Stata, 2010).

7.4 Results

7.4.1 Dietary fibre intake in the UKWCS

As may be seen in Table 7.1, mean intake of AOAC-fibre among cohort women was 38.3g/day while mean dietary NSP intake was 25.6g/day. Mean intake of insoluble and soluble NSP were respectively 16.0 (6.0) g/day, and 10.8 (3.9) g/day. The AOAC-fibre intake is higher than NSP intake by 33% which in actual terms equal to 12.7g/day.

Table 7.1 Dietary AOAC-fibre, IDF and IDF intakes in the UKWCS (n=34,454)

Dietary intake	Mean (SD) g/day	Median (IQR) g/day
AOAC-fibre (g/d)	38.3(14.4)	36.4(18.3)
AOAC-fibre density (g/1000kcal/d)	17.3(4.7)	16.9(6.2)
NSP (g/d)	25.6(9.8)	24.2(12.5)
NSP density (g/1000kcal/d)	11.6(3.3)	11.3(4.4)
Insoluble fibre (g/d)	16.0(6.0)	15.1(8.6)
Soluble fibre (g/d)	10.8(3.9)	10.3(4.8)

Table 7.2 demonstrates the mean (95%CI) of dietary fibre intake (AOAC-fibre and NSP) consumed by the UKWCS participants based on major baseline characteristics. Variations in amount consumed from NSP and AOAC-fibre were observed across different factors. For example smoking status, BMI categories, socioeconomic status, self-reported vegetarian, dietary supplement consumption, UK fruit and vegetable recommendations, self-reported legumes intake and menopausal status. With fruit and vegetables intake, women who consumed less than 400g/day of fruit and vegetables (which is less than the UK recommendations) reported dietary AOAC-fibre and NSP intakes of 25.9g/day and 17.1g/day respectively. However, overestimation of dietary AOAC-fibre intake using FFQ as dietary assessment method should be considered. Mean intakes were 37.1g/day from AOAC-fibre and 24.9g/day from NSP for overweight and obese women, which is slightly but statistically significantly lower than women with BMI of less than 25kg/m^2 (39.1g/day and 26.1g/day respectively. Women who did not report being vegetarian or vegan consumed 36.4g/day from AOAC-fibre intake and 24.3g/day from NSP intake.

Table 7.2 Mean (95%CI) of dietary fibre intakes expressed in g/day and g/1000kcal/day for UKWCS participants by major characteristics

Variable	AOAC-fibre g/day	NSP g/day	P value	AOAC-fibre g/1000kcal/day	NSP g/1000kcal/day	P value
Smoking status						
No	38.9(38.7, 39.1)	25.9(25.8, 26.1)	<0.01	17.0(17.0, 17.1)	11.5(11.5, 11.6)	<0.01
Yes	33.6(33.2, 34.1)	22.5(22.2, 22.7)		15.1(14.9, 15.3)	10.1(10.0, 10.2)	
BMI						
BMI <25	39.1(38.8, 39.3)[1]	26.1(25.9, 26.2)	<0.01	17.1(17.0, 17.2)[1]	11.4(11.4, 11.4)	<0.01
BMI ≥25	37.1(36.9, 37.4)	24.8(24.6, 24.9)		16.5(16.4, 16.6)	11.0(10.9, 11.1)	
Socio-economic class						
Professional	38.8(38.6, 39.1)	25.9(25.8, 26.1)	<0.01	17.0(16.9, 17.1)	11.4(11.3, 11.4)	<0.01
Intermediate	37.6(37.3, 37.9)	25.2(24.9, 25.4)		16.6(16.5, 16.7)	11.1(11.1, 11.2)	
Manual	37.1(36.7, 37.6)	24.7(24.4, 25.1)		16.4(16.2, 16.5)	10.9(10.8, 11.0)	
Self-report vegetarian or vegan status						
No	36.4(36.3, 36.6)	24.3(24.2, 24.5)	<0.01	16.0(15.9, 16.0)	10.7(10.6, 10.7)	<0.01
Yes	43.4(43.1, 43.7)	28.9(28.7, 29.1)		19.2(19.1, 19.3)	12.8(12.7, 12.8)	
Fruit and vegetables intake g/day						
<400g/d	25.9(25.7, 26.1)	17.1(17.0, 17.2)	<0.01	13.5(13.3, 13.5)	8.9(8.8, 8.9)	<0.01
≥400g/d	42.9(42.4, 42.7)	28.5(28.4, 28.6)		18.1(17.9, 18.1)	12.1(12.1, 12.1)	
Legumes consumer						
Yes	39.6(39.5, 39.8)	26.4(26.3, 26.6)	<0.01	17.2(17.1, 17.3)	11.5(11.4, 11.6)	<0.01
No	32.1(31.7, 32.5)	21.5(21.2, 21.7)		15.1(14.9, 15.2)	10.1(9.9, 10.2)	
Dietary supplement use						
No	36.5(36.2, 36.7)	24.4(24.2, 24.6)	<0.01	16.2(16.1, 16.2)	10.8(10.7, 10.8)	<0.01
Yes	39.5(39.3, 39.7)	26.4(26.2, 26.5)		17.3(17.2, 17.3)	11.5(11.5,11.6)	
Menopausal status at baseline						
Pre	38.1(37.9, 38.4)	25.4(25.3, 25.6)	<0.05	16.7(16.6, 16.7)	11.1(11.0, 11.2)	<0.01
Post	38.4(38.1, 38.7)	25.7(25.5, 25.8)		17.0(16.9, 17.1)	11.4(11.3, 11.5)	

7.4.2Agreement between AOAC-fibre and NSP intakes in the UKWCS

A strong positive correlation coefficient was observed between the AOAC-fibre and NSP intake estimates ($r = 0.9$; $P < 0.01$). A scatter diagram of dietary fibre intake using two methods is presented in Figure 7.1 showing a clear positive relationship. To assess the magnitude of agreement between the two analytical methods the Kappa analysis method was applied (Viera and Garrett, 2005). The agreement concerning the classification of women into the same quintile on the basis of AOAC-fibre and NSP intakes was 84% ($K = 0.8$). A weighed Kappa analysis was used to compare adjacent quintiles as a measure of partial agreement. Among the 34,454 women, the agreement between the two analytical methods in estimating total dietary fibre intakes was seen with a weighed Kappa value of 0.9. This is considered almost perfect agreement (Viera and Garrett, 2005) and indicates that a majority of the women were allocated to the same quintile by both analytical methods. This result suggests that the majority of high AOAC-fibre consumers (women who consumed more than or equal 50g/day) were also high NSP-fibre consumers (women who consumed more than 39g/day), and the majority of low AOAC-fibre consumers (women who consumed less than 21g/day) were also low NSP-fibre consumers (women who consumed less than 13g/day). No subjects were misclassified into quintiles at the opposite extremes, confirming the strong agreement of the dietary fibre intake values derived by these two analytical methods (as shows in Table 7.3). Percentages of women in the adjacent quintiles of AOAC-fibre and NSP intakes were only 8-11%.

Analysis of main fibre sources also showed high degree of agreement with Kappa value ranging between 0.5 (for potato group) and 0.9 (for cereal high in fibre, fruits and nuts) , and weighted kappa ranging between 0.7 (potato group) to 0.9 (cereal, fruit, vegetables, legumes and nuts). Ranges of Kappa were seen for food groups that contain lesser amount of fibre (Appendix H).

Figure 7.1 Relationship between AOAC-fibre and NSP intakes in women

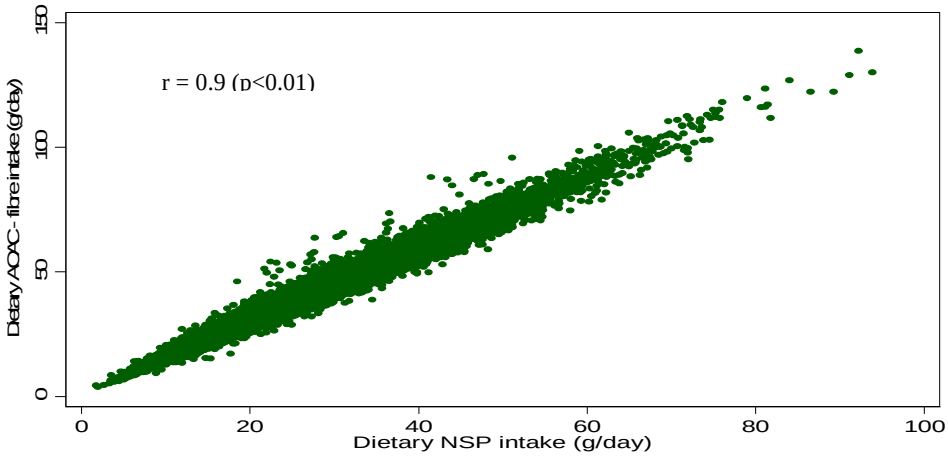


Table 7.3 Number of participants across dietary fibre intake quintiles estimated from AOAC-fibre and NSP values

		NSP intake					
		Lowest NSP	2nd	3rd	4th	Highest NSP	Total(n)
AOAC-fibre intake	Lowest AOAC-fibre	6,489	570	2	0	0 (0%)[1]	7,061
	2nd	574	5,730	748	5	1	7,058
	3rd	1	762	5,519	742	7	7,031
	4th	0	2	773	5,707	499	6,981
	Highest AOAC-fibre	0 (0%)[1]	0	1	536	5,786	6,323
	Total (n)	7,064	7,064	7,043	6,990	6,293	34,454

[1] n (%) of women who were misclassified into extreme quintiles in comparison to total number of participant

A regression analysis of dietary fibre intakes was used to predict the AOAC derived dietary fibre intake if the NSP dietary fibre value was known. This analysis showed that 98% of variability can be explained by the AOAC method. For every 1 g increase in NSP value, a 1.43 g/day increase in AOAC dietary fibre value ($p<0.001$) was estimated.

7.4.3Food groups in relation to AOAC-fibre, NSP intakes and types of fibres in the UKWCS.

As shown in Table 7.4, the five main sources of dietary fibre (both AOAC-fibre and NSP), accounting for more than three quarters of intake, were cereals, fruit, vegetables, legumes and potatoes. Within the cereal group, wholemeal bread contributed 7.4% and 6.1% respectively. Muesli and porridge account for 2.2% and

1.7% of total AOAC-fibre intake and 1.4% and 0.5% of total NSP intake. Fruits and vegetables contributed equally to total AOAC-fibre and NSP intakes and accounted for almost 40% of total AOAC-fibre and NSP intakes (39.3% and 39.6%; respectively). The legume group contributed to 8.9% and 9.5% of total AOAC-fibre and NSP intakes which is higher than cereal “source of fibre” group. Nuts were small contributor to fibre intake generally, providing 1.7% and 1.8% of total AOAC-fibre and NSP intakes.

Main five sources of IDF, contributing the majority of the intake, were cereals (18.8%), vegetables (12.9%), fruits (10.1%), legumes (5.2%) and potatoes (4.7%). Main food sources of SDF, were cereal (9.8%), vegetables (9.9%), potatoes (6.3%) and legumes (4.3%). Within the cereal foods, the main food group that contributed to total IDF and SDF were wholemeal bread (6.6% and 1.1%) and bran-based breakfast cereals (4.8% and 1.3%).

Table 7.4 Percentage of contribution of food types to average total daily AOAC-fibre, NSP, SDF and IDF intakes among cohort women (out of 100%)

Type of food	AOAC-fibre intake % (95%CI)	NSP intake % (95%CI)	SDF intake%(95%CI)	IDF intake %(95%CI)
Cereals high in fibre	17.9(17.7,18.0)	19.9(19.6, 19.9)	4.3(4.2, 4.4)	16.0(15.9, 16.1)
Wholemeal bread	7.4(7.4, 7.6)	6.1(6.0, 6.2)	1.1(1.0, 1.2)	6.6(6.4, 6.7)
All Bran and bran flakes	5.2(5.1, 5.2)	3.5(3.5, 3.5)	1.3(1.3, 1.4)	4.8(4.7, 4.8)
Muesli	2.2(2.1, 2.3)	1.4(1.4, 1.5)	0.8(0.7, 0.8)	2.0(2.0, 2.1)
Cereals source of fibre	7.8(7.4, 7.8)	6.6(6.6, 6.7)	3.0(3.0, 3.1)	4.9(4.8, 4.9)
Brown bread	2.7(2.5, 2.7)	1.9(1.6, 1.9)	0.9(0.8, 0.9)	1.9(1.8, 1.9)
Porridge	1.7(1.6, 1.8)	0.5(0.4, 0.5)	0.9(0.8, 1.0)	1.1(1.0, 1.1)
Spaghetti wholemeal pasta	1.3(1.2, 1.3)	0.9(0.7, 0.9)	0.3(0.2, 0.4)	1.1(1.0, 1.2)
Other cereals	5.4(5.4, 5.5)	4.9(4.8, 4.9)	2.5(2.5, 2.6)	2.6(2.6, 2.7)
White pasta	1.1(1.0, 1.2)	0.9(0.8, 0.9)	0.7(0.6, 0.8)	0.6(0.5, 0.6)
Non-sugar cereals	0.4(0.3, 0.4)	0.2(0.0, 0.2)	0.1(0.08, 0.2)	0.1(0.09, 0.1)
Buns/pastries	0.4(0.2, 0.5)	0.1(0.09, 0.1)	0.1(0.09, 0.1)	0.1(0.08, 0.1)
Potatoes	10.3(10.2, 10.4)	10.9(10.7, 10.9)	6.3(6.2, 6.4)	4.7(4.6, 4.8)
Fruit	19.8(19.7, 19.9)	18.9(18.7, 18.9)	8.9(8.9, 8.9)	10.1(10.0, 10.2)
Vegetables	19.5(19.4, 19.6)	20.7(20.6, 20.8)	9.9(9.9, 10.0)	12.9(12.8, 13.0)
Legumes	8.9 (8.8, 9.0)	9.5(9.5, 9.6)	4.3(4.2, 4.3)	5.2(5.1, 5.3)
Nuts	1.7 (1.7, 1.8)	1.8(1.8, 1.9)	0.6(0.5, 0.6)	1.2(1.1, 1.2)
Milk and milk products	0.6(0.5, 0.6)	0.1(0.09, 0.1)	0.05(0.05, 0.08)	0.1(0.08, 0.1)
Meat and meat products	1.9 (1.9, 2.0)	1.3(1.2, 1.3)	0.8(0.8, 0.9)	0.9(0.8, 0.9)
Fish and fish dishes	0.5 (0.4, 0.5)	0.2(0.2, 0.3)	0.1(0.1, 0.2)	0.1(0.1, 0.2)
Beverages	0.4(0.3, 0.4)	0.7(0.6, 0.7)	0.1(0.08, 0.1)	0.6(0.5, 0.6)
Alcoholic beverages	0(0)	0(0)	0(0)	0(0)
Eggs and eggs dishes	0.3 (0.2, 0.3)	0.3(0.2, 0.3)	0.08(0.08, 0.09)	0.2(0.1, 0.2)
Fat spread	0(0)	0(0)	0(0)	0(0)
Savoury snacks	0.9 (0.8, 0.9)	0.5(0.4, 0.5)	0.6(0.6, 0.7)	0.9(0.5, 0.6)
Miscellaneous foods	0.8 (0.7, 0.8)	0.9(0.8, 0.9)	0.5(0.4, 0.5)	0.1 0.1, 0.2)
Sugar and confectionary	1.0(0.9, 1.0)	0.5(0.4, 0.5)	0.1(0.08, 0.1)	0.4(0.3, 0.5)

Table 7.5 Dietary NSP and AOAC-fibre intakes (mean (SD), mean difference g/d and %) from all food groups included in the baseline FFQ

No	Fibre intake from food groups	AOAC-fibre intake g/day		NSP intake g/day		Mean difference	
		Mean	SD	Mean	SD	g/day	%
1	Cereals and cereals products	12.5	7.3	8.2	4.9	4.3	34
1.a	Cereal high in fibre	7.4	6.6	5.5	4.8	1.9	25
1.b	Cereal source of fibre	3.0	2.9	1.8	1.7	1.2	40
1.c	Other cereal	1.8	1.3	1.1	0.8	0.7	44
2	Vegetables	7.5	4.5	5.2	2.8	2.3	31
3	Fruit	7.9	6.0	4.7	3.3	3.2	37
4	Potatoes	3.6	2.2	2.8	1.6	0.8	22
5	Legumes	3.4	2.6	2.3	1.6	1.1	32
6	Nuts	0.7	1.1	0.45	0.77	0.25	36
7	Sugars, preserves and snacks	0.3	0.5	0.3	0.4	0	0
8	Sauces, soups and miscellaneous	0.4	0.5	0.1	0.2	0.3	75
9	Meat and meat products	0.6	0.7	0.3	0.3	0.3	50
10	Milk and milk products	0.2	0.3	0	0	0.2	100
11	Non-alcoholic beverages	0.1	0.1	0.1	0.3	0	0
12	Fish and fish dishes	0.2	0.2	0.1	0.1	0.1	50
13	Eggs and eggs dishes	0.1	0.2	0.1	0.1	0	0
14	Fat and oils	0	0	0	0	0	0
15	Alcoholic beverages	0	0	0	0	0	0

The mean difference between the dietary fibre intakes calculated using the AOAC and Englyst methods ranged from 0.2–4.3 g/day (presented in Table 7.5). The greatest differences between the AOAC-fibre and NSP-fibre intakes were observed in the cereals group, where a high amount of other non-digestible substances content was anticipated. This reflects a considerable amount of indigestible materials other than NSP in the cereal group that needs to be considered when estimating the total dietary fibre intake.

Also high variations between AOAC-fibre and NSP from vegetables, fruit and legume groups were observed in comparing to the remaining food groups that contain lesser amount of fibre. Generally, this indicates a great amount of indigestible materials other than NSP in the main fibre sources. A small variation observed in the nuts group in term of quantity may suggest a negligible amount of indigestible material other than NSP in the AOAC-fibre intake.

As shown in Table 7.6 correlations were generated between intakes of food groups in g/day against NSP or AOAC-fibre to explore how different food groups contribute to fibre intake overall. Pearson's correlations coefficients (*r*; 95%CI) between total dietary fibre intake and different food groups were found to be highly significant. Stronger correlations were found between dietary fibre intake (AOAC-fibre and NSP) and intakes of cereals, vegetables, fruit, legumes (r = 0.61,

0.67, 0.62, 0.55 and *r* = 0.58, 0.67, 0.61, 0.55; respectively). Within the cereal group, cereals high in fibre group were found to be strongly correlation with AOAC-fibre and NSP intake (r=0.55 and 0.59; respectively). While, other cereal group weakly correlated with NSP (*r*=0.13) and AOAC-fibre intakes (*r*=0.16). Within legumes group, dried legumes, which included chickpeas, lentils, baked beans, red kidney beans, have a slightly stronger correlation with dietary fibre intakes compared to fresh legumes including peas and green beans (*r* = 0.45 and 0.38 with AOAC-fibre and *r* = 0.43 and 0.41 with NSP intake). Intakes of other food groups were weakly correlated with dietary fibre intakes.

Table 7.6 Correlation between dietary fibre intakes expressed as NSP and AOAC-fibre within food groups among cohort women (n=34,454)

Type of food group	Pearson's correlation r (95%CI)	
	AOAC-fibre intake	NSP intake
Cereals and cereals products	0.61(0.60, 0.61)	0.58(0.57, 0.58)
Cereals high in fibre	0.55(054, 0.56)	0.59(0.58, 0.60)
Cereals source of fibre	0.40(0.39, 0.42)	0.35(0.34, 0.36)
Other cereals	0.16(0.15, 0.17)	0.13(0.12, 0.14)
Potatoes	0.28(0.27, 0.29)	0.27(0.26, 0.28)
Fruit	0.62 (0.61, 0.63)	0.61(0.60, 0.62)
Vegetables	0.67(0.66, 0.67)	0.67(0.66, 0.67)
Legumes	0.55(0.54, 0.56)	0.55(0.54, 0.56)
Dried legumes	0.45(0.43, 0.46)	0.43(0.41, 0.43)
Fresh legumes	0.38(0.37, 0.39)	0.41(0.39, 0.42)
Nuts	0.29(0.28, 0.30)	0.28(0.27, 0.29)
Milk and milk products	0.13(0.12, 0.14)	0.12(0.11, 0.13)
Meat and meat products	-0.08(-0.09, -0.07)	-0.09(-0.10, -0.08)
Fish and fish dishes	0.14(0.13, 0.15)	0.14(0.13, 0.15)
Non-alcoholic beverages	0.14(0.12, 0.14)	0.14(0.12, 0.15)
Alcoholic beverages	-0.06(-0.07, -0.05)	-0.05(-0.06, -0.04)
Eggs and eggs dishes	0.11(0.10, 0.12)	0.10(0.09, 0.11)
Fat spread	0.06(0.05, 0.07)	0.06(0.05, 0.07)
Savoury snacks	0.02(0.01, 0.03)	0.02(0.01, 0.03)
Miscellaneous foods	0.20(0.19, 0.21)	0.20(0.19, 0.21)
Sugar and confectionary	0.06(0.05, 0.07)	0.04(0.03, 0.06)

7.4.4Predictors of AOAC-fibre and NSP intakes in the UKWCS

The number of women who participated in the UKWCS is large which makes small effects statistically significant due to high statistical power. The highly significant levels may not be relevant; however our results assess the size of the effect. Dietary predictors of high AOAC-fibre intake are provided in Table 7.7. Increment unit was chosen as 10 grams for all food groups to compare across different sources of fibre. With every 10g increment in consumption of food from cereal, potatoes, nuts, legumes, fruit and vegetables groups, the NSP and AOAC-fibre intake is increased significantly. While with every 10 gram increment of dairy,

fat spread, sugar and alcoholic beverages consumption, the intake of NSP and AOAC-fibre decreases significantly. No association was observed between food such as meat and non-alcoholic beverages and NSP and no association between non-alcoholic beverages and greater intake of AOAC-fibre.

Table 7.7 Energy adjusted Beta coefficients (95%CI) of dietary fibre intake expressed as NSP and AOAC-fibre with every 10 grams consumed from food groups

Food intake	AOAC-fibre (g/day)		NSP intake (g/day)	
	Regression coefficient[1] (95%CI)	P	Regression coefficient[1] (95%CI)	P
Cereals high in fibre	0.76(0.76, 0.78)	<0.01	0.60(0.60, 0.61)	<0.01
Cereals source of fibre	0.31(0.30, 0.32)	<0.01	0.14(0.13, 0.15)	<0.01
Other cereals	0.07(0.06, 0.08)	<0.01	0.03(0.02, 0.04)	<0.01
Potatoes	0.19(0.18, 0.19)	<0.01	0.13(0.12, 0.13)	<0.01
Nuts and seeds	0.54(0.51, 0.57)	<0.01	0.34(0.32, 0.37)	<0.01
Legumes	0.55(0.54, 0.56)	<0.01	0.39(0.38, 0.39)	<0.01
Fruits	0.22(0.21,0.22)	<0.01	0.15(0.14, 0.15)	<0.01
Vegetables	0.22(0.22, 0.23)	<0.01	0.16(0.15, 0.16)	<0.01
Dairy and dairy products	-0.014(-0.166, -0.011)	<0.01	-0.01(-0.12, -0.008)	<0.01
Fat spreads	-0.424(-0.465, -0.384)	<0.01	-0.23(-0.26, -0.20)	<0.01
Meat and meat alternatives	-0.014(-0.019, -0.008)	<0.01	-0.0008(-0.004, 0.003)	0.65
Sugars and confectionary	-0.113(-0.128, -0.098)	<0.01	-0.08(-0.09, -0.07)	<0.01
Non-alcoholic beverages	-0.0003(-0.0011, 0.0004)	0.30	0.0005(-0.0006, 0.001)	0.07
Alcohol beverages	-5.0(-5.4, -4.6)	<0.01	-2.4(-2.7, -2.16)	<0.01
Miscellaneous	0.03(0.23, 0.03)	<0.01	0.02(0.01, 0.03)	<0.01

As shown in Table 7.7, the strongest dietary predictors of the high AOAC-fibre intake were: cereals ‘high in fibre’, pulses, nuts and seeds, cereals ‘source of fibre’ and lastly vegetables and fruit groups. With every 10g increment of fibre rich the cereals group, dietary AOAC-fibre and dietary NSP intakes increased by 0.76g/day and 0.60g/day respectively.

Similar dietary predictors were found with NSP intake however the vegetables and fruits groups were stronger predictors than the cereals sources of fibre group. This can be explained by the variation in the different ways of measuring dietary fibre in foods. The strongest dietary predictors of low dietary fibre intake were found to be: alcoholic beverages group; fat spreads group; sugars and confectionary group; and lastly the meat and meat alternatives group. With every 1 unit (10ml) increase in alcohol consumption, the daily AOAC-fibre and NSP intakes reduced greatly by 5g and 2.4g respectively among cohort women.

Multiple regressions were carried out again with higher increment in the food group to predict the amount of NSP and AOAC-fibre consumed from main food sources. The multivariable regression model included total energy intake and main fibre sources of AOAC-fibre and NSP intakes. Every unit increment of food group

consumed was 80g/day which is equivalent to a standard portion of food. Table 7.8 shows an example of food intake in each food group that is equivalent to 80g/day. However, nuts and seed group was expressed for each 10g/day increment (equivalent to 10 whole pistachio or cashew nuts).

Table 7.8 the portion of food that equivalent to 80g/day based on some food groups

Fibre source	80g/day equivalent to
Fruit group	one recommended portion[1]
Vegetable group	one recommended portion[1]
Legumes group	one recommended portion[1]
High fibre cereal group	two large thick slices of wholemeal bread
Other cereal group	two thick large slice of white bread
Cereals source of fibre	two crumpets
Potatoes group	small jacket potato without skin

[1](National Health Services)

Table 7.9 shows that the greatest increment of AOAC-fibre intake was seen with cereals high in fibre (β = 6.04), legumes (β = 4.40) and cereals source of fibre (β = 2.41). While, greatest increment of NSP intake was seen from cereals high in fibre (β = 4.78), legumes (β = 3.09) and vegetables (β = 1.27). Within the cereals group, the increment of AOAC-fibre and NSP is greatest with consumption of cereal high in fibre and smallest increment was with other cereal group. This reflects that consumption of dietary fibre from different subtypes of cereal was not greatly affected by analytical method.

Overall, similarity in the NSP and AOAC-fibre dietary predictors were observed. This indicates that analytical aspects may play a minor role when fibre intake is predicted from different food groups. However, effect size may vary slightly.

Table 7.9 Energy adjusted regression coefficients (95%CI) of dietary fibre intake expressed as NSP and AOAC-fibre with every 80 grams consumed from food groups except nut and seeds (10g/day)

Food intake	AOAC-fibre (g/day)		NSP intake (g/day)	
	Regression coefficient[1] (95%CI)	p value	Regression coefficient[1] (95%CI)	p value
Cereals high in fibre	6.04(5.98, 6.11)	<0.01	4.78(4.74, 4.83)	<0.01
Cereals source of fibre	2.41(2.36, 2.46)	<0.01	1.10(1.07, 1.14)	<0.01
Other cereals	0.49(0.43, 0.54)	<0.01	0.15(0.11, 0.19)	<0.01
Potatoes	1.45(1.41, 1.49)	<0.01	1.02(0.99, 1.05)	<0.01
Nuts and seeds	0.51 (0.48, 0.55)	<0.01	0.32(0.29, 0.34)	<0.01
Legumes	4.40(4.32, 4.48)	<0.01	3.08(3.02, 3.13)	<0.01
Fruits	1.78(1.75, 1.80)	<0.01	1.15(1.13, 1.16)	<0.01
Vegetables	1.73(1.71, 1.75)	<0.01	1.27(1.25, 1.29)	<0.01

[1] multivariate regression model adjusted for food groups listed above and other food groups (dairy and dairy products, fat spreads, meat and meat dishes, eggs and egg dishes, fish and fish dishes, sugars and confectionary, non-alcoholic beverages, alcohol beverages and miscellaneous).

Non-dietary related factors were also examined to predict high dietary fibre intake as a continuous variable. Factors such as, BMI, physical activity, smoking, age, dietary supplement consumption, self-reported vegetarian status and socioeconomic class were examined. Coefficients of regression of dietary fibre intake were predicted for each factor accounting for other factors listed in Table 7.10.

As reported previously, 28% defined themselves as vegetarians. Overall, vegetarians' status was strongest predictor of fibre intake. Women who described themselves as vegetarian or vegan consumed 7 g/day more AOAC-fibre than non-vegetarians (95% CI: 7.92, 7.45; P < 0.01), while NSP intake was 4.75 g/day higher among vegetarians than non-vegetarians (95% CI: 4.56, 4.94; P < 0.01). Lower daily AOAC-fibre and NSP intakes were observed among smokers compared to non-smokers by 3.79g/day and 2.54g/day respectively. Compared with the women in the professional social class, consumption of AOAC-fibre and NSP among the intermediate social class was less by 0.7 g and 0.5g respectively. While, women in the routine and manual social class consumed 1 gram less and 0.75g less of AOAC-fibre and NSP intakes in comparison to the highest social class. Obese women consumed almost 1 g/day of AOAC-fibre and approximately half a gram of NSP less than women with low BMI (less than 25kg/m^2). The findings identified that women who were younger, smoke, non-vegetarian, from a low social class, high BMI were more likely to have a lower fibre intake.

Table 7.10 Predictors of high fibre intake calculated by AOAC-fibre and NSP intakes

		Coefficient regression (95% CI)[1]	
Variable		AOAC-fibre intake (g/d)	NSP-fibre intake (g/d)
Age (5 year)		0.59(0.52, 0.66)	0.41(0.36, 0.45)
BMI (kg/m^2)	< 25	1	1
	25-30	-0.52(-0.80, -0.25)	-0.37(-0.56, -0.17)
	> 30	-0.97(-1.33, -0.62)	-0.68(-0.93, -0.42)
Physical activity (15min/d)		0.43(0.37, 0.49)	0.31(0.26, 0.35)
Smoker at baseline		-3.79(-4.17, -3.42)	-2.54(-2.81, -2.27)
Vegetarian		7.19(6.92, 7.45)	4.75(4.56, 4.94)
Dietary supplement users		-1.85(-2.09, -1.62)	-1.20(-1.37, -1.03)
Socio-economic class	Professionals*	1	1
	Intermediate	-0.69(-0.96, -0.43)	-0.45(-0.63, -0.26)
	Routine and manual	-1.11(-1.52, -0.69)	-0.75(-1.05, -0.46)

* Reference group; [1]adjusted for each other variables.

7.4.5Characteristics of highest versus lowest AOAC-fibre and NSP consumers in the UKWCS

Women were assigned to five equal quintiles according to AOAC-fibre intake and NSP, where the mean daily dietary fibre intake was 21g and 13.7g in 1st quintile, 30g and 19.9g in 2nd quintile, 36.9g and 24.5g in 3rd quintile, 45g and 30.2g in 4th quintile and 61.3g and 39.5 in 5th quintile respectively. Comparing the highest versus lowest fibre consumers across the AOAC-fibre quintiles, high AOAC-fibre consumers were more likely to be non-smokers (93%) as compared with the low AOAC-fibre consumers (82%), and were also more likely to be older, vegetarian, and have higher socioeconomic status (professional and managerial class).

In terms of dietary characteristics, women in the highest AOAC-fibre quintile consumed approximately five more vegetable portions per day than women in the lowest quintile (8 versus 3 servings, respectively; $P < 0.01$) and triple the number of fruit portions per day as compared with the lowest consumers (9 versus 3 portions, respectively; $P < 0.01$). High AOAC-fibre consumers reported lower intakes of protein (15%), fat (31%), saturated fat (9.6%), and alcohol, but higher carbohydrate intakes (58%) as compared to low AOAC-fibre consumers ($p<0.01$). Table 7.11 and 7.12 show similar pattern in the characteristics across AOAC-fibre and NSP quintiles

Generally, dietary fibre intakes expressed in g/day and g/1000kcal/day and obtained as NSP or AOAC-fibre were all have similar characteristics across the fibre quintiles.

Table 7.11 Baseline characteristics and nutrient intakes across AOAC-fibre and NSP quintiles

		AOAC-fibre intake				*NSP intake*			
Dietary fibre quintile		*1st*	*3rd*	*5th*	*P value*	*1st*	*3rd*	*5th*	*P value*
Mean (SD) intake g/day		21.0 (4.1)	36.9 (2.0)	61.3 (10.9)	<0.01	13.7(2.8)	24.5(1.4	39.5(4.3)	<0.01
Range g/day		≤26.4	>33.4 -≤ 40.5	>50.3	0.01	≤17.4	>22.2-≤27.1	>33.8	
No. of participants		7060	7043	6293		7057	7031	5722	
Lifestyle and socio-demographic characteristics									
Age (year)		51.9 (9.2)	52.5 (9.3)	52.4 (9.3)	0.8	51.9(9.3)	52.3(9.2)	52.4(9.3)	<0.01
BMI (kg/m^2)		24.9 (4.5)	24.4 (4.3)	23.9(4.3)	0.01	24.9(4.5)	24.3(4.3)	23.9(4.0)	<0.01
Physical activity (MET/Week)		17.8(11.1)	16.8 (11.1)	19.2 (12.3)	0.01	11.7(12.3)	13.8(12.4)	19.1(12.4)	<0.01
Self-reported vegetarian (%)		16	26	42	0.01	16	26	41	
Socio-economic status (%)	Professional	61	63	67	0.01[1]	61	62	66	<0.01[1]
	Intermediate	29	28	25		29	28	25	
	Routine and manual	10	9	8		10	9	8	
Educational level (%)	No education record	20	16	16	<0.01	20	16	16	<0.01[1]
	O-level	33	30	29		33	31	29	
	A-level	22	25	26		22	25	25	
	Degree	24	27	29		24	27	29	
Smoking status (%)		18	9	7	0.01	18	9	8	<0.01
Dietary supplement users (%)		50	58	65	0.01	50	58.1	65	<0.01
Dietary characteristics									
Energy intake (kcal)		1643 (431)	2221(466)	2864 (516)	0.01	1672(452)	2219(483)	2799(509)	<0.01
Carbohydrates (% E)		51.7	54.5	58	0.01	51.7	54.4	57.8	<0.01
Protein intake (%E)		16.7	15.9	15.2	0.01	16.6	15.9	15.1	<0.01
Fat intake (% E)		35	33.1	30.9	0.01	35.1	33.2	30.7	<0.01
Sat fat intake (% E)		12.9	11.5	9.6	0.01	13	11.4	9.7	<0.01
Alcohol (units/day)		0.9 (1.1)	0.8 (0.9)	0.7 (0.9)	0.01	0.55(1.3)	0.58(1.2)	0.49(1.1)	<0.01
Vitamin C (mg/day)		105.3 (45.8)	164.7 (56.0)	255(94)	0.01	104.3(44.5)	164.8(55.4)	242.6(79.5)	<0.01
Iron (mg/day)		10.8 (2.7)	17.4 (3.8)	26.3 (7.3)	0.01	10.9(2.8)	17.4(3.9)	25.4(5.8)	<0.01
Calcium (mg/day)		874 (289)	1128 (308)	1400 (357)	0.01	883.4(295)	1129(315)	1373.5(347)	<0.01
Vegetable portions/day		3.0 (1.5)	4.9 (2.3)	7.9 (3.6)	0.01	2.9(1.4)	4.9(1.9)	7.5(2.9)	<0.01
Fruit portions/day		2.6(1.6)	4.7(2.4)	9.1(6.2)	0.01	2.6(1.6)	4.8(2.4)	8.1(4.7)	<0.01

[1]overall p value

7.5 Discussion

The objective of this study was to examine the relationship between dietary fibre intakes derived using the AOAC and Englyst methods in participants of the UKWCS. A high degree of correlation was expected since NSP is the dominant fibre type measured by both methods of analysis. AOAC-fibre values were consistently significantly higher than NSP-fibre values for all food items included in the dietary fibre intake calculation because resistant starch and lignin were included in the AOAC-derived dietary fibre as compared to the Englyst method (Wolters *et al.*, 1992).

High daily dietary fibre intake was observed in the studied population compared to the data from the NDNS (Food Standards Agency, 2010). However, there is still an argument around how much of this additional fibre is due to FFQ methodology and how much is due to health consciousness, or underreporting by NDNS participants. These results suggest that the studied population includes primarily women with healthy lifestyles. Daily AOAC-fibre intake in the UKWCS averaged 38.3 g, this is higher than the recommendation proposed by the WHO/FAO (World Health Organization, 2003). The average daily intake of NSP was 25 g, which also exceeds current UK dietary recommendations (Department of Health, 1991). Mean daily AOAC-fibre intake was significantly higher, by 32%, than daily NSP-fibre intake. This is reflected in WHO/FAO recommended daily intake, which for adults are respectively 25g and 20g (Lunn and Buttriss, 2007). The significant higher values of AOAC-fibre is due to the presence of components other than NSP (Englyst *et al.*, 2007).

The high level of NSP intake found in the cohort is perhaps not so surprising when we consider that sample was chosen to include a large proportion of vegetarians who consume high amounts of dietary fibre in compared to meat eaters (Cade *et al.,* 2004a). The advantage of UKWCS is to have a wide range of dietary exposure to investigate the protective effect of diet on health outcomes. Another probable reason for having high NSP estimates may partly be due to overestimation of dietary intake using FFQ data (Calvert *et al.*, 1997). Considering these limitations and appreciating that the data are not providing us with actual intakes, but rather indications of levels of dietary fibre intake. When cohort women classified into five equal groups based on their dietary NSP and AOAC-fibre intake expressed in g/day and g/1000kcal/day, similarity in the trends were observed with the socio-

demographic, lifestyle and dietary characteristics. Due to the large sample size of the UKWCS (high statistical power) a small effect of fibre intake examined was statistically significant.

This is the first epidemiological study to examine the influence of analytical methods on estimated dietary fibre intake by exploring the degree of agreement and correlation between the AOAC and Englyst methods with respect to subject classification using individual food items. The broad range of dietary fibre intakes in the studied population, with a large proportion of vegetarians, facilitates evaluation of fibre rich food sources using both analytical methods. Conversely, the studied population appears to be more health conscious than average, and may not reflect dietary fibre intakes among the general population.

More recent adaptations to the AOAC fibre analysis method, which now incorporates greater levels of resistant starch and non-digestible oligosaccharides, would have generated potentially greater differences in categorization of the women into fibre intake quintiles.

It was reported in a review by Englyst (Englyst *et al.*, 2007), that the AOAC method is highly dependent on food processing and therefore variability is expected from one product to another. This variability can affect the dietary fibre intake estimation in the present study where each of the 217 food items in the FFQ may represent several types of food with different preparation methods such as canned, boiled, grilled, or fresh while evidence showed that cooking methods have no effect on the NSP content (Reistad and Frolich, 1984).

A significant result of this study is the strong agreement between the AOAC and Englyst methods in classifying high and low dietary fibre consumers, which is supported by the absence of subjects misclassified into extreme opposite quintiles. This suggests that classifications of high or low dietary fibre intakes are comparable when calculated using these two different analytical methods. High fibre consumers are classified similarly using each fibre analysis method.

A potential limitation of the study was the restricted number of AOAC-fibre values in the UK Food Composition Tables, resulting in half of the matched food items being derived from the USDA nutrient database (U.S Department of Agriculture, 2010). This may have affected the precision with which food items were identified and, subsequently, the mean difference between the AOAC- and NSP-fibre values. Although half of the AOAC-fibre values were derived from

USDA which may have an impact on the precision of AOAC-fibre intake estimation, the results showed a strong agreement between the two analytical methods.

One-third of the total dietary fibre intake calculated using both analytical methods was contributed by cereals and cereal products, while two-thirds came from fruits and vegetables, including legumes and potatoes.

The FFQ is frequently used in epidemiological research because it is considered to be a cost-effective dietary assessment tool (Willett, 1998). Dietary fibre conversion factor may provide a basis for calculation of AOAC-fibre intake from NSP-fibre intake, which may help in improving the comparability between the UK studies with other studies that used AOAC-fibre estimation when pooled evidence is examined for dose-response relationship.

A conversion factor of 1.33 has been suggested for derivation of AOAC fibre intake using NSP values (Lunn and Buttriss, 2007) Another study by Reistad and Frolich (1984) found a range between 1.1 to 1.4 from measured four vegetables. The current analysis showed a conversion factor of 1.43 which is slightly higher than previously established ratio (Lunn and Buttriss, 2007). Our results were derived from the intakes of 217 food items, whereas the previously reported ratio of 1.33 was derived from analysis of only 115 food items (Peattie *et al.*, 1983). However, with half of the values for AOAC-fibre content in foods derived from non-UK resources, this finding may restrict the use of current conversion factor.

Because the present study is the first to examine the relationship between these calculation methods in an epidemiological application, additional investigation is needed to establish whether or not this ratio is applicable across a broad range of population groups. The estimated ratios for each food group may vary, which may be explained by differences in the amounts of indigestible materials mainly resistant starch (RS) measured by the AOAC-fibre values for each food group (Englyst *et al.*, 2007). Specifically, certain food groups may be a richer source of resistant starch. It had been reported by Englyst *et al.* (2007) that up to 5% of starch in most of cereal products is RS while other types of food such as legumes are richer in RS (10-20% of total starch). The amount of RS is highly dependent on food processing (cooking methods..etc). It should be recognized that current AOAC methods (985.29 and 991.43) capture only part of the spectrum of total RS that may be better quantified by newer AOAC method described by McCleary (2007). Resistant starch health

benefits have been reported (Buttriss and Stokes, 2008). The great benefits of resistant starch was seen on colonic function mediated by production of short chain fatty acids while seem to have smaller beneficial effect on lipid and glucose markers (Nugent, 2005). More robust evidence has been found supporting the impact of NSP consumption such as promoting a regular bowel habit (Buttriss and Stokes, 2008). It has been argued that not all resistant substances included in the dietary fibre definition could demonstrate each physiological effect with a similar magnitude (Buttriss and Stokes, 2008). Therefore, Englyst suggested that for nutrition research, detail information on specific food component will be more helpful for future epidemiology studies to establish the intakes and health effects of fibre substances other than NSP (Englyst *et al.*, 2007). But without adequate databases, it is not possible to do this.

In the UK, women's diets generally were found to be low in fibre (Ruxton and Derbyshire, 2010) with a mean intake of 12.6g/day that was significantly lower than men and even lower that the recommended dietary intake (Henderson *et al.*, 2002) . The mean dietary intake of NSP reported in the most recent NDNS rolling programme remained the same (12.8g/day) (Department of Health, 2012).

In addition to determine the fibre rich sources, this can be useful for implementation of dietary guidelines. Dietary and non-dietary factors that significantly predict high fibre consumers were identified in this chapter, and may help to identify best approaches to increase fibre intake from different food groups. Non-dietary factors identification may help to identify target populations that need health promotion to encourage high fibre intake in future interventional plans. It is theoretically possible, that depending on the dietary pattern of an individual, a very high intake of foods rich in resistant starch may make a disproportionate contribution to AOAC fibre intake, and may mean the highest AOAC fibre consumer would not be similarly categorized as a high NSP consumer. Food groups which are particularly high in resistant starch include potatoes and potato products, bread, legumes and certain breakfast cereals such as cornflakes (Sajilata *et al.*, 2006). The current findings showed similar positive associations between different food intakes and dietary fibre intakes expressed as NSP and AOAC-fibre.

7.6 Conclusion

As anticipated, mean AOAC-derived dietary fibre intake was significantly higher than the NSP intake. The weighed kappa (κ) analysis found a strong

agreement in terms of allocation to quintiles using the two fibre analysis methods. This indicates that the majority of the cohort participants were placed in the same fibre quintile, regardless of the method of fibre analysis used. None of the consumers were misclassified in the opposite extreme quintiles. Each 1g increase in NSP was associated with a 1.43g increase in AOAC dietary fibre. This is somewhat higher than a previously published ratio of 1.3. These analyses suggest that for the conduct of a meta-analysis that compares risk of disease in the highest versus the lowest quintiles of dietary fibre intake the method of fibre analysis employed is of minor importance. Women in the highest dietary fibre intake group (above 51g/day of AOAC-fibre and above 39g/d of NSP) will be categorized as such regardless of the method used. However, for a dose-response meta-analysis, pooling of results using different fibre analysis methods would be inappropriate. Dietary and dietary related predictors of high AOAC-fibre and NSP intakes were found to be overall similar. However, the different types of food consumed may have varied effect on magnitude of NSP and AOAC-fibre intakes. Younger women, who smoke, have high BMI, from low social class, are non-vegetarians were lower fibre consumer, and thus should be considered for forthcoming health promotion that encourages high fibre consumption.

Chapter 8: Dietary fibre intake and risk of T2DM among British women

8.1 Introduction

As described in the literature review (Chapter 2), the tremendous burden of high prevalence and incidence of T2DM on health status, economy and quality of life over time resulted in the establishment of guidelines for prevention (World Health Organization, 2003). Growing evidence on health benefits of dietary fibre and the importance of the primary prevention of the chronic diseases suggested the need for a re-assessment of the dietary fibre goal in the UK (Lunn and Buttriss, 2007). This may help in gaining the benefit of the optimum effect of dietary fibre intake in the prevention action plan. On the other hand, dietary fibre is a heterogeneous combination of compounds and evidence on physiochemical properties of dietary fibre may partly explain the variation in the physiological effect on human health (Jenkins *et al.*, 2002). Several potential mechanisms suggested the link between dietary fibre intake and glucose response was described in section 4.4.2.

Inconsistency was observed in prospective studies that examined the effect of dietary fibre intake on the risk of T2DM among women (Hopping *et al.*, 2010, Salmeron *et al.*, 1997b). Where, recent cohort showed significant effect of dietary fibre intake on men than women (Hopping *et al.*, 2010). This suggests possible differential effect of dietary fibre on risk of diabetes by gender. In addition, lack of prospective studies that examine the effect of dietary fibre intake on T2DM on women particularly in UK suggests the needs for further exploration. Also, a limited number of cohort studies have examined the effect of soluble and insoluble dietary fibre intakes on risk of T2DM. Thus, further research is required to investigate the effect of dietary fibre fractions on the risk of T2DM in women.

As discussed in chapter 4, very few studies have looked at NSP intake in relation to T2DM and as UK food composition tables provide only NSP values thus UK studies usually use NSP values rather than AOAC-fibre values. This study considers one of the first studies that looked at the effect of dietary fibre intake estimated by AOAC method and risk of T2DM among women in the UK. Therefore, this chapter aims to examine the association between total dietary fibre intake and risk of T2DM in the UKWCS. The examined dietary fibre intake was obtained from

two different measurement methods (AOAC-fibre intake and NSP intake). Another aim of the current chapter is to determine whether intakes of soluble and insoluble fibre can protect from developing T2DM among cohort women. Another main focus of the current study is to examine the association between the intakes of dietary fibre from specific food sources and the risk of T2DM.

8.2 Method

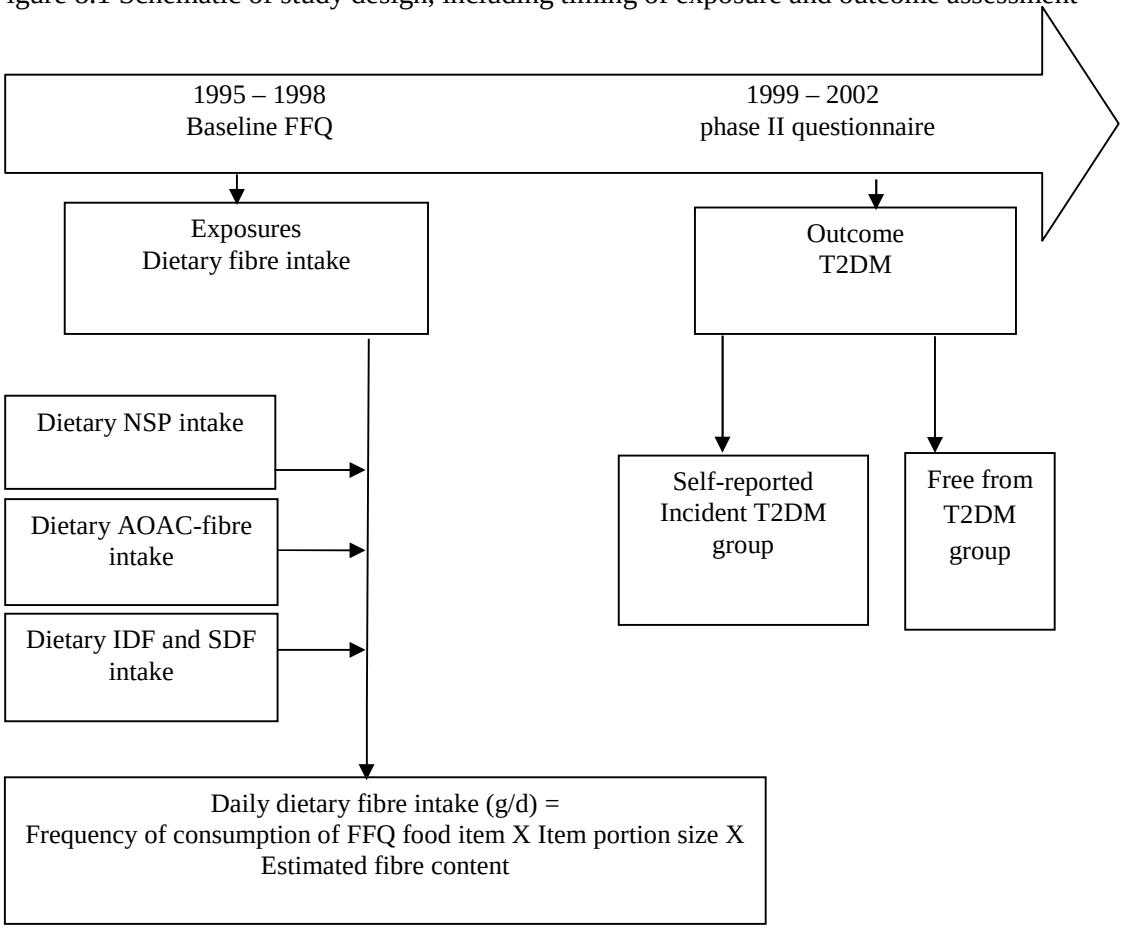
8.2.1 Study population

As described in chapter 6, the UKWCS (Cade *et al.*, 2004a) is a population based prospective study that includes mainly white women, aged 35-69, with a wide range of dietary patterns. This gives an optimal opportunity to determine risk of nutrition related chronic diseases in a population with varied degree of dietary exposure.

8.2.2 Outcome measure

Selection of cases of diabetes was undertaken as demonstrated in chapter 6. From 12,149, only 11,982 women were included as non-cases after excluding participants who reported unfeasible total energy intake (<500kcal/day and >5000 kcal/day) (n=53). In addition to 114 incidents were identified for the current chapter. Figure 8.1 summarises the timescale of baseline and phase II data capture used for the current analyses. Dietary variables were obtained from the baseline FFQ and outcome (T2DM status) was obtained from baseline and phase II. Overall, the total number of women included in the analysis was 12,096.

Figure 8.1 Schematic of study design, including timing of exposure and outcome assessment



8.2.3Exposure measures

8.2.3.1 Total dietary fibre intake

From Figure 8.1, total dietary fibre intake was estimated from baseline FFQ and expressed as AOAC-fibre and NSP (g/day). AOAC-fibre values were obtained as described in chapter 6. Dietary fibre intake was expressed as estimated intakes in g/day and fibre density in g/1000kcal/day.

8.2.3.2 Insoluble and soluble fibre intake

Estimated dietary insoluble and soluble fibre intakes were calculated from the baseline FFQ. The fibre values were obtained from 5th edition of the UK composition tables was used as more relevant time-frame of baseline than 6th edition (Holland *et al.*, 1991b). Insoluble fibre values of foods and food products were composed of cellulose, insoluble non-cellulosic polysaccharides and lignin while soluble fibre values included the soluble non-cellulosic polysaccharides. Details on the recently obtained data for AOAC-fibre intake are available in chapter 6.

8.2.3.3 Dietary fibre sources by main food groups

FFQ food items were allocated to food groups along the lines used in the UK food composition tables (Food Standards Agency, 2002) and are shown in Table 8.1. Selected 118 foods and food products from the baseline FFQ were assigned to five food groups as being the highest contributors to total dietary fibre intake. Fibre intakes from cereals, fruit, vegetables, legumes, and nuts groups were expressed in grams per day.

Table 8.1 list of food items from the baseline FFQ based on main food grouping in McCance and Widdowson's **(2002)**

No	Food groups	No items	List of FFQ food item
1	Cereals	37	wholemeal bread average, papadums, crispbread, muesli, all bran, Weetabix, oat, wheat germ, brown bread average, chapattis, tortillas, cream crackers, porridge, sugar coated cereals , wholemeal pasta, burgar, cereal bar, chocolate biscuits, fruitcake, buns, scones, fruit pie, white bread, pitta, white pasta, white rice, brown rice, macaroni cheese, couscous, non-sugar cereal, wild rice, plain biscuits, sandwich biscuits sponge cakes, sponge puddings,
2	Fruit	24	apple, avocado, bananas, grapes, kiwi, mango, orange, papaya, pears, pineapple, apricot, melon, nectarines, peaches, plums, raspberries, currants, rhubarb, strawberries, dates, fig, prunes, mixed dried fruits, currants & sultanas
3	Vegetables	38	quron, textured vegetable protein, veg chilli, mixed beans casserole, stir fry veg, veg dishes, veg pizza, beetroot, broccoli, Brussels, cabbage, carrots, cauliflower, celery, coleslaw, low cal coleslaw, courgettes, cucumber, garlic, lettuce, leeks mushrooms, olives, peppers, swede, sweet corn, tomatoes, turnip, watercress boiled potato, jacket potato, potato salad chips, roasted potato barley
4	Legumes	10	lentils, chickpeas, hummus, baked beans, red kidney beans, bean sprout, blacked eyed beans, butter beans, green beans, peas
5	Nuts and seeds	9	Peanut butter, nut pate, peanuts, cashews, pecans, sunflower seeds, Bombay mix, peanuts/pistachio, mix nuts

8.2.4Other variables

The baseline FFQ included questions on socio-demographic factors such as age, ethnicity, socioeconomic status and educational level. Diabetes related risk factors such as smoking status and physical activity and anthropometric factors such as calculated BMI (kg/m^2) from self-reported weight and height, waist circumference (cm) and calculated weight change (different between weight at age of 20 years and present weight) and expressed in % and Kg that were generated from information provided by women were also obtained from the baseline questionnaire. Family history of diabetes was also obtained from the baseline FFQ. Magnesium intake (mg/day) was obtained from the baseline FFQ. Details on variables of interest will be demonstrated later in section.8.3.3.

8.3 Statistical analyses considerations

8.3.1 Outlier detection in the UKWCS

Distribution of each variable of interest was checked for any possible errors such as typing error. Continuous variables were checked by box plots and histograms to look for any potential outliers while frequency was checked for the categorical variables. Scatter plots were also useful to detect potential outliers by evaluating the consistency of the continuous variables for dietary fibre and other variables. After data checking, characteristics of cases and non-cases were examined separately, aiming to gain an understanding of the studied population. It has been advised by Willett (2012) that exclusion of extreme energy intake values will counteract for the under and over-reporters.

Detection of outliers is an important issue aiming to avoid bias in estimation of dietary exposure and to have estimates reflecting the main body of data. Outliers can occur due to measurement error or inaccurate data entry or can be due to natural variation within the studied population (Willett, 1998). Evidence shows that potential over reporting and under reporting lies with total energy intakes of more than 3500 kcal/day and below 500 kcal/day for women (Willett, 1998). The purpose of limiting studied participants on the basis of their energy intakes was to minimize the measurement error that can affect the relationship between the disease outcome and nutrient intake. This is because energy intake is the best measure that has narrow physiological predictable range (Willett, 1998).

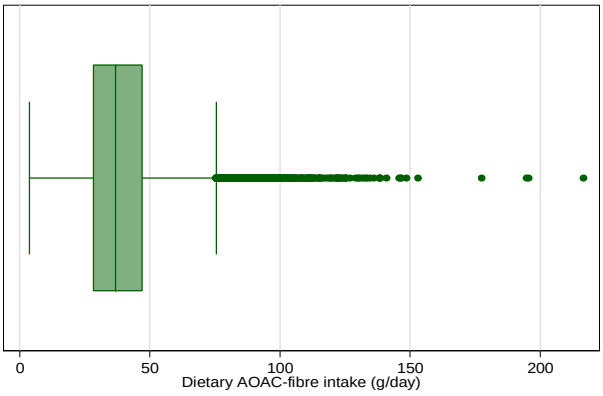
Statistical methods were used for outliers' detection. Univariate statistical methods basically examine data distribution through minimum and maximum values and scatter plots may be used for normally distributed data. Box plots were also used as another exploratory approach as they may help in identifying potential outliers. The cut-off point for maximum amount of dietary AOAC-fibre was calculated as follow. The inter-quartile range (IQR) equals the 75^{th} quartile (Q3) minus the 25^{th} quartile (Q1) and assumes that potential outliers are at any data point below the lower threshold = Q1 – 1.5×IQR or above the upper threshold = Q3 + 1.5×IQR.

Outliers were identified using the following steps:
25% percentile = 28.3g/day
75% percentile = 47.4g/day
Inter-quartile range = 19.1g/day
1.5 X IQR = 28.6g/day
Upper threshold = 47.4 + 28.6 = 76.0g/day

When women with AOAC-fibre consumption above 76g/day were assumed to be potential outliers, this accounted for 2.5% (n=303) of the total cohort. The box plot of dietary AOAC-fibre values as shown in Figure 8.2 illustrates dietary AOAC-fibre values present above the upper threshold considered as potential outliers. Potential outliers were then used in the subgroup statistical analyses as a way to examine the relationship between dietary fibre intake and risk of diabetes among women with AOAC-fibre intake below 76g/day as well as examine the relationship among all women (including women with AOAC-fibre intake above 76g/day) and compared the results. The findings confirm that conclusion is not sensitive when AOAC-fibre intake above 76 included in the analysis. This indicates that including women who consumed very high AOAC-fibre intake in the analysis have no effect on the studied relationship. From this, it have been recommended to not exclude participants (Willett, 2012).

Another reason for retaining women with AOAC-fibre intake above 76g/day was that this resulted in elimination of three women with incident diabetes which can greatly affect the number of cases in the T2DM group and therefore affect the relationship between dietary fibre intake and risk of T2DM. It have been reported previously that outliers may result from natural variation (Laurikkala *et al.*, 2000). Therefore energy restriction (500 – 5000 kcal/day) was the only exclusion criteria applied.

Figure 8.2 Distribution of dietary AOAC-fibre intake among all participants



8.3.2 Model development

In the current study, logistic regression method was carried out to investigate the effect of baseline dietary fibre intake on risk of incident T2DM. So, logistic models with T2DM incidence as the dependent variable and dietary fibre intake as a

primary predictor variable were constructed. Covariates were also included in the logistic model. To address the question of whether the observed association between specific nutrient intake and a disease outcome is reliable, adjustment for confounders is an important step to obtain valid estimates of the exposure-outcome relationship (Kleinbaum and Klein, 2010). A confounder is a variable that is associated with the exposure and is also an independent risk factor of the disease outcome (Willett, 1998). Several approaches have been used in nutritional epidemiological research to identify potential confounders (Willett, 1998).

Confounders identified from previous robust evidence, an approach which is also referred to as *a priori confounders,* is suggested as being an important approach (Willett, 1998). Previous studies provide information on the independent variables being a predictor of disease occurrence and also linked to the exposure (Kirkwood and Sterne, 2003). Further potential confounders may be a possible concern because argument regarding their link to exposure and outcome was also considered. Effect modifiers are variables that maybe assumed to modify the estimate effect size or direction of the association and were also evaluated in this chapter (Kirkwood and Sterne, 2003).

Another approach helping to detect potential confounders which can be used with the previous approach is the univariate and multivariate linear regression analysis. This aims to examine the relationship of the outcome (T2DM) with each independent variable (exposure of interest) in the studied population (Kirkwood and Sterne, 2003). The purpose is to explore whether the independent variable was significantly associated with T2DM and on the other hand whether this variable shows changes in the estimated effect size of dietary fibre intake.

Initially, an unadjusted logistic model was used out to explore the relationship between dietary fibre intake and risk of T2DM in the UKWCS. This is useful as later on covariates may be added to the model and results can reflect the influence of the adjusted confounders on the relationship in term of direction and effect size. Confounders included in the logistic model reduce the contribution of variation to other than dietary exposure of interest which is the aim of the study. The number of covariates included in the model was considered carefully because if high number of covariate included then it may reduce variation in the dietary exposure of interest and lead to estimation bias which, can be seen by wide confidence intervals (Willett,

1998). On the other hand, if some potential confounders were not included in the model then estimation bias from residual confounders is more likely to occur.

Age may be considered as an example to illustrate the selection of potential confounders by statistical approaches and background scientific knowledge. Evidence has shown that age is a strong risk factor for T2DM, and evidence has shown an association between dietary fibre intake and age. A linear regression model was carried out with T2DM as the outcome measure and age as an independent variable to determine whether age is associated with T2DM independent of dietary fibre intake. On the other hand, linear regression analyses were also carried out to examine the relationship between age (independent variable) and dietary fibre intake. Age was considered as a confounder, as analyses showed a significant association with T2DM independent to dietary fibre and significant association between age and dietary fibre intake in separate regression analysis.

Overall, the examined dietary variable (age) was independently associated with dietary fibre intake and with T2DM, this variable was considered as cofounder in the model with the respect to previous evidence. However, an independent variable should not considered as a confounder if it is assumed to be on the causal pathway (Kirkwood and Sterne, 2003) because adjustment of a suspected intermediate variable will result in estimate bias.

8.3.2.1 Rationale of modelling strategy

This strategy aimed to provide information on how variables were chosen for the logistic regression models. Steps were followed to select the most appropriate variables which aimed to obtain a valid effect estimate and at the same time avoid over adjustment. All variables of interest were derived from the baseline FFQ.

8.3.3Identification of potential confounders

As explained before, one must control for confounders when building a regression model which examines the association between the exposure (dietary fibre intake) and the outcome of interest (risk of T2DM) (Willett, 1998). *A Priori* confounders and regression analyses approaches were considered for building the current models with other important considerations aiming to have best model with reliable estimated effect.

Three types of variables were considered for regression models, *a priori* confounders, evidence based confounders, and effect modifiers. *A priori* confounders are variables believed to confound the relationship between dietary

fibre intake and risk of diabetes such as ethnicity (Hopping *et al.*, 2010, Stevens *et al.*, 2002) and gender (Stevens *et al.*, 2002, Schulze *et al.*, 2007, Montonen *et al.*, 2003). The population in the current study were women and 98% were white therefore those variables were not considered as covariates in the analyses. Only BMI was considered to be effect modifier, since it was to modify the estimated effect size of dietary fibre intake with the risk of diabetes in previous study (Colditz *et al.*, 1992, Krishnan *et al.*, 2007). Variables of interest explored in the present study were divided into dietary and non-dietary measures.

8.3.3.1 Non dietary measures

8.3.3.1.1 Age

Age was considered as a confounder from previous cohort studies that have examined the association between risk of T2DM and dietary fibre intake. From literature review in chapter 2, evidence suggests a positive association between age and risk of T2DM. On the other hand, prospective evidence in chapter 4 showed participants with high dietary fibre intake were more likely to be older. In addition, the regression analysis of the UKWCS data showed age (OR=1.06; 95%CI 1.04, 1.08) was positively associated with the risk of T2DM (see Table 8.2). Also, for every 10 year increase in age, dietary fibre increased by 0.5g/day (95%CI: 0.2, 0.8; $p<0.01$). Therefore age was considered as a covariate in all adjusted models.

8.2 Univariate and age-adjusted odds ratios (95% CIs) of potential risk factors for T2DM in the UKWCS

Characteristics		Unadjusted OR (95%CI)	P	Age adjusted OR (95%CI)	P
Age (per year)		1.06(1.04 -1.08)	<0.01	-	-
Waist circumference (cm)		1.07(1.05 -1.08)	<0.01	1.07(1.05 -1.08)	<0.01
Weight change from age of 20 years (%)		1.06(1.05 -1.08)	<0.01	1.06(1.04 -1.07)	<0.01
Weight change from age of 20 years (Kg)		1.07(1.06 -1.08)	<0.01	1.07(1.05 -1.08)	<0.01
Baseline BMI		1.12(1.09 -1.15)	<0.01	1.11(1.07 -1.15)	<0.01
BMI categories	Normal weight (≥18 - <25)	1	<0.01	1	<0.01
	Overweight (≥25 - <30)	3.83(2.31 -6.36)		3.37(2.01 -5.6)	
	Obese (≥30)	13.4(8.37 -21.32)		11.9(7.3 -19.1)	
Socio-economic status	Professional/managerial	1	0.53	1	0.5
	Intermediate	0.95(0.60 -1.49)		0.87(0.55 -1.37)	
	Routine/manual	1.40(0.73- 2.65)		1.32(0.70 -2.52)	
Education level	No education	1	<0.01	1	0.03
	O-level	0.58 (0.34 -0.97)		0.83(0.48 -1.41)	
	A-level	0.34 (0.19 -0.63)		0.46(0.25 -0.85)	
	Degree	0.36 (0.20 -0.64)		0.53(0.29 -0.96)	
Marital status	Married	1	0.01	1	0.52
	Divorced	0.96(0.46 -2.00)		0.93(0.45 -1.94)	
	Widowed	2.54(1.40 -4.60)		1.39(0.74 -2.61)	
	Single	1.46(0.79 -2.70)		1.42(0.77 -2.63)	
Common dietary pattern	Meat group	1	<0.01	1	0.1
	Fish group	0.46(0.23 -0.92)		0.55(0.27 -1.11)	
	Vegetarian group	0.49(0.26 -0.90)		0.62(0.33 -1.15)	
Self-reported vegetarian or vegan		0.51(0.32 -0.81)	<0.01	0.63(0.39 -1.00)	0.05
Smoking		1.52(0.86 -2.67)	0.14	1.70(0.94 -3.06)	0.07
Physical activity (MET)		0.98(0.96 -1.00)	0.16	0.98(0.96 -1.00)	0.1
Alcohol intake (g/d)		0.98(0.95 -1.00)	0.08	0.98(0.96 -1.01)	0.29
Energy intake		0.99(0.99 -1.00)	0.8	0.99(0.99 -1.00)	0.6
Magnesium intake (mg/day)		0.99(0.99 -1.00)	0.07	0.99(0.99 -1.00)	0.052
Dietary supplement use		0.63(0.43 -0.93)	0.02	0.63(0.43 -0.93)	0.02
Family history of diabetes		4.15(2.65 -6.52)	<0.01	4.50(2.86 -7.01)	<0.01

8.3.3.1.2 Socioeconomic status and educational level

Socioeconomic status was investigated in terms of educational level, income and occupational status in 6147 participants involved in the Alameda Country Study (Maty *et al.*, 2005). This study showed that incidence of diabetes was inversely associated with socioeconomic status based on income, occupation and education among the US studied population in unadjusted models. However, the hazard ratio was significantly higher with a lower education level but not with income or occupation in a demographic (age, gender, ethnicity) adjusted model (Maty *et al.*, 2005). The non-significant association between occupation and risk of diabetes was explained by the possibility of having high proportion of younger age in the studied population who may not achieve the occupational abilities. Another study including national health surveys from eight European countries reported that diabetes was more prevalent in the lower education group (OR = 1.60 (95%CI: 1.45,1.80) among

men and women in comparison to the higher education group. The odds ratio for diabetes in women was 2.19 (95%CI: 1.82, 2.63) and in men was 1.30 (95%CI: 1.11, 1.51), and therefore the influence of education level may be stronger in women than men. On the other hand, the odds ratio for DM (OR= 1.26, 95%CI: 0.98–1.62) in the UK was not statistically significant comparing lowest vs. highest education levels (Dalstra *et al.*, 2005). An earlier UK observational study concluded that the prevalence of diabetes was significantly inversely related to socioeconomic status which was based on a deprivation score (including employment and income parameters) (Connolly *et al.*, 2000).

In the current study, socioeconomic status variable was based on the employment status as in the National Statistics Socio-economic Classification. The Wald test was used to assess the main overall effect of socioeconomic status on the risk of diabetes. This showed that there was no significant main effect of socioeconomic status on the risk of diabetes ($p=0.5$) even after age adjustment. However, the educational level variable did show a significant effect on the risk of diabetes in the age adjusted model ($p<0.03$). Table 8.3 shows that women with a degree have an odds ratio of 0.36 (95%CI: 0.2, 0.64) in comparison to non-educated women. On the other hand, current analyses showed a significant overall effect of educational level and socioeconomic status on dietary fibre intake among the UKWCS ($p<0.01$). Women with a degree consume 1.4g of dietary fibre daily more than non-educated women ($p<0.01$). A significant moderate correlation was seen between socioeconomic status and educational level with a Spearman rank correlation of 0.41 ($p<0.01$). If its assumed that both educational level and socioeconomic status represent similar information then an unreliable effect estimate is suspected if both were included in the model (this may also be considered as potential over-adjustment) (Kleinbaum and Klein, 2010).

Another important point to consider when selecting covariates is the amount of missing data. A high number of missing values in the variable will more likely result in bias because of low statistical power therefore, variables with a large number of missing values was eliminated. In this case, the educational level variable had 675 missing values more than socioeconomic status where 5 of them were from diabetic cases. Finally, socioeconomic status was modelled.

8.3.3.1.3 Lifestyle variables

Prospective evidence had found a significant relationship between smoking and risk of T2DM (see chapter 2). A temporal relationship has been reported as smoking precedes diabetes incidence. Also dose-response association have been reported as the relationship was stronger for heavy smokers in comparison to light smokers which suggests an element of causation. However, this is not yet well established (Willi *et al.*, 2007). On the other hand, low dietary fibre intake is often associated with other unhealthy behaviours that favour diabetes, such as, smoking, lack of physical activity and high alcohol consumption as shown in chapter 4.

From Table 8.3, self-reported smoking was found to have no significant association with risk of T2DM after adjustment for age (OR= 1.70; 95%CI: 0.94, 3.06) in comparison to non-smokers. The insignificance results probably related to small number of both cases and smokers in the UKWCS. On the other hand, the age-adjusted analysis showed that women who smoked at baseline consumed 5.5g/day less dietary fibre than non-smokers (95%CI: -6.5,-4.5; $p<0.01$). Therefore, smoking was modelled, as evidence showed an association between smoking and risk of diabetes as well as with dietary fibre.

The metabolic equivalents tasks (METs) variable, which represents physical activity level, was analysed in relation to diabetes and dietary fibre intake as it is one of the important risk factors for diabetes that needs to be considered. Physical activity level was included in the model as convincing evidence is available that low levels of physical activity increase the risk of T2DM (from chapter 2). Cohort studies from chapter 4, showed that participants with high fibre intake were more likely to engage in physical activity in comparison to participants with low dietary fibre intake, which supports considering physical activity as a covariate in the model.

Linear regression analyses in the current study showed that physical activity (per 1 unit increase in MET) was not statistically significant in reducing the risk of T2DM (OR=0.98; 95%CI: 0.96, 1.00) however a significant positive association was found between dietary fibre intake and physical activity ($p<0.01$) even with age adjustment. As physical activity increased, dietary fibre intake also increased (p-trend<0.01). Smoking status and physical activity variables were included in the logistic models.

8.3.3.1.4 Other variables

Previous evidence consistently supports an association between family history of diabetes (especially amongst first degree relatives) and the risk of developing diabetes, because of underlying genetic susceptibility. Indeed it has been suggested as a public health tool for screening (Harrison *et al.*, 2003). Few studies have looked at the relationship between family history of diabetes and dietary behaviours. A cross-sectional study from a DIRECT project showed that participants free from diabetes with a positive family history of diabetes were more likely to consume 5 or more servings of fruits and vegetables per day than those with no family history (Baptiste-Roberts *et al.*, 2007). This reflects high dietary fibre intake among those with a positive family history of diabetes may be due to a high level of awareness relating to healthy behaviours in comparison to those with a negative family history.

On the other hand, a previous study reported that participants with a family history of diabetes were more likely to have health protective behaviours (such as weight reduction, diet, engaging in exercise) in addition to which they were more likely to report being screened for diabetes (Forsyth and Goetsch, 1997).

Recent evidence reported a differential effect of dietary fibre intake on the risk of T2DM based on genetic background, as dietary fibre intake was positively associated with risk of diabetes among participants who were a specific type of gene (TCF7L2 rs7903146 gene) carrier while an inverse association was found between dietary fibre intake and risk of T2DM among participants who were CC genotype carriers but not among CT or TT genotype carriers (Hindy *et al.*, 2012). It may be assumed that a positive family history is a proxy for increased genetic susceptibility, but also could be shared environment factors such as diet.

The current analysis in the UKWCS showed that women with a positive family history of diabetes were at four times higher risk of diabetes than women with no family history (OR=4.15; 95%CI: 2.65, 6.52) (see Table 8.3). The relationship was clearer after age adjustment; the risk of diabetes was 4.5 times higher among women with a positive family history than other women with no family history of diabetes (95%CI: 2.86, 7.01; p<0.0). Cohort women with a positive family history of diabetes consumed less dietary AOAC-fibre (~1g/day) in comparison with women with a negative family history which was marginally significant (p=0.06). Only 22% of women who had diabetes reported a positive family history of it; 76% of diabetic cases reported a negative family history. In

another words, a small number of diabetic women who reported positive family history in comparison to diabetic women with negative family history of diabetes (25 vs. 89) may affect the estimated association between dietary fibre intake and family history of diabetes. Finally, family history of diabetes variable was included in model 4.

8.3.3.2 Dietary measures

8.3.3.2.1 Energy intake

The total energy intake variable was expressed as kcal per day, and was derived from the baseline FFQ. The aim for the current study is to investigate the importance of nutrient intake in relation to the risk of the disease therefore energy adjustment was usually considered. Variation in the total energy intake between people creates variation in the intakes of nutrients that are unrelated to dietary composition for the reason that positive associations were found between most nutrients and the total energy intake. In another words, people may have high fibre intake just because they eat more. In addition to this, if assumed that energy intake indirectly affects health outcome (Willett, 2012) then in this case, the association between nutrient of interest (dietary fibre) and risk of diabetes may be confounded by energy intake. Therefore, energy intake was considered as a potential confounder.

It has been assumed that dietary fibre intake has a biological influence on T2DM risk therefore avoiding high variation in nutrient intake as a result of variation in energy intake. Thus energy adjustment would eliminate the influence of variation in energy related factors. For example, as an individual with a large body size is assumed to consume a high energy intake, thus a certain amount of dietary fibre intake will have less effect than on an individual with a small body size. In another situation, assumption regards an individual with a small body size who is physically active most probably consumes high total energy. Thus it was suggested that adjustment of physical activity, BMI and energy intake would be appropriate in this case (Willett, 1998). Different statistical methods for energy adjustment were reported by Willett (1998). Nutrient density method and absolute dietary fibre intake adjusted for total energy intake were concluded for the logistic regression analyses in this study.

Dietary fibre density as AOAC-fibre density and as NSP density were calculated by dividing total dietary fibre intake (total AOAC-fibre or NSP dietary

intake) by total energy intake. The nutrient density method is also used in the current study as another way to adjust for energy intake.

8.3.3.2.2 Anthropometric variables

From chapter 2, obesity is one of the main independent risk factors for the development of T2DM. In terms of BMI as a measurement of obesity, a combined risk of T2DM was 7.19 (95%CI: 5.74, 9.00) for obese participants and overweight participants had a relative risk of 2.99 (95%CI: 2.42, 3.72) compared to participants with a healthy BMI (Abdullah *et al.*, 2010). Additionally, other evidence found weight gain from early adulthood increased risk of T2DM in compared to weight gain as later life as demonstrated in chapter 2.

Evidence from chapter 4 suggested that high fibre consumers were likely to have lower BMI. An earlier cross-sectional study found that fibre intake was inversely associated with body mass index after adjusting for sex, age, education level and income ($p<0.01$) (Alfieri *et al.*, 1995). When this relationship was examined in the Seven Countries Study that included men aged 40 ± 59, it was found that dietary fibre was inversely associated with skin fold thickness and BMI (Kromhout *et al.*, 2001). Over 12 years of follow up, a prospective study including 74,000 US female nurses, aged 38–63 years, found those with the highest increase in the dietary fibre intake gained less weight (1.5 kg less) than those with the smallest increase (the median change in fibre intake over period of time= 8.90 vs. -3.40g/day; $p<0.01$). Increase in dietary fibre intake over a period of time was significantly associated with less weight gain and risk of weight gain was lower among those with increased dietary fibre intake in relation to least change of dietary fibre intake (OR = 0.66 95%CI: 0.58, 0.74; $p<0.01$) (Liu *et al.*, 2003). A large cohort study of European men and women reported total dietary fibre was inversely associated with weight and waist circumference (Du *et al.*, 2010). Overall, BMI, weight changes and waist circumference (WC) were preliminarily considered as *a priori* confounders.

In the UKWCS dataset, BMI, WC and weight change variables were derived from the following questions. "*Approximately how much do you weigh at present? In stones, pounds or kg or don't know*", "*What is your present height? In ft, inches or centimetres or don't know*", "*What is your present waist size? In inches or centimetres or don't know*" and "*Approximately how much did you weigh when you were 20 years old? In stones, pounds or kg or don't know*". BMI was calculated by the formula:

BMI (kg/m^2) = weight /height^2

The weight change variable was expressed in kg and as the percentage which was generated from the difference between self-reported weight at the age of 20 years and weight when completing the baseline FFQ from the above questions.

Table 8.3 demonstrates the odds ratios of T2DM with BMI, WC and weight changes (%) by logistic univariate regression analyses. Waist circumference (OR= 1.07, 95%CI, 1.05, 1.08), BMI (OR=1.06; 95%CI 1.05, 1.08) and weight change (OR=1.07; 95%CI 1.06, 1.08) were all positively associated with risk of diabetes. All variables showed significant results even after age adjustment (OR=1.07; 95%CI: 1.05, 1.09), (OR=1.06; 95%CI: 1.04, 1.07) and (OR=1.11; 95%CI: 1.07, 1.14) respectively. A review by Slavin (2005) reported inverse associations between dietary fibre intake and obesity in the interventional studies and high fibre intake helped to prevent obesity.

Obesity variables were assessed for high correlation using Pearson's correlation as shown in Table 8.3 as all reflect body fatness. Results showed a moderate to high positive significant correlation between BMI and weight changes (r = 0.57; 95%CI: 0.56, 0.58), BMI and waist circumference (WC) (r = 0.63; 95% CI: 0.62 to 0.64), and WC with weight changes (r = 0.48; 95%CI: 0.47, 0.49).

8.3 Pearson's correlation between obesity variables in the UKWCS

	BMI (Kg/m^2)	Weight changes (%)	WC (cm)
BMI (Kg/m^2)	1.0	-	-
Weight change (%)	0.57 <0.01	1.0	-
WC (cm)	0.63 <0.01	0.48 <0.01	1.0

The extent of missing data was also considered before including the variable in the model. The WC variable was excluded because of the high number of missing values in diabetic cases (missing values n=34) and in non-diabetic cases (missing values n= 2085). BMI and weight changes were left for further evaluation.

From a recent cohort study of 20,187 participants, it was reported that weight gain (of more than 3 units of BMI) during adulthood even with acceptable BMI

(<21) at the baseline increased the risk of diabetes by almost 8-fold RR=7.68; 95%CI: 4.72, 12.50) in comparison with those who gained 0.5 units (Oguma *et al.*, 2005). This suggests that weight change is a strong predictor of the risk of diabetes.

In the UKWCS dataset, women with diabetes at phase II reported an adulthood mean weight gain of 17 kg (SD =15.1) while women free of diabetes had an average weight gain of 7 kg from age of 20 until the phase II questionnaire. Decision was made regards weight changes variable to be included in the model.

8.3.3.2.3 Alcohol intake variable

Evidence from chapter 2 found a U-shaped relationship between alcohol intake and the risk of T2DM and greater intake found to have no significant increase in risk of T2DM in comparison to non-consumers. This relationship was also significant in studies that included women.

Prospective evidence from chapter 4 reported an inverse association between intake of dietary fibre and alcohol intake. In the UKWCS data, regression analysis showed that risk of diabetes was not associated with alcohol intake even with age adjustment (p=0.29) (see Table 8.3). Overall, alcohol intake is still a potential confounder as literature evidence showed the association between T2DM and alcohol intake and suggested an association of dietary fibre intake with alcohol intake.

8.3.3.2.4 Fat intake variable

US prospective evidence in chapter 2 showed total fat intake was associated with increased risk of diabetes; however, after BMI adjustment, the association was no longer statistically significant. This may suggest that fat intake acts through obesity in the causal pathway towards the development of T2DM.

Regression analysis of the UKWCS data showed that the risk of diabetes was not significantly associated with fat intake (OR=0.98, 95%CI: 0.95, 1.01; p=0.3) even after age adjustment. On the other hand, for every 1% increase in energy from fat intake, there was a 0.6g decrease in dietary fibre intake (95%CI: -0.67, - 0.62; p<0.01). In addition to adjustment of weight changes, the total energy intake was also considered in the model which is highly correlated with fat intake (r= 0.9, p<0.01) thus fat intake was not modelled.

8.3.3.2.5 Other dietary variables

The two meta-analyses from chapter 2 drew similar conclusions, that magnesium intake was inversely associated with T2DM.

In the current study, a strong positive correlation between dietary fibre intake and dietary magnesium intake was seen (r=0.90; p<0.01). No significant associations were seen between magnesium intake and risk of T2DM among cohort women. However, the magnesium variable was not considered for further adjustment. This is because magnesium intake is highly correlated with dietary fibre intake, thus to show a separate effect is hard as it will be subjected to estimation bias. High fibre foods (cereal, fruit, vegetables and legumes) are primary source of magnesium then, adjustment of magnesium may result into adjustment of fibre.

Characteristics of dietary supplementation users were demonstrated in previous studies and reported that participants who uses dietary supplementations were more likely to be female, non-smokers, highly educated and light drinkers (De Jong *et al.*, 2003). This indicates that supplement use is in the line with other healthy behaviours. Data from the US Department of Agriculture and Diet and Health knowledge survey (Sebastian *et al.*, 2007) reported that smoking and vegetarian status were significant predictors of supplement use among women. In an addition, a healthier diet was more likely to be present among supplement users than non-users and this finding was also reported previously in the UKWCS (Kirk *et al.*, 1999).

Current analysis showed a significant T2DM risk reduction with the use of dietary supplements (OR=0.63; 95%CI: 0.43, 0.93) even with age adjustment. There is limited evidence that evaluates the relationship between the use of dietary supplement and risk of diabetes (Xu *et al.*, 2011) and because of the great variation in the amount and type of supplement use, evidence may not be comparable. Dietary supplements use was eliminated from the model.

Other variables of interest are the self-reported vegetarian or vegan variable and the common dietary pattern variable. The self-reported vegetarian or vegan variable was derived from the questions '*Would you describe yourself as a vegetarian?' and 'Would you describe yourself as vegan?*'. The common dietary pattern variable was defined in previous study (Cade *et al.*, 2004a) and classified women into four main dietary groups (meat, oily fish, fish and vegetarian). In the current study almost one third of women defined themselves as vegetarian or vegan, while in the dietary pattern categories, 18% of women were in the vegetarian group. From the logistic regression, women who reported being vegetarian or vegan had a significant 49% risk reduction of T2DM (OR=0.51; 95%CI: 0.32, 0.81) and marginally significant risk reduction was seen with age adjustment in comparison to

non-vegetarians (OR=0.63; 95%CI: 0.39, 1.00). Common dietary patterns were a significant predictor of the risk of T2DM, however this relationship was not significant after age adjustment which may be explained by the younger mean age for vegetarians (49 years) than the meat-eating group (53 years). As expected, dietary fibre intake was significantly higher among vegetarian than non-vegetarians. Vegetarian status was not considered in the model because of two main reasons. Adjustment of vegetarian status may result in adjustment of dietary exposure of interest (dietary fibre) as women who reported being vegetarian consume higher amount of dietary fibre in comparison with non-vegetarians. It is interesting to examine whether the association between dietary fibre intake and risk of T2DM among women differ among different dietary pattern groups? However, number of participants in each group was small which limits any further analysis (non-cases/cases: non-vegetarian group 8009/91 and vegetarians 3974/23)

Overall, dietary supplement use, vegetarian status and magnesium intake were not included in the logistic regression models. A concern regarding over adjustment that may leave small variation in the exposure (dietary fibre) while other covariates may account for almost all variation, which results in a wide uninformative 95% confident interval (Willett, 1998) was also considered.

8.3.4Investigating linearity in the regression model

Fang with his colleagues (2009) encourage running linearity assessment for continuous variables before running any logistic regression for binary outcomes. In linear regression analysis, the null hypothesis assumes the relationship is linear. In case if the relationship is not linear then categorization of the predictors is recommended (Peacock and Kerry, 2007). Another study illustrated one of the disadvantages of categorization is the presence of residual confounders after categorizing the predictor variables which may result in incomplete control of covariates in the model and may lead to biased estimation (Becher, 1992). Linearity was assessed for continuous variables such as age and alcohol intake before they were included in the adjusted logistic regression analysis. This aimed to avoid an incorrect conclusion because there is in fact a non-linear association (Kirkwood and Sterne, 2003). After categorization of non-linear relationship, a likelihood ratio test was performed to compare between the log likelihood for the model that included continuous variables such as age and the likelihood ratio of categorised variable such as age quintiles. Significant results reject the null hypothesis of having a linear

relationship between exposure and outcome and indicate non-linear relationship. Age and risk of diabetes association appeared linear with non-significant likelihood ratio test (p=1.0) therefore age was included in the model as a continuous variable. The likelihood ratio test showed a significant result for alcohol intake (<0.01), indicating a nonlinear relationship with the risk of diabetes. Therefore alcohol intake was split into tertiles and included in the final model as a categorical variable.

To avoid the influence of over adjustment on the estimated effect, the number of covariates added in the model was estimated roughly by the number of observations in the dataset as suggested previously (Kirkwood and Sterne, 2003). For example, if ten variables were modelled, then at least 100 observations should be present in each variable considering each group as a separate variable. Thus only alcohol intake was split into three equal groups as evidence pointed towards a nonlinear relationship with the risk of T2DM.

There is no specific rule on the maximum number of covariates that can be included in the model; however this depends largely on the studied population as more covariates can be included in a large study than a small one (Kirkwood and Sterne, 2003). The number of participants in each group was considered in the current analyses were considering covariate selection. Thus model 1 was adjusted for age and model 2 was adjusted for age, weight change, and total energy intake. Model 3 was adjusted for model 2 covariates plus: smoking, physical activity (METs), alcohol intake (g/day), and socioeconomic status (professional, intermediate and routine classes). Model 4 was adjusted for all the above covariates plus family history of diabetes. Table 8.4 summarises reasons for inclusion or exclusion of variables to fit the best model to examine the effect of dietary fibre on the risk of T2DM.

In summary, regression analyses were carried out to examine the association of each potential confounder with the risk of T2DM and with dietary fibre intake. Variables that showed a significant association with dietary fibre intake and risk of diabetes in the UKWCS database were considered for adjustment. Age, weight changes and socioeconomic status were considered for adjustment. Other variables such as smoking, physical activity, total energy intake and alcohol intake that were not found to have a significant relation with the dietary fibre and risk of diabetes in the UKWCS data but were well-established via prior knowledge as discussed above were also considered as covariates in the fully adjusted model.

Table 8.4 Evidence and justification for selected covariates in the model development

Variable	Prior confounder		Association in UKWCS data[1]		Other reasons[2]	In model
	T2DM[3]	DF[4]	T2DM[3]	DF[4]		
age	Wild (2004)	Schulze(2007) Schulze(2004) Steven(2002) Meyer(2000) Krishnan(2007)	(+)	(+)	No	Yes
socio-economic status	Connolly (2000)		(-)	(+)	Highly correlated with education	Yes
Educational level	Dastra (2005) Maty (2005)	Schulze(2007) Steven(2002) Hopping(2010) Meyer(2000)	(+)	(+)	High missing values	No
Smoking	Tonstad (2009) Willi (2007)	Steven(2002) Montonen(2003) Krishnan(2007) Meyer(2000) Schulze(2004) Schulze(2007)	(-)	(+)	No	Yes
Physical activity	Jeon (2007)	Meyer(2002) Steven(2002) Schulze(2004) Schulze(2007)	(-)	(-)	No	Yes
BMI	Abdullah (2010)	Salvin(2004) Krombout(2001)	(+)	(+)	High correlation	No
Weight change[5]	Oguma (2005) Schienkiewtz (2006) Vazquez (2007)	Schienkiewitz(2006) Koh-Banerjeel(2004)	(+)	(+)	Strong predictor	Yes
WC	Vazquez (2007) Wannamettee (2010)	Krishnan(2007) Schulze(2004) Schulze(2007)	(+)	(+)	Missing values[6]	No
Alcohol intake	Baliumas (2009)	Krishnan(2007) Schulze(2004) Schulze(2007)	(-)	(+)		Yes
Family history diabetes	Harrison (2003)	Schulze(2004) Krishnan(2007)	(+)	(-)		Yes

[1] Regression analyses carried out in the UKWCS (p value = (+) indicate significant and (–)indicate not significant), [2]Reasons for exclusion such as missing values in variable and suspected Multicollinearity between variables, [3]type 2 diabetes mellitus, [4]Dietary fibre, [5]weight change is the difference between 20 years of age and completion of baseline FFQ. [6]Missing data: DM cases = 34 and non-cases= 2,085.

8.3.5Current study in the systematic review

The final result from the multivariate model in the current study was included in the meta-analysis which was carried out in chapter 4. The odds ratio and 95%CI of comparing the lowest versus highest dietary AOAC-fibre intake was included.

8.3.6Statistical analysis

Characteristics of women are presented in Table 8.5 as median or geometric means (interquartile ranges) or means (SD) for continuous variables and frequencies (percentages) for categorical variables. Normality tests were carried out before using parametric tests (t-test) and non-parametric tests (Mann-Whitney test) aiming to examine significant differences between cases and non-cases as appropriate. Significant differences between categorical variables were examined using chi squared tests and the analysis of covariance test. Women's characteristics in Table 8.4 were examined for potential trends across dietary AOAC-fibre quintiles using one-way analysis of variance and chi-squared tests where appropriate.

Logistic regression analysis was used to estimate the odds ratios (OR) and 95% confidence intervals of incident T2DM associated with the dietary fibre intake in the UKWCS. Logistic models were created for each of the four primary exposures NSP (g/d), AOAC-fibre (g/d), AOAC-fibre density (g/1000kcal/day) and NSP-density (g/1000kcal/day) intakes.

Dietary fibre intake quintiles were built into logistic regression analysis models which were adjusted for a range of potential confounders. The lowest quintile is presented as a reference.

All potential confounders were derived from the baseline questionnaire. These were: age (years), weight change (%), METs (kcal/kg/hr), alcohol intake (categorical variable (tertiles); g/day), total energy intake, self-reported smoking status (yes, no), socioeconomic status (professional and managerial social class, intermediate social class, routine and manual social class based on UK national socioeconomic status classification) and family history of diabetes (yes, no).

The relationship between dietary fibre intake as a continuous variable with risk of T2DM was examined after adjustment of potential confounders aiming to avoid the potential of losing information through categorization (Kirkwood and Sterne, 2003) and to ensure the greatest statistical power provided by the continuous variable (Willett, 1998). However, because categorization has advantages in visualizing the variation in the number of cases and non-cases by amount of dietary intake and can limit the influence of possible outlying data, dietary fibre intakes were divided into five quintiles to facilitate comparison between the risk of T2DM in the highest dietary fibre quintile and lowest dietary fibre quintile. The first

quintile represents women with lowest dietary AOAC-fibre intake and the fifth quintile represents women with highest dietary AOAC-fibre intake.

Odds ratios estimated by logistic regression analyses were used to the estimate risk of T2DM relating to insoluble and soluble dietary fibre intakes (g/d). The intakes of dietary fibre sources which included cereal fibre, vegetable fibre, fruit fibre, legumes fibre and fibre from nuts and seeds were also examined in relation to the risk of T2DM. All statistical analyses were carried out using STATA version 12 (Corp-Stata, 2010). All statistically significant results were based on 2 sided tests.

8.3.7Sensitivity analyses

Sensitivity analyses were performed on the UKWCS data to explore the relationship between dietary fibre intake and the risk of T2DM within subgroups in the categorical variables. Subgroup analysis was carried out to estimate the effect of dietary fibre intake on the risk of diabetes among overweight and obese women. BMI as an effect modifier was investigated previously (Colditz *et al.*, 1992, Krishnan *et al.*, 2007) because a suspected modification effect on the relationship between dietary fibre on the risk of diabetes was assumed. In the current study, because of small number of diabetic cases, BMI was divided into two groups ($BMI<25kg/m^2$ and $BMI\geq25kg/m^2$). This aimed to evaluate whether dietary fibre intake effect on risk of T2DM might be different among women with normal BMI compared to those who are overweight and obese. Therefore the assumption that BMI may modify the effect of dietary fibre intake on the risk of diabetes was made. Magnitude of association between risk of T2DM and obesity was reported in meta-analysis of prospective studies, the risk of T2DM was 7 times higher (95% CI: 6.47, 8.28) among obese participants and almost 3 times higher (95% CI: 2.57, 3.32) among overweight participants in comparison to participants with normal BMI independent to age, family history of diabetes and physical activity (Abdullah *et al.*, 2010).

8.4 Results:

8.4.1Characteristics of the studied population

Out of 12,096 UKWCS participants with a mean age of 52.4 (9.1) years, 114 incident cases of self-reported T2DM were identified. Table 8.5 provides detailed description of cases and non-cases. There were non-significant differences between cases and non-cases for socio-demographic, lifestyle and dietary characteristics except for age, BMI, weight change, alcohol intake, vegetarian status, educational

level and family history of diabetes mellitus. In Table 8.5, age was significantly higher among diabetic cases than non-diabetic cases (57.4 years vs. 50.4, respectively). Diabetic cases had significantly higher BMI (28.5 vs. 23.6, respectively) and weight change (17.2kg vs. 6.3kg, respectively) than non-diabetic cases. A higher percentage of diabetic women were found in the non-vegetarians group (80% vs. 66%) and a smaller number of diabetic women had a degree in comparison to number of women in the non-cases group (23% vs. 32%).

Diabetic participants were less likely to consume alcohol (mean intake 1.6g/day vs. 5.3g/day respectively) in comparison to non-cases. The mean difference of alcohol consumption was 3.7g/day less among diabetic cases in comparison to non-diabetic women. A larger number of diabetic women reported a positive family history of diabetes in comparison to non-diabetic women (22% vs. 6% respectively). A higher number of smokers was found among diabetic participants in comparison to participants free from diabetes (12% vs. 8% respectively). However, the overall difference was not statistically significant (p=0.14).

Lower total energy intake and dietary fibre intake were found among diabetic cases. In addition, insoluble and soluble fibre intake was also lower among diabetic women in comparison to non-diabetic women however the differences were not statistically significant.

Diabetic and non-diabetic women tended to consume similar proportions of macronutrients in their diet (54% from carbohydrates, 16% from protein, 33% from fat). Even for fruit consumption, diabetic women consumed an equal number of portions per day as non-diabetic women while a slightly but not significantly lower number of portions of vegetables per day were consumed in comparison to non-diabetic women (4.3 vs. 4.6; p=0.1). No significant differences were observed between prevalent (n= 152) and incident diabetic (n=114) women in term of dietary fibre intake, alcohol intake, total energy intake, self-reported vegetarianism, smoking status, macronutrient intake however prevalent cases tended to have statistically significant lower BMI in comparison to incident of cases (27.8kg/m^2 vs. 29.2kg/m^2). However, some prevalent cases may be type 1 diabetes mellitus.

Table 8.5 Baseline characteristics of T2DM cases and non-cases

All values were mean (SD) unless indicated		**T2DM**			***p***
		Non cases	**cases**	**Total**	
Number of cases/ non-cases		11,982	114	12,096	
Age (years)[1]		50.4(13.5)	57.4(12.1)	51.1(14.1)	<0.01
BMI (kg/m^2)[1]		23.6(4.3)	28.5(8.2)	23.8(4.4)	<0.01
Weight change (kg)[1,2]		6.3(10.4)	17.2(16.3)	6.3(10.8)	<0.01
Total energy intake (kcal)		2281(662)	2265.9(655)	2264(683)	0.68
Carbohydrate intake (E%)		54.8(6.3)	54.7(6.5)	54.8(6.4)	0.85
Protein intake (E%)		15.6(2.6)	16.1(2.9)	15.9(2.7)	0.04
Fat intake (E%)		33.2(5.8)	32.6(6.3)	33.1(5.8)	0.34
Saturated fat intake (E%)		11.3(3.2)	11.6(3.4)	11.4(3.2)	0.31
Fruit intake (portion per day)[3]		4.3(3.8)	4.3(4.1)	4.3(3.8)	0.97
Vegetable intake (portion per day)[3]		4.6(3.3)	4.3(4.3)	4.6(3.4)	0.10
AOAC-fibre g/day		38.7(15.2)	40.1(15.1)	39.0(15.4)	0.36
AOAC-fibre density (g/1000kcal)		17.7(4.6)	17.0(4.4)	17.3(4.7)	0.25
NSP (g/day)		26.8(10.4)	26.1(10.3)	26.0(10.5)	0.54
NSP density (g/1000kcal)		11.8(3.3)	11.4(3.1)	11.5(3.3)	0.42
NSP-Soluble fibre (g/d)		11.2(4.1)	11.0(4.2)	11.0(4.2)	0.61
NSP-insoluble fibre (g/d)		16.9(6.9)	16.1(6.8)	16.3(7.1)	0.27
Magnesium intake (mg/day)		456 (140)	433 (133)	445(141)	0.05
Ethanol (g/d)[1]		5.3(11.2)	1.6(9.4)	5.1(11.1)	0.03
Physical activity (MET)[3]		14.0(12.6)	12.3(15.4)	13.7(13.1)	0.05
Smoking status (%)	Smoker	8.4	12.3	8.5	0.14
	Non smoker	91.6	87.7	91.5	
Self-reported Vegetarian (%)	Yes	33	20	33	<0.01
	No	66	80	67	
Socio-economic status (%)	Professional/ managerial	67	66	67	0.53
	Intermediate	26	24	25	
	Routine and manual	7	10	7	
Educational achievement (%)	No educational records	12	24	12	<0.01
	O-level	30	35	30	
	A-level	26	18	26	
	Degree	32	23	32	

[1] median (IQR), [2]Weight changes (kg) is the difference between self-reported weight at age of 20 years – weight at baseline FFQ values reported, [3] values are geometric mean (IQR)

All dietary fibre intakes in UKWCS participants were high and similar in cases and non-cases (see Table 8.3). As may be seen in Table 8.6, women with high dietary AOAC-fibre intake were more likely to be older, vegetarian, non-smoker, well educated, from high social class, have lower BMI, and engage in physical activity and less likely to have gained weight and to have a positive family history of diabetes. Energy intake from protein, fat and saturated fat intakes decreased with

increasing intake of AOAC-fibre, whereas energy intake from carbohydrates was positively associated with dietary fibre intake. Fruit consumption expressed as portions per day was three times higher among women in the highest dietary AOAC-fibre quintile in comparison to women in the lowest quintile. A similar pattern was observed with vegetable intake which was positively associated with dietary AOAC-fibre intake. Briefly, this indicates as dietary fibre intake increases, lifestyle behaviours and dietary factors tend to be healthier.

Table 8.6 Baseline characteristics with increasing AOAC-fibre intake, values are mean (SD) or frequency (%)

Variables		Dietary AOAC-fibre quintiles					P trend
		1st	2nd	3rd	4th	5th	
Number of participants (n)		2057	2370	2502	2573	2594	
Diabetes cases (n)		22	30	16	23	23	0.22
AOAC fibre (g/day)		21.8(6.0)	30.1(3.4)	36.8(3.5)	44.8(4.8)	58.9(12.9)	<0.01
NSP (g/day)		14.3(4.1)	19.9(2.6)	24.5(2.7)	30.0(3.5)	39.6(8.9)	<0.01
AOAC-fibre (g/1000kcal/d)		11.6(2.1)	14.6(1.2)	16.9(1.1)	19.4(1.4)	23.4(3.5)	<0.01
NSP (g/1000kcal/day)		8.5(3.1)	10.2(3.2)	11.3(3.4)	12.4(3,5)	14.2(4.0)	<0.01
Age (years)		50.4(13)	50.8(14)	51.4(14)	51.3(14)	52.5(13)	<0.01
BMI (kg/m^2)		23.7(4.6)	23.4(4.1)	23.3(4.1)	23.2(4.3)	22.9(4.3)	<0.01
Weight change (kg)		7.2(10.4)	6.3(10.7)	6.3(10.4)	6.3(9.9)	5.8(10.4)	<0.01
Soluble fibre (g/day)		6.4(1.8)	8.7(1.4)	10.4(1.6)	12.5(1.9)	16.3(3.9)	<0.01
Insoluble fibre (g/day)		8.4(2.7)	12.2(2.0)	15.4(2.3)	19.1(2.8)	25.6(6.0)	<0.01
Alcohol intake (g/day)		5.7(13.2)	5.1(11.7)	5.3(10.7)	5.1(10.7)	4.3(10.0)	<0.01
Magnesium intake (mg/day)		290(62)	371(55)	431(60)	499(67)	640(121)	<0.01
Energy intake (kcal/day)		1614(554)	1950(565)	2173(623)	2434(657)	2899(859)	<0.01
Carbohydrate intake (%E)		51.7(6.7)	53.5(5.9)	54.5(5.9)	55.7(5.8)	57.8(6.1)	<0.01
Protein intake (%E)		16.7(3.0)	16.1(2.7)	15.9(2.6)	15.5(2.5)	15.1(2.4)	<0.01
Fat intake (%E)		35.1(5.9)	33.8(5.5)	33.2(5.5)	32.4(5.5)	30.9(5.7)	<0.01
Saturated fat intake (%E)		12.9(3.4)	12.0(3.1)	11.5(3.0)	10.8(2.9)	9.7(2.8)	<0.01
Fruit intake (portion/d)		2.7(1.6)	3.9(2.1)	4.8(2.4)	5.9(3.0)	9.1(6.1)	<0.01
Vegetable intake(portion/d)		3.0(1.5)	4.1(1.7)	4.9(1.9)	5.9(2.1)	8.0(3.5)	<0.01
Physical activity (MET)		12.2(12.2)	13.3(12.2)	14.5(12.4)	15.1(13.0)	16.9(14.4)	<0.01
Vegetarian (%)		19	25	28	38	46	<0.01
F/H of diabetes[1] (%)		2.3	2.4	2.5	2.7	2.9	0.35
Smoking status (%)	non smoker	85	90	93	94.3	94.4	<0.01
	smoker	15	9.9	6.6	5.7	5.6	
Socio-economic status (%)	Professional	62.6	64.6	65	67.9	69.1	<0.01
	Intermediate	28.4	26	26.9	24.8	23.9	
	Routine	8.6	9.4	8	7.1	6.9	
	No record	15.6	13.8	14	11.9	13	<0.01
Educational level (%)	O-level	33.1	30.6	30.3	27.8	27.8	
	A-level	23.8	25.7	24.8	26.8	28	
	Degree	27.4	29.7	30.6	33.3	31	

[1] Positive family history of diabetes. MET is metabolic equivalent task is defined as the ratio of work metabolic rate to a standard resting metabolic rate of 1.0.(4.184 kJ) kg .h , I MET is considered a resting metabolic rate obtained during quiet sitting.

8.4.2Total dietary fibre intake and risk of T2DM

Initial unadjusted logistic models were carried out to determine the effect of AOAC-fibre and NSP intakes expressed as (g/day) and energy adjusted logistic

models to determine effect of AOAC-fibre and NSP intakes (expressed as g/1000kcal/day) on the risk of T2DM shown in Figure 8.3 and 8.4. This suggested that the probability of having T2DM decreased with increased dietary fibre intake among UKWCS. Table 8.7 presents the logistic regression analyses which examined the effect of a daily increment of 5 grams of dietary fibre, which is equivalent to AOAC-fibre content of 2 medium slices of wholemeal bread or 80grams of baked beans or a portion of boiled peas (85g) with the risk of T2DM and daily 5grams increment in NSP which is equivalent to NSP content of 2.5 slices of wholemeal bread. In all the models, odds ratios were less than one for dietary intakes from both the AOAC-fibre and NSP derived values however these associations did not reach statistically significant levels. A marginally significant protective effect on the development of T2DM was observed with every 2.5g/1000kcal/day increment of AOAC-fibre (OR=0.91; 95%CI: 0.82, 1.00; p=0.07) independent of age.

This suggests the potential of a lower risk of developing T2DM with higher AOAC-fibre intake independent of age, however this association was attenuated with further adjustment for weight change and smoking. This suggests that other factors may have a greater influence on risk of T2DM than dietary fibre.

Table 8.7 Odds ratios (95%CI) for incidence of T2DM risk with every increment of dietary fibre intake derived by two analytical methods among UKWCS

Nutrient	OR (95%CI)	P value
AOAC-fibre intake 5g/day		
Model 1[1]	0.96(0.90, 1.02)	0.21
Model 2[2]	0.98(0.92, 1.04)	0.59
Model 3[3]	0.99(0.92, 1.05)	0.78
Model 4[4]	0.99(0.92, 1.06)	0.83
AOAC-fibre density intake 2.5 g/1000kcal/day		
Model 1[1]	0.91 (0.82, 1.00)	0.07
Model 2[2]	0.96(0.86, 1.07)	0.50
Model 3[3]	0.96(0.87, 1.07)	0.51
Model 4[4]	0.96(0.86, 1.07)	0.50
NSP intake 5 g/day		
Model 1[1]	0.95 (0.86, 1.04)	0.29
Model 2[2]	0.98(0.89, 1.07)	0.69
Model 3[3]	0.99(0.90, 1.09)	0.92
Model 4[4]	0.99(0.90, 1.09)	0.95
NSP density intake 2.5 g/1000kcal/day		
Model 1[1]	0.89(0.77, 1.03)	0.14
Model 2[2]	0.96(0.82, 1.12)	0.63
Model 3[3]	0.96(0.83, 1.12)	0.65
Model 4[4]	0.96(0.83, 1.15)	0.67

[1]M1 = age; [2]M2 = M1 plus total energy intake, weight changes (%) and smoke (yes/no); [3]M3 = M2 plus socioeconomic status, alcohol intake and METs, [4]M4 = M3 plus family history of diabetes.

Figure 8.3 Predicted risk of T2DM with dietary fibre intake expressed as AOAC-fibre and NSP intakes (g/day)

(A) (B)

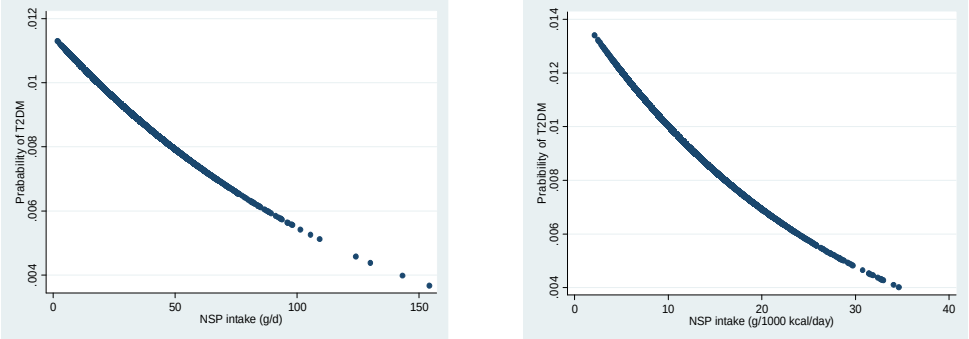
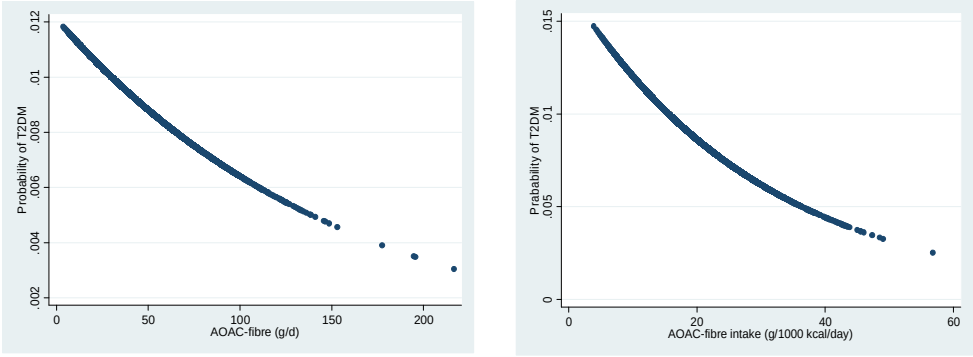


Figure 8.4 Predicted risk of T2DM with dietary fibre intake expressed as AOAC-fibre and NSP intakes (g/1000kcal/ day)

(C) (D)



The risks of T2DM tend to decrease with increasing intakes of AOAC-fibre expressed in g/day and g/1000kcal/day in the age-adjusted models. Similarly, lower probability of developing T2DM across NSP quintiles in comparison to lowest NSP group with a p for trend=0.55 were observed. Inconsistency in odds ratios across NSP density quintiles in Table 8.8 was observed in age-adjusted and multivariate models. Overall, the effect of total dietary fibres intake estimated from the two analytical methods on the risk of T2DM is inconclusive. Logistic regression analyses were also carried out with tertile rather than quintile categorization and the results were similar.

Weight variable was created previously for vegetarian status (Cade *et al.*, 2010). In this variable, the weighting was based on the probability of being sampled to account for high percentage of vegetarians in the cohort and to obtain results that representative to the general population. The results from weighted logistic regression analyses for total dietary fibre effect (continuous and categorical) on the risk of T2DM were similar to the original results in all logistic models.

Table 8.8 Odds ratios (95%CI) for incidence of T2DM by quintile of intakes of dietary AOAC-fibre, NSP, AOAC-fibre density and NSP density among UKWCS

Fibre intake		Median (IQR)	Range (g/d)	T2DM OR (95%CI)				
				Cases[1]	Model 1[2]	Model 2[3]	Model 3[4]	Model 4[5]
AOAC-fibre (g/day)	Q1	21.8(6.0)	<26.4	22	1	1	1	
	Q2	20.1(3.4)	26.4-33.4	30	1.14(0.65, 1.98)	1.35(0.76, 2.49)	1.49(0.80, 2.76)	1.57(0.84, 2.93)
	Q3	36.8(3.5)	33.5-40.5	16	0.55(0.29, 1.06)	0.66(0.31, 1.38)	0.76(0.36, 1.62)	0.77(0.36, 1.64)
	Q4	44.8(4.8)	40.5-50.3	23	0.77(0.43, 1.40)	0.97(0.48, 1.96)	1.04(0.50, 2.16)	1.10(0.53, 2.29)
	Q5	58.9(12.9)	>50.5	23	0.76(0.42, 1.37)	0.96(0.42, 2.17)	0.99(0.43, 2.29)	1.03(0.44, 2.39)
	P value			114	0.21	0.59	0.78	0.83
AOAC-fibre density (g/1000 kcal /day)	Q1	11.6(2.1)	<13.3	22	1	1	1	1
	Q2	14.6(1.2)	13.3-15.7	24	0.93(0.52, 1.67)	1.09(0.60, 2.00)	1.19(0.64, 2.20)	1.22(0.66, 2.26)
	Q3	16.9(1.1)	15.8-18.1	23	0.83(0.46, 1.50)	1.06(0.58, 1.96)	1.03(0.54, 1.94)	1.03(0.54, 1.95)
	Q4	19.4(1.4)	18.2-21.0	22	0.75(0.41, 1.36)	0.90(0.48, 1.71)	0.92(0.48, 1.78)	0.96(0.49, 1.85)
	Q5	23.4(3.5)	>21.0	23	0.77(0.42, 1.39)	1.03(0.55, 1.94)	1.03(0.54, 1.96)	1.06(0.55, 2.02)
	P value				0.07	0.50	0.51	0.50
NSP (g/day)	Q1	14.3(4.0)	<17.4	22	1	1	1	1
	Q2	19.9(2.3)	17.4-22.2	26	0.98(0.55, 1.74)	1.25(0.68, 2.29)	1.35(0.72, 2.53)	1.40(0.74, 2.65)
	Q3	24.5(2.4)	22.3-27.1	18	0.62(0.33, 1.17)	0.72(0.35, 1.48)	0.84(0.40, 1.74)	0.86(0.41, 1.79)
	Q4	30.0(3.2)	27.2-33.8	25	0.84(0.47, 1.49)	1.11(0.56, 2.19)	1.19(0.59, 2.41)	1.24(0.61, 2.52)
	Q5	39.6(8.8)	>33.8	23	0.76(0.42, 1.37)	1.06(0.48, 2.33)	1.11(0.49, 2.49)	1.15(0.51, 2.61)
	P value				0.29	0.69	0.92	0.95
NSP density (g/1000 kcal/day)	Q1	7.5 (1.5)	<8.7	19	1	1	1	1
	Q2	9.6 (0.8)	8.7-10.5	22	1.03(0.55,1.92)	1.22(0.64, 2.31)	1.30(0.68, 2.50)	1.28(0.66, 2.47)
	Q3	11.3(0.8)	10.5-12.1	27	1.16(0.65, 2.09)	1.56(0.85, 2.87)	1.56(0.83, 2.92)	1.59(0.84, 2.99)
	Q4	13.0(1.0)	12.1-14.2	26	1.00(0.55, 1.83)	1.30(0.69, 2.44)	1.35(0.70, 2.58)	1.37(0.71, 2.63)
	Q5	15.9(2.4)	>14.2	20	0.78(0.41, 1.47)	0.98(0.49, 1.95)	0.98(0.48, 1.98)	0.98(0.49, 1.99)
	P value				0.0.14	0.0.63	0.65	0.67

[1] Diabetic cases [2]M 1 = age [3]M 2 = M 1 plus total energy intake, weight changes (%) and smoke (yes/no), [4]M3 = M2 plus socioeconomic status, alcohol intake and METs, [5]M4 = M3 plus family history of diabetes. Note: adjustment for energy intake was not included in fibre density models.

8.4.3Sources of fibre and risk of T2DM

With respect to dietary fibre sources, the risk of T2DM is reported in Table 8.9. High cereal fibre intake was significantly associated with a lower incidence of diabetes in the age-adjusted model (OR = 0.86; 95%CI: 0.75, 0.99, p=0.03). In other words, for every 5 grams' increase in cereal AOAC-fibre that is equivalent to two slices of wholemeal bread, the risk of T2DM was reduced by 14%. In the further adjusted models, the increase in cereal fibre intake did not significantly decrease the risk of T2DM among women. A lower risk of diabetes was also observed with increased intakes of legumes fibre (OR=0.90; 95%CI: 0.61, 1.31; p=0.59) and vegetable fibre (OR=0.93; 95%CI: 0.76, 1.15) in the age-adjusted model and (OR= 0.93; 95%CI: 0.62, 1.38; p=072; OR= 1.04; 95%CI: 0.84, 1.28) for legumes fibre and vegetable fibre respectively in the full-adjusted model but these values were not statistically significant. Dietary fibre intakes from fruit and from nuts and seeds did not show any significant association with the risk of T2DM in any of the logistic models. In conclusion, all odds ratios for cereal fibre are less than one however; the other fibre sources tend to show less consistent direction of association.

Table 8.9 Odds ratios (95%CI) for incidence of T2DM with every increment of intakes of dietary fibre from cereal, vegetables, fruits, legumes and nuts among UKWCS

Dietary fibre sources	OR (95%CI)	P value
Cereal fibre intake (5g/day)		
Model 1[1]	0.86(0.75, 0.99)	0.03
Model 2[2]	0.90(0.76, 1.06)	0.22
Model 3[3]	0.87(0.74, 1.04)	0.13
Model4[4]	0.88(0.74, 1.04)	0.16
Vegetable fibre intake (5g/day)		
Model 1[1]	0.93(0.76, 1.15)	0.55
Model 2[2]	1.00(0.81, 1.24)	0.95
Model 3[3]	1.03(0.84, 1.27)	0.73
Model 4[4]	1.04(0.84, 1.28)	0.70
Fruit fibre intake (5g/day)		
Model 1[1]	1.01(0.88, 1.17)	0.78
Model 2[2]	1.07(0.92, 1.25)	0.33
Model 3[3]	1.07(0.91, 1.25)	0.38
Model 4[4]	1.08(0.92, 1.26)	0.33
Legumes fibre intake (5g/day)		
Model 1[1]	0.90(0.61, 1.31)	0.59
Model 2[2]	0.92(0.61, 1.38)	0.69
Model 3[3]	0.92(0.62, 1.37)	0.69
Model 4[4]	0.93(0.62, 1.38)	0.72
Fibre from nuts and seeds intake (5g/day)		
Model 1[1]	0.89(0.74, 1.08)	0.27
Model 2[2]	1.00(0.85, 1.19)	0.92
Model 3[3]	1.00(0.84, 1.18)	0.97
Model 4[4]	0.99(0.83, 1.17)	0.91

[1]M 1 = age; [2]M 2 = M 1 plus total energy intake, weight changes (%) and smoke (yes/no); [3]M3 = M2 plus socioeconomic status, alcohol intake and METs; [4]M4 = M3 plus family history of diabetes. Note: adjustment for energy intake was not included in fibre density models

During a 4 year follow-up, inverse associations between the intakes of dietary fibre sources and the risk of T2DM were observed generally across the dietary fibre sources quintiles. However, p values for trends did not reach statistically significant levels in the logistic models reported in Table 8.10. Across all sources no evidence of a trend can be seen. In multivariate adjusted models, the point estimates are rather inconsistent across quintiles with possible exception of cereal fibre. Overall, dietary fibres from main fibre sources were not associated with the risk of T2DM in the age and fully-adjusted models.

Table 8.10 Odds ratios (95%CI) for incidence of T2DM by quintile of intakes of dietary fibre sources among UKWCS

Fibre intake		Median (IQR)	Range	Cases[1]	T2DM OR (95%CI) Model 1[2]	Model 2[3]
Cereal fibre	Q1	4.4 (2.2)	<6.3	29	1	1
	Q2	8.0 (1.7)	6.3-9.8	20	0.59(0.33, 1.04)	0.75(0.40, 1.39)
	Q3	11.5(1.8)	9.9-13.5	21	0.56(0.32,1.00)	0.66(0.35, 1.25)
	Q4	15.8(2.4)	13.6-18.5	23	0.59(0.34, 1.02)	0.62(0.31, 1.22)
	Q5	22.7(6.2)	>18.5	21	0.52(0.30, 0.93)	0.67(0.33, 1.37)
	p value			114	0.14	0.66
Vegetable fibre	Q1	1.1(1.8)	<3.9	28	1	1
	Q2	8.6(1.0)	3.9-5.8	22	0.59(0.33, 1.04)	0.67(0.36, 1.26)
	Q3	10.8(1.2)	5.9-7.9	16	0.44(0.24, 0.82)	0.67(0.35, 1.28)
	Q4	13.5(1.6)	7.9-10.9	19	0.52(0.29, 0.92)	0.80(0.42, 1.53)
	Q5	18.6(5.1)	>10.9	29	0.77(0.46, 1.29)	1.05(0.57, 1.95)
	p value			114	0.05	0.48
Fruit fibre	Q1	2.2(1.5)	<3.5	23	1	1
	Q2	4.6(1.5)	3.5-5.6	17	0.56(0.30, 1.06)	0.71(0.36, 1.40)
	Q3	6.6(1.1)	5.7-7.8	18	0.55(0.29, 1.03)	0.66(0.33, 1.33)
	Q4	9.3(1.6)	7.8-11.3	29	0.82(0.47, 1.43)	1.18(0.63, 2.19)
	Q5	14.7(6.0)	>11.3	27	0.78(0.44, 1.37)	1.02(0.52, 1.99)
	p value			114	0.28	0.34
Legumes fibre	Q1	1.0(0.6)	<1.5	27	1	1
	Q2	1.9(0.4)	1.5-2.3	21	0.75(0.42, 1.33)	0.73(0.38, 1.38)
	Q3	2.6(0.4)	2.4-3.3	19	0.67(0.37, 1.21)	0.73(0.38, 1.41)
	Q4	3.9(0.8)	3.4-5.1	26	0.91(0.52, 1.57)	1.06(0.58, 1.95)
	Q5	6.9(2.6)	>5.2	21	0.85(0.47, 1.51)	0.90(0.46, 1.74)
	p value			114	0.70	0.69
Fibre from nuts and seeds	Q1	0.08(0.06)	<0.12	27	1	1
	Q2	0.1(0.05)	0.13-0.23	15	0.56(0.29, 1.05)	0.61(0.30, 1.23)
	Q3	0.3(0.1)	0.24-0.50	23	0.90(0.51, 1.59)	1.09(0.58, 2.04)
	Q4	0.6(0.2)	0.51-1.0	15	0.58(0.30, 1.10)	0.81(0.40, 1.63)
	Q5	1.8(1.6)	>1.0	19	0.65(0.36, 1.17)	0.98(0.49, 1.93)
	p value			114	0.23	0.54

[1] diabetic cases [1]M1 = age; [2]M 2 = M1 plus total energy intake, weight changes (%) and smoking (yes/no), socioeconomic status, alcohol intake and METs, family history of diabetes.

8.4.4Dietary soluble and insoluble fibres and risk of T2DM

The odds of incident T2DM in association with insoluble and soluble fibre intake are shown in Table 8.11. The risk of T2DM tended to be lower with higher insoluble fibre intake; however this did not reached the significance level. No

significant association was found between increase in soluble dietary fibre intake and risk of T2DM in any of the models. Furthermore, the odds ratios across insoluble and soluble dietary fibre quintiles were inconsistent (Table 8.12).

Table 8.11 Odds ratios (95%CI) for incidence of T2DM with every increment of intakes of insoluble and soluble dietary fibre among UKWCS

Types of dietary fibre	OR (95%CI)	P value
Insoluble fibre intake g/day		
Model 1[1]	0.89(0.77,1.03)	0.13
Model 2[2]	0.94(0.82, 1.15)	0.47
Model 3[3]	0.96(0.83, 1.11)	0.63
Model 4[4]	0.95(0.78, 1.15)	0.63
Soluble fibre intake g/day		
Model 1[1]	0.91(0.72,1.15)	0.46
Model 2[2]	0.96(0.76, 1.22)	0.76
Model 3[3]	1.00(0.79, 1.27)	0.95
Model 4[4]	1.01(0.80,1.29)	0.89

[1]M1 = age; [2]M2 = M 1 plus total energy intake, weight changes (%) and smoke (yes/no); [3]M3 = M2 plus socioeconomic status, alcohol intake and METs; [4]M4 = M3 plus family history of diabetes. Note: adjustment for energy intake was not included in fibre density models

Table 8.12 Odds ratios (95%CI) for incidence of T2DM by quintile of intakes of soluble and insoluble dietary fibre among UKWCS

Dietary intake		Median (IQR)	Range	Cases[1]	Diabetes mellitus OR (95%CI) Model 1[2]	Model 2[3]
Insoluble fibre	Q1	8.4(2.6)	<10.4	26	1	1
	Q2	12.1(1.6)	10.5-13.7	21	0.66(0.37, 1.52)	0.72(0.39,1.32)
	Q3	15.3(1.6)	13.8-17.1	18	0.53(0.29, 0.98)	0.58(0.29,1.12)
	Q4	19.1(2.2)	17.2-21.6	26	0.72(0.42, 1.25)	0.82(0.43,1.54)
	Q5	25.6(5.7)	>21.6	23	0.63(0.36, 1.11)	0.74(0.35,1.55)
	p for trend			114	0.31	0.58
Soluble fibre	Q1	8.5(3.5)	<7.6	22	1	1
	Q2	12.2(3.7)	7.6-9.5	30	1.22(0.70, 2.13)	1.38(0.70,2.51)
	Q3	15.1(4.1)	9.5-11.4	13	0.48(0.24, 0.96)	0.57(0.27,1.22)
	Q4	18.7(4.6)	11.5-14.0	22	0.80(0.44, 1.46)	1.00(0.49,2.00)
	Q5	25.0(7.3)	>14.0	27	0.99(0.56, 1.75)	1.29(0.59,2.81)
	p for trend			114	0.07	0.12

[1] diabetic cases [1]M 1 = age [2]M 2 = M 1 plus total energy intake, weight changes (%) and METs.

8.4.5Dietary fibre intake and risk of T2DM among women based on BMI

Sensitivity analysis was carried out to examine whether we can visualize a trend in the association between dietary AOAC-fibre and risk of diabetes among women with BMI<25kg/m^2 and women with BMI≥25kg/m^2. Table 8.13 provides the number of diabetic and non-diabetic cases in each group where chi square test showed a significance difference between the groups ($p<0.01$). Two thirds of diabetic cases were either overweight or obese and only 25% had a BMI of less than 25kg/m^2 while 68% of women free from diabetes had a BMI of less than 25kg/m^2 and only third of them were either overweight or obese.

Table 8.13 Distribution of women (N, %) based on DM and BMI categories

Group	BMI <25kg/m^2	BMI ≥25kg/m^2
T2DM incident group (n, %)	29 (25)	85 (75)
Non-T2DM group (n, %)	8,122 (68)	2,860 (32)

Table 8.14 presents the odds ratios of T2DM among overweight and obese women, with every 5g increment in total dietary AOAC-fibre intake. The risk of T2DM was non-significantly reduced with increasing dietary fibre intake. However, the small number of diabetic cases in each group limits the interpretation of the current findings. In the current study the interaction term was not considered in the logistic models mainly because of the possibility of losing statistical power with stratification due to the small number of cases. In addition, no significant modification effect of BMI on the relationship between dietary fibre and risk of diabetes (p=0.64) was seen in the current study. However, sensitivity analysis was carried out to examine whether the association between dietary AOAC-fibre and risk of diabetes among women with BMI<25kg/m^2 and women with BMI≥25kg/m^2 differ, as evidence showed significant different between obese and non-obese.

Table 8.14 Odds ratios (95%CI) for incidence of T2DM by quintile of intakes of dietary AOAC-fibre intakes (expressed as g/day and g/1000kcal/day) among UKWCS

		BMI<25kg/m^2		BMI≥25kg/m^2	
		cases/non cases	OR (95%CI)[a]	cases/non cases	OR(95%CI)[a]
AOAC-fibre intake (g/day)	Q1	5/1,278	1	17/757	1
	Q2	6/1,560	1.11(0.32, 3.81)	24/780	1.37(0.69, 2.69)
	Q3	5/1,682	0.96(0.25, 3.64)	11/804	0.59(0.26, 1.36)
	Q4	8/1,757	1.27(0.32, 5.04)	15/793	0.78(0.34, 1.74)
	Q5	5/1,757	1.25(0.24, 6.36)	18/726	0.84(0.33, 2.13)
	p for trend		0.99		0.23
	per 5g/day [b]		1.02(0.85, 1.22)		0.96(0.87, 1.07)
	p value		0.81		0.52
AOAC-fibre intake (g/1000kcal/day)	Q1	4/1,267	1	18/755	1
	Q2	5/1,510	1.14(0.30, 4.30)	19/798	1.05(0.54, 2.05)
	Q3	6/1,689	1.06(0.28, 4.02)	17/781	0.88(0.44, 1.77)
	Q4	6/1,785	1.03(0.27, 3.96)	16/788	0.76(0.37, 1.57)
	Q5	8/1,871	1.52(0.44, 5.20)	15/738	0.91(0.44, 1.85)
	p for trend		0.94		0.91
	per 2.5 g/day [b]		1.02(0.84, 1.25)		0.91(0.81, 1.05)
	p value		0.79		0.24

[a] Adjusted to age, total energy intake, smoking status, socioeconomic status, alcohol intake, physical activity (MET). [b]5 grams of AOAC-fibre equivalent to two slices of wholemeal breads [c] equivalent to one slice of wholemeal bread

8.4.6UKWCS findings added in the Forest plot of the comprehensive review

Original meta-analysis in the comprehensive review (chapter 4) was used. Pooled estimate that examined risk of T2DM with the intake of total dietary fibre was 0.97 (95%CI: 0.93, 1.02). The current findings in this chapter particularly AOAC-fibre intake (g/day) and risk of T2DM comparing extreme quintiles from model 4 was included with the other studies in the meta-analysis. This resulted in a total of 11 cohort studies that examined the relationship between TDF intake and risk of T2DM. Estimated risks were obtained from the fully adjusted models for all cohort studies. Figure 8.5 demonstrates similar findings in chapter 4 where lack of association between TDF intake and risk of T2DM comparing highest versus lowest consumers independent of potential confounders (combined risk = 0.97; 95%CI: 0.93, 1.01) with moderate heterogeneity was seen. Another pooled estimate from the current findings of the UKWCS and other two cohorts (Sluijs *et al.*, 2010b, Barclay *et al.*, 2007) that reported dietary fibre intake as continuous variable showed estimate risk= 0.93 (95%CI; 0.87, 1.01; I^2=49%; p=0.13).

Figure 8.5 Estimate risk for the association between total dietary fiber consumption and risk of T2DM for individual cohort studies and all cohort studies combined.

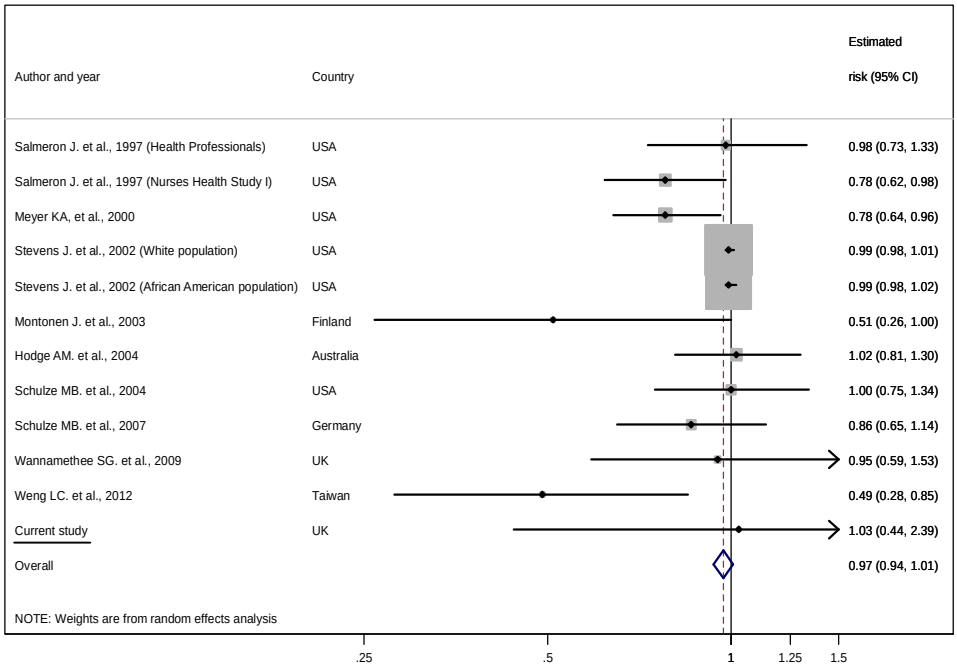
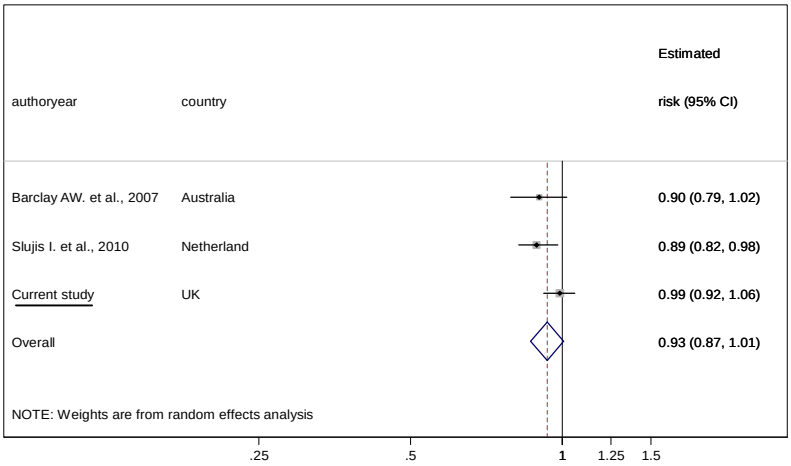


Figure.8.6 Estimate risk with every unit increment in TDF including the UKWCS



8.5 Discussion

The current results are pointing towards a lower risk of developing T2DM in UKWCS participants who consume diets high in fibre particularly cereal fibre. This study was carried out because of two main reasons. Few studies have examined the AOAC-fibre intake in relation to risk of T2DM prospectively in the UK among women. Also inconclusive findings regarding the role of dietary fibre sources on risk of T2DM led to the current analysis. Prior studies have noted the importance of high dietary fibre intake on reducing the risk of T2DM (Sluijs *et al.*, 2010b, Meyer *et al.*, 2000, Salmeron *et al.*, 1997b). Prospective evidence reported the beneficial effect of insoluble rather than soluble fibre on development of T2DM.

8.5.1Intakes of total dietary fibre, key fibre sources and types of fibre in relation to the risk of T2DM

For the first time, current study examined the relationship between risk of T2DM and dietary fibre intake expressed as AOAC-fibre and NSP intakes. The results of this study did not show any significant association between the risk of T2DM with dietary intake of AOAC-fibre or NSP.

. The present study showed high cereal AOAC-fibre intake was associated with decreased incidence of T2DM independent of age however this relationship was attenuated with additional adjustment for weight change, physical activity, total energy, smoking, socioeconomic status and alcohol intake. In another words, with every 5 grams increases in cereal fibre, the risk of diabetes reduced by 14% independent of age. A similar effect size was reported in a previous cohort study among the white US population (Stevens *et al.*, 2002). Most cohort studies showed a

protective effect of cereal fibre on the risk of diabetes, while for those few studies which did not find a significant association between cereal intake and risk of diabetes, this may be due to small sample sizes (Colditz *et al.*, 1992, Barclay *et al.*, 2007) or underestimation of cereal fibre intake from a small number of food items in the FFQ (Hodge *et al.*, 2004). In our study the 37 food items which were assigned to the cereals group rule out the possibility of the latter explanation.

Six out of seven previous prospective studies among women participants have demonstrated significant associations between high cereal fibre intake (more than 5.9g/d; more than 7.3g/day; more than 7.5g/day in 2 studies; and more than 10.2g/4184J/day) and the low risk of T2DM by 18%, 36%, 28%, and 12% (Krishnan *et al.*, 2007, Meyer *et al.*, 2000, Hopping *et al.*, 2010, Schulze *et al.*, 2004a). In our study women in the 5th quintile who consumed more than 18.5g/day of cereal fibre, had a 48% risk reduction (OR= 0.52; 95%CI: 0.30, 0.93) compared to women who consumed less than 6.3g/day in the 1st quintile. However our data showed that the trend with increase in cereal fibre intake was not statistical significant (p=0.14).

β-glucans are soluble cereal fibres that are naturally present in foods and added to foods. It was demonstrated in a previous review that intervention with high intake of β-glucans from oat and barley showed significant reductions in post-prandial blood glucose response compared to control in healthy participants (European Food Safety Authority, 2011). That may be explained by high viscosity of fibre which slows the rate of digestion and absorption of glucose (Würsch and Pi-Sunyer, 1997, Battilana *et al.*, 2001). However epidemiological studies observed the protective effect of insoluble fibre rather than soluble fibre.

A recent randomized controlled trial examined the short term effect of purified insoluble cereal fibre consumption of (31.2g/day) for three days among overweight and obese women and found an improvement in whole body insulin sensitivity (Weickert *et al.*, 2006). Despite the lower intake of insoluble fibre in prospective studies, the results suggest beneficial effects of insoluble fibre on the development of T2DM.

In the current study, fibre from vegetables, fruits, legumes and from nuts and seeds was not associated with diabetes risk in UKWCS participants. However the lack of evidence for a linear relationship may less likely to be due to insufficient amount of fibre intakes from legumes, nuts and seeds sources. This because the amount of fibre consumed from the above sources is higher than average UK

women and higher than cohorts who reported significant associations (Schulze *et al.*, 2004a). In contrast, the findings of the current study are similar to prospective evidence demonstrated in chapter 4. Where none of the studies showed an effect of high legumes consumption on the risk of T2DM, nine out of 11 studies reported no effect of vegetables fibre and 10 of the 11 cohorts reported no effect of fruit fibre on the risk of T2DM.

After including current study into other cohort in the meta-analysis, findings remained similar and protective effect of total dietary fibre on risk of T2DM was seen. However, high heterogeneity should be considered.

Cohort women have health conscious characteristics where the average dietary AOAC-fibre intake exceeded the recommended intake (Lunn and Buttriss, 2007). This may suggest that there are no further benefits of dietary fibre regarding the risk of diabetes above the recommended intake.

Another possible reason is the overestimation of dietary intake by FFQs which has been previously reported (Willett, 2012). Participants may tend to over-report food consumption with FFQ (Calvert *et al.*, 1997) . The small number of incident cases of diabetes may also result in less power and failing to reject the null hypothesis of no significant difference between diabetic and non-diabetic groups. This may be because women with diabetes cases were self-reported or time frame too short. This allows the possibility of including undiagnosed diabetic cases in the non-cases group.

8.5.2 Strengths and limitations

This study had several important strengths; it is a prospective study that addressed the issue of recall bias. The UKWCS was designed to have a broad variation in dietary patterns to allow investigation of disease in relation to diet. The baseline FFQ had been validated previously in a subsample of women (Spence *et al.*, 2002). Moderate stability of dietary patterns in the UKWCS was reported in a previous analysis (Greenwood *et al.*, 2003). This limits the possibility of estimated bias relating to great changes in diet consumed by the studied population.

On the other hand, unavoidable measurement errors are expected when assessing diet by FFQ which is a common dietary assessment tool used in cohort studies (Cade *et al.*, 2004b).

Adjustment of covariates was limited in the current analyses to those with existing strong evidence on the relationship with the variables of interest aiming to avoid over adjustment that may lead to estimation bias (Kirkwood and Sterne, 2003). In our studied population, the risk of T2DM was four times higher among overweight women and the risk raised progressively with higher BMI. This reflects the powerful relationship between T2DM and obesity. Previous pooled estimate from prospective evidence reported that the risk of T2DM is higher among overweight and highest among obese women in comparison to normal weight. The strong association between obesity and T2DM may reflect that obesity is such a dominant cause that effect of fibre may be relatively weak in comparison, and fibre may act to decrease the risk by reducing weight gain. So fibre effect could be adjusted out when the weight change variable was included in the multivariate model.

The influence of residual confounders in the current study and in any cohort study may lead to a smaller difference than actually exists in the risk estimation. Pooled data from a recent meta-analysis reported significant positive associations between glycaemic index (GI) and glycaemic load (GL) with the risk of T2DM (GI RR=1.40, 95% CI: 1.23, 1.59; GL RR= 1.27, 95% CI: 1.12, 1.45) (Barclay *et al.*, 2008). Unfortunately, these variables were not adjusted in the current models because of unavailability of data.

It appears that observed non-significant inverse associations between total dietary AOAC-fibre intake and risk of T2DM in current study may be partly explain by type II error. Type II error may be explained by the measurement errors related to the dietary assessment method used. overestimation of dietary fibre intake was suspected as participants tended to overestimate dietary intake by FFQ (Spence *et al.*, 2002). Misreporting of vegetables and fruit using the FFQ was reported in previous study (Calvert *et al.*, 1997). This is because FFQ may not consider the exact portion eaten as some food items eaten by less than an estimated average portion.

A possible reason for the observed high amount of dietary AOAC-fibre intake in the UKWCS could be attributed to the high proportion of vegetarians in the studied population or may be because women in the current cohort were more likely to have healthier lifestyles (majority of women were non-smokers and more likely to

engage in physical activity) where, their mean BMIs lower in comparison to the average mean BMI for UK women of similar age (Henderson *et al.*, 2002).

The small number of identified cases may possibly relate to asymptomatic undiagnosed diabetes which could underestimate the true association or if we assume women with healthier behaviours were more likely to be screened and participate in the study then the pattern of low diabetes risk with healthier food or nutrients tended to be underestimated.

Possible explanations for some of the results were misclassification and selection bias. Diagnosis of diabetes was obtained from self-reported question in the phase II questionnaire. Previous evidence showed a high level of confirmation (95%) was reported when medical records were checked for sampled participants (Krishnan *et al.*, 2007). However this does not rule out misclassification as rechecking usually done for diabetic cases and not for other participants

In the current study, the mean dietary AOAC-fibre difference of 1.4 (95%CI: -1.3, 4.2; p=0.32) was observed between diabetic and non-diabetic groups. A post hoc test was carried out aiming to estimate the minimum important effect size that should be present to detect a significant association which can be useful in planning future study (Walters, 2009). The calculated minimum important effect of 4.1g/day at power of 80% and at a significance level of 0.05 in a two-sided test was found for the UKWCS. Confidence intervals are useful for interpreting results in situations where no significant associations are observed (Levine and Ensom, 2001). From the current study, the odds ratio of diabetes was 0.99 (95%CI: 0.92, 1.06) with every 5g/day increment of AOAC-fibre in the fully adjusted model. The value ‘1’ included in the confidence interval indicates that the possibility of no difference between the two groups cannot be ruled out. As discussed before, the lack of association may be attributed to many aspects (no relationship, misclassification or inadequate power).

Failure to differentiate between single nutrient end point effects from other highly correlated nutrients is another limitation of the cohort study. Future research on causal associations between fibre intake and the risk of T2DM from interventional trials can be helpful with more emphasis on different dietary fibre sources and type of dietary fibre consumed. One possible way that may help to increase the number of cases is by identify participants with a preclinical condition, in this case, pre-diabetes. This also may provide useful evidence of causal relationship (Willett, 2012). However, diagnosis of pre-diabetes condition which

also referred to metabolic syndrome needs data on blood pressure, blood triglycerides concentration, blood high-density lipoprotein cholesterol concentration, fasting glucose concentration, and WC measurement to meet the criteria of diagnosis of metabolic syndrome (Alberti *et al.*, 2009). Unfortunately, these data were not available in the current cohort study.

Future prospective studies should be carried out to evaluate the effect of dietary fibre fractions and fibre sources on the development of T2DM. Further evaluation on whether the relationship between risk of T2DM and dietary fibre varied among obese and normal individuals will be useful for future interventional trials. Future work should focus on the long term clinical trials with dietary fibre fractions intervention among high risk population. This will provide evidence on the effectiveness of the dietary fibre intake on measured blood glucose, insulin and even H_{A1C} levels.

8.6 Conclusion

Inverse associations between dietary fibre and risk of T2DM among UKWCS did not reach the significance level and only greater cereal fibre intake was associated with reduced risk of diabetes over an average period of 4 years independent of age. No association between insoluble fibre intake or soluble fibre and the risk of T2DM was observed. As the studied population was found to have healthier characteristics, this may suggest there is no further beneficial effect of dietary fibre against risk of diabetes in health conscious individuals who are high fibre consumers compared to general population but this does not rule out the beneficial effect of dietary fibre per se.

Chapter 9: Legumes intake and the risk of T2DM in the UKWCS

9.1 Introduction

A number of different dietary components have been explored in connection with prevention of T2DM (as shown in chapter 2). Different food intake and risk of T2DM have been explored in previous studies also demonstrated in chapter 2. The association between risk of T2DM and wholegrain intake where extensively evaluated in previous prospective studies and reported consistent significant beneficial effect of wholegrain intake on the risk of T2DM as discussed in chapter 2. Studies exploring fruit and vegetable intake in relation to risk of T2DM have been inconclusive, although a recent meta-analysis reported a potential of leafy green vegetables benefit effect on risk of T2DM (Cooper *et al.*, 2012). Few prospective studies which have examined the association between legumes intake and the risk of T2DM (Villegas *et al.*, 2008). Lack of data on legume consumption among British women as well as lack of prospective studies that examine the effect of legume consumption and risk of T2DM in British women had led to the current study.

The health benefits of legumes intake have been demonstrated previously (Venn and Mann, 2004). Legumes are one of the recommended foods in the UK healthy eating guidelines (Food Standards Agency, 2007). Eating more plant-based foods is advised by several dietary guidelines and legumes were offered as one of the main sources (Leterme, 2002).

Legume consumption varies among European countries and is generally lower than in Asian populations (Schneider, 2002) . Dried legumes such as chickpeas, faba beans and lentils are considered to be a small UK market that targets minority groups such as African or Asian immigrants (Schneider, 2002). However, it has been reported in a review that demands for peas in the UK was 61,000 tons in 1999. On the other hand, because people may favour healthy and balanced diets, it had been suggested that the trend in increasing legume consumption may improve over time (Schneider, 2002). UK dietary surveys have a limited data on legume consumption (Henderson *et al.*, 2002) and there is a need to quantify the amount consumed in the UK to investigate the relation with health outcome (Derbyshire, 2011).

One of the challenges in the epidemiological studies is the difficulties in separating the effect of nutrients because of high correlation between nutrients consumed. As people eat food, not nutrients and different components of food may have different function this support the investigation between the food intake and risk of disease in prospective studies. Therefore, the aim of this study is to examine whether a high intake of legumes lowers the risk of T2DM among UK women. As there is a high proportion of vegetarians in the UKWCS who are likely to consume more fruit, vegetables and legumes, this gives the advantage of exploring legumes intakes in relation to health outcomes. In addition, the presence of a wide range of dietary patterns in the UKWCS facilitates the comparison between women with different amount of legumes consumption. Exploring the socio-demographic, dietary and lifestyle characteristics of high legumes consumers in the UKWCS may help to identify the characteristics of target populations for future randomized controlled trials. Also this chapter may help in generating further hypothesis regards potential compounds in food that may protect from disease outcome (Willett, 2012) for future research.

9.2 Method

9.2.1Studied population

The UKWCS population explored were described in previous chapter 6 (Cade *et al.*, 2004a). Data analyses were carried out among women who completed baseline FFQ and phase II questionnaires after an average of 4.1 years of follow-up. Women with unfeasible total energy intake excluded. This left 12,096 women from the UKWCS for the present analyses (114 women developed T2DM and 11,982 women remained free from diabetes).

*9.2.2Inc*id*ence of T2DM*

Selection of cases was demonstrated in chapter 6. Incidence of T2DM was self-reported in the phase II questionnaire.

9.2.3Non-dietary measures

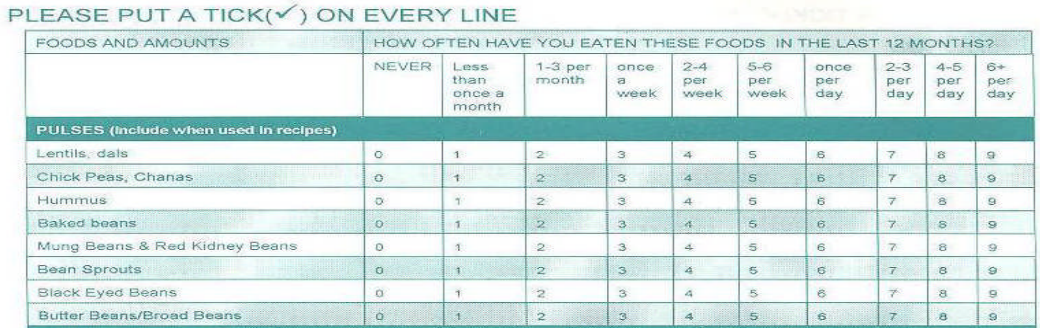
Socio-demographic, anthropometric measures, lifestyle characteristics and other factors (family history of diabetes and self-reported vegetarians) were obtained from baseline FFQs. Details on the measured variables reported in chapter 6.

9.2.4Estimating intakes of legumes in the UKWCS

Average daily legume intakes expressed in g/day were calculated from ten food items listed in the baseline FFQ as shown in Figure 9.1. Daily legume intake (g/day) was estimated by multiplying frequency data for the 10 legume FFQ intakes by standards portion sizes as described in chapter 6.

Intakes were divided into dried and fresh legumes based on the usual cooking methods of consumption. Dried legumes were those that were generally stored in a dried state, and then consumed after long boiling or canning such as lentils, chickpeas, baked beans and red kidney beans, whereas the fresh group was mostly consumed as fresh or frozen vegetables and sometimes canned, such as peas and green beans. As shown in Table 9.1, 8 dried and 2 fresh legumes were studied. A list of food items can be seen in Table 9.1. Women were classified into tertiles on the basis of total legumes consumption: low legumes intake group; medium legumes intake and high legumes intake group.

Figure.9.1 Frequency of consumption of selected legumes from baseline FFQ



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
PULSES (Include when used in recipes)										
Lentils, dals	0	1	2	3	4	5	6	7	8	9
Chick Peas, Chanas	0	1	2	3	4	5	6	7	8	9
Hummus	0	1	2	3	4	5	6	7	8	9
Baked beans	0	1	2	3	4	5	6	7	8	9
Mung Beans & Red Kidney Beans	0	1	2	3	4	5	6	7	8	9
Bean Sprouts	0	1	2	3	4	5	6	7	8	9
Black Eyed Beans	0	1	2	3	4	5	6	7	8	9
Butter Beans/Broad Beans	0	1	2	3	4	5	6	7	8	9

Table.9.1 list of legumes considered for analyses

Types of legumes	Food items
Dried legumes	Lentils Blacked eyed beans Chickpeas Hummus Baked beans Mung beans and red kidney beans Bean sprouts Butter beans and broad beans
Fresh legumes	Peas Green beans

9.2.5Statistical analyses

Logistic regression analysis was used to determine the relationship of baseline legume intake to risk of incident T2DM. As described in chapter 8 (model development section), three models were run, unadjusted, age-adjusted logistic

model, finally a multivariable-adjusted model, which additionally included, smoking status, socioeconomic status (professional, intermediate and manual), alcohol intake (g/day), physical activity level, weight change and family history of diabetes. Logistic regression analyses were expressed in odds ratios and 95% confident intervals (95%CI) were carried out. The aim was to examine the odds ratios of T2DM with the dietary intakes of total, dried and fresh legumes in the UKWCS. An estimated effect was calculated with every 40g increment of dietary legumes which is equivalent to a half portion of the UK five a day dietary recommendation (National Health Services).

Descriptive analysis of legumes consumption was carried out for diabetic and non-diabetic women. Comparisons of dietary, lifestyle, socio-demographic and other relevant characteristics among women with low, medium and high legume intakes were presented using analysis of variance (ANOVA) for continuous variables and the chi-square test for categorical characteristics after normality tests were undertaken. Odds ratios of T2DM were determined for dried, fresh and total legumes intakes across tertiles and women in low legumes consumption group considered as the reference.

9.3 Results

9.3.1 Dietary legumes intakes among British women

The median intake of dried legumes was 31g/day and the median fresh legume intake was 21g/day. The overall daily median total legume intake was 60.7g among the studied population (Table 9.2). Intake of dried legumes was significantly higher among women free from diabetes compared with women that developed diabetes, while fresh legume and subsequently total legume intakes did not differ significantly between both groups (Table 9.3).

Table 9.2 Legumes intake (g/d) of cohort women

Food intakes	Legumes intake (g/day)	
	Mean (SD)	Median (IQR)
Dried legume	41.8(39.6)	31(34.1)
Fresh legumes	30.8(26.7)	21(31.6)
Total legumes	72.6(50.7)	60.7(55.3)

Table 9.3 Mean legume intakes among diabetes and non-diabetes women

Food intakes g/day	Diabetic cases	non-diabetes cases	P
Dried legumes	23.2(35.3)	30.2(34.8)	0.02
Fresh legumes	26.8(40.9)	23.2(30.3)	0.08
Total legumes	58.1(58.9)	60.6(55.0)	0.50

All values were geometric means (IQR)

Characteristics of 12,096 cohort participants by total legumes, dried legumes and fresh legumes intakes are shown in Table 9.4.

Overall, women with high dried legume intakes tended to be younger, less likely to gain weight, more likely to have lower BMI, more likely to report being vegetarian, more likely to be from the highest social class and more likely to be supplement users. Mixed picture of women characteristics in the high fresh legume consumption group can be seen. Women who consumed high fresh legumes tended to be older, with slightly higher BMI, more likely to gain weight but more likely to be non-smokers and engage in physical activity.

Table9.4 Characteristics of women by tertiles of intake of total legumes, dried legumes and fresh legumes

Variables		Dried legumes intake			Fresh legumes			Total legumes		
		1st	2nd	3rd	1th	2th	3rd	1st	2nd	3rd
Number of participants (no.)		3,747	4,017	4,218	3,927	4,234	3,821	3,710	4,173	4,099
Number of T2DM cases (no.)		55	22	37	30	42	42	38	33	43
Legumes intake										
Range (g/d)		<21	21-42	>42	<15	15-37	>37	<46	46-80	>80
Median (IQR) intake		12.5(10.4)	31.2(9.4)	67.6(47.1)	10.5(7.4)	21.0(14.2)	60.0(18.0)	31.7(17.0)	61.5(16.7)	113.1(47.7)
Age (years)		53.6(9.0)	51.7(8.7)	50.0(8.5)*	51.2(8.9)	51.2(8.6)	52.7(8.9)*	52.6 (9.0)	51.7(8.7)	50.0(8.5)*
BMI (kg/m^2)		24.2(4.4)	23.9(3.8)	23.9(4.5)*	23.8(4.3)	24.1(4.0)	24.1(4.5)*	23.9(4.3)	24.0(3.9)	24.1(4.5)
Weight change (kg)[2]		7.2(9.4)	7.1(9.4)	6.7(9.6)	6.6(9.5)	7.2(9.2)	7.1(9.6)*	6.8(9.3)	7.1(9.4)	6.9(9.7)
Self-reported vegetarian (%)		16	28	53*	34	32	33	21	31	46*
Smoker (%)		9	8	9	10	8	8*	10	8	8*
Socio-economic status (%)	Professional	65	67	69*	68	67	66	67	68	67
	Intermediate	27	26	23	25	26	26	26	25	25
	Routine	7	7	8	7	8	8	7	6	8
	No education	14	12	12	12	12	13*	13	11	13*
Educational level (%)	O-level	31	29	29	28	30	31	29	29	31
	A-level	26	27	25	25	26	27	26	27	26
	Degree	28	31	34	34	31	29	32	33	30
Dietary supplement use %		58	60	64*	59	59	62*	57	59	63*
AOAC- fibre (g/day)		33.1(12.1)	38.2(13.1)	48.3(15.7)*	34.3(13.1)	39.2(13.3)	47.1(16.3)*	31.4 (11.3)	38.4 (12.0)	49.7 (15.7)*
Legumes AOAC-fibre (g/d)		1.5(0.9)	2.8(0.9)	6.2(2.9)*	2.4(2.3)	3.4(2.3)	4.8(3.0)*	1.4(0.6)	2.8(0.6)	5.6 (2.9)*
Carbohydrate intake (%E)		53.8(6.5)	54.4(6.0)	56.0(6.2)*	55.0(6.5)	54.5(6.1)	54.9(6.4)*	54.2(6.2)	54.4(6.1)	55.8(6.3)*
Protein intake (%E)		16.1(2.7)	15.7(2.6)	15.1(2.5)*	15.4(2.6)	15.6(2.6)	15.7(2.6)*	15.8(2.7)	15.6(2.6)	15.3(2.6)*
Fat intake (%E)		33.6(5.9)	33.4(5.6)	32.5(5.7)*	33.1(6.0)	33.3(5.6)	33.0(5.7)*	33.5(5.9)	33.5(5.5)	32.5(5.8)*
Saturated fat intake (%E)		12.1(3.4)	11.5(3.1)	10.4(2.9)*	11.5(3.3)	11.4(3.1)	11.0(3.1)*	12.0(3.4)	11.5(3.0)	10.5(3.0)*
Energy intake (kcal/d)		2067(597)	2245(623)	2506(684)*	2081(606)	2286(630)	2479(691)*	2012(585)	2252(588)	2553(693)*
Alcohol intake (unit/d)[1]		0.45(1.13)	0.49(1.16)	0.49(1.12)*	0.43(1.08)	0.49(1.15)	0.50(1.20)*	0.45(1.12)	0.51(1.17)	0.47(1.11)*
Family history of diabetes (%)		6.6	6.6	6.2	6.0	6.7	6.6	6.4	6.6	6.5
Physical activity (METs)		16.1(10.6)	16.7(11.0)	18.1(11.5)*	15.4(10.4)	16.9(11.0)	18.8(11.5)*	15.4 (10.2)	16.8 (10.9)	18.7 (11.7)*
Dietary patterns	meat group	84	70	46*	65	66	66*	78	68	53*
	fish group	8	14	22	16	15	15	11	14	20
	vegetarian	8	16	32	20	20	19	11	18	27

All values mean (SD) unless stated otherwise * ANOVA p<0.01, [1] log transformation preformed and values expressed as geometric mean (IQR). [2]Weight changes was calculated by subtracting weight (Kg) at age of 20 years from the present weight (Kg).

9.3.2Legumes intake and the risk of T2DM

From Table 9.5, Odds ratios of T2DM (95%CI) were determined for the intakes of total, dried and fresh legumes. With every 40 g/day increase in dried legumes, the risk of T2DM reduced by 23% (OR=0.77, 95%CI: 0.61, 0.99) in the unadjusted model. However, this inverse association was attenuated after adjustment for age, total energy intake, smoking status, weight changes in percentage, socioeconomic status, alcohol intake, physical activity and family history of diabetes. No significant associations were found between intakes of total legumes intake or fresh legumes with the risk of T2DM in age and fully adjusted models.

Table 9.5 Association between the incidence of T2DM and intake of total legumes, dried and fresh legumes (g/d) for 12,096 women

Nutrient	Unadjusted OR 95%CI	*P*	Age adjusted OR 95%CI	*P*	Fully adjusted OR[1] 95%CI	*P*
Dried legumes 40 g/day[2]	0.77(0.61, 0.99)	0.04	0.83(0.65, 1.06)	0.14	0.86(0.67, 1.10)	0.24
Fresh legumes 40 g/day[2]	1.22(1.00, 1.12)	0.04	1.16(0.94, 1.42)	0.14	1.11(0.90, 1.38)	0.29
Total legumes 40g/day[2]	0.95(0.82, 1.12)	0.60	0.60(0.83, 1.14)	0.76	0.98(0.83, 1.15)	0.85

[1] Adjusted of age, total energy intake, smoking status, weight changes in percentage, socioeconomic status, alcohol intake, physical activity and family history of diabetes. [2]40grams per day equivalent to half portion of 5 a day fruit and vegetables(National Health Services).

Across the dried legumes intake tertiles (Table 9.6), significant inverse associations were observed between the risk of T2DM and dried legumes intake independent of age, total energy intakes, weight changes, smoking, socioeconomic status, alcohol intakes and family history of diabetes with OR = 0.85(95%CI: 0.52, 0.84, p for trend= 0.03). Intakes of fresh legumes and total legumes showed no significant association with the risk of T2DM in all models. Further adjustment of total dietary fibre and magnesium intake did not alter the association between the risk of diabetes and legumes intake.

Table 9.6 Odds ratios and 95% confidence intervals of T2DM according to tertiles of total, dried and fresh legumes intakes among 12,096 British women

Legume intake			T2DM	OR (95%CI)	
Type	Tertile	Range g/day	Cases	Age adjusted model	Fully adjusted model[1]
Dried legumes	1	< 21	55	1	1
	2	21 - 42	22	0.41 (0.25, 0.68)	0.49 (0.29, 0.84)
	3	≥ 42	37	0.72 (0.47, 1.12)	0.85 (0.52, 0.89)
	P value[2]		114	<0.01	0.03
Fresh legumes	1	< 15	30	1	1
	2	15– 37	42	1.31 (0.82, 2.1)	1.41 (0.84, 2.37)
	3	≥ 37	42	1.29 (0.80, 2.1)	1.51 (0.89, 2.59)
	P value[2]		114	0.46	0.27
Total legumes	1	< 46	38	1	1
	2	46 – 80	33	0.82 (0.51, 1.31)	0.89 (0.53, 1.49)
	3	≥ 80	43	1.12 (0.71, 1.74)	1.33 (0.80, 2.22)
	P value[2]		114	0.41	0.25

[1]adjusted for age, total energy intake, weight change %, smoking status, socioeconomic status, alcohol intake, METs and family history of diabetes. [2]Overall P-value.

9.4 Discussion

The current study is one of the first studies to examine the effect of different types of legumes on the risk of T2DM in the UK women. This adds to the knowledge base regarding the estimated effect of different types of legumes on the risk of T2DM. In general, the mean intake of legumes among the studied population was relatively high compared with mean intake among women from the NDNS report (Henderson *et al.*, 2002). This gives a good opportunity to explore the effect of legume intakes in relation to the risk of T2DM. However this may limit the generalization of the results, as the percentage of vegetarians in the UKWCS is higher than UK population (28% vs. 2%) (Bates *et al.*, 2011).

The results point to a potentially protective effect of high legumes intake on the development of T2DM, mainly through consumption of dried type of legumes but not the high fresh legumes intake. Women who consumed at least a half portion per day of dried legumes based on the UK recommendation (National Health Services) have a significantly lower risk of T2DM (15%) independent of potential confounders. non-significant estimated effect of total legumes intake on the risk of T2DM was observed in the current analysis.

Current study has shown an non-significant dose-response relationship between intake of dried legumes and risk of T2DM while significance protective effect when the comparison was made between extreme quintiles. This may relate to wide variation of dried legumes consumption between women across quintiles. It was found that high proportion (40%) of women who consume 1-2 portions per day of dried legumes in the 3rd tertile in addition to 40% of the women in the 1st tertile

consume less than a quarter of portion per day (0-10g/day) and 5% of women in 1st tertile were non-consumers.

These findings are supported by two previous epidemiological studies (Feskens *et al.*, 1991, Villegas *et al.*, 2008). Among men (n=338), in a 20-year follow up of the Finnish and Dutch cohorts of the Seven Countries study, legume and vegetable intakes were inversely associated with 2-h glucose level after adjustment for age, BMI and energy intake (Feskens *et al.*, 1995) and this was supported by an earlier study (Feskens *et al.*, 1991) of 175 men and women which showed that increased legume intake was associated significantly with lower incidence of glucose intolerance measured by using oral glucose tolerance tests after 4 year of follow up (OR=0.4, 95%CI: 0.18, 0.90).

The Shanghai Women's Health Study, also examined the association between total legumes intake and risk of diabetes among Chinese women and found a protective effect of total legumes intake with a multivariate adjusted relative risk of 0.62 (95%CI: 0.51, 0.74). These authors also reported that intake of other legumes (other than soybeans) of mean 37.1g/day in comparison to mean 5.6g/day intake associated with a relative risk of 0.76 (95%CI: 0.64, 0.90) independent to age, total energy intake, WHR, BMI, smoking, alcohol intake, vegetable intake, dietary fibre, physical activity, income level, educational level, occupation, and hypertension (Villegas *et al.*, 2008). In contrast, in the Nurses' Health Study (Bazzano *et al.*, 2008), no association was found between legume intake and the risk of diabetes among women aged 38-63 years followed for 18 years. However, the median intake of legume was reported as 0.17 servings/day (median intake in Q1=0.07 and Q5=0.45 servings/day).

A possible explanation for the current findings might be that women who consume high dried legumes were more likely to adhere to a healthier lifestyle as they were more likely to be vegetarians and more likely to have lower BMI. In addition to this, they tended to be younger and from the highest social class. These factors were previously shown to be linked to lower risk of T2DM. Thus the potential influence of factors that link with women who consumed high amounts of dried legumes may still be possible because of residual confounders' effects that are unavoidable despite inclusion of these variables in the model.

One of the main strength in the UKWCS is the presence of high proportion of self-reported vegetarians that allows assessing the risk of disease (T2DM). The

amount of dried legumes consumed was higher than the amount of fresh legume consumed (41.8g/day vs. 30.8g/day; respectively), this may be adequate to show the beneficial effect of dried legumes on the risk of T2DM. The mean intake of total legumes in the current study is even higher than mean intake of total legumes for the Chinese women in the previous study (35.9g/day) which also showed a significant protective effect among high legumes consumers (Villegas *et al.*, 2008).

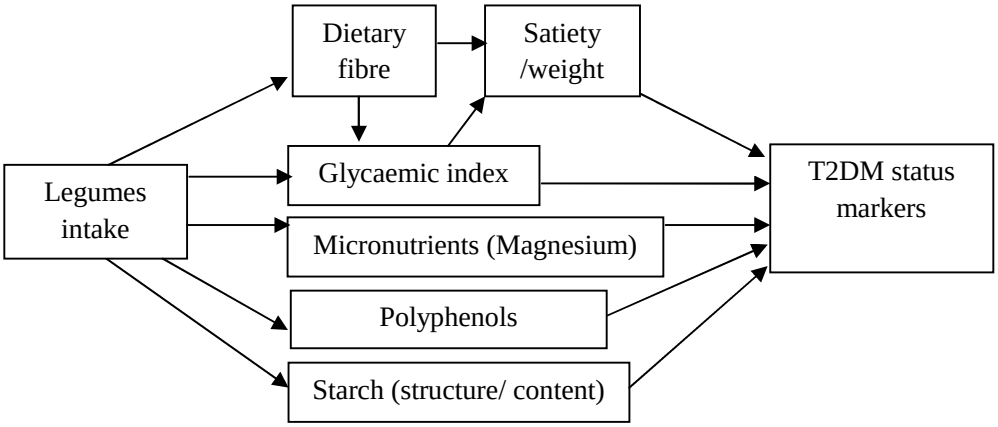
Legume may have an impact on diabetes through number of potential mechanisms. These include acting or via fibre content, starch and type of starch, impact of weight management via mechanisms associated with satiety and polyphenols contents. Flow chart 9.2 illustrates potential plausible mechanisms through which high intake of dried legumes could reduce markers of diabetes status.

There is a good biological plausibility from trials data with legume consumption manipulated with clear effect on measured blood glucose levels. Jenkins (1980) demonstrated that dried legumes among normal participants have the lowest postprandial glucose level in comparison to other types of food (e.g. grains). The presence of dietary fibre and types of starch, that is highly resistant to human enzymes digestion were suggested compounds that may result into a slow rate of intestinal digestion and absorption. Therefore, may have a role in regulate blood glucose. The low glycemic index characteristics of legumes was also suggested to have a role in improving glucose response (Jenkins *et al.*, 1981). It was reported that diet rich in grain, fruit, vegetables and legumes has an anti-inflammatory effect. Observational evidences summarized by Galland (2010) showed that C-reactive protein, a biomarker of inflammation, was significantly reduced by greater intake of dietary fibre. Also impact of dietary fibre on satiety was reported in previous review, where the plausible mechanism related to its viscosity and bulking properties (Slavin and Green, 2007). In addition to, dietary fibre effect on satiety through gut hormones was reported. Ghrelin[14] and glucagon-like peptide[15] hormonal levels were altered with greater consumption of dietary fibre, which showed to affect subsequent meal consumption (Weickert and Pfeiffer, 2008).

[14] Ghrelin is peptide hormone which stimulate hunger and produced by gastric mucosal cells and pancreatic cells (Shintani *et al.*, 2001).

[15] Glucagon like peptide is secreted by intestinal epithelial cells dependents on the presence of food in the lumen of the small intestine which has antihyperglycemic effect, enhanced satiety and reduced energy intake (Flint *et al.*, 1998).

Figure 9.2 Suggested mechanisms through which high intake of legumes could reduce T2DM status markers



Pooled results from 10 randomized controlled trials showed that intake of legumes ranging between 90 – 377g/day for a period of at least 1 week up to four months resulted in significant reduction in fasting blood glucose (−0.82, 95% CI −1.36 to −0.27) and fasting blood insulin (−0.49, 95% CI −0.93 to −0.04) among normoglycemic and diabetic participants (Sievenpiper *et al.*, 2009). This evidence strongly supports the beneficial effect of dried legumes such as chickpeas on fasting blood glucose levels.

Dried legumes contain higher amount of starch in comparison to fresh legumes (Food Standards Agency, 2002). For example, cooked red kidney beans contain 12-14% starch while boiled green beans contain 3% of starch. Type of starch structure and other characteristics in legumes compared with cereal starch was demonstrated in previous review (Guillon and Champ, 2002). It was reported that the low digestibility of legume starch in compared to cereal starch is likely due to the high amylose content in legumes which have high capacity to retrograde than starch with low amylose content. Experimental studies measuring starch in lentils and corn reported a higher proportion of slow digestible starch and resistant starch in cooked lentil compared to corn, while higher proportion of rapid digestible starch was found in corn than lentil (Chung *et al.*, 2009).

Legume fibre contains heterogeneous dietary fibre with varied chemical structures (cellulose, hemicelluloses, pectin, gums, mucilage, resistant starch and lignin). In general, delayed gastric emptying, prolongation in small intestine absorption and production of short chain fatty acids post fibre fermentation are all physiological effect of fibre that have a positive effect on glycemic response

(Jenkins *et al.*, 1987). This is supported by some epidemiological studies showing an inverse association between total dietary fibre intake and risk of T2DM (Schulze *et al.*, 2007). However, findings from the systematic review in chapter 4, demonstrate that prospective studies which examined legumes fibre on the risk of diabetes did not support the beneficial effect of legumes fibre intake on the risk of diabetes which was explained by authors due to low intake of legumes fibre among the studied populations. Similarly, the findings in chapter 8, where no effect of legumes fibre on the risk of T2DM in the UKWCS was seen even though the large proportion were vegetarians and consumption of legumes was high.

Indirect effect of dietary fibre on the risk of diabetes is one of the possible explanations. The risk reduction of T2DM with dried legumes intake can also be explain by the relationship between the beneficial effect of high fibre diet on decreasing satiety which may act through body weight reduction and consequently be linked to a lower risk of T2DM (Weickert and Pfeiffer, 2008). However evidence on legumes fibre is still lacking.

Dried legumes also contain variable amounts of polyphenolic compounds (1.53 – 6.56 mg gallic acid equivalents g^{-1}. For example, red kidney beans where dark pigmentation is because of the presence of high polyphenolic substances) (Oomah *et al.*, 2011). Negative correlation between polyphenols in legumes and blood glucose response was reported previously (Thompson *et al.*, 1984). This may have a role in diabetes prevention.

A recent comprehensive review on bioactive compounds hypothesized a potential health-related protective mechanism in whole grain cereal other than fibre components (Fardet, 2010). Dried legumes are a rich source of micronutrients such as magnesium which was found to have an inverse association with T2DM risk in prospective studies (Dong *et al.*, 2011). Intracellular magnesium has a key role in insulin action and found to be low in concentration among T2DM which suggested to be possible mechanism that may explain the inverse association between risk of T2DM and dietary Magnesium (Barbagallo *et al.*, 2003). Some cross-sectional studies have reported serum magnesium and intracellular magnesium were significantly lower in pre-diabetic-patients in comparison to healthy individuals (Lima *et al.*, 2009) as well as in T2DM participants (Resnick *et al.*, 1993). This is also supported by evidence from interventional studies which observed a significant reduction in fasting blood glucose post high magnesium supplementation (Song *et*

al., 2006). Additionally, magnesium supplementation for 3 months reduced insulin resistance significantly in non-diabetic participants (Guerrero-Romero *et al.*, 2004). However, prospective studies on the dietary magnesium intake in different sources particularly legumes in relation to the risk of T2DM is still limited and required further research.

Evidence from the National Health and Nutrition Examination Survey for American adults aged $\geq$ 19 years found that consumption of dry beans and peas such as kidney beans, black beans, chickpeas, lentils and other types resulted in a nutrient profile that was higher in fibre, protein, folate, zinc, iron, and magnesium and lower in saturated fat and total fat in comparison to non-consumers (Mitchell *et al.*, 2009). This may reflect the underlying reason of high quality diet with high intake of dry legumes.

In the current study the mean (SD) intakes of dietary fibre, zinc, iron and vitamins from dried legumes were higher in comparison to fresh legumes intake. These compounds may have a link to T2DM directly or indirectly. However this is quite complicated and need further research to examine the association between dietary bioactive compounds from legumes on health related outcomes prospectively in populations with high legumes consumption.

Further research is required to evaluate whether different types of legumes have health beneficial effects which may help in supporting the current dietary recommendations. It have been reported in earlier UK survey that baked beans more likely be eaten by men than women (48% and 41% respectively) (NDNS, 2002). This may suggest whether the association between legumes intake and risk of T2DM differ between men and women which required further research on men. Another potential future research in the UKWCS data is to split vegetable intake and to examine different types of vegetable intake and polyphenols contents in vegetables and legumes that may have an impact on the risk of T2DM. Further prospective evidence on the dose-response relationship between legumes intake and risk of diabetes will help in identify the potential amount of legumes that has beneficial effect on developing T2DM.

9.5 Conclusion

Due to limited prospective evidence in legumes intake area, the main purpose of the current study was to determine the effect of legumes intake on the risk of T2DM among women. The above findings indicate that diets with high intakes of

dried legumes may be associated with a decreased risk of T2DM which is supported by previous prospective studies. This suggests that dried legumes have potential beneficial nutritional properties (e.g. higher fibre, magnesium and polyphenols) in relation to diabetes prevention. Because dried legumes represent a small market in the UK, this information could be used to develop targeted interventions aimed at increasing dried legumes intake and to examine health related outcomes.

Chapter 10: Dietary fibre intake obtained from different dietary methods in the UKWCS.

10.1 Introduction

Chapter 4 showed non-significant association between total dietary fibre intake and main fibre sources intakes with the risk of T2DM after adjustment for potential confounders. Whether these non-significant findings were partly related to biases inherent within the dietary assessment methods used described in this chapter. Measurement of diet is usually subjected to misreporting (Willett, 2012). The FFQ approach is a common dietary assessment method used in studies with a large sample sizes such as cohort studies because it is cheap, easy to deliver and simple to process and analyse in comparison to other methods such as food diaries. On the other hand, FFQs depend on individual memory and thus poor recall is likely to happen. Other sources of bias of FFQs are the composition of the fixed food item list and the portion size selected (Willett, 2012). An earlier comparison study carried out on 160 women aged 50-65 years who provided dietary information using 16 days of weighed diary records and other different methods such as FFQ, 24hrs recall and 7 days diary, reported that mean dietary fibre intake from 4 day food diary was similar to weighed record (Bingham *et al.*, 1994). This gives the advantage of considering the food diary as the most appropriate reference method for validation studies in nutritional epidemiology research.

Furthermore, food diaries have been recommended as one of the reference methods that can be used when a validation study with FFQ is required because of the source of bias is highly independent (Cade *et al.*, 2004b). Comparison between different dietary methods is important in epidemiological studies, particularly studies which focus on diet-disease relationships. This is because the risk estimate of the disease outcome in relation to nutrient intake should consider FFQs' ability to rank participants rather than provide estimation of absolute dietary intake (Willett, 2012). An earlier UKWCS preliminary validation study reported reasonable correlations between FFQ and 4 day diaries for selected micronutrient intakes (Spence *et al.*, 2002).

Stability of dietary pattern obtained from baseline FFQ was previously assessed by a second FFQ in the UKWCS; half of the studied women maintained the same dietary pattern. However, for nutrient intakes, the kappa statistic ranged between

0.18 to 0.21 for total energy, vitamin C and energy from fat. This suggested poor stability for the selected nutrients (Greenwood *et al.*, 2003). However repeatability of FFQ based on fibre intake was not reported in the study. The current chapter compares daily dietary fibre intake recorded by 4 days diary with dietary fibre intake estimated by baseline and repeated FFQ. To tackle the issue of measurement error that may explain the non-significant findings in chapter 8, comparisons between NSP intakes obtained from different dietary assessment methods permits assessment whether high NSP consumers based on FFQs were similar to high NSP consumers based on food diaries.

10.2 Aims and objectives

- Compare between baseline FFQs and repeated FFQs to describe the extent of reproducibility of baseline FFQs in terms of consistency of measuring dietary fibre by the same person at different times.
- Compare between food diaries and FFQs to examine the relative validity of the FFQ. Evaluate how well the dietary assessment methods agree on determining dietary fibre intake by examining the association between dietary fibre intake obtained from diary and FFQ in the UKWCS. This was done through two main approaches:
  - Assessing the reproducibility of FFQ-fibre intake by comparing the total dietary fibre intakes obtained from food diary and FFQ (either baseline or repeated) in the UKWCS.
  - Comparing main fibre sources intakes obtained from baseline FFQ and food diary.
- Evaluate the ability of the FFQ used in the UKWCS to rank individuals based on dietary fibre intake by examining whether women were allocated similarly into categories based on dietary fibre intakes derived by these different methods.
- Finally, identify the percentage of women who were misclassified into extreme NSP quintiles.

10.3 Methodology

10.3.1 Studied population

Description of the UKWCS is in chapter 6. The time scale of women who completed different dietary assessment methods in the UKWCS is also demonstrated

in Figure 6.2 in chapter 6. Figure 10.1 represents the average differences in time between the completed FFQ at baseline, repeated FFQ and food diary. Time periods were obtained from the mean difference in women's age when completed forms were received. Of the 1,918 women who completed a repeated FFQ, only 20% (384) of them completed a food diary (which had subsequently been fully coded). Food diaries were completed 1.6 years before the repeat FFQ. After 4.8 (SD = 0.4) years on average from the baseline FFQ, food diaries were received. This indicates that the time gap between repeated FFQs and food diaries was shorter than the time gap between baseline FFQs and food diaries.

Figure 10.1 time range between each dietary assessment method used in the UKWCS

Women who did not complete at least 3 days of diary records were excluded from the analysis (9 women). Therefore, the comparison between food diary and repeated FFQ was carried out on 382 women. Finally, the comparison between food diary and baseline FFQ was carried out on 1870 women.

10.3.2 Dietary fibre intake

The daily intake of NSP was derived from 3 dietary assessment sources. The daily NSP intake from the baseline FFQ is referred to as baseline FFQ-fibre intake. Daily NSP intake derived from 4 days semi-weighed food diary is referred to as diary-fibre intake. Finally dietary NSP intake recorded by repeated FFQ is referred to as repeated FFQ-fibre.

10.3.2.1 Food frequency questionnaire

The self-administered FFQ at baseline which was repeated between the year 2000 and the year 2001. Details on baseline and repeated FFQ as well as the dietary NSP intake calculation were demonstrated in chapter 6.

Fibre from cereal, vegetables, fruit and nuts contributed to 95% of total dietary fibre intake in the UKWCS (chapter 7). The main food sources of fibre were obtained based on McCance and Widdowson's food grouping as listed in the

baseline FFQ (Food Standards Agency, 2002) (Table 10.1). Intakes of four fibre sources from baseline FFQ were calculated for each participant (g/day).

Table 10.1 Food items in main fibre sources (categories based on food grouping used by the UK food composition tables (Food Standards Agency, 2002))

Food group	Food item
Cereal and cereal products	White bread average, brown bread average, wholemeal bread average, chapatis, papadums, tortillas, pitta, crispbread, crackers average, breakfast cereals, porridge, sugar cereals, non-sugar cereals, muesli, all bran, branflakes, cream crackers/ biscuit cereals, rice & pasta, white pasta, spaghatti wholemeal pasta, white rice, brown rice, wild rice, macaroni cheese, barley, oats, bulgar wheat, wheat germ, couscous, biscuits, sweets and puddings, plain biscuits, chocolate biscuits, full coated sandwich, cream biscuits, fruitcake, sponge cake, buns/pastries, scones/pancakes/muffins/ croissants, fruit pies, tarts, crumbles, sponge puddings
Vegetables	Beetroot, broccoli/greens, brussels, cabbage, carrots, cauliflower, celery, coleslaw, courgettes, cucumber, garlic, green/runner beans, leeks, lettuce, mushrooms, aubergine, olives, parsnips, peas, peppers, swede, sweetcorn, tomatoes, turnip, watercress, potatoes, jacket potatoes, old potatoes roasted, potato salad, Quorn, Textured Vegetable Protein, veg chilli, mixed bean casserole/ratatouille, stir fry veg, veg dishes, veg pizza, vegetable pate, lentils, chickpeas, hummaus, baked beans, red kidney beans, beansprouts, black eyed beans, butter beans
Fruit	Apple, avocado, bananas, grapes, kiwi, mango, oranges, satsumas, grapefruit, papaya, pears, pineapple, apricot, melon, average nectarines, peaches, plums, raspberries, currants - red/white /black, rhubarb, strawberries, dates, dried, figs, prunes, mixed dried fruit, currants/raisins/sultanas
Nuts and seeds	peanuts/pistachios, cashews, pecans, sunflower seeds, peanut butter, mixed nuts/raisins, nut pate

10.3.2.2 Four days food diary

Food diaries were completed by (n=12,453) cohort women between 1999 and 2002. Details of the food diary format, mailing and coding are in chapter 6. Nutrient intakes were obtained for 15% of women who completed food diary. Full dietary information was available electronically for only 1,883 women that had been fully coded in DANTE program and whose data were available for analysis. Nutrient intakes were calculated as the average of intakes over 3 or 4 days' recordings depending on completion. Fibre from cereal, vegetables, fruit and nuts were also derived from the diary similar to the main food sources identified by the baseline FFQ.

10.3.3 Statistical analyses

10.3.3.1 Total dietary fibre intake

Characteristics of women who completed different dietary methods were explored and these are illustrated in Table 10.3. Distributions of all relevant variables were examined by histograms. Paired t-tests were used to examine whether the mean differences in dietary intakes for selected nutrients and NSP intake

(expressed in g/day and g/1000kcal) reported in FFQs and food diaries were statistically significant.

Initially, correlation coefficients were obtained between FFQs and 4 day food diaries. Pearson's correlation was performed for normally distributed data while Spearman rank correlation was used for not normally distributed data. Because correlation only measures the strength of association, the Bland-Altman method was also used to measure the extent of agreement between two dietary assessment methods. The Bland-Altman plot is a graphical approach to compare FFQ-fibre intake and diary-fibre intake by plotting the difference of two measurements against the mean of both measures for each subject (Bland and Altman, 1986). This method permits assessment of the agreement between two measurements, where the mean difference between then is the estimated bias and the 95% limits of agreement are the mean difference $\pm$ 2 SD.

The ability of the FFQ to rank cohort women was assessed by determining the percentage of women who were correctly classified into the same group and the percentage of women who were grossly misclassified. Therefore, women were categorized into 5 quintiles based on dietary fibre consumption derived from FFQs and diaries, with the first dietary fibre quintile referring to women with the lowest dietary fibre intake and the fifth quintile those with the highest intake. Cross-tabulation with percentages of women classified into same or adjacent and extreme groups of dietary fibre quintiles were obtained for the three comparisons using the Kappa statistic (K) (Landis and Koch, 1977). The aim of this statistical approach is to determine how well the examined dietary methods agreed on allocating women to the same quintile based on dietary fibre intake beyond the probability of chance. Furthermore, the weighed Kappa statistic (K_w) was also obtained to reduce the chance of misclassification by counting adjacent categories as partial agreement (Viera and Garrett, 2005). Kappa values range between 0 to 1 where the value of 0 means no agreement between the tested methods and the value of 1 means that the agreement between the two methods is perfect (by implication therefore, both methods measure exactly the same thing). Specifically, this statistical approach helped to determine whether the women who were categorized in to the high NSP group by FFQs are similar to women who were categorized in to the high NSP group by 4 days diaries records in the UKWCS.

Pearson's correlations and Kappa statistics analysis was carried out to examine whether the agreement between dietary fibre assessment methods differ based on self-reported vegetarian status.

10.3.3.2 Intakes from main dietary fibre sources

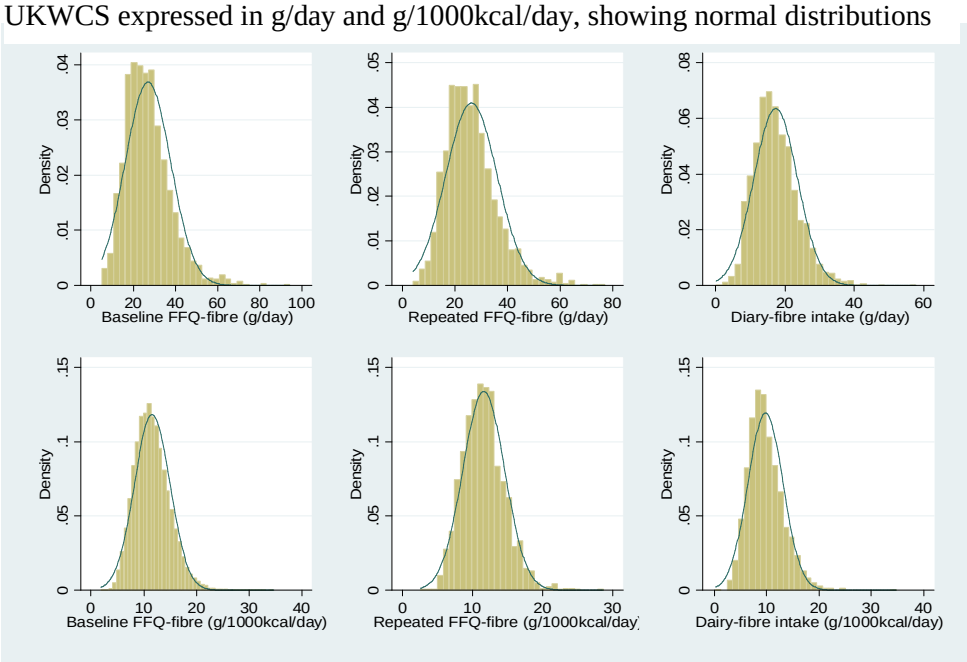
Main fibre sources were also of interest, therefore fibre from cereal, vegetables, fruit and nuts were derived from baseline FFQs and food diaries. For fibre sources analysis, 1,879 women who completed both the food diary and FFQ were available. Comparisons between the four main sources as well as insoluble and soluble fibre intakes derived from baseline FFQs and food diaries were carried out.

Total dietary fibre variables were normally distributed (as shown by the histograms in Figure 10.2), however for distributions that were positively skewed (vitamin C, iron, calcium and folate) the log transformation was carried out before the two samples t-test. If normal distribution was not achieved with log transformation, then a non-parametric test (Wilcoxon sign-rank test) was performed and medians and interquartile ranges (IQR) were stated. Spearman's rank correlations were also carried out when data were not normally distributed. This aimed to measure the strength of association between fibre sources obtained by the two dietary assessment methods. Bland-Altman plots were also performed after log transformation to assess the extent of the agreement between the FFQ-fibre and diary-fibre intakes for different food sources. Women were split into 5 quintiles based on their dietary fibre sources and the degree of agreement between women's categories using two dietary methods was calculated using Kappa statistics. Un-weighed Kappa and weighted Kappa were both calculated. Table 10.2 presents the interpretations of agreement for Kappa as described by Landis and Koch (1977).

Table 10.2 Kappa statistic interpretation

Kappa statistic	Strength of agreement
• 0.00	Poor (no agreement)
• 0.00-0.20	Slight
• 0.21-0.40	Fair
• 0.41-0.60	Moderate
• 0.61-0.80	Substantial
• 0.81-1.00	Almost Perfect

Figure 10.2 Baseline FFQ-fibre, repeated FFQ-fibre and diary-fibre in the UKWCS expressed in g/day and g/1000kcal/day, showing normal distributions



10.4 Results

10.4.1 Studied population

The characteristics of the women were explored and showed that, the percentage of women in BMI categories, smoking status, social class, educational achievement and the dietary supplement use did not differ greatly among women who completed different dietary assessments methods, however women who completed repeated FFQ actually have a lower percentage of vegetarians in compared with women in other dietary methods. This indicates similarity in most of women's characteristics therefore studied samples were relatively comparable (Table 10.3).

Table 10.3 Baseline characteristics of women who completed different dietary assessment methods in the UKWCS

Characteristics of women		Baseline FFQ	Repeated FFQ	Semi-weighed food diary
Number of participants		35,372	1,918	1,883
Age (years)		52(9)	56(9)	54(9)
BMI categories (%)	Underweight	1	2	2
	Normal	62	52	63
	Overweight	24	28	22
	Obese	13	17	13
Smoker (%)		10	7	9
Vegetarian (%)		72	47	70
Professional socio-economic status (%)		66	67	65
Degree achievement (%)		27	36	28
History of any medical illnesses (%)		27	23	31
Dietary supplementation use (%)		58	69	62

10.4.2 Comparisons between baseline FFQ-fibre, repeated FFQ-fibre and diary-fibre intakes in the UKWCS

Mean daily dietary intake of fibre from FFQs were significantly higher by 36% (baseline) and by 28% (repeated) than food diaries with a mean difference of 9.7 and 7.3g/day respectively. The majority of nutrient intakes were also significantly higher in FFQs than diaries. In comparison to baseline FFQs, dietary nutrient intakes from repeated FFQs tended to be statistically significantly lower than baseline FFQs except for total energy intake, vitamin C and iron; however the mean difference was relatively small. From Table 10.4, whilst differences in intakes (due to large sample size) by method were often statistically significant, actual differences were not generally clinically meaningful, particularly comparing the 2 FFQ databases.

Table 10.4 Comparison between dietary intakes of women who completed different dietary assessment methods

Dietary variables All values were mean (SD) unless stated otherwise	Baseline FFQ	Repeated FFQ	Difference	Food diary	Baseline FFQ	Difference	Food diary	Repeated FFQ	Difference
No of participants	1918	1918	-	1870	1870	-	384	384	-
NSP (g/d)	28.4(11.4)	26.3(9.7)	2.1[1]	17.3(6.3)	27(10.8)	-9.7[1]	19.0(6.8)	26.4(8.4)	-7.3[1]
NSP density (g/1000kcal/d)	12.6(3.3)	11.6(2.9)	1.0[1]	9.7(3.3)	11.9(3.3)	-2.1[1]	10.7(3.5)	11.7(2.9)	-1.0[1]
Energy intake include alcohol (kcal/d)	2277(736)	2280(664)	-2.5	1873(448)	2339(721)	-465[1]	1812(409)	2286(626)	-473[1]
Carbohydrate intake (E%)	56(6.3)	55(6.8)	1.0[1]	51(7.9)	55(6.4)	-4.0[1]	52(7.8)	55(6.7)	-2.2[1]
Protein intake (E%)	15(2.7)	14(2.5)	0.6[1]	16(3.5)	16(2.7)	-0.06	15(3.4)	14(2.4)	0.6[1]
Fat intake (E%)	33(6.0)	32(5.9)	1.2[1]	33(6.8)	33(5.8)	-0.07	33(7.1)	32(6.2)	0.8[1]
Saturated fat intake (E%)	11(3.2)	10(3.2)	0.48[1]	12(3.7)	11(3.2)	-0.53[1]	11(3.8)	10(3.3)	0.9[1]
MUPA (E%)	11(2.4)	10(2.3)	0.36[1]	10(2.7)	11(2.3)	-0.4[1]	10(2.7)	10.3(2.3)	-0.3[1]
PUFA (E%)	7(1.8)	6(1.3)	0.3[1]	5(2.2)	6(1.8)	-0.8[1]	5(2.3)	6(1.8)	-0.6[1]
Vitamin C[2]	159.4(99.7)	158.7(96.3)	0.7	108.3(81)	160(92)	-51.7[1]	114(82)	159(96)	-45[1]
Iron [2]	16.6(9.1)	17.2(8.8)	-0.6	12.5(4.8)	17.3(8.4)	-4.8[1]	12.3(4.8)	17.4(8.8)	-5.1[1]
Folate[2]	413(139)	391(122)	21.2[1]	294(111)	408(139)	-114.3[1]	304(125)	398(114)	-93.7[1]
Calcium[2]	1163(406)	1089(385)	73.9[1]	874(295)	1170(393)	-296.2[1]	883(319)	1094(371)	-211.0[1]

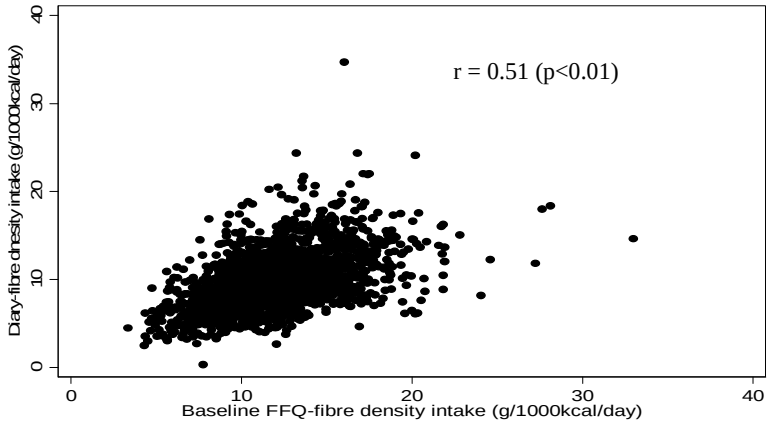
[1] p value<0.01; [2] log transformation carried out and value listed is geometric mean (IQR)

The baseline FFQ was moderately correlated with the repeated FFQ in unadjusted fibre intake (r=0.62, p<0.01) and in energy adjusted fibre intake (r=0.63, p<0.01). Significant correlations were also seen between diary-fibre intake expressed as g/1000kcal/day and FFQs (r= 0.51for baseline FFQ-fibre and r= 0.47 for repeated FFQ-fibre) as shown in Table 10.5. Figure 10.3 demonstrates graphically the positive trend of association between diary-fibre and baseline FFQ-fibre.

Table 10.5 Pearson's correlation coefficients (95%CI) of dietary fibre intake obtained by FFQ and food diary

		Diary-fibre	Baseline FFQ-fibre
Unadjusted correlation (r)	Diary-fibre	1	
	Baseline FFQ-fibre	0.36 (0.32, 40)	1
	Repeated FFQ-fibre	0.37(0.28, 0.45)	0.62(0.59, 0.64)
Energy adjusted correlation (r)	Diary-fibre	1	
	Baseline FFQ-fibre	0.51(0.47, 54)	1
	Repeated FFQ-fibre	0.47(0.39, 0.54)	0.64(0.61, 0.66)

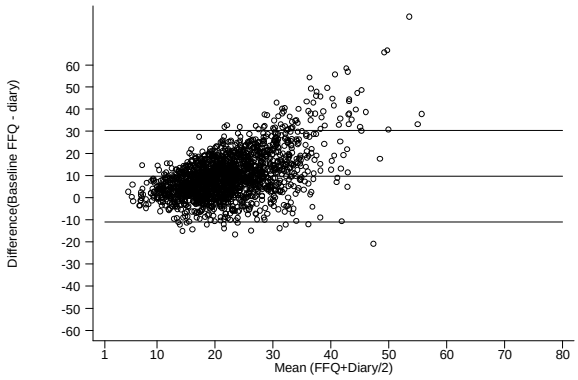
Figure 10.3 Scatter plot between energy adjusted NSP intakes (g/1000kcal/day) estimated by FFQ and food diary



Bland-Altman plots Figure 10.4 and 10.5 show that the differences between FFQ-fibre and diary-fibre increased as the mean of dietary fibre intake from the two methods increased. This illustrates visually there were more large positive differences than large negative ones. Furthermore, there was a tendency for greater positive mean differences with greater mean dietary fibre intake. There was a clear tendency for women to overestimate their dietary fibre intake by FFQ by an average of 9.7g/day in the baseline FFQ and by 7.3g/day in the repeated FFQ with 95% limit of agreement -11.1 to 30.4 g/day and -9.9 to 24.6 g/day respectively in comparison to food diaries. In Figure 10.6 the Bland-Altman plot comparing baseline and repeat

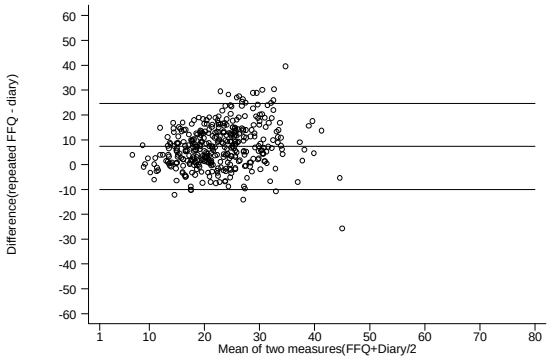
FFQ, shows a smaller average difference (-2.1g/day). However, the 95% limit of agreement was wide ranging between -20.7 to 16.4g/day.

Figure 10.4 Bland-Altman plot between FFQ-fibre and diary-fibre intake in the UKWCS (n=1870)



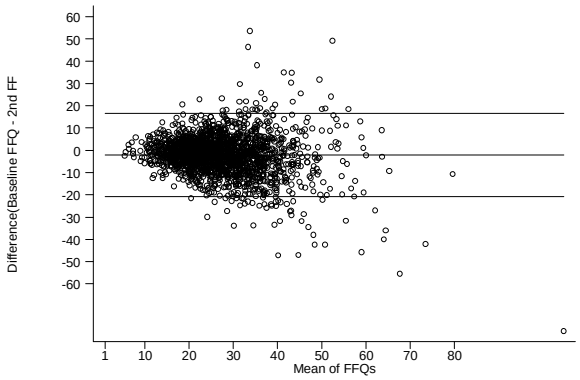
Bland-Altman comparison of baseline FFQ-fibre and diary-fibre intake (g/day)
95% limits of agreement = -11.052 to 30.390
Mean difference: 9.669 (95%CI: 9.19, - 10.139)

Figure 10.5 Bland-Altman plot between FFQ-fibre and diary-fibre intake in the UKWCS (n=382)



Bland-Altman comparison of repeated FFQ-fibre and diary-fibre (g/day)
95% limits of agreement = -9.932 to 24.621
Mean difference: 7.345 (95%CI: 6.475 to 8.214)

Figure 10.6 Bland-Altman plot between FFQ-fibre and diary-fibre intake in the UKWCS (n=1918)



Bland-Altman comparison of baseline FFQ-fibre and repeated FFQ-fibre (g/day)
95% limits of agreement =-20.793 to 16.488
Mean difference: -2.153 (CI -2.570 to -1.735)

Furthermore, women were divided into 5 equal sized categories based on their dietary fibre intake (from high to low dietary fibre intake) obtained from FFQs and food diaries and the Kappa statistic was used to assess the extent of agreement between dietary fibre groups. Table10.6 presents the percentages of the agreements and Kappa values obtained for diary-fibre intakes and repeated-fibre intakes, diary-fibre and baseline fibre intakes as well as baseline-fibre and repeated-fibre intakes. The agreement tended to be fair between the different methods of recording dietary fibre intakes.

The weighed Kappa value (K=0.45) between baseline-fibre and repeated-fibre (Classified as moderate agreement by Viera and Garrett (2005)) reflects the high repeatability of FFQ despite the average 3 years gap in completion.

On the other hand, the degrees of agreement were found to be fair between diary-fibre and FFQ-fibre intake (with baseline-fibre, K_w= 0.25 and with repeated-fibre, K_w=0.22). Dietary fibre density showed slight improvement in Kappa values as well as in percentage of agreement. It has been reported that Kappa values are affected by the number of categories (Willett, 1998). However, in the same population, with a lesser number of dietary fibre groups (3 and 4 groups), kappa values remained similar to values for 5 fibre groups, as shown in Table 10.6.

Table 10.6 Comparison between dietary assessment methods using Kappa and weighed Kappa statistics

Comparisons	Dietary fibre (g/day)			Dietary fibre density (g/100kcal/day)		
	K (%)[1]	K_w (%)[2]	Agreement	K (%)[1]	K_w (%)[2]	Agreement
Diary-fibre vs. BL-FFQ-fibre[3]	0.12 (29%)	0.25 (70%)	Slight to fair	0.16 (33%)	0.35 (74%)	Slight to fair
Diary-fibre vs. R- FFQ-fibre[4]	0.10 (29%)	0.22 (69%)	Slight to fair	0.15 (33%)	0.32 (73%)	Slight to fair
BL-FFQ-fibre[3] vs.R-FFQ-fibre[4]	0.25 (40%)	0.45 (78%)	Fair to moderate	0.22 (38%)	0.42 (77%)	Fair to moderate

[1]kappa statistic (% women classified into same quintile), [2]weighed Kappa statistics (percentage of women classified into same or adjacent quintile), [3]BL-FFQ-fibre is fibre intake from baseline FFQ, [4]R-FFQ-fibre is fibre intake from repeated FFQ

Similar findings (Table 10.6) were seen when Kappa statistic was carried out after excluding women who changed their diet at the baseline.

One of the main aims of this chapter was to assess whether the FFQ was reasonably able to rank women in similar groups in comparison to the reference method (food diary). Tables 10.7, 10.8 and 10.9 showed low percentages of women who were grossly misclassified. Baseline and repeated FFQs classified 82% and 84% of women into the same or across 2 adjacent quintiles of dietary fibre intake

and only 2% of women were misclassified into extreme quintiles in comparison with the food diary method. Around one third of women were allocated to the same quintile in the diary and FFQ comparisons. As expected, higher percentage of women (40%) were correctly classified in the same quintile with the FFQ comparisons (Table 10.10).

Table 10.7 Number of participants across dietary fibre intake quintiles obtained by food diary and repeated FFQ

		4 food diary					
		Lowest NSP	2nd	3rd	4th	Highest NSP	Total(n)
Repeated FFQ	Lowest NSP	21	20	7	11	5(1.3%)[1]	64
	2nd	11	14	16	24	18	83
	3rd	8	17	12	12	20	69
	4th	5	12	18	30	29	94
	Highest NSP	3(0.7%)[1]	13	8	17	33	74
	Total (n)	48	76	61	94	105	384

[1] n (%) of women who were misclassified into extreme quintiles in comparison to total number of participants

Table 10.8 Number of participants across dietary fibre intake quintiles obtained by the baseline FFQ and food diary

		4 food diary					
		Lowest NSP	2nd	3rd	4th	Highest NSP	Total(n)
Baseline FFQ	Lowest NSP	115	91	60	48	20(1.1%)[1]	374
	2nd	84	91	83	71	46	375
	3rd	66	83	83	87	56	375
	4th	46	57	71	89	109	372
	Highest NSP	24(1.3%)[1]	52	78	78	142	374
	Total (n)	375	374	375	373	373	1870

[1] n (%) of women who were misclassified into extreme quintiles in comparison to total number of participants

Table 10.9 Number of participants across dietary fibre intake quintiles obtained by the baseline and repeated FFQ

		Repeated FFQ					
		Lowest NSP	2nd	3rd	4th	Highest NSP	Total(n)
Baseline FFQ	Lowest NSP	158	58	18	9	6(0.3%)[1]	249
	2nd	121	126	62	37	11	357
	3rd	60	104	110	83	29	386
	4th	30	69	114	120	83	416
	Highest NSP	15(0.78%)[1]	27	78	135	254	510
	Total (n)	384	284	383	384	383	1,918

[2] n (%) of women who were misclassified into extreme quintiles in comparison to total number of participants

Table 10.10 Cumulative percentage agreement between diary fibre intakes derived from FFQ at baseline and repeated, and 4 days food diary

Percentage of agreement	Baseline-fibre vs. Repeated-fibre	Diary-fibre vs. Repeated-fibre	Diary-fibre vs. Baseline-fibre
Exact	40%	28%	29%
Within 1 adjacent quintile	80%	65%	66%
Within 2 adjacent quintile	97%	85%	84%

10.4.3 Comparisons between main fibre sources and fibre fractions obtained from baseline FFQ and food diary

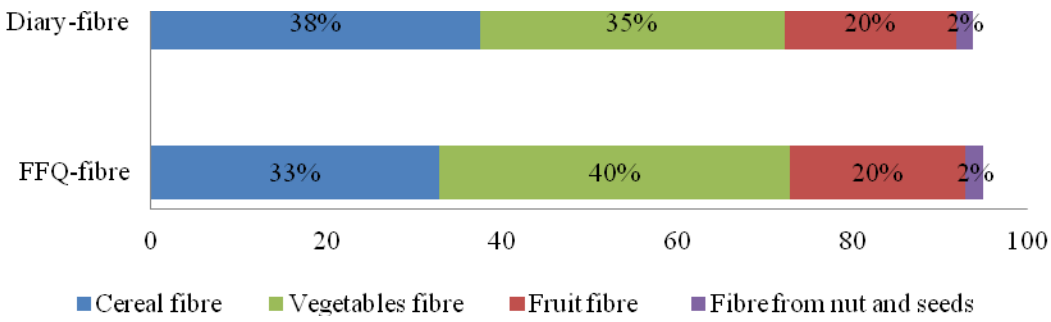
Cereal, fruit, vegetables and nuts were found to be the main contributors of dietary fibre intake in chapter 7. The main sources of fibre as well as the soluble and insoluble fibre derived by FFQ and food diary are shown in Table 10.11. Four food groups contributed to 95% and 94% of the total dietary fibre intake obtained by baseline FFQs and food diaries. As demonstrated in Figure 10.7, the main food source of diary-fibre was the cereal group, while the main fibre source of FFQ-fibre was the vegetables group in the studied population. For all sources, dietary fibre intake obtained from the baseline FFQ was significantly higher than diary-fibre intake ($p<0.01$) as shown in Table 10.11. Fibre from vegetables and fruit showed the second highest percentage of differences between the FFQ and diary methods (50% and 38%) after fibre from nuts and seeds (100%). The lowest percentage of difference was fibre from cereal (25%).

Table 10.11 Intakes of main fibre sources derived by baseline FFQ and food diary in the UKWCS (n=1870)

Fibre source	Dietary fibre intake (g/day)			
	FFQ baseline	4 day diary	% of difference	p value
Cereal fibre	7.2(6.8)	5.4(4.6)	25%	<0.01
Vegetables fibre [1]	9.8(6.0)	5.2(3.7)	50%	<0.01
Fruit fibre	4.2(4.1)	2.6(2.8)	38%	<0.01
Fibre from nut and seeds[2]	0.2(0.45)	0 (0.35)	100%	<0.01
Soluble dietary fibre[3]	11.3(4.3)	6.4(2.4)	44%	<0.01
Insoluble dietary fibre[3]	17.0(7.3)	8.3(3.5)	52%	<0.01

All values are geometric mean (IQR) unless stated otherwise, [1]Vegetables include legumes and potatoes; [2] median (IQR); [3]mean (SD)

Figure 10.7 Percentage of the main dietary fibre intake contributors using baseline FFQ and food diary



All correlation coefficients between fibre sources and types of fibre estimated by food diaries and both FFQs were found to be statistically significant ($p<0.1$). The highest correlations were observed between baseline FFQ and diary

methods for cereal fibre, fruit fibre and fibre from nuts and seeds while vegetable fibre showed a weak but significant correlation between both dietary assessment methods.

High number of missing values in the UK food composition tables for soluble and insoluble fibre which mainly affects diary-fibre intake, while for baseline FFQ fibre fractions were generated for missing values in the baseline FFQ data as shown in chapter 6, therefore, further analysis of insoluble and soluble fibre intakes between baseline FFQ and diary fibre intakes was not performed.

Table 10.12 Pearson's correlation coefficients (r) and 95%CI between total NSP, dietary fibre fractions and fibre sources determined by baseline FFQ and food diary among 1870 women

Dietary fibre intake	Correlation coefficient (95%CI)	Description
Total NSP	0.36(0.32, 0.40)	Moderate correlation
NSP density	0.50(0.47, 0.54)	Moderate correlation
Cereal fibre[1]	0.42(0.38, 0.46)	Moderate correlation
Vegetables fibre[1]	0.26(0.23, 0.31)	Weak correlation
Fruit fibre[1]	0.44(0.46, 0.48)	Moderate correlation
Fibre from nuts[1]	0.37(0.33, 0.41)	Moderate correlation

[1]Spearman rank correlation coefficient and 95%CI

The Kappa method was used to measure the strength of agreement for fibre sources. For cereal fibre categories, 31% of women were classified in the same fifths giving a Kappa of 0.14, a slight agreement. However, 71% of women were classified in the same or adjacent fifths giving a weighed Kappa of 0.28, a fair agreement. Similarly, fruit fibre and fibre from nuts and seeds all showed a slight agreement in Kappa and fair agreement in weighed Kappa as illustrated in Table 10.12. But the strength of agreement for vegetable fibre remained "slight" with the weighed kappa according to the interpretation provided by Landis and Koch (1977).

Table 10.13 Kappa and weighed Kappa agreement between main fibre sources groups derived from baseline FFQ and food diary

Fibre source	Kappa			Weighed Kappa		
	Women[1] (%)	K	Description	Women[2] (%)	K	Description
Cereal	31	0.14	Slight agreement	71	0.28	Fair agreement
Vegetables	26	0.08	Slight agreement	67	0.17	Slight agreement
Fruit	34	0.18	Slight agreement	72	0.31	Fair agreement
Nuts and seeds	31	0.13	Slight agreement	64	0.24	Fair agreement

[1]Percentage of women classified in the same quintile, [2] Percentage of women classified into the same or adjacent quintile.

From Tables 10.13, the percentages of women classified into the same or adjacent quintile in cereal fibre and fruit fibre were the highest. The number of women who were allocated in the same or adjacent quintile based on vegetable fibre intake was slightly less in comparison to groupings based on cereal and fruit. The

lowest percentage of agreement was seen when women were classified based on intake of fibre from nuts and seeds. Two to four percent of women were grossly misclassified into the extreme fibre quintiles in cereal, fruit and vegetables groups, while a slightly higher percentage of misclassification in opposite categories was observed in the nuts fibre quintiles (9%). This indicates that women who were classified into the highest cereal, vegetables and fruit fibre quintile by the FFQ method were more likely to be in the highest cereal, vegetables and fruit fibre quintile by the diary method. However for nuts, it seems that infrequent consumption of nuts by women results in relatively lower percentages of women allocated in the same and adjacent quintiles.

Table 10.14 Cumulative percentage of agreement in intake from main fibre sources between baseline FFQ and food diary

Percentage of agreement	Cereal fibre	Vegetables fibre	Fruit fibre	Fibre from nuts and seeds
Exact	32%	27%	35%	27%
Within 1 adjacent quintile	68%	62%	69%	55%
Within 2 adjacent quintile	89%	84%	89%	74%
Grossly misclassified	3%	4%	2%	9%

From the above observations, the weak agreement between FFQ-fibre and diary-fibre from vegetables may be attributed to the great difference in estimating vegetables intake by FFQ. Women were split into two main categories based on self-reported vegetarian status at the baseline. The aim was to assess whether the degree of agreement between dietary fibre measured by different methods will differ among vegetarians and non-vegetarians. Table 10.15 shows that higher weighed Kappa values were calculated among non-vegetarians in comparison to vegetarians as well as correlations between different dietary methods being slightly stronger among non-vegetarians than vegetarians, however women allocated in the opposite quintile were relatively similar.

Table 10.15 Degree of agreement and Pearson's correlations between different dietary methods in the UKWCS

Vegetarian status	Dietary methods	Dietary fibre intake (g/day)				Dietary fibre intake (g/1000kcal/day)			
		K (%)[1]	K_w (%)[2]	Miss[3]	r(n)[4]	K (%)[1]	K_w (%)[2]	Miss[3]	r(n)[3]
Vegetarian	BL-FFQ[5] vs. Diary	0.10 (29%)	0.22 (70%)	2%	0.31 (563)	0.13 (34%)	0.30 (75%)	0.7%	0.44 (563)
	R-FFQ[6] vs. Diary	0.10 (28%)	0.20 (69%)	2%	0.31 (194)	0.18 (36%)	0.29 (73%)	0.5%	0.41 (194)
	FFQs[7]	0.21 (37%)	0.41 (77%)	1%	0.58 (968)	0.20 (38%)	0.38 (77%)	0.5%	0.61 (968)
Non-vegetarian	BL-FFQ[5] vs. Diary	0.11 (29%)	0.24 (70%)	2%	0.35 (1307)	0.16 (33%)	0.34 (74%)	0.8%	0.49 (1307)
	R-FFQ[6] vs. Diary	0.11 (29%)	0.22 (69%)	2%	0.41 (188)	0.12 (29%)	0.33 (73%)	2%	0.50 (188)
	FFQs[7]	0.28 (42%)	0.49 (79%)	1%	0.65 (950)	0.23 (38%)	0.44 (77%)	0.9%	0.64 (950)

[1]Kappa value (% of agreement) [2] weighed Kappa value (% of agreement) [3]Gross-misclassification [4]correlation coefficient (no of participants) [5]baseline FFQ-fibre [6]repeated FFQ-fibre [7]baseline and repeated FFQs

10.5 Discussion

A significant overestimation of dietary fibre intake by the FFQ method in comparison to the food diary method was observed. This is supported by previous studies which reported higher dietary fibre intake by FFQ in comparison to other dietary methods such as diet history and weighed records (Grootenhuis *et al.*, 1995, Bingham *et al.*, 1994). Bingham *et al.* (1994) suggested that for many nutrients this was attributed to inappropriate frequency, rather than portion size selection by the researchers. Comparisons between dietary fibre intakes derived from different dietary methods provided useful information on the ability of an FFQ to rank subjects based on their dietary fibre intake in comparison to the reference method. The observed overestimated dietary fibre intakes in FFQs in comparison to food diaries suggest the presence of systematic bias in reporting of dietary fibre intake in questionnaires in relation to food diaries. However the FFQ is still the best approach for prospective studies especially with large sample sizes (Cade *et al.*, 2004b); this is because of several advantages of the FFQ. For example, the FFQ is cheap, less time consuming and easy to process which leads to superior preference over inviting participants to record their dietary intake in detail in a more accurate way. No dietary assessment method is without limitations. However, the food record is an appropriate method to be a reference method (Cade *et al.*, 2004b) to assess FFQ validity because its bias is largely independent of the FFQ method.

10.5.1 Correlations and differences in NSP intake obtained from two dietary assessment methods

In this analysis, the overall correlation between repeated dietary assessment methods usually ranged between 0.5 to 0.7 as reported previously (Cade *et al.*, 2002). Similarly the current study showed a correlation coefficient of 0.62 between the baseline FFQ and repeated FFQ. The energy adjusted correlation between diary-fibre and baseline-fibre was similar to the diary-fibre and repeated-fibre correlation which was also within the expected range. Similar to current results, a preliminary validation study showed other nutrient intakes (carbohydrates, protein, fat, calcium, iron, vitamin A and vitamin C) recorded in food diaries in the UKWCS have higher correlations with the repeated FFQ than baseline FFQ (Spence *et al.*, 2002). This may relate to the shorter time interval between repeat FFQ and food diary in comparison to baseline FFQ and diary.

A small mean difference between baseline FFQs and repeated FFQs was seen (2g/day). This may be explained by either under-estimation or a true change in diet over time (3.1 years period). In the current chapter, participants reported higher dietary fibre intake from the baseline FFQ (mean difference of 36%) and from the repeated FFQ (mean difference of 28%) in comparison to the food diary. Low mean difference (3g/day) was reported in a previous study when FFQs were compared to diet history(Grootenhuis *et al.*, 1995). This may relate to a shorter time period between the two methods (4 weeks interval) however, it may not be ideal to compare with the current study because of different reference method used. Another validation study showed a higher relative difference between NSP intake measured by FFQ and weighed records (53%) among women aged between 19-58 years (Masson *et al.*, 2003). Also, an overestimation of NSP intake in FFQ was reported by Bingham *et al.* (1994) in comparison to weighed records among British women aged between 50-65 years.

In general, the overestimation of nutrient intakes reported by Bingham *et al.* (1994) was thought more likely due to selection of higher frequency categories in Cambridge FFQ in compared to weighed records. In particular, mean intake of vegetables was higher in Cambridge FFQ than 7 days diary and 16 days weighted records (386g/day vs. 246g/day and 272g/d respectively). The portion sizes in the Cambridge FFQ were obtained from the weighed records. This reflects that overestimation may relate to the higher frequency of vegetable consumption rather

than portion size. On the other hand, mean vegetable intakes reported by the Oxford FFQ was 406g/day which is almost double the amount reported by 7 days dairy (246g/day). The portion sizes in the Oxford FFQ were obtained from an average of several resources (dietary surveys and Crawley 1988) and reported portions were larger than Cambridge FFQ. This suggests the influence of the frequency of consumption and portion size on estimated highest vegetable consumption in Oxford FFQ in compared to food diary.

In the UKWCS, the vegetable intake was on average 334 g/day by baseline FFQ and 218 (115) g/day by diaries. This indicates that vegetable intake was overestimated by 35% in baseline FFQ compared to diaries among UKWCS while overestimation in Oxford FFQ was 40% and 30% in Cambridge FFQ compared to diaries. Consumption of fruit was on average 307g/day from FFQ and 228 g/day from diaries in the UKWCS. Overestimation of fruit consumption among women in the UKWCS was higher by 25% in FFQ than diaries. While estimated fruit consumption by Oxford FFQ (average = 219g/day) higher by 9% and by Cambridge FFQ (average = 231g/day) by 6% than diaries (average = 198g/day).

Compared to weighed records (274g/day), the Oxford FFQ underestimated cereal consumption (199 g/day), while 7 days diary overestimated cereal consumption (246 g/day) significantly. However, comparing cereal consumption reported by food diary, the Oxford FFQ underestimated cereal consumption while Cambridge FFQ reported almost similar cereal consumption. This may suggest that portion size may have a greater impact rather than frequency of consumption in estimated cereal intake. The frequency of consumption and portion size are important factors relate to the overestimation in the FFQs.

10.5.2 Agreement between NSP intakes derived from food diaries and FFQs

The 95% limits of agreement of fibre intake illustrated by Bland-Altman plots were wide (-11.1 to 30.3g/day between baseline FFQ-fibre and diary-fibre; and -9.9 to 24.6 g/day between repeated FFQ-fibre and diary-fibre). It may suggest that the observed differences in NSP measurements may relate to the food composition tables used for analysis. For example, some loss of vegetable fibres in food diaries may be due to non-inclusion of mixed dishes, containing vegetables in vegetable group in the food tables used to analyse the diaries and there are relatively few vegetables dishes in the FFQ. An earlier study reported a mean difference of 0.07g

when dietary fibre intake was determined by questionnaire and dietary history, 4 weeks apart for healthy women and men (Cantwell *et al.*, 2005) however this study was not intended to assess dietary fibre intake and no further details on the FFQ used was provided. The high mean difference from the current chapter may partly relate to the longer time interval between the two dietary fibre assessment methods (more than a year).

A perfect agreement should not be expected between any two different dietary assessment methods such as between diary and FFQ due to the impact of several factors. Measurement errors may occur with each method as the diary method may not represent the long term dietary intake. Other factors such as the effect of a time gap between completed FFQs and food diaries and within-individual variations should also be considered.

The agreement between diary-fibre is similar with baseline-fibre and repeated-fibre even though the average time gap between baseline and diary was 4.8 years while the time gap between diary-fibre and repeated FFQ-fibre was 1.6 years. This may reflect that the time gaps in the studied population may have a minor role especially as the studied population included a high proportions of vegetarians which have previously been shown to have a more stable diet (Greenwood *et al.*, 2003).

In the current chapter, after excluding women who reported a change in diet at baseline, the association between diary-fibre and FFQ-fibre did not alter. This may be explained by high missing answers and when women who answered this question are compared, it can be seen that there is no great difference in the dietary fibre intake between women who reported yes or no to the diet change question and this may not have substantially affected the overall dietary intakes.

10.5.3 Classification of women based on their NSP intakes obtained from diaries and FFQs

The high percentage of women who were similarly classified by the FFQ in comparison to the food diary (as reference method) and the low percentage of women who were misclassified in the current study were supported by previous studies (Brunner *et al.*, 2001, Grootenhuis *et al.*, 1995). In the Whitehall II longitudinal study of British men and women, 32-44% of women were classified in the same categories and gross misclassification was only 3% (Brunner *et al.*, 2001). In another validation study between fibre estimated by FFQ and dietary history

reported among the elderly, more than half of participants were in the same dietary fibre category and only 2.7% were classified into the opposite group (Grootenhuis *et al.*, 1995). This may indicate the acceptable ranking of UKWCS participants with dietary fibre intake obtained by FFQs which may mean that women were allocated into same or adjacent quintiles by different method and a valid risk estimate was more likely to be obtained.

Misclassification of women was reported previously and it was found that in comparison to weighed records, the percentage of women who were misclassified into extreme groups was less or equal to 10% for most nutrients intakes (Bingham *et al.*, 1994). In particular, 1-6% of women were misclassified into extreme dietary NSP intake when comparisons were made between FFQs and 7 days food diaries (Bingham *et al.*, 1994). In addition, 39% and 51% of women were allocated to the same quartile using the FFQ method and 7 days diary method respectively in comparison to 16 days weighed record method.

Ward *et al.* (2011) reported an inconsistency in associations between dietary fibre and coronary heart disease in the prospective EPIC-Norfolk study, which was possibly due to the use of different dietary assessment methods to derive dietary fibre intake. In an age adjusted logistic model, the risk of CHD was significantly associated with dietary fibre intake derived from the 7 days diary method but not with dietary fibre intake derived from FFQs among men. Furthermore, there was an argument by Bingham and Day (2006) on the great impact of measurement errors related to the use of FFQ in estimating dietary intake which may explain the non-significant findings in risk estimates of breast cancer with fat intake.

Food diaries provide detail of dietary data including portions eaten which minimize errors from the researcher estimating portion size while FFQ are potentially subjected to bias related to portion size. On the other hand, diaries reflect a shorter period of time in comparison to FFQs which measure long term diet and diary may also fail to consider seasonal variation (Willett, 2012).

10.5.4 Main fibre sources: Contributions, correlations and agreements between food diaries and baseline FFQs

Cereal, vegetables, fruit and nuts were the main fibre contributors in both dietary assessment methods (baseline FFQ and food diary). These findings are in

line with Bingham et al study who also reported similar NSP contributors of by FFQ, food diary and 24 hours recall methods.

However the percentage of fibre attributed to vegetables was higher using FFQs than food diaries and the percentage of fibre attributed to cereal was higher using food diaries than FFQs. This may reflect the overestimation of vegetables by FFQs as vegetable portion size consumed cannot be determined accurately or due to multiple vegetable items on the list, encouraging women to over-report consumption. However, cereal foods may be easier to count in frequency, and portion sizes are usually easier to predict than vegetables which may explain the poor correlations between FFQs and diaries in estimating vegetable fibre in relation to cereal fibre.

The majority of fibre sources showed moderate correlations when estimated by FFQ and diary methods. Fair agreements were seen between FFQ and diary methods for all fibre sources. The percentages of women who were classified into the same or adjacent categories based on fibre sources were higher in cereal, fruit and vegetables fibre based groupings than nuts. On the other hand, the percentage of women who were misclassified into extreme groups based on cereal, vegetables and fruit did not exceed 4% but reached 9% in fibre from nuts and seeds. This may be because some women who reported nut intake in FFQs did not record the same in the diaries possibly due to infrequent consumption. This resulted in a high number of women allocated in the lowest quintile by the diary method which led to high misclassification.

The comparisons between FFQs and food diaries were carried out on a subsample of the UKWCS, so that only 5% of the initial population and 15% of diary-completers were used. Hence, the average data reported may not represent the true mean dietary fibre intake. However, the aim of this chapter is to assess whether FFQ able to rank women by fibre intake.

A recent nested case-control study, using seven UK cohort studies, including UKWCS data, reported inverse association between dietary fibre intake obtained from food diaries and risk of colorectal cancer (OR=0.61; 95%CI: 0.42, 0.90) but non-significant association between dietary fibre intake obtained from FFQ and risk of colorectal cancer (OR= 0.88; 95%CI: 0.56, 1.38) comparing extreme quintiles (Dahm *et al.*, 2010). The findings in the current chapter show fair agreement between baseline FFQs and food diaries. This may partly explain the non-significant association between dietary fibre intake and risk of T2DM in chapter 8.

Future comparison studies to explore the relationship between physiological measurement and estimated dietary fibre intake are required. Faecal weight measurement has been suggested as a potential marker of dietary fibre intake (Bates *et al.*, 1997). This is because of the positive linear relationship between faecal weight and dietary fibre consumption (one gram of pectin intake, increases stool output by 1.3g and intake from fibre-rich sources such as fruit, vegetables and bran increases the stool output by 4.9 - 5.7 g/day) which was reported in a previous comprehensive review by Cummings (2001). However, to obtain a valid estimation, it has been suggested researchers should measure the faecal weight for at least 5 days for each subject due to wide day to day variation. Biological variation in term of transit time, stool size, frequency and weight were other factors suggested to have a role in the relationship between dietary fibre intake and faecal weight measurement (Bates *et al.*, 1997).

10.6 Conclusion

Studies investigating whether diet in terms of nutrient intakes has an effect on disease risk require a large number of participants, which had led to the use of more feasible dietary assessment method (FFQ). FFQ is the most convenient dietary method that usually provides habitual dietary intakes over a long period of time in prospective studies, whereas the food diary method requires great resources such as funds, time and trained staff. It is important to know the relationship between what is measured and what is expected to be the best measure that is closest to the truth, even though no measure of diet is without error. Measurement error related to the dietary assessment tool used is more of a concern as prospective evidence showed differences in the risk estimate when dietary fibre intake obtained from food diary and FFQ. The results of this investigation show large difference between diary-fibre and FFQ-fibre with fair agreement. In general, therefore, it seems that non-significant association between incidence of T2DM and intake of dietary fibre in the UKWCS may more likely to be differ when dietary fibre obtained from diaries. The present findings provide additional evidence with respect to lower level of agreement for vegetable fibre which supported by improvement in the degree of agreement among non-vegetarians population compared with vegetarians. These findings provide a potential future research to explore validity of FFQ-fibre intake using bowel habits data from the UKWCS.

Chapter 11: Final discussion

Most studies in the field of dietary fibre have focussed on either dietary fibre analysis or dietary fibre intake in relation to diseases; this thesis for the first time explored the analytical aspect of dietary fibre in food, particularly legumes, and then focused on the epidemiological aspects of dietary fibre through analyses of the UKWCS data. Working with dietary data, particularly dietary fibre intake demands knowledge of the analytical methodology and definition. Chapter 5 dealt with dietary fibre analysis in legumes commonly consumed in the UK using the Association of Official Analytical Chemists (AOAC) analytical method. The findings in chapter 6, 7, 8, 9, and 10 were obtained from analyses of the UKWCS data. The UKWCS is one of the largest population-based prospective studies in the UK that was designed to assess relationships between diet and chronic disease. The majority of the analyses on the UKWCS data presented in this thesis has not been done before.

This thesis contributes to the evidence base on the relationship between dietary fibre intake and risk of T2DM in middle-aged women by examining the risk of T2DM in UKWCS participants in relation to:

- Intake of total dietary fibre obtained from existing NSP values and added data that were obtained to estimate AOAC-fibre intake in the UKWCS.
- Key sources and types (insoluble and soluble) fibre

For the first time, chapter 9 examined the relationship between legume consumption and risk of T2DM in the UKWCS which contribute to the evidence base on the association between legume intake and risk of T2DM. Chapter 10 provided an opportunity for determine the degree of agreement between FFQ-fibre intake and diary-fibre intake as a way to explore potential bias related to dietary assessment method (FFQ).

11.1 Dietary fibre analysis

The results indicate that AOAC-fibre values are dependent on the cooking method as the results showed that AOAC-fibre content for canned legumes were significantly lower by 31% compared to AOAC-fibre content in boiled legumes. Additionally, the determined ratio of 1:1.43 for AOAC-fibre: NSP could be used as a tool for NSP to AAOC-fibre approximation for the legume group in future studies. The variation in dietary fibre analysis methods as demonstrated in chapter 5 gives an example of some challenges focused by epidemiologists as described recently

(Westenbrink *et al.*, 2013). Therefore, a new way of investigation was undertaken in the second part of this thesis, which focused on examining the relationship between the two analytical methods in estimating dietary fibre in the UKWCS with a view to facilitate analyses of the relationship between dietary fibre intake and the risk of T2DM. This provides a good example of the impact of applying dietary fibre values obtained from different dietary fibre analytical methods on the relationship between dietary fibre intake and risk of T2DM using a large population based prospective study.

.

11.2 Impact of dietary fibre measurement methods on prospective study

The impact of dietary fibre analysis on estimated dietary fibre intake was of concern especially for epidemiologists interested in ecological studies. Therefore, chapter 7 added new prospective evidence focused on the correlation and degree of agreement between the AOAC-fibre and NSP intakes based on categorization and characteristics of women as well as the main fibre contributors.

As expected, AOAC-fibre intake was significantly higher than the NSP intake by an average of one third. The strong agreement between the two fibre analysis methods shows that the majority of the cohort women were located in the same fibre group, irrespective of the analytical fibre method used. None of the participants was grossly misclassified. With each 1g/day increment in the NSP intake, the AOAC-fibre intake increased by 1.43g/day. This is slightly higher than a previously published ratio of 1.3, which may relate to higher number of food items in the current data. One-third of the AOAC-fibre and one third of NSP intakes were from cereals and cereal products, while two-thirds came from fruits and vegetables, including legumes and potatoes.

Women in the highest fibre categories were most likely to be older, active, have a low BMI, to be vegetarians, educated and non-smokers irrespective of the analytical method used. This is in line with other previous studies (Schulze *et al.*, 2007, Montonen *et al.*, 2003, Stevens *et al.*, 2002, Meyer *et al.*, 2000, Weng *et al.*, 2012). Generally, the lifestyle characteristics of studied cohort women were following the UK recommendations.

The main findings in chapter 7 showed that for carrying out a meta-analysis that compares risk of disease between extreme quintiles of dietary fibre intake, the

laboratory method of fibre analysis used has a minimal role. High fibre consumers will be classified as such irrespective of the fibre method applied. However, for a dose-response meta-analysis, combined risk estimates using different fibre analysis methods would be unacceptable because of significant variation in the amount of dietary fibre consumed.

A new dietary fibre AOAC method was carried recently by McCleary *et al.* (2010). However, the future concern regards the complexity in applying the latest AOAC fibre analysis method, which now includes a greater level of resistant starch and non-digestible oligosaccharides which would have created theoretically major differences in categorization of the women into fibre intake quintiles. This method will add greater financial and labour burdens from re-analysing all food in the food composition data across countries, aiming to improve and standardize the nutrients databank and facilitate future comparisons between studies.

One source of weakness in the obtained AOAC-fibre values which could affect the measurements of dietary fibre intake relate to lack of AOAC-fibre values in the UK food composition tables. Half of the matched food items were obtained from the USDA nutrient database. This may have affected the accuracy with which food items were identified. Even then, the results showed a strong agreement between the two analytical methods.

The findings of UKWCS analyses showed neither AOAC-fibre, or NSP intakes were significantly associated with T2DM incidence. This is in the line with previous studies (Hopping *et al.*, 2010, Schulze *et al.*, 2004a) and with the current meta-analysis in chapter 4. Only high cereal fibre intake (5g/day equivalent to 2 slices of wholemeal bread) was associated with a reduced risk of T2DM (14%) over an average period of 4.1 years independent of age.

Since 1999, the Food Standard Agency in UK recommended the use of AOAC-fibre values for food labelling purposes and this may add a concern as no clear link between the food label and the guideline (Food Standards Agency, 2002). The use of AOAC method to measure dietary fibre content in food products partly aim to be in line with majority of the countries. This may derives to consider updating UK food composition database and dietary fibre recommendation. Additionally, the new AOAC (2009.01) method include greater amount of dietary fibre components (resistant starch and resistant oligosaccharides) in their values (McCleary, 2007). Resistant starch is considerably affected by different influential

factors that may contribute to a variation in the nutrient content. Whether to include dietary fibre values measured by AOAC method or to have a new column in the table with values for other than NSP content is possible.

11.3 Dietary fibre intake and risk of T2DM

The main advantages of the UKWCS data are that their prospective nature limited recall bias and the wide range of dietary pattern provided an opportunity to examine disease risk. T2DM is a chronic disease characterized by slow, gradual progression from normal glucose levels to persistent hyperglycaemia. This process is mainly due to insulin action dysfunction. Critically, findings from the literature review were inconclusive regarding the effect of carbohydrate intake on the risk of T2DM. This was partly because of the complexity in determining nutrient intake. The association was more relate to nature of the food (liquid vs. solid) and to digestible carbohydrate content (whether it is low or high GI food) rather than total intake.

In relation to the risk of T2DM, results of the comprehensive cohort review (Chapter 4) showed no overall effect of diets high in fibre intake on the risk of T2DM comparing extreme quintiles. Gender, follow-up duration and possibly study origin were some of the potential sources of heterogeneity. Only people with the highest cereal fibre intake experienced an 25% risk reduction in T2DM in comparison to the lowest consumers in addition to significant dose-response relationship (combined risk =0.96; 95%CI: 0.93, 0.99). Recently, it was suggested that aleurone cell walls present in cereal food which are rich in phenolic acids but not present in fruit fibre may be the active compound and may be one of the plausible mechanisms that might explain the strong association between cereal fibre and the risk of T2DM (S. Lillioja et al. 2013).

Specifically, insoluble fibre intake was significantly associated with a reduction in the risk of T2DM but not with SDF intake. The lack of association between non-cereal fibre sources and the risk of T2DM may suggest that specific fibre types/sources are more or less effective in terms of influencing diabetes status markers. However, other bioactive compounds may be of more benefit than dietary fibre in health conscious women. This potential hypothesis was derived after the results presented in chapter 9 which showed evidence of a preventive effect with a greater consumption of dried legumes. Women who consumed at least a half portion per day of dried legumes had a 15% risk reduction independent of known risk

factors of diabetes in comparison to women who consumed less than a quarter portion per day. This is in line with other studies (E. Feskens, Bowles, C.H. and Kromhout, D. 1991; R. Villegas et al. 2008). D. Jenkins et al. (1983) demonstrated a significant reduction of postprandial glucose concentrations after legumes consumption in comparison to cereal food among diabetic participants and this may possibly be attributed to specific properties of legumes such as a high content of viscous soluble dietary fibre and high amounts of slow digestible starch (U. Lehmann and Robin, F. 2007), C-type of legume starch structure (K.S. Sandhu and Lim, S.T. 2008) and polyphenol content (L. Thompson et al. 1984) may explain the protective effect of legumes

The lack of associations between dietary fibre intake, fibre from main food sources and fibre fraction with the risk of T2DM independent of potential confounders may have to several explanations. Generally, women in the UKWCS were quite health conscious as baseline recruitment aimed for one third of the studied population to be vegetarian, thus it was expected that they would exhibit healthy dietary patterns. From the review in chapter 4, evidence suggested that dietary fibre intake could be an indicator of healthy characteristics. For example women who were non-smokers, vegetarians, had low BMI and were more physically active were more likely to be in a high fibre quintile group as demonstrated in chapter 4 and 7. These health conscious characteristics were also seen when analysis was carried out in chapter 10 for dietary fibre estimated from food diaries. From food diaries, NSP intake on average was found to be 17.3 (6.3) g/day. This may suggests that those women probably consumed enough dietary fibre to protect them from diabetes or may be that voluntary participants tend to be healthier so it may take them longer to develop diabetes which refereed to health worker effect. But this does not exclude the other beneficial effects of dietary fibre per se.

A broader picture on prospective evidence that focused on the relationship between dietary fibre intake and risk of T2DM was examined. The final conclusion from pooled analysis in chapter 8 remained as non-significant association between the risk of T2DM in people with high fibre intake. Additionally, a lack of association between dietary fibre intake and the risk of T2DM was found in pooled cohorts reported a dose-response relationship.

Lack of significant associations of dietary fibre intake on the risk of T2DM in the UKWCS does not rule out other beneficial effect of diet in particular, fibre

intake in the studied population. The hypothesis-oriented approach for dietary pattern generated by Ashton *et al.* (2013) referred to as the T2DM prevention index score was examined in relation to the risk of T2DM. The main results showed a healthy dietary pattern including high intakes of fruit, vegetables, wholegrain, legumes and nuts with a low intake of meat, sugar, alcohol and poultry, was associated with significant decreased risk of T2DM in the UKWCS. Another analysis carried out by Alrefaai *et al.* (2013) showed the odds of being constipated after an average of 4 years was significantly reduced (12%-16%) among women in the highest dietary fibre quintile in comparison to women in the lowest dietary fibre quintile independent of potential confounders.

The findings in chapter 8 were not meant to represent the UK population but generally the UKWCS was meant to include participants with a wide range of dietary exposures to help in studying the relationship between diet and disease. A previously generated variable (Taylor *et al.*, 2007) was used in the logistic regression analysis where the large proportion of vegetarians in the UKWCS was considered in the analysis to show results representative to UK population however, the non-significant association between total dietary fibre intake and risk of T2DM remained the same. Also this variable was used in chapter 9 and the results showed lack of associations between legume intake (total, fresh legume and dried) and risk of T2DM.

Despite the wide variation in dietary fibre intake across the UKWCS participants, the women were well-nourished, this means that there may be an inadequate number of women with very low dietary fibre intake (only 3% from the total population of women consumed less than 18g/day of NSP) to show a decreased risk of T2DM in such women.

Another possible explanation of no significant effect is may be due to type II errors which occur once the null hypothesis fails to be rejected and the result likely to be a false negative. Type II errors can be minimized by increasing the power and increasing the precision of dietary measurement (Kirkwood and Sterne 2003). The small number of self-reported diabetes cases may be attributed to under reporting due to undiagnosed diabetes cases that were misclassified in the diabetes free group. Identifying new cases and undergoing the confirmation process could have been undertaken by contacting local GPs in subsample if time and money had been available. The small number of cases may also be attributed to the small percentage

of minority ethnic groups in the UKWCS who are more vulnerable to T2DM than the white population.

Additionally women who are believed to have healthier lifestyle probably are more likely to be screened in medical clinics. This may result in underestimating the number of women with diabetes who also voluntarily participate in a study this called health worker effect (Li and Sung, 1999). Another potential bias may be due to the lack of diet monitoring over the studied period which leaves the possibility of true change or fluctuation in the dietary pattern that may also affect the risk of T2DM in the studied population. However, the moderate stability of dietary pattern reported in a subsample from the UKWCS (Greenwood *et al.*, 2003) may lessen the possibility of great influence of dietary changes over time.

The possibility of residual confounders also may hide the true association between dietary fibre intake and the risk of T2DM. The effect of dietary fibre on the T2DM incidence was also subject to measurement errors related to the FFQ. However, the wide variation in the dietary exposure is one of the advantages of the UKWCS population. How much weight women gained may have a greater impact on the risk of T2DM, where the dietary fibre effect size is relatively small. The massive over riding relationship between obesity and T2DM may hide the effect of dietary fibre on the risk of T2DM.

Multiple testing is one of the limitations that can lead to the production of false significant associations due to chance (Bland and Altman, 1995). The suggested solution for this problem is to lower the significance level (P-value) for in the subgroup analyses. However, if an assumption was made to test the main exposure of interest which is the total dietary fibre intake in relation to T2DM, and other analyses explored type of fibre and source of fibre, then this possibly may not refer to multiple testing.

Incompleteness of data because of either loss of follow up or death of participants or both were suggested difficulties in identifying incidence of disease in cohort studies (Kelsey, 1996). From the baseline, only 40% of women completed phase II. In general, despite the large number of drop outs, there was a considerable number of participants who completed phase II because of the large number recruited at baseline. However, this did not help in identifying a large number of diabetic cases. High number of women left the question of diabetes diagnosis blank (1,081 (8% of total women 14,172) of women missed the question in phase II and

19% (207) of them missed the question at baseline and phase II) which also could affect the final number of incident of diabetes included in chapter 8 and 9 analyses.

Sample size calculations should be performed prior to data collection, as type I and type II errors can happen when statistical tests are performed to examine the null hypothesis (Kirkwood and Sterne, 2003). Data processing bias such as coding errors, data entry and calculating mistakes are factors which can affect the precision of the measured variable. However, the UKWCS data entry was validated through double data entry by separate professional staff.

As discussed in chapter 2, the lack of a national registry for diabetes in the UK may restrict the availability of data on the new onset of diabetes. In the UK, diagnosis of diabetes is usually carried out by the local general practitioner (GP). In the UKWCS, a large number of women completed phase II, thus tracing participants through contacting the local GP is impossible and impractical for single researcher.

One of the main strengths of this study is that recall bias is not of great concern in prospective studies as dietary data is collected before the onset of the disease.

Another important finding in this thesis was the fair agreement between dietary FFQ-fibre intake and diary-fibre intake. This is may be due to a poor agreement of fibre intake from vegetable and nuts which was demonstrated in chapter 10. Measurement bias related to the use of FFQs was suggested as a potential source of heterogeneity in the meta-analysis and possible explanation of non-significant findings between total dietary fibre intake and risk of T2DM in chapter 8. The validation study in chapter 10 showed weak correlation and low agreement between vegetable fibre intake obtained by baseline FFQs and diaries. In addition, the overestimation of vegetable fibre was high with a mean difference of 50% when comparing FFQs to diaries. This indicates relatively low validity of FFQ particularly for vegetable intake.

11.4 Impact of dietary fibre assessment methods in prospective study

The findings from chapter 10 showed fair agreement between FFQ-fibre and diary-fibre intakes with a very small number of women were grossly misclassified which is similar to previous studies. It can be appreciated that FFQ reasonably rank women as it should be when compared with reference method (Food diary) and FFQ

still remain the most convenient method to study large population. However further modifications in the FFQ may help in improving level of accuracy of the method.

The main fibre source by diary was the cereal group, while the main fibre source by FFQ was the vegetables group. The highest difference between FFQ-fibre and diary-fibre was observed in nuts and vegetables (100% and 50% respectively). This may be attributed to infrequent nuts consumption which was shown by having zero nut intakes in diaries. Further analysis showed that vegetable fibre intakes estimated by diaries and FFQs were in poor agreement which could be attributed to over reporting in the FFQ. As slight improvement in the agreement between FFQ-fibre and diary-fibre was observed when the agreement was determined among non-vegetarians only.

The lack of association between fibre sources such as vegetable fibre and fruit fibre and legumes in prospective studies may be due to poor dietary assessment method applied (FFQ). An earlier study by Jenkins *et al.* (1977) reported that intake of pectin reduced post-prandial glucose response significantly. So there is conflict in the results from epidemiological and experimental studies. The strong evidence on the protective effect of cereal fibre in cohort studies could be because cereal fibre is easier to measure than fruit and vegetables by dietary assessment methods (mostly FFQ). It could also be that participants can provide intake of cereal food more accurately with respect to the portion consumed than fruit and vegetable consumption. In the UKWCS, women consumed a high amount of vegetables, but still no effect was found on the risk of T2DM which may be due to poor dietary vegetable fibre estimation. Or it may be that there is no effect of dietary vegetable fibre on low risk people who already consume high level of fibre.

Weak correlations and agreement between the FFQ and diary methods may relate to vegetable fibre which was largely overestimated by FFQs and this can possibly be attributed to the high proportion of vegetarians in the studied population. This may suggested that the findings in chapter 8 regarding the non-significant effect of dietary fibre on the risk of T2DM may remain the same with estimated dietary fibre obtain from food diary.

Whilst discussion of the ratio of AOAC-fibre: NSP may seem rather academic, there is some evidence that research groups are employing ratios to convert between fibre methods. For example, Parkin (2011) estimated the attributable proportion of colorectal cancer of 12.2% that could be prevented by increasing NSP intake to

18g/day. This was based on the effect estimated by Dahm *et al.* (2010): a relative risk = 0.84 (0.71 to 1.00) and the WCRF relative risk of 0.9 per 10g/day increment in dietary fibre. Calculated equivalent was reported as a reduction in risk of 2.9% per 1g of fibre. A measure of 18g/day of NSP was assumed to be equivalent to 23g/day, of total fibre which was made based on an ambiguous ratio of 1:1.28 (AOAC-fibre: NSP). It was unclear how the authors had arrived at this value. This reflects the importance of the ratio used to convert NSP values into AOAC-fibre values for further calculations in epidemiological studies.

11.5 Future research

11.5.1 Laboratory future research

The AOAC method has been modified several times, the latest modification labelled as 2009.01 (McCleary *et al.*, 2010), which combines AOAC 895.29 (Prosky *et al.*, 1985) measured fibre value and low molecular weight dietary fibre value (resistant starch and non-digestible oligosaccharides) to meet the approved definition by the Codex Alimentarius Commission (2008). Further experimental investigations are needed to compare and assess the variation in the dietary fibre values in a range of foods that are commonly consumed in the UK particularly those foods which have high amounts of resistant starch and non-digestible oligosaccharides such as legumes using the new AOAC method (2009.01), classical AOAC (985.29) and Englyst method. The findings in chapter 5 provide insights for future research in Kuwait. Dietary fibre values were obtained by the AOAC method for 32 Kuwaiti dishes in a previous study (Dashti *et al.*, 2003). However, a wider range of traditional food requires dietary fibre measurement and re-analysis of common Kuwaiti dishes using the new method will provide fibre values that can be used for future epidemiological studies in Kuwait (my country of origin). Also, incorporating new dietary fibre values into national food composition data will be an interesting area of research in the future.

The findings in chapter 5 suggest that the AOAC-fibre: NSP ratio is dependent on the cooking method. This means that the boiled legume ratio may be more suitable for studies which focus on minority ethnic groups in the UK (Church *et al.*, 2006) where boiled legumes are mostly consumed, compared to the rest of the UK general population which is more likely to consume canned legumes (Schneider, 2002). It is useful to understand the importance of the impact of cooking methods on dietary fibre analysis as the AOAC procedure is dependent on these. The

preliminary analysis of legumes may raise an interest on whether a ratio between the two analytical methods is an appropriate tool for epidemiological studies. However, more research on the AOAC-fibre: NSP ratio derived from a wide range of food items needs to be undertaken to explore the association between AOAC-fibre and NSP more clearly.

11.5.2 Epidemiological future research

11.5.2.1 UKWCS future research

The validation study from chapter 10 provides useful data to be incorporated into the regression calibration method which can correct for nutrient-based measurement error as suggested by Rosner and Gore (2001). The regression calibration approach is commonly used to predict nutrient record intake as a function of FFQ intake and to model the risk of disease with the predicted values obtained aiming for a corrected risk estimate. It would be interesting to examine whether risk estimates differ after applying the regression calibration approach.

Stratified analyses of the data used for chapter 10 could be undertaken in the future to assess the impact of different lifestyle factors on the agreement between FFQ-fibre and diary-fibre intake.

Initially when the cohorts were conceived in early 1990s, individual ethical approval was obtained from 174 local ethic committees in the UK because there was no central system. Women who completed the baseline questionnaire and provided accurate NHS or GP information could be traced to identify more diabetes events in the UKWCS after obtaining another ethical approval as individual consent from participants is impractical. The potential of two datasets referred to as the UK General Practice Research Database (GPRD) (Walley and Mantgani, 1997) and Health Improvement Network (THIN) (Ruigómez *et al.*, 2010) can be explored and used to link diabetes events. These databases provide medical information for chronic disease for research purposes.

Future meta-analysis aiming for dose-response relationship between dietary fibre intake and the risk of T2DM, can provide more robust evidence for dietary guidelines aiming for primary prevention of T2DM.

11.5.2.2 Other future research

More broadly, research is needed to determine the effect of dietary fibre intake on the risk of T2DM among young populations. Since evidence shows that T2DM is diagnosed at a younger age (Rosenbloom *et al.*, 1999), then cohort studies in young

populations with a long follow up period could be useful in examining the early dietary fibre intake in relation to pre-diabetes and diabetes in later life. This will help in assessing the temporal relation if repeated dietary measurement over time is considered.

Recently, a high prevalence of diabetes (21%) in Kuwait (20-79 years) was reported in a 2011 International Diabetes Federation report (Whiting *et al.*, 2011). This indicates the huge future economic and health impact of diabetes. One of the main objectives of the department of Food Science and Nutrition in the Ministry of Health of Kuwait is to tackle public health problems such as obesity and diabetes from dietary and lifestyle aspects. An early randomized control trial, the Diabetes Prevention Study in Finland, showed a successful dietary and lifestyle intervention, that resulted in significant weight loss over the first year on high risk population (Eriksson *et al.*, 1999) and reduction in diabetes incidence after a median of 4 years (Lindstrom *et al.*, 2006). If this is an example for future study in Kuwait then this idea could be applied with some considerations regarding the target population, the most suitable dietary assessment method, recruitment and sample size in the study design stage.

There is a need for dietary assessment methods characterized by high accuracy in measuring dietary intake in order to examine associations between diet and disease outcomes. The use of internet-based dietary data collection such as using online 24-hour recall and FFQs or using mobile applications for future cohort studies may be possible, especially for targeting younger populations. This new technology may decrease the time spend for data collection and coding by the researcher.

The biomarkers of nutrient intake can be used to validate the accuracy of estimated dietary intake and can reflect the consumption of nutrient intake. However for dietary fibre, no such biomarker has been reported in previous prospective studies. However, stool samples as an NSP marker, were suggested previously (Bates *et al.*, 1997). Stool collection for at least 5 days could yield a valid estimate. This was suggested as a marker for NSP, as for every 1 gram increment of NSP, the stool weight increased by 5 grams. However, this method has several limitations: for example, it may be seen as unacceptable by subjects; it has high biological variation in term of transit time and faecal weight, and also cannot be used in a large number of participants.

Another potential future research using bowel habits data from the UKWCS dataset to explore validity of fibre intake using FFQ. This may achieve by using questions on the frequency of bowel habits captured at the baseline in relation to dietary fibre intake. Evidence from the UKWCS data suggest with every 5g increment in dietary NSP intake, the risk of constipation (defined by frequency of bowel) reduced by 16 % (Alrefaai *et al.*, 2013).

It would be interesting to assess the effects of dietary fibre intake on the risk of T2DM prospectively in Kuwait, as a high prevalence of diabetes has been reported (Whiting *et al.*, 2011) and because no cohort study has been carried out in Kuwait. It would be a good opportunity to start a prospective study using self-reported diabetes to identify new onsets of cases. However, the use of biochemical tests such as blood glucose samples and Hb_{A1C} measurements would be more robust in establishing a greater degree of accuracy of diabetes diagnosis minimising the bias from misclassification. On the other hand, the high cost and practicality should be evaluated carefully during the study design step.

FFQ is the commonly used dietary assessment method in prospective studies however modification of existing FFQ should be considered to cover traditional food consumed in Kuwait. A validation study for subsample using another method such as food records can help in assessing the reliability and validity of FFQs.

Dietary assessment method can be improved in several ways. More food items especially fibre rich sources such as legumes may need to be included in the FFQ particularly for participants with relatively high legume consumption. The results from chapter 5 suggest that cooking methods can be used as a guide for the researcher when fibre values from food tables are assigned to a food item, although cooking methods may play a minor role in NSP values in comparison to AOAC-fibre values as demonstrated in chapter 5.

In summary, it is useful to incorporate food laboratory analysis into the epidemiological studies in a way that will measure dietary fibre content in the food consumed by participants. However this has a high cost and time burden especially if the sample size is quite large. The result from chapter 5 suggested the effect of cooking methods on the dietary fibre content using the AOAC method. This observation can be incorporated into epidemiological research by obtaining dietary information on the cooking methods which may help in minimizing the estimation

error in the epidemiological research particularly among vegetarians and high legumes consumers.

11.5.3 Public health implications

Dietary fibre found to have several beneficial effects on the risk of diseases. Consistent evidence showed only high intake of cereal fibre reduced the risk of T2DM. Despite the evidence in chapter 4, and chapter 8, that neither greater intakes of total dietary fibre, or fruit, vegetables and legumes fibre, appear to reduce the risk of T2DM, the public health message to encourage high dietary fibre intake through consumption of fruit and vegetables needs to be maintained. The lack of association between a single nutrient in fruit, vegetables and legumes such as dietary fibre doesn't mean other potential compounds have no role with regard to glycaemia. Recent evidence suggested a possible beneficial role of other bioactive components (such as polyphenols) present in most plants including fruit and vegetables on glycaemia (Williamson, 2013).

Since evidence suggests a high intake of dietary fibre reduces the risk of other chronic diseases such as heart disease (Pereira *et al.*, 2004), stroke (Threapleton *et al.*, 2012) and colorectal cancer (Aune *et al.*, 2011), encouraging high dietary fibre consumption needs to be maintained for other health benefits.

The results from chapter 5 showed that significant effect of cooking methods on dietary fibre content which can be used as a guide for consumers and practitioners, since most of legumes consumed in the UK are canned, may be more dried legume intake could be encouraged.

11.6 In summary

Type 2 diabetes mellitus is a common public health problem that requires primary prevention. Potential mechanism for dietary fibre action on diabetes markers was reported previously. For the first time this thesis investigated different aspects of dietary fibre using laboratory and epidemiological approaches. Dietary fibre analysis of legumes commonly consumed in UK indicate that AOAC-fibre values are highly dependent on the cooking method as the results showed that AOAC-fibre content for canned legumes were significantly lower by one third compared to boiled legumes. A non-significant lower risk of developing T2DM was seen in British women who consume diets high in fibre, this was significant for age adjusted cereal fibre. However, the lack of statistical significance overall, suggests

either a weak relationship with fibre, nutrient related measurement error or may be a reflection of a relatively small number of cases.

In a sample of health-conscious women, greater dietary fibre intake may indicate no additional benefit in term of diabetes prevention. Although greater consumption of dried legumes contributed to lower diabetes risk.

References

Abdullah, A., Peeters, A., de Courten, M. & Stoelwinder, J., 2010. The magnitude of association between overweight and obesity and the risk of diabetes: a meta-analysis of prospective cohort studies. *Diabetes Research of Clinical Practice,* 89**,** 309-19.

AbuMweis, S., Jew, S. & Ames, N., 2010. β-glucan from barley and its lipid-lowering capacity: a meta-analysis of randomized, controlled trials. *European Journal of Clinical Nutrition,* 64**,** 1472-1480.

Adeghate, E., Schattner, P. & Dunn, E., 2006. An update on the etiology and epidemiology of diabetes mellitus. *Annals of the New York Academy of Sciences,* 1084**,** 1-29.

Afzal, S., Bojesen, S.E. & Nordestgaard, B.G., 2013. Low 25-Hydroxyvitamin D and risk of type 2 diabetes: A prospective cohort study and meta-analysis. *Clinical Chemistry,* 59**,** 381-391.

Aggarwal, A., Monsivais, P. & Drewnowski, A., 2012. Nutrient intakes linked to better health outcomes are associated with higher diet costs in the US. *PLoS One,* 7**,** e37533.

Ahmadi-Abhari, S., Luben, R.N., Powell, N., Bhaniani, A., Chowdhury, R., Wareham, N.J., Forouhi, N.G. & Khaw, K.-T., 2013. Dietary intake of carbohydrates and risk of type 2 diabetes: the European Prospective Investigation into Cancer-Norfolk study. *British Journal of Nutrition***,** 1-11.

Ainsworth, B.E., Haskell, W.L., Leon, A.S., Jacobs, D.R., Montoye, H.J., Sallis, J.F. & Paffenbarger, R.S., 1993. Compendium of physical activities: classification of energy costs of human physical activities. *Medicine and Science in Sports and Exercise,* 25**,** 71-80.

Alberti, K., Eckel, R.H., Grundy, S.M., Zimmet, P.Z., Cleeman, J.I., Donato, K.A., Fruchart, J.-C., James, W.P.T., Loria, C.M. & Smith, S.C., 2009. Harmonizing the Metabolic Syndrome A Joint Interim Statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation,* 120**,** 1640-1645.

Alfieri, M.A.H., Pomerleau, J., Grace, D.M. & Anderson, L., 1995. Fiber intake of normal weight, moderately obese and severely obese subjects. *Obesity Research,* 3**,** 541-547.

Aller, R., de Luis, D.A., Izaola, O., La Calle, F., del Olmo, L., Fernandez, L., Arranz, T. & Hernandez, J.M.G., 2004. Effect of soluble fiber intake in lipid and glucose levels in healthy subjects: a randomized clinical trial. *Diabetes Research and Clinical Practice,* 65**,** 7-11.

Alrefaai, L., Cade, J. & Burley, V., 2013. Dietary fibre intake and constipation in the UK Women's Cohort Study. *Proceedings of the Nutrition Society,* 72**,** E287.

Alvarez, J.A. & Ashraf, A., 2009. Role of vitamin D in insulin secretion and insulin sensitivity for glucose homeostasis. *International Journal of Endocrinology,* 2010.

American Diabetes Association (ADA). 2008. Nutrition recommendations and interventions for diabetes. *Diabetes Care,* 31**,** S61-78.

Anderson, J. & Bridges, S., 1988. Dietary fiber content of selected foods. *American Journal of Clinical Nutrition,* 47**,** 440-447.

Anderson, N. & Clydesdale, F., 1980. Effects of processing on the dietary fiber content of wheat bran, pureed green beans, and carrots. *Journal of Food Science,* 45**,** 1533-1537.

Andersson, A., Tengblad, S., Karlström, B., Kamal-Eldin, A., Landberg, R., Basu, S., Åman, P. & Vessby, B., 2007. Whole-grain foods do not affect insulin sensitivity or markers of lipid peroxidation and inflammation in healthy, moderately overweight subjects. *The Journal of Nutrition,* 137**,** 1401-1407.

AOAC. 2001. *Report of the Dietary Fiber Definition Committee to the Board of Directors of AACC International* [Online]. Available: http://www.aaccnet.org/grainbin/definitiondietaryfiber.asp [Accessed 6 May 2011].

Apata, D.F., 2008. Effect of cooking methods on available and unavailable carbohydrates of some tropical grain legumes. *African Journal of Biotechnology,* 7**,** 2940-2945.

Arngrímsson, S.Á., McAuley, E. & Evans, E.M., 2009. Change in body mass index is a stronger predictor of change in fat mass than lean mass in elderly black and white women. *American Journal of Human Biology,* 21**,** 124-126.

Ashton, L., Cade, J.E. & Burley, V.J., 2013. A type 2 diabetes mellitus prevention index predicts incident diabetes in the UK Women's Cohort Study. *Proceedings of the Nutrition Society,* 72**,** E257

Asp, N.G., 1987. Dietary fibre-Definition, chemistry and analytical determination *Molecular Aspects of Medicine,* 9**,** 17-29.

Association of Official Analytical Chemist. 1995. Official Methods of Analysis, 991.43 Arlington VA, : AOAC.

Aune, D., Chan, D.S., Lau, R., Vieira, R., Greenwood, D.C., Kampman, E. & Norat, T., 2011. Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. *BMJ: British Medical Journal,* 343.

Baliunas, D.O., Taylor, B.J., Irving, H., Roerecke, M., Patra, J., Mohapatra, S. & Rehm, J., 2009. Alcohol as a Risk Factor for Type 2 Diabetes A systematic review and meta-analysis. *Diabetes Care,* 32**,** 2123-2132.

Baptiste-Roberts, K., Gary, T.L., Beckles, G.L., Gregg, E.W., Owens, M., Porterfield, D. & Engelgau, M.M., 2007. Family history of diabetes, awareness of risk factors, and health behaviors among African Americans. *Journal Information,* 97.

Barbagallo, M., Dominguez, L.J., Galioto, A., Ferlisi, A., Cani, C., Malfa, L., Pineo, A., Busardo, A. & Paolisso, G., 2003. Role of magnesium in insulin action, diabetes and cardio-metabolic syndrome X. *Molecular Aspects of Medicine,* 24**,** 39-52.

Barclay, A.W., Flood, V.M., Rochtchina, E., Mitchell, P. & Brand-Miller, J.C., 2007. Glycemic index, dietary fiber, and risk of type 2 diabetes in a cohort of older Australians. *Diabetes Care,* 30**,** 2811-2813.

Barclay, A.W., Petocz, P., McMillan-Price, J., Flood, V.M., Prvan, T., Mitchell, P. & Brand-Miller, J.C., 2008. Glycemic index, glycemic load, and chronic disease risk—a meta-analysis of observational studies. *American Journal of Clinical Nutrition,* 87**,** 627-637.

Bastard, J.-P., Jardel, C., Bruckert, E., Blondy, P., Capeau, J., Laville, M., Vidal, H. & Hainque, B., 2000. Elevated levels of interleukin 6 are reduced in serum and subcutaneous adipose tissue of obese women after weight loss. *Journal of Clinical Endocrinology & Metabolism,* 85**,** 3338-3342.

Bates, B., Lennox, A., Bates, C. & Swan, G., 2011. National Diet and Nutrition Survey. Headline Results from Years 1 and 2 (Combined) of the Rolling Programme (2008/2009–2009/2010). *Department of Health and Food Standards Agency: London.*

Bates, C.J., Margetts, B.M. & Nelson, M., 1997. Biochemical markers of nutrient intake. *Design Concepts in Nutritional Epidemiology.*

Battilana, P., Ornstein, K., Minehira, K., Schwarz, J., Acheson, K., Schneiter, P., Burri, J., Jeâquier, E. & Tappy, L., 2001. Original Communication Mechanisms of action of β-glucan in postprandial glucose metabolism in healthy men. *European journal of clinical nutrition,* 55**,** 327-333.

Bazzano, L.A., Li, T.Y., Joshipura, K.J. & Hu, F.B., 2008. Intake of fruit, vegetables, and fruit juices and risk of diabetes in women. *Diabetes Care,* 31**,** 1311-1317.

Becher, H., 1992. The concept of residual confounding in regression models and some applications. *Statistics in Medicine,* 11**,** 1747-1758.

Beck, N.H., Henriksen, J.E., Vaag, A. & Hother, N.O., 1995. Pathophysiology of non-insulin-dependent diabetes mellitus (NIDDM). *Diabetes Research and Clinical Practice,* 28, Supplement**,** S13-S25.

Bell, G.I., 1991. Molecular defects in diabetes mellitus. *Diabetes,* 40**,** 413-422.

Ben-Gal, I., 2005. Outlier detection. *Data Mining and Knowledge Discovery Handbook.* Springer.

Betteridge, V., 2009. Dietary fibre: An evolving definition? *Nutrition Bulletin,* 34**,** 122-125.

Bewick, V., Cheek, L. & Ball, J., 2004. Statistics review 11: Assessing risk. *Critical Care,* 8**,** 287.

Bialostosky, K., Wright, J.D., Kennedy-Stephenson, J., McDowell, M. & Johnson, C.L., 2002. Dietary intake of macronutrients, micronutrients, and other dietary constituents: United States 1988-94. *Vital and Health Statistics. Series 11, Data from the national health survey***,** 1.

Bingham, S., Gill, C., Welch, A., Day, K., Cassidy, A., Khaw, K., Sneyd, M., Key, T., Roe, L. & Day, N., 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**,** 619-644.

Bingham, S.A. & Day, N., 2006. Commentary: Fat and breast cancer: time to re-evaluate both methods and results? *International Journal of Epidemiology,* 35**,** 1022-1024.

Bland, J.M. & Altman, D.G., 1995. Multiple significance tests: the Bonferroni method. *BMJ: British Medical Journal,* 310**,** 170.

Bland, M.J. & Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet,* 327**,** 307-310.

Boeing, H., Bechthold, A., Bub, A., Ellinger, S., Haller, D., Kroke, A., Leschik-Bonnet, E., Müller, M.J., Oberritter, H. & Schulze, M., 2012. Critical review: vegetables and fruit in the prevention of chronic diseases. *European Journal of Nutrition***,** 1-27.

Boggs, D.A., Rosenberg, L., Ruiz-Narvaez, E.A. & Palmer, J.R., 2010. Coffee, tea, and alcohol intake in relation to risk of type 2 diabetes in African American women. *American Journal of Clinical Nutrition,* 92**,** 960-966.

Bogracheva, T.Y., Morris, V., Ring, S. & Hedley, C., 1998. The granular structure of C ‐ type pea starch and its role in gelatinization. *Biopolymers,* 45**,** 323-332.

Bombelli, M., Facchetti, R., Sega, R., Carugo, S., Fodri, D., Brambilla, G., Giannattasio, C., Grassi, G. & Mancia, G., 2011. Impact of body mass index and waist circumference on the long-term risk of diabetes mellitus, hypertension, and cardiac organ damage. *Hypertension,* 58**,** 1029-1035.

Bravo, L., Abia, R. & Sauracalixto, F., 1994. Polyphenols as dietary fiber associated compounds - Comparative-study on in-vivo and in-vitro properties. *Journal of Agricultural and Food Chemistry,* 42**,** 1481-1487.

Brennan, M.A., Derbyshire, E.J., Brennan, C.S. & Tiwari, B.K., 2012. Impact of dietary fibre ‐ enriched ready ‐ to ‐ eat extruded snacks on the postprandial glycaemic response of non ‐ diabetic patients. *Molecular nutrition & food research,* 56**,** 834-837.

British Nutrition Foundation. 1990. *Complex carbohydrates in foods: The report of the British Nutrition Foundation's Task Force,* UK, Chapman and Hall for British Nutrition Foundation.

Brown, L., Rosner, B., Willett, W.W. & Sack, F.M., 1999 Cholesterol-lowering effects of dietary fiber: a meta-analysis. *American Journal of Clinical Nutrition,* 69**,** 30-42.

Brunner, E., Juneja, M. & Marmot, M., 2001. Dietary assessment in Whitehall II: comparison of 7 day diet diary and food-frequency questionnaire and validity against biomarkers. *British Journal of Nutrition,* 86**,** 405-414.

Burkitt, D., 1975. *Refined carbohydrate foods and disease,* Academic Press.

Buttriss, J. & Stokes, C., 2008. Dietary fibre and health: an overview. *Nutrition Bulletin,* 33**,** 186-200.

Cade, J., Burley, V. & Greenwood, D., 2004a. The UK Women's Cohort Study: comparison of vegetarians, fish-eaters and meat-eaters. *Public Health Nutrition,* 7**,** 871-878.

Cade, J., Burley, V., Warm, D., Thompson, R. & Margetts, B., 2004b. Food-frequency questionnaires: a review of their design, validation and utilisation. *Nutrition Research Reviews,* 17**,** 5-22.

Cade, J., Frear, L. & Greenwood, D., 2006. Assessment of diet in young children with an emphasis on fruit and vegetable intake: using CADET–Child and Diet Evaluation Tool. *Public Health Nutrition,* 9**,** 501-508.

Cade, J., Thompson, R., Burley, V. & Warm, D., 2002. Development, validation and utilisation of food-frequency questionnaires: A review. *Public Health Nutrition,* 5**,** 567-588.

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**,** 431-438.

Cade, J.E., Taylor, E.F., Burley, V.J. & Greenwood, D.C., 2010. Common Dietary Patterns and Risk of Breast Cancer: Analysis From the United Kingdom Women's Cohort Study. *Nutrition and Cancer-an International Journal,* 62**,** 300-306.

Calvert, C., Cade, J., Barrett, J. & Woodhouse, A., 1997. Using cross-check questions to address the problem of mis-reporting of specific food groups on Food Frequency Questionnaires. UKWCS Steering Group. United Kingdom Women's Cohort Study Steering Group. *European Journal of Clinical Nutrition,* 51**,** 708.

Camire, M., Cho, S., Craig, S., Devrie, J., Gordon, D., Jones, J., Li, B., Lineback, D., Prosky, L. & Tungland, B., 2001. The definition of dietary fiber. *Cereal Foods World,* 46**,** 112-124.

Cantwell, M.M., Gibney, M.J., Cronin, D., Younger, K.M., O'Neill, J.P., Hogan, L. & Flynn, M., 2005. Development and validation of a food-frequency questionnaire for the determination of detailed fatty acid intakes. *Public Health Nutrition,* 8**,** 97-107.

Chang, M.C. & Morris, W.C., 1990. Effect of Heat-treatment on chemical-analysis of dietary fiber *Journal of Food Science,* 55**,** 1647-&.

Chaplin, S., 2005. *Type 2 diabetes: prevention and management,* ILSI Europe.

Charles, M.A., Balkau, B., Vauzelle-Kervröedan, F., Thibult, N. & Eschwege, E., 1996. Revision of diagnostic criteria for diabetes. *Lancet,* 348**,** 1657-1658.

Chawla, R. & Patil, G., 2010. Soluble dietary fiber. *Comprehensive Reviews in Food Science and Food Safety,* 9**,** 178-196.

Cho, S. & Dreher, M.L., 2001. *Handbook of dietary fiber,* CRC Press.

Choi, S.H., Kim, T.H., Lim, S., Park, K.S., Jang, H.C. & Cho, N.H., 2011. Hemoglobin A1c as a Diagnostic Tool for Diabetes Screening and New-Onset Diabetes Prediction A 6-year community-based prospective study. *Diabetes Care,* 34**,** 944-949.

Chung, H.-J., Liu, Q. & Hoover, R., 2009. Impact of annealing and heat-moisture treatment on rapidly digestible, slowly digestible and resistant starch levels in native and gelatinized corn, pea and lentil starches. *Carbohydrate Polymers,* 75**,** 436-447.

Church, S., Gilbert, P. & Khokhar, S., 2006. Ethnic groups and foods in Europe. *Synthesis Report*.

Church, S.M., 2006. The history of food composition databases. *Nutrition Bulletin,* 31**,** 15-20.

Codex Alimentarius Commission. 2008. Report of the 30th session of the codex committee on nutrition and foods for special dietary uses, Cape Town, South Africa, 3–7 November 2008. ALINORM 09/32/26.

Colditz, G.A., Manson, J.E., Stampfer, M.J., Rosner, B., Willett, W.C. & Speizer, F.E., 1992. Diet and risk of clinical diabetes in women *American Journal of Clinical Nutrition,* 55**,** 1018-1023.

Colditz, G.A., Willett, W.C., Rotnitzky, A. & Manson, J.E., 1995. Weight gain as a risk factor for clinical diabetes mellitus in women. *Annals of Internal Medicine,* 122**,** 481-486.

Connolly, V., Unwin, N., Sherriff, P., Bilous, R. & Kelly, W., 2000. Diabetes prevalence and socioeconomic status: a population based study showing increased prevalence of type 2 diabetes mellitus in deprived areas. *Journal of Epidemiology and Community Health,* 54**,** 173-177.

Cook, R.J., 1998. Kappa. *Encyclopedia of biostatistics*.

Cooper, A., Forouhi, N., Ye, Z., Buijsse, B., Arriola, L., Balkau, B., Barricarte, A., Beulens, J., Boeing, H. & Büchner, F., 2012. Fruit and vegetable intake and type 2 diabetes: EPIC-InterAct prospective study and meta-analysis. *European Journal of Clinical Nutrition,* 66**,** 1082-1092.

Corp-Stata. 2010. Statitsitcal software:Release 11. 12 ed.: College Station, TX.

Cosgrove, D.J., 2005. Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology,* 6**,** 850-861.

Crawley, H., 1993. *Food Portion Sizes,* London HMSO.

Cummings, J. & Englyst, H., 1991. Measurement of starch fermentation in the human large intestine. *Canadian journal of physiology and pharmacology,* 69**,** 121-129.

Cummings, J. & Stephen, A., 2007. Carbohydrate terminology and classification. *European Journal of Clinical Nutrition,* 61**,** S5-S18.

Cummings, J.H., 2001. The effect of dietary fiber on fecal weight and composition. *CRC Handbook of Dietary Fiber in Human Nutrition,* 3**,** 183-252.

Cust, A., Skilton, M., Van Bakel, M., Halkjaer, J., Olsen, A., Agnoli, C., Psaltopoulou, T., Buurma, E., Sonestedt, E. & Chirlaque, M., 2009. Total dietary carbohydrate, sugar, starch and fibre intakes in the European Prospective Investigation into Cancer and Nutrition. *European Journal of Clinical Nutrition,* 63**,** S37-S60.

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., 2010. Dietary fiber and colorectal cancer risk: a nested case–control study using food diaries. *Journal of the National Cancer Institute,* 102**,** 614-626.

Dalstra, J.A., Kunst, A.E., Borrell, C., Breeze, E., Cambois, E., Costa, G., Geurts, J., Lahelma, E., Van Oyen, H. & Rasmussen, N., 2005. Socioeconomic

differences in the prevalence of common chronic diseases: an overview of eight European countries. *International Journal of Epidemiology,* 34**,** 316-326.

Dashti, B., Al-Awadi, F., Khalafawi, M.S., Sawaya, W. & Al-Amiri, H., 2003. Soluble and insoluble dietary fibre in thirty-two Kuwaiti dishes. *Food Chemistry,* 83**,** 557-561.

De Jong, N., Ocke, M.C., Branderhorst, H.A. & Friele, R., 2003. Demographic and lifestyle characteristics of functional food consumers and dietary supplement users. *British Journal of Nutrition,* 89**,** 273-282.

Deeks, J.J., Higgins, J. & Altman, D.G., 2008. Analysing Data and Undertaking Meta‐Analyses. *Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series***,** 243-296.

Deharveng, G., Charrondiere, U.R., Slimani, N., Southgate, D.A.T. & Riboli, E., 1999. Comparison of nutrients in the food composition tables available in the nine European countries participating in EPIC. *European Journal of Clinical Nutrition,* 53**,** 60-79.

DeMan, J.M., 1999. *Principles of food chemistry 3rd edition,* USA, Library of Congree Cataloging in -Publication data.

Department of Health. 1991. *Dietary Reference Values for Food and Energy and Nutrients for the United Kingdom: Report of the Panel on Dietary Reference Values of the Committee on Medical Aspects of Food Policy.,* London, The Stationary Office

Department of Health. 2001. National sevice framework for diabetes: standards.

Department of Health. 2012. National Diet and Nutrition Survey: Headline Results from Years 1, 2 and 3 (combined) of the Rolling Programme 2008/09 – 2010/11. Department of Health and Food Standard Agency.

Department of Health, D., 2011. National Diet and Nutrition Survey: Headline results from year 1 and 2 (combined) of the rolling programmed 2008/9-2009/10. UK: Department of Health and Food Standard Agency.

Derbyshire, E., 2011. The nutritional value of whole pulses and pulse fractions. *Pulse Foods: Processing, Quality and Nutraceutical Applications***,** 363.

DeVries, J.W., 2004. Dietary fiber: The influence of definition on analysis and regulation. *Journal of AOAC International,* 87**,** 682-706.

DeVries, J.W. & Rader, J.I., 2005. Historical perspective as a guide for identifying and developing applicable methods for dietary fiber. *Journal of AOAC International,* 88**,** 1349-1366.

Diabetes UK. 2009. Diabetes in the UK 2009: Key statistics on diabetes.

Diabetes UK. 2011. Evidence-based nutrition guidelines for the prevention and management of diabetes.

Diabetes UK. 2012. *Diabetes in the UK 2012: Key statistics on diabetes* [Online]. Available: www.diabetes.org.uk/Professionals/Publications-reports-and-resources/Reports-statistics-and-case-studies/Reports/Diabetes-in-the-UK-2012

Dilis, V. & Trichopoulou, A., 2009. Nutritional and health properties of pulses. *Mediterranean Journal of Nutrition and Metabolism,* 1**,** 149-157.

Dong, J.-Y., Xun, P., He, K. & Qin, L.-Q., 2011. Magnesium Intake and Risk of Type 2 Diabetes Meta-analysis of prospective cohort studies. *Diabetes Care,* 34**,** 2116-2122.

Drewnowski, A., Renderson, S.A., Driscoll, A. & Rolls, B.J., 1997. The Dietary Variety Score: Assessing Diet Quality in Healthy Young and Older Adults. *Journal of the American Dietetic Association,* 97**,** 266-271.

Du, H., Boshuizen, H.C., Forouhi, N.G., Wareham, N.J., Halkjær, J., Tjønneland, A., Overvad, K., Jakobsen, M.U., Boeing, H. & Buijsse, B., 2010. Dietary fiber and subsequent changes in body weight and waist circumference in European men and women. *American journal of clinical nutrition,* 91**,** 329-336.

Eastwood, M.A. & Morris, E.R., 1992. Physical properties of dietary fiber that influence physiological function: a model for polymers along the gastrointestinal tract. *The American journal of clinical nutrition,* 55**,** 436-442.

Eckel, R.H., Grundy, S.M. & Zimmet, P.Z., 2005. The metabolic syndrome. *Lancet,* 365**,** 1415-1428.

Eerlingen, R.C. & Delcour, J.A., 1995. Formation, analysis, structure an properties of type-III enzyme resistant starch. *Journal of Cereal Science,* 22**,** 129-138.

Englyst, H., Bingham, S., Runswick, S., Collinson, E. & Cummings, J., 1988. Dietary fibre (non‐starch polysaccharides) in fruit, vegetables and nuts. *Journal of Human Nutrition and Dietetics,* 1**,** 247-286.

Englyst, H., Wiggins, H.S. & Cummings, J.H., 1982. Determination of the Non-Starch Polysaccharides in plant foods by Gas-liquid-Chromotography of constituent sugars as alditol acetates. *Analyst,* 107**,** 307-318.

Englyst, H.N., Anderson, V. & Cummings, J.H., 1983. Starch and non‐starch polysaccharides in some cereal foods. *Journal of the Science of Food and Agriculture,* 34**,** 1434-1440.

Englyst, H.N. & Cummings, J.H., 1988. Improved method for measurement of dietary fiber as non-starch polysaccharides in plant foods. *Journal of AOAC International,* 71 808-814.

Englyst, H.N., Quigley, M.E., Englyst, K.N., Bravo, L. & and Hudson, G.J., 1996. 'Dietary fibre' Measurement by the Englyst NSP procedure Measurement by the AOAC Prosky Procedure Explanation of the differences; Report of a Study Commossioned by the Ministry of Agriculture, FIsheries and Food *Journal of Association Public Analysts,* 32**,** 1-38.

Englyst, K.N. & Englyst, H.N., 2005. Carbohydrate bioavailability. *British Journal of Nutrition,* 94**,** 1-11.

Englyst, K.N., Liu, S. & Englyst, H.N., 2007. Nutritional characterization and measurement of dietary carbohydrates. *European Journal of Clinical Nutrition,* 61**,** S19-S39.

Ericson, U., Sonestedt, E., Gullberg, B., Hellstrand, S., Hindy, G., Wirfalt, E. & Orho-Melander, M., 2013. High intakes of protein and processed meat

associate with increased incidence of type 2 diabetes. *British Journal of Nutrition,* 109**,** 1143-1153.

Eriksson, J., Lindström, J., Valle, T., Aunola, S., Hämäläinen, H., Ilanne-Parikka, P., Keinänen-Kiukaanniemi, S., Laakso, M., Lauhkonen, M. & Lehto, P., 1999. Prevention of Type II diabetes in subjects with impaired glucose tolerance: the Diabetes Prevention Study (DPS) in Finland Study design and 1-year interim report on the feasibility of the lifestyle intervention programme. *Diabetologia,* 42**,** 793-801.

Esposito, K., Kastorini, C.M., Panagiotakos, D.B. & Giugliano, D., 2010a. Prevention of type 2 diabetes by dietary patterns: a systematic review of prospective studies and meta-analysis. *Metabolic Syndrome and Related Disorders,* 8**,** 471-476.

Esposito, K., Maiorino, M.I., Ceriello, A. & Giugliano, D., 2010b. Prevention and control of type 2 diabetes by Mediterranean diet: a systematic review. *Diabetes Research and Clinical Practice,* 89**,** 97-102.

European Commission. 2008. Commission Directive 2008/100/EC of 28 October 2008 amending Council Directive 90/496/EEC on nutrition labelling for foodstuffs as regards recommended daily allowances, energy conversion factors and definitions. *Official Journal of the European Union,* 51**,** L285.

European Food Information Resource e-search. 2010. *EuroFIR eSearch Prototype* [Online]. Available: http://esearch.eurofir.org/ [Accessed 11 May 2011].

European Food Safety Authority. 2010. Scientific Opinion on Dietary Reference Values for carbohydrates and dietary fibre. *EFSA Journal,* 8**,** 1462.

European Food Safety Authority, E., 2011 Scientific Opinion on the substantiation of health claims related to beta-glucans from oats and barley and maintenance of normal blood LDL-cholesterol concentrations (ID 1236, 1299), increase in satiety leading to a reduction in energy intake (ID 851, 852), reduction of post-prandial glycaemic responses (ID 821, 824), and “digestive function” (ID 850) pursuant to Article 13(1) of Regulation (EC) No 1924/2006. *European Food Standard Agency Journal,* 9 2207-2228.

Fang, J., Austin, P.C. & Tu, J.V., Test for linearity between continuous confounder and binary outcome first, run a multivariate regression analysis second. Proceedings of the SAS Global Forum, 2009. Citeseer, 22-25.

FAO. 1998. Carbohydrates in human nutrition. Rome, Italy

Fardet, A., 2010. New hypotheses for the health-protective mechanisms of whole-grain cereals: what is beyond fibre? *Nutrition Research Reviews,* 23**,** 65-134.

Feskens, E., Bowles, C.H. & Kromhout, D., 1991. Carbohydrate intake and body mass index in relation to the risk of glucose intolerance in an elderly population. *American Journal of Clinical Nutrition,* 54**,** 136-140.

Feskens, E.J., Virtanen, S.M., Räsänen, L., Tuomilehto, J., Stengård, J., Pekkanen, J., Nissinen, A. & Kromhout, D., 1995. Dietary factors determining diabetes and impaired glucose tolerance: a 20-year follow-up of the Finnish and Dutch cohorts of the Seven Countries Study. *Diabetes Care,* 18**,** 1104-1112.

Feskens , E.J.M. & Kromhout , D., 1990. Habitual dietary intake and glucose tolerance in euglycaemic men: the Zutphen Study. *International Journal of Epidemiology,* 19**,** 953-959.

Flint, A., Raben, A., Astrup, A. & Holst, J.J., 1998. Glucagon-like peptide 1 promotes satiety and suppresses energy intake in humans. *Journal of Clinical Investigation,* 101**,** 515.

Food and Nutrition Board. 2001. *Dietary Reference Intakes: Proposed Definition of Dietary Fiber*, The National Academies Press.

Food and Nutrition Board. 2005. Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids (Macronutrients). Washington, DC: The National Academies Press.

Food Standards Agency. 2002. *McCance and Widdowson's The Composition of Foods, Sixth summary Edition*, Cambridge: Royal Society of Chemistry.

Food Standards Agency. 2007. FSA nutrient and food based guidelines for UK. Food Standard Agency

Food Standards Agency, F., 2010. *National Diet Nutrition Survey: headline results from year 1 (2008/2009)* [Online]. Available: http://tna.europarchive.org/20110116113217/http://www.food.gov.uk/science/dietarysurveys/ndnsdocuments/ndns0809year1 [Accessed 22 May 2011].

Forsyth, L.H. & Goetsch, V.L., 1997. Perceived threat of illness and health protective behaviors in offspring of adults with non-insulin-dependent diabetes mellitus. *Behavioral Medicine,* 23**,** 112-121.

Fung, T.T., McCullough, M., van Dam, R.M. & Hu, F.B., 2007. A prospective study of overall diet quality and risk of type 2 diabetes in women. *Diabetes Care,* 30**,** 1753-1757.

Fung, T.T., Schulze, M., Manson, J.E., Willett, W.C. & Hu, F.B., 2004. Dietary patterns, meat intake, and the risk of type 2 diabetes in women. *Archives of Internal Medicine,* 164**,** 2235-2240.

Furda, I., 2001. The Crude Fiber Method. *CRC handbook of dietary fiber in human nutrition.* CRC PressI Llc.

Galland, L., 2010. Diet and inflammation. *Nutrition in Clinical Practice,* 25**,** 634-640.

Goff, L.M., Griffin, B.A., Lovegrove, J.A., Sanders, T.A., Jebb, S.A., Bluck, L.J. & Frost, G.S., 2013. Ethnic differences in beta-cell function, dietary intake and expression of the metabolic syndrome among UK adults of South Asian, black African-Caribbean and white-European origin at high risk of metabolic syndrome. *Diabetes and Vascular Disease Research*.

Goñi, I., García-Diz, L., Mañas, E. & Saura-Calixto, F., 1996. Analysis of resistant starch: a method for foods and food products. *Food Chemistry,* 56**,** 445-449.

González, E.M., Johansson, S., Wallander, M.A. & Rodríguez, L.G., 2009. Trends in the prevalence and incidence of diabetes in the UK: 1996–2005. *Journal of Epidemiology and Community Health,* 63**,** 332-336.

Green, C., 2001. Fibre in enteral nutrition. *Clinical Nutrition,* 20**,** 23-39.

Greenwood, D., Cade, J., Draper, A., Barrett, J., Calvert, C. & Greenhalgh, A., 2000. Seven unique food consumption patterns identified among women in the UK Women's Cohort Study. *European Journal of Clinical Nutrition,* 54**,** 314-320.

Greenwood, D., Gilthorpe, M., Golding, C. & Cade, J., Stability over time of dietary patterns in the UK Women's Cohort Study. Proceedings-Nutrition Society of London, 2003. CABI Publishing; 1999, 89A-89A.

Grootenhuis, P.A., Westenbrink, S., Sie, C.M., De Neeling, J.N.D., Kok, F.J. & Bouter, L.M., 1995. A semiquantitative food frequency questionnaire for use in epidemiologic research among the elderly: Validation by comparison with dietary history. *Journal of Clinical Epidemiology,* 48**,** 859-868.

Guerrero-Romero, F., Tamez-Perez, H.E., González-González, G., Salinas-Martínez, A.M., Montes-Villarreal, J., Treviño-Ortiz, J.H. & Rodríguez-Morán, M., 2004. Oral Magnesium supplementation improves insulin sensitivity in non-diabetic subjects with insulin resistance. A double-blind placebo-controlled randomized trial. *Diabetes & Metabolism,* 30**,** 253-258.

Guillon, F. & Champ, M., 2000. Structural and physical properties of dietary fibres, and consequences of processing on human physiology. *Food Research International,* 33**,** 233-245.

Guillon, F. & Champ, M.M.J., 2002. Carbohydrate fractions of legumes: uses in human nutrition and potential for health. *British Journal of Nutrition,* 88**,** S293-S306.

Halton, T.L., Liu, S., Manson, J.E. & Hu, F.B., 2008. Low-carbohydrate-diet score and risk of type 2 diabetes in women. *The American journal of clinical nutrition,* 87**,** 339-346.

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

Harding, A.-H., Day, N.E., Khaw, K.-T., Bingham, S., Luben, R., Welsh, A. & Wareham, N.J., 2004. Dietary Fat and the Risk of Clinical Type 2 Diabetes The European Prospective Investigation of Cancer-Norfolk Study. *American Journal of Epidemiology,* 159**,** 73-82.

Harrison, T.A., Hindorff, L.A., Kim, H., Wines, R., Bowen, D.J., McGrath, B.B. & Edwards, K.L., 2003. Family history of diabetes as a potential public health tool. *American Journal of Preventive Medicine,* 24**,** 152-159.

Henderson, L., Gregory J & G, a.S., 2002. The National Diet and Nutrition Survey: adults aged 19 to 64 years Volume 1: Types and quantities of foods consumed. *In:* TSO (ed.). London

Hex, N., Bartlett, C., Wright, D., Taylor, M. & Varley, D., 2012. Estimating the current and future costs of Type 1 and Type 2 diabetes in the UK, including direct health costs and indirect societal and productivity costs. *Diabetic Medicine,* 29**,** 855-862.

Higgins, J.P., Thompson, S.G., Deeks, J.J. & Altman, D.G., 2003. Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal,* 327**,** 557.

Hindy, G., Sonestedt, E., Ericson, U., Jing, X.-J., Zhou, Y., Hansson, O., Renström, E., Wirfält, E. & Orho-Melander, M., 2012. Role of TCF7L2 risk variant and dietary fibre intake on incident type 2 diabetes. *Diabetologia,* 55**,** 2646-2654.

Hipsley, E.H., 1953. Dietary fibre and pregnancy toxaemia. *British Medical Journal,* 2**,** 420-422.

Hodge, A.M., English, D.R., O'Dea, K. & Giles, G.G., 2004. Glycemic index and dietary fiber and the risk of type 2 diabetes. *Diabetes Care,* 27**,** 2701-2706.

Holland, B., Brown, J. & Buss, D.H., 1993. *Fish and fish products: The third supplement to McCance & Widdowson's: the Composition of Foods (5th Edition),* Roylal Society of Chemistry.

Holland, B., Unmin, I.D. & Buss, D.H., 1991a. *Vegetables, Herbs and Spices: The fifith supplement to McCance & Widdowson's: The Composition of Foods (4th edition),* Royal Society of Chemisty.

Holland, B., Unmin, I.D. & Buss, D.H., 1992a. *Fruit and Nuts: the first supplement to McCance & Widdowson's: The composition of Foods (5th Edition),* Royal Society of Chemisty.

Holland, B., Unwin, I.D. & Buss, D.H., 1988. *Cereals and Cereal products: The third supplement to McCance & Widdowson's The Composition of Foods (4th Edition),* Royal Society of Chemistry.

Holland, B., Unwin, I.D. & Buss, D.H., 1989. *Milk products and eggs: The fourth supplement to McCance & Widdowson's: the Composition of Foods (4th Edition),* Royal Society of Chemistry

Holland, B., Welch, A.A. & Buss, D.H., 1992b. *Vegetable Dishes: The Second supplement to McCance & Widdowson's: The Composition of Food (5th Edition),* Royal Society of Chemistry.

Holland, B., Welch, A.A., Unwin, I.D., Buss, H.D., Paul, A.A. & Southgate, D.A.T., 1991b. *McCance and Widdowson's the Composition of Foods* Cambridge , UK, Royal Society of Chemistry

Hoover, R., Hughes, T., Chung, H.J. & Liu, Q., 2010. Composition, molecular structure, properties, and modification of pulse starches: A review. *Food Research International,* 43**,** 399-413.

Hopping, B.N., Etber, E., Grandinetti, A., Verheus, M., Kolonel, L.N. & Maskarinec, G., 2010. Dietary Fiber, Magnesium, and Glycemic Load Alter Risk of Type 2 Diabetes in a Multiethnic Cohort in Hawaii. *Journal of Nutrition,* 140**,** 68-74.

Howlett, J.F., Betteridge, V.A., Champ, M., Craig, S.A., Meheust, A. & Jones, J.M., 2010. The definition of dietary fiber–discussions at the Ninth Vahouny Fiber Symposium: building scientific agreement. *Food & Nutrition Research,* 54.

Hu, F.B. & Malik, V.S., 2010. Sugar-sweetened beverages and risk of obesity and type 2 diabetes: epidemiologic evidence. *Physiology & Behavior,* 100**,** 47-54.

Ioannidis, J. & Lau, J., 1999. Pooling research results: benefits and limitations of meta-analysis. *The Joint Commission Journal on Quality Improvement,* 25**,** 462.

Japan Diabetes Society. 1988. Diabetes mellitus in twins: a cooperative study in Japan. *Diabetes Research and Clinical Practice,* 5**,** 271-280.

Jenkins, D., Wolever, T., Taylor, R.H., Barker, H., Fielden, H., Baldwin, J.M., Bowling, A.C., Newman, H.C., Jenkins, A.L. & Goff, D.V., 1981. Glycemic index of foods: a physiological basis for carbohydrate exchange. *American Journal of Clinical Nutrition,* 34**,** 362-366.

Jenkins, D.J., Jenkins, A.L., Wolever, T.M., Collier, G.R., Rao, A.V. & Thompson, L.U., 1987. Starchy foods and fiber: reduced rate of digestion and improved carbohydrate metabolism. *Scandinavian Journal of Gastroenterology,* 22**,** 132-141.

Jenkins, D.J., Kendall, C.W., Axelsen, M., Augustin, L.S. & Vuksan, V., 2000. Viscous and nonviscous fibres, nonabsorbable and low glycaemic index carbohydrates, blood lipids and coronary heart disease. *Current Opinion in Lipidology,* 11**,** 49-56.

Jenkins, D.J., Wolever, T.M., Leeds, A.R., Gassull, M.A., Haisman, P., Dilawari, J., Goff, D.V., Metz, G.L. & Alberti, K.G., 1978. Dietary fibres, fibre analogues, and glucose tolerance: importance of viscosity. *British Medical Journal,* 1**,** 1392-1394.

Jenkins, D.J.A., Kendall, C.W.C., Vuksan, V., Vidgen, E., Parker, T., Faulkner, D., Mehling, C.C., Garsetti, M., Testolin, G., Cunnane, S.C., Ryan, M.A. & Corey, P.N., 2002. Soluble fiber intake at a dose approved by the US Food and Drug Administration for a claim of health benefits: serum lipid risk factors for cardiovascular disease assessed in a randomized controlled crossover trial. *American Journal of Clinical Nutrition,* 75**,** 834-839.

Jenkins, D.J.A., Leeds, A.R., Gassull, M.A., Cochet, B. & Alberti, K.G.M.M., 1977. Decrease in Postprandial Insulin and Glucose Concentrations by Guar and Pectin. *Annals of Internal Medicine,* 86**,** 20-23.

Jenkins, D.J.A.W., T. M. S.; Taylor, R. H.; Barker, H. M.; Fielden, H. 1980. Exceptionally low blood glucose response to dried beans: comparison with other carbohydrate foods. *British Medical Journal* 281 578-580

Jeon, C.Y., Lokken, R.P., Hu, F.B. & Van Dam, R.M., 2007. Physical Activity of Moderate Intensity and Risk of Type 2 Diabetes A systematic review. *Diabetes Care,* 30**,** 744-752.

Jing, Y., Han, G., Hu, Y., Bi, Y., Li, L. & Zhu, D., 2009. Tea consumption and risk of type 2 diabetes: a meta-analysis of cohort studies. *Journal of General Internal Medicine,* 24**,** 557-562.

Joint FAO/WHO Food Standards Programme. 2011. Report of the thirty second session of the Codex committee on nutrition and foods for special dietary uses Available: http://www.cclac.org/documentos/CCNFSDU/2011/1%20Alinorm/REP11_NFe.pdf.

Jones Jr, J., 1991. *Kjeldahl method for nitrogen determination*.

Kahn, S.E., Hull, R.L. & Utzschneider, K.M., 2006. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature,* 444**,** 840-846.

Kao, W., Puddey, I.B., Boland, L.L., Watson, R.L. & Brancati, F.L., 2001. Alcohol consumption and the risk of type 2 diabetes mellitus: atherosclerosis risk in communities study. *American Journal of Epidemiology,* 154**,** 748-757.

Kelsey, J.L., 1996. *Methods in observational epidemiology 2 nd*, Oxford University Press.

Kirk, S.F.L., Cade, J.E., Barrett, J.H. & Conner, M., 1999. Diet and lifestyle characteristics associated with dietary supplement use in women. *Public health nutrition,* 2**,** 69-73.

Kirkwood, B.R. & Sterne, J.A.C., 2003. Essential medical statistics: measurement error: assessment and implication. *Massachusetts: Blackwell Science Ltd.* 2nd ed.

Kleinbaum, D. & Klein, M., 2010. Logistic regression: A Self Learning Text–. Springer.

Knowler, W.C., Barrett-Connor, E., Fowler, S.E., Hamman, R.F., Lachin, J.M., Walker, E.A., Nathan, D.M. & null. 2002. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England Journal of Medicine,* 346**,** 393-403.

Knudsen, K.E.B. & Hessov, I., 1995. Recovery of Inulin from Jerusalem artichoke (Helianthus-Tuberosus L) in the small intestine of man. *British Journal of Nutrition,* 74**,** 101-113.

Kolattukudy, P.E., 1981. Structure, Biosynthesis, and Biodegradation of cutin and suberin. *Annual Review of Plant Physiology and Plant Molecular Biology,* 32**,** 539-567.

Krishnan, S., Rosenberg, L., Singer, M., Hu, F.B., Djousse, L., Cupples, L.A. & Palmer, J.R., 2007. Glycemie index, glycemie load, and cereal fiber intake and risk of type 2 diabetes in US black women. *Archives of Internal Medicine,* 167**,** 2304-2309.

Kromhout, D., Bloemberg, B., Seidell, J., Nissinen, A. & Menotti, A., 2001. Physical activity and dietary fiber determine population body fat levels: the Seven Countries Study. *International Journal of Obesity,* 25**,** 301-306.

Kutos, T., Golob, T., Kac, M. & Plestenjak, A., 2003. Dietary fibre content of dry and processed beans. *Food Chemistry,* 80**,** 231-235.

Landis, J.R. & Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics***,** 159-174.

Larsson, S. & Wolk, A., 2007. Magnesium intake and risk of type 2 diabetes: a meta - analysis. *Journal of Internal Medicine,* 262**,** 208-214.

Laurikkala, J., Juhola, M., Kentala, E., Lavrac, N., Miksch, S. & Kavsek, B., Informal identification of outliers in medical data. Proceedings of the 5th International Workshop on Intelligent Data Analysis in Medicine and Pharmacology, 2000. Citeseer, 20-24.

Lee, S.C. & Prosky, L., 1992. Dietary fiber analysis. *Cereal Foods World,* 37**,** 765.

Lee, S.C., Prosky, L. & Devries, J.W., 1992. Determination of total, soluble, and insoluble dietary fiber in foods - Enzymatic Gravimetric Method, MES-TRIS buffer - Collaborative study. *Journal of AOAC International,* 75**,** 395-416.

Leterme, P., 2002. Recommendations by health organizations for pulse consumption. *British Journal of Nutrition,* 88**,** 239-242.

Levine, M. & Ensom, M.H., 2001. Post hoc power analysis: an idea whose time has passed? *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy,* 21**,** 405-409.

Li, B.W., Andrews, K.W. & Pehrsson, P.R., 2002. Individual sugars, soluble, and insoluble dietary fiber contents of 70 high consumption foods. *Journal of Food Composition and Analysis,* 15**,** 715-723.

Li, B.W. & Cardozo, M.S., 1993. Simplified enzymatic-gravimetric method for total dietary fiber in legumes compared with a modified AOAC method. *Journal of Food Science,* 58**,** 929-932.

Li, B.W., ZhenKun, Z. & Jekot, J.J., 1997. Effect of lipid extraction methods on total dietary fiber and nonstarch polysaccharide contents of selected nuts and seeds. *Journal of AOAC International,* 80**,** 98-101.

Li, C.-Y. & Sung, F.-C., 1999. A review of the healthy worker effect in occupational epidemiology. *Occupational Medicine,* 49**,** 225-229.

Li, C. & Uppal, M., 2010. Canadian Diabetes Association National Nutrition Committee Clinical Update on Dietary Fibre in Diabetes: Food Sources to Physiological Effects. *Canadian Journal of Diabetes,* 34**,** 355-361.

Lillioja, S., Neal, A.L., Tapsell, L. & Jacobs, D.R., 2013. Whole grains, type 2 diabetes, coronary heart disease, and hypertension: Links to the aleurone preferred over indigestible fiber. *BioFactors,* 39**,** 242.

Lima, M.d., L, Cruz, T., Rodrigues, L.E., Bomfim, O.v., Melo, J., Correia, R., Porto, M., Cedro, A. & Vicente, E.z., 2009. Serum and intracellular magnesium deficiency in patients with metabolic syndrome--evidences for its relation to insulin resistance. *Diabetes Research and Clinical Practice,* 83**,** 257.

Lindstrom, J., Ilanne-Parikka, P., Peltonen, M., Aunola, S., Eriksson, J.G., Hemio, K., Hamalainen, H., Harkonen, P., Keinanen-Kiukaanniemi, S., Laakso, M., Louheranta, A., Mannelin, M., Paturi, M., Sundvall, J., Valle, T.T., Uusitupa, M., Tuomilehto, J. & Finnish Diabetes Prevention Study, G., 2006. Sustained reduction in the incidence of type 2 diabetes by lifestyle intervention: follow-up of the Finnish Diabetes Prevention Study. *Lancet,* 368**,** 1673-1679.

Lindström, J., Louheranta, A., Mannelin, M., Rastas, M., Salminen, V., Eriksson, J., Uusitupa, M. & Tuomilehto, J., 2003. The Finnish Diabetes Prevention Study (DPS) Lifestyle intervention and 3-year results on diet and physical activity. *Diabetes care,* 26**,** 3230-3236.

Liu, S., Willett, W.C., Manson, J.E., Hu, F.B., Rosner, B. & Colditz, G., 2003. Relation between changes in intakes of dietary fiber and grain products and changes in weight and development of obesity among middle-aged women. *American Journal of Clinical Nutrition,* 78**,** 920-927.

Liu, S.M., Buring, J.E., Sesso, H.D., Rimm, E.B., Willett, W.C. & Manson, J.E., 2002. A prospective study of dietary fiber intake and risk of cardiovascular disease among women. *Journal of the American College of Cardiology,* 39**,** 49-56.

Livesey, G., Taylor, R., Livesey, H. & Liu, S., 2013. Is there a dose-response relation of dietary glycemic load to risk of type 2 diabetes? Meta-analysis of prospective cohort studies. *American Journal of Clinical Nutrition,* 97**,** 584-596.

Lobo, R.A., 2008. Metabolic syndrome after menopause and the role of hormones. *Maturitas,* 60**,** 10-18.

Lunn, J. & Buttriss, J.L., 2007. Carbohydrates and dietary fibre. *Nutrition Bulletin,* 32**,** 21-64.

Malik, V.S., Popkin, B.M., Bray, G.A., Després, J.-P. & Hu, F.B., 2010a. Sugar-sweetened beverages, obesity, type 2 diabetes mellitus, and cardiovascular disease risk. *Circulation,* 121**,** 1356-1364.

Malik, V.S., Popkin, B.M., Bray, G.A., Després, J.-P., Willett, W.C. & Hu, F.B., 2010b. Sugar-Sweetened Beverages and Risk of Metabolic Syndrome and Type 2 Diabetes A meta-analysis. *Diabetes Care,* 33**,** 2477-2483.

Mañas, E., Bravo, L. & Saura-Calixto, F., 1994. Sources of error in dietary fibre analysis. *Food Chemistry,* 50**,** 331-342.

Mañas, E. & Saura-Calixto, F., 1993. Ethanolic precipitation: A source of error in dietary fibre determination. *Food Chemistry,* 47**,** 351-355.

Mann, J. & Cummings, J., 2009. Possible implications for health of the different definitions of dietary fibre. *Nutrition, Metabolism and Cardiovascular Diseases,* 19**,** 226-229.

Mann, J., Cummings, J., Englyst, H., Key, T., Liu, S., Riccardi, G., Summerbell, C., Uauy, R., Van Dam, R. & Venn, B., 2007. FAO/WHO scientific update on carbohydrates in human nutrition: conclusions. *European Journal of Clinical Nutrition,* 61**,** S132-S137.

Manson, J.E., Stampfer, M.J., Colditz, G.A., Willett, W.C., Rosner, B., Hennekens, C.H., Speizer, F.E., Rimm, E.B. & Krolewski, A.S., 1991. Physical activity and incidence of non-insulin-dependent diabetes mellitus in women. *Lancet,* 338**,** 774-778.

Marconi, E., Ruggeri, S., Cappelloni, M., Leonardi, D. & Carnovale, E., 2000. Physicochemical, nutritional, and microstructural characteristics of chickpeas (Cicer arietinum L.) and common beans (Phaseolus vulgaris L.) following microwave cooking. *Journal of Agricultural and Food Chemistry,* 48**,** 5986-5994.

Marlett, J.A., 1992. Content and composition of dietary fiber in 117 frequently consumed foods. *Journal of the American Dietetic Association,* 92**,** 175-186.

Marlett, J.A. & Longacre, M.J., 1996. Comparison of in vitro and in vivo measures of resistant starch in selected grain products. *Cereal Chemistry,* 73**,** 63-68.

Marshall, M., 2010. Ash Analysis. *In:* Nielsen (ed.) *Food Analysis.* Springer US.

Masson, L., McNeill, G., Tomany, J., Simpson, J., Peace, H., Wei, L., Grubb, D. & Bolton-Smith, C., 2003. Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic. *Public Health Nutrition,* 6**,** 313-321.

Maty, S.C., Everson-Rose, S.A., Haan, M.N., Raghunathan, T.E. & Kaplan, G.A., 2005. Education, income, occupation, and the 34-year incidence (1965–99)

of type 2 diabetes in the Alameda County Study. *International Journal of Epidemiology,* 34**,** 1274-1281.

Mayer-Davis, E.J., D'Agostino Jr, R., Karter, A.J., Haffner, S.M., Rewers, M.J., Saad, M. & Bergman, R.N., 1998. Intensity and amount of physical activity in relation to insulin sensitivity. *JAMA: Journal of the American Medical Association,* 279**,** 669-674.

Mayer-Davis, E.J., Monaco, J.H., Hoen, H.M., Carmichael, S., Vitolins, M.Z., Rewers, M.J., Haffner, S.M., Ayad, M.F., Bergman, R.N. & Karter, A.J., 1997. Dietary fat and insulin sensitivity in a triethnic population: the role of obesity. The Insulin Resistance Atherosclerosis Study (IRAS). *American Journal of Clinical Nutrition,* 65**,** 79-87.

McCane, D., Hanson, R.L., Charles, M.-A., Jacobsson, L.T., Pettitt, D.D., Bennett, P.H. & Knowler, W.C., 1994. Comparison of tests for glycated haemoglobin and fasting and two hour plasma glucose concentrations as diagnostic methods for diabetes. *British Medical Journal,* 308**,** 1323-1328.

McCleary, B. & Prosky, L., 2001. *Advanced dietary fibre technology,* Blackwell Science Ltd.

McCleary, B.V., 2007. An integrated procedure for the measurement of total dietary fibre (including resistant starch), non-digestible oligosaccharides and available carbohydrates. *Analytical and Bioanalytical Chemistry,* 389**,** 291-308.

McCleary, B.V., De Vries, J.W., Rader, J.I., Cohen, G., Prosky, L., Mugford, D.C., Champ, M. & Okuma, K., 2010. Determination of total dietary fiber (CODEX definition) by enzymatic-gravimetric method and liquid chromatography: Collaborative study. *Journal of AOAC International,* 93**,** 221-233.

McCullough, M.L., Feskanich, D., Stampfer, M.J., Giovannucci, E.L., Rimm, E.B., Hu, F.B., Spiegelman, D., Hunter, D.J., Colditz, G.A. & Willett, W.C., 2002. Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance. *American Journal of Clinical Nutrition,* 76**,** 1261-1271.

McDougall, G.J., Morrison, I.M., Stewart, D. & Hillman, J.R., 1996. Plant cell walls as dietary fibre: Range, structure, processing and function. *Journal of the Science of Food and Agriculture,* 70**,** 133-150.

Meeuwsen, S., Horgan, G. & Elia, M., 2010. The relationship between BMI and percent body fat, measured by bioelectrical impedance, in a large adult sample is curvilinear and influenced by age and sex. *Clinical Nutrition,* 29**,** 560-566.

Meisinger, C., Löwel, H., Thorand, B. & Döring, A., 2005. Leisure time physical activity and the risk of type 2 diabetes in men and women from the general population. *Diabetologia,* 48**,** 27-34.

Mertens, D., 2003. Challenges in measuring insoluble dietary fiber. *Journal of Animal Science,* 81**,** 3233-3249.

Meyer, K.A., Kushi, L.H., Jacobs, D.R. & Folsom, A.R., 2001. Dietary fat and incidence of type 2 diabetes in older Iowa women. *Diabetes Care,* 24**,** 1528-1535.

Meyer, K.A., Kushi, L.H., Jacobs, D.R., Slavin, J., Sellers, T.A. & Folsom, A.R., 2000. Carbohydrates, dietary fiber, and incident type 2 diabetes in older women. *American Journal of Clinical Nutrition,* 71**,** 921-930.

Mitchell, D.C., Lawrence, F.R., Hartman, T.J. & Curran, J.M., 2009. Consumption of dry beans, peas, and lentils could improve diet quality in the US population. *Journal of the American Dietetic Association,* 109**,** 909-913.

Mitri, J., Muraru, M.D. & Pittas, A.G., 2011. Vitamin D and type 2 diabetes: a systematic review. *European Journal of Clinical Nutrition,* 65**,** 1005-1015.

Mongeau, R. & Brassard, R., 1989. A comparison of three methods for analyzing dietary fibre in 38 foods *Journal of Food Composition and Analysis,* 2**,** 189-199.

Montonen, J., Järvinen, R., Knekt, P., Heliövaara, M. & Reunanen, A., 2007. Consumption of sweetened beverages and intakes of fructose and glucose predict type 2 diabetes occurrence. *Journal of Nutrition,* 137**,** 1447-1454.

Montonen, J., Knekt, P., Jarvinen, R., Aromaa, A. & Reunanen, A., 2003. Whole-grain and fiber intake and the incidence of type 2 diabetes. *American Journal of Clinical Nutrition,* 77**,** 622-629.

Morrish, N.J., Wang, S.L., Stevens, L.K., Fuller, J.H. & Keen, H., 2001. Mortality and causes of death in the WHO multinational study of vascular disease in diabetes. *Diabetologia,* 44**,** S14-S21.

Muley, A., Muley, P. & Shah, M., 2012. Coffee to reduce risk of type 2 diabetes? A systematic review. *Current Diabetes Reviews,* 8**,** 162-168.

Murakami, K., Okubo, H. & Sasaki, S., 2005. Effect of dietary factors on incidence of type 2 diabetes: a systematic review of cohort studies. *Journal of Nutritional Science and Vitaminology,* 51**,** 292.

Mussatto, S.I. & Mancilha, I.M., 2007. Non-digestible oligosaccharides: A review. *Carbohydrate Polymers,* 68**,** 587-597.

National Health Services. *Five a Day* [Online]. Available: http://www.nhs.uk/LiveWell/5ADAY/Pages/5ADAYhome.aspx [Accessed 16.06 2013].

National Institute of Health and Care Excellence. 2012. Preventing type 2 diabetes: risk identification and interventions for individuals at high risk.

National Institutes of Health. *Medline Plus* [Online]. U.S.A. Available: http://www.nlm.nih.gov/medlineplus/ [Accessed 20th March 2013].

NDNS. 2002. The National Diet & Nutrition Survey: adults aged 19 to 64 years. Her Majesty's Stationery Office (HMSO).

Nettleton, J.A., Hivert, M.-F., Lemaitre, R.N., McKeown, N.M., Mozaffarian, D., Tanaka, T., Wojczynski, M.K., Hruby, A., Djoussé, L. & Ngwa, J.S., 2013. Meta-Analysis Investigating Associations Between Healthy Diet and Fasting Glucose and Insulin Levels and Modification by Loci Associated With Glucose Homeostasis in Data From 15 Cohorts. *American Journal of Epidemiology,* 177**,** 103-115.

Nettleton, J.A., Lutsey, P.L., Wang, Y., Lima, J.A., Michos, E.D. & Jacobs, D.R., 2009. Diet soda intake and risk of incident metabolic syndrome and type 2

diabetes in the Multi-Ethnic Study of Atherosclerosis (MESA). *Diabetes Care,* 32**,** 688-694.

Nugent, A.P., 2005. Health properties of resistant starch. *Nutrition Bulletin,* 30**,** 27-54.

O'Shea, N., Arendt, E.K. & Gallagher, E., 2012. Dietary fibre and phytochemical characteristics of fruit and vegetable by-products and their recent applications as novel ingredients in food products. *Innovative Food Science & Emerging Technologies,* 16**,** 1-10.

Oakenfull, D., 2001. Physical chemistry of dietary fiber. *CRC handbook of dietary fiber in human nutrition***,** 33-44.

Office for National Statistics. 2012. Population Ageing in the United Kingdom, its Constituent Countries and the European Union. UK.

Official Methods of Analytical Chemists, A., 1995. AOAC Official Method 991.43 Total, Soluble, and Insoluble Dietary Fibre in Foods. *AOAC Official Methods of Analysis.*

Oguma, Y., Sesso, H.D., Paffenbarger, R.S. & Lee, I.M., 2005. Weight change and risk of developing type 2 diabetes. *Obesity Research,* 13**,** 945-951.

Oh, Y.N. & Grundleger, M.L., 1990. Improvement in soluble fiber content of wheat fiber through enzymic modification. *Journal of Agricultural and Food Chemistry,* 38**,** 1142-1145.

Olson, B.H., Anderson, S.M., Becker, M.P., Anderson, J.W., Hunninghake, D.B., Jenkins, D.J., LaRosa, J.C., Rippe, J.M., Roberts, D.C. & Stoy, D.B., 1997. Psyllium-enriched cereals lower blood total cholesterol and LDL cholesterol, but not HDL cholesterol, in hypercholesterolemic adults: results of a meta-analysis. *Journal of Nutrition,* 127**,** 1973-1980.

Oomah, B., Patras, A., Rawson, A., Singh, N. & Compos-Vega, R., 2011. Chemistry of pulses. *Pulse Foods: Processing, Quality and Nutraceutical Applications***,** 9-55.

Orozco, L.J., Buchleitner, A.M., Gimenez-Perez, G., Roqué, I.F.M., Richter, B. & Mauricio, D., 2008. Exercise or exercise and diet for preventing type 2 diabetes mellitus. *Cochrane Database Systematic Review,* 3.

Palafox‐Carlos, H., Ayala‐Zavala, J.F. & González‐Aguilar, G.A., 2011. The role of dietary fiber in the bioaccessibility and bioavailability of fruit and vegetable antioxidants. *Journal of Food Science,* 76**,** R6-R15.

Parkin, D.M., 2011. Cancers attributable to consumption of alcohol in the UK in 2010. *Br J Cancer,* 105**,** S14-S18.

Peacock, J.L. & Kerry, S., 2007. *Presenting medical statistics from proposal to publication: a step-by-step guide,* Oxford University Press.

Peattie, M.E., Buss, D.H., Lindsay, D.G. & Smart, G.A., 1983. Reorganization of the British total diet study for monitoring food constituents from 1981. *Food and Chemical Toxicology,* 21**,** 503-507.

Pereira, M.A., O'Reilly, E., Augustsson, K., Fraser, G.E., Goldbourt, U., Heitmann, B.L., Hallmans, G., Knekt, P., Liu, S.M., 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**,** 370-376.

Perera, A., Meda, V. & Tyler, R.T., 2010. Resistant starch A review of analytical protocols for determining resistant starch and of factors affecting the resistant starch content of foods. *Food Research International,* 43**,** 1959-1974.

Priebe, M., van Binsbergen, J., de Vos, R. & Vonk, R., 2008. Whole grain foods for the prevention of type 2 diabetes mellitus (Review).

Prosky, L., Asp, N.-G., Schweizer, T.F., DeVries, J.W., Furda, I. & Lee, S.C., 1994. Determination of soluble dietary fiber in foods and food products: collaborative study. *Journal of AOAC International,* 77**,** 690.

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**,** 677-679.

Prosky, L., Asp, N.G., Schweizer, T.F., Devries, J.W. & Furda, I., 1992. Determination of insoluble and soluble dietary fiber in foods and food-products-Collebrative study. *Journal of AOAC International,* 75**,** 360-367.

Rehm, J., Baliunas, D., Borges, G.L., Graham, K., Irving, H., Kehoe, T., Parry, C.D., Patra, J., Popova, S. & Poznyak, V., 2010. The relation between different dimensions of alcohol consumption and burden of disease: an overview. *Addiction,* 105**,** 817-843.

Rehman, Z.U. & Shah, W.H., 2005. Thermal heat processing effects on antinutrients, protein and starch digestibility of food legumes. *Food Chemistry,* 91**,** 327-331.

Reistad, R. & Frolich, W., 1984. Content and composition of dietary fiber in some fresh and cooked norwegian vegetables. *Food Chemistry,* 13**,** 209-224.

Resnick, L., Altura, B., Gupta, R., Laragh, J., Alderman, M. & Altura, B., 1993. Intracellular and extracellular magnesium depletion in type 2 (non-insulin-dependent) diabetes mellitus. *Diabetologia,* 36**,** 767-770.

Rimm, E.B., Manson, J.E., Stampfer, M.J., Colditz, G.A., Willett, W.C., Rosner, B., Hennekens, C.H. & Speizer, F.E., 1993. Cigarette smoking and the risk of diabetes in women. *American Journal of Public Health,* 83**,** 211-214.

Risérus, U., Willett, W.C. & Hu, F.B., 2009. Dietary fats and prevention of type 2 diabetes. *Progress in Lipid Research,* 48**,** 44.

Roberfroid, M. & Salvin, L.J., 2001. Resistant Oligosaccharides. *In:* Cho & Dreher (eds.) *Handbook of Dietary fibre* Switzerland.

Romaguera, D., Norat, T., Wark, P., Vergnaud, A., Schulze, M., van Woudenbergh, G., Drogan, D., Amiano, P., Molina-Montes, E. & Sanchez, M., 2013. Consumption of sweet beverages and type 2 diabetes incidence in European adults: results from EPIC-InterAct. *Diabetologia,* 56**,** 1520-1530.

Rose, D. & Pevalin, D.J., 2003. *A researcher's guide to the national statistics socio-economic classification,* SAGE Publications Ltd.

Rosenbloom, A.L., Joe, J.R., Young, R.S. & Winter, W.E., 1999. Emerging epidemic of type 2 diabetes in youth. *Diabetes Care,* 22**,** 345-354.

Rosner, B. & Gore, R., 2001. Measurement error correction in nutritional epidemiology based on individual foods, with application to the relation of diet to breast cancer. *American Journal of Epidemiology,* 154**,** 827-835.

Ruigómez, A., Martín‐Merino, E. & Rodríguez, L.A.G., 2010. Validation of ischemic cerebrovascular diagnoses in the health improvement network (THIN). *Pharmacoepidemiology and Drug Safety,* 19**,** 579-585.

Ruxton, C. & Derbyshire, E., 2010. Women's diet quality in the UK. *Nutrition Bulletin,* 35**,** 126-137.

Sajilata, M.G., Singhal, R.S. & Kulkarni, P.R., 2006. Resistant starch - A review. *Comprehensive Reviews in Food Science and Food Safety,* 5**,** 1-17.

Salmeron, J., Ascherio, A., Rimm, E.B., Colditz, G.A., Spiegelman, D., Jenkins, D.J., Stampfer, M.J., Wing, A.L. & Willett, W.C., 1997a. Dietary fiber, glycemic load, and risk of NIDDM in men. *Diabetes Care,* 20**,** 545-550.

Salmeron, J., Hu, F.B., Manson, J.E., Stampfer, M.J., Colditz, G.A., Rimm, E.B. & Willett, W.C., 2001. Dietary fat intake and risk of type 2 diabetes in women. *American Journal of Clinical Nutrition,* 73**,** 1019-1026.

Salmeron, J., Manson, J.E., Stampfer, M.J., Colditz, G.A., Wing, A.L. & Willett, W.C., 1997b. Dietary fiber, glycemic load, and risk of non-insulin-dependent diabetes mellitus in women. *JAMA: Journal of the American Medical Association,* 277**,** 472-477.

Salmerón, J.H., Frank, B., Manson, J.E., StampferMeir, J., A., C., RimmEric, B. & Willett, W.C., 2001. Dietary fat intake and risk of type 2 diabetes in women. *American Journal of Clinical Nutrition,* 73**,** 1019-1026.

Sandhu, K.S. & Lim, S.T., 2008. Digestibility of legume starches as influenced by their physical and structural properties. *Carbohydrate Polymers,* 71**,** 245-252.

Sarwar, N., Gao, P., Seshasai, S., Gobin, R., Kaptoge, S., Di Angelantonio, E., Ingelsson, E., Lawlor, D., Selvin, E. & Stampfer, M., 2010. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. *Lancet,* 375**,** 2215-22.

Sattar, N., Wannamethee, S. & Forouhi, N., 2008. Novel biochemical risk factors for type 2 diabetes: pathogenic insights or prediction possibilities? *Diabetologia,* 51**,** 926-940.

Saura-Calixto, F., 2010. Dietary fiber as a carrier of dietary antioxidants: an essential physiological function. *Journal of Agricultural and Food Chemistry,* 59**,** 43-49.

Saura-Calixto, F., García-Alonso, A., Goni, I. & Bravo, L., 2000. In vitro determination of the indigestible fraction in foods: an alternative to dietary fiber analysis. *Journal of Agricultural and Food Chemistry,* 48**,** 3342-3347.

Scanlon, P., 2008. The English national screening programme for sight-threatening diabetic retinopathy. *Journal of Medical Screening,* 15**,** 1-4.

Schienkiewitz, A., Schulze, M.B., Hoffmann, K., Kroke, A. & Boeing, H., 2006. Body mass index history and risk of type 2 diabetes: results from the European Prospective Investigation into Cancer and Nutrition (EPIC)–Potsdam Study. *American Journal of Clinical Nutrition,* 84**,** 427-433.

Schneider, A.V., 2002. Overview of the market and consumption of pulses in Europe. *British Journal of Nutrition,* 88**,** 243-250.

Schulze, M.B., Liu, S.M., Rimm, E.B., Manson, J.E., Willett, W.C. & Hu, F.B., 2004a. Glycemic index, glycemic load, and dietary fiber intake and incidence of type 2 diabetes in younger and middle-aged women. *American Journal of Clinical Nutrition,* 80**,** 348-356.

Schulze, M.B., Manson, J.E., Ludwig, D.S., Colditz, G.A., Stampfer, M.J., Willett, W.C. & Hu, F.B., 2004b. Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. *JAMA: Journal of the American Medical Association,* 292**,** 927-934.

Schulze, M.B., Schulz, M., Heidemann, C., Schienkiewitz, A., Hoffmann, K. & Boeing, H., 2007. Fiber and magnesium intake and incidence of type 2 diabetes - A prospective study and meta-analysis. *Archives of Internal Medicine,* 167**,** 956-965.

Schulze, M.B., Schulz, M., Heidemann, C., Schienkiewitz, A., Hoffmann, K. & Boeing, H., 2008. Carbohydrate intake and incidence of type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. *British Journal of Nutrition,* 99**,** 1107-1116.

Scientific Advisory Committee of Nutrition, S., 2008. *SACN statement on dietary fibre* [Online]. Available: http://www.sacn.gov.uk/pdfs/final_draft_sacnstatement_on_dietary_fibre_for_website.pdf [Accessed 20th March 2013].

Sebastian, R.S., Cleveland, L.E., Goldman, J.D. & Moshfegh, A.J., 2007. Older adults who use vitamin/mineral supplements differ from nonusers in nutrient intake adequacy and dietary attitudes. *Journal of the American Dietetic Association,* 107**,** 1322-1332.

Selvendran, R.R., 1984. The plant cell wall as a source of dietary fiber: chemistry and structure. *American Journal of Clinical Nutrition,* 39.

Shintani, M., Ogawa, Y., Ebihara, K., Aizawa-Abe, M., Miyanaga, F., Takaya, K., Hayashi, T., Inoue, G., Hosoda, K. & Kojima, M., 2001. Ghrelin, an endogenous growth hormone secretagogue, is a novel orexigenic peptide that antagonizes leptin action through the activation of hypothalamic neuropeptide Y/Y1 receptor pathway. *Diabetes,* 50**,** 227-232.

Sievenpiper, J.L., Kendall, C.W.C., Esfahani, A., Wong, J.M.W., Carleton, A.J., Jiang, H.Y., Bazinet, R.P., Vidgen, E. & Jenkins, D.J.A., 2009. Effect of non-oil-seed pulses on glycaemic control: a systematic review and meta-analysis of randomised controlled experimental trials in people with and without diabetes. *Diabetologia,* 52**,** 1479-1495.

Singh, B. & Saxena, A., 2010. Surrogate markers of insulin resistance: A review. *World Journal of Diabetes,* 1**,** 36.

Slavin, J., 2003a. Impact of the proposed definition of dietary fiber on nutrient databases. *Journal of Food Composition and Analysis,* 16**,** 287-291.

Slavin, J., 2003b. Why whole grains are protective: biological mechanisms. *Proceedings of the Nutrition Society,* 62**,** 129-134.

Slavin, J., 2004. Whole grains and human health. *Nutrition Research Reviews,* 17**,** 99-110.

Slavin, J. & Green, H., 2007. Dietary fibre and satiety. *Nutrition Bulletin,* 32**,** 32-42.

Slavin, J.L., 2005. Dietary fiber and body weight. *Nutrition,* 21**,** 411-418.

Slavin, J.L., Martini, M.C., Jacobs, D.R. & Marquart, L., 1999. Plausible mechanisms for the protectiveness of whole grains. *American Journal of Clinical Nutrition,* 70**,** 459s-463s.

Slimani, N., Deharveng, G., Unwin, I., Southgate, D., Vignat, J., Skeie, G., Salvini, S., Parpinel, M., Møller, A. & Ireland, J., 2007. The EPIC nutrient database project (ENDB): a first attempt to standardize nutrient databases across the 10 European countries participating in the EPIC study. *European Journal of Clinical Nutrition,* 61**,** 1037-1056.

Sluijs, I., Beulens, J.W., Spijkerman, A.M., Grobbee, D.E. & van der Schouw, Y.T., 2010a. Dietary intake of total, animal, and vegetable protein and risk of type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition (EPIC)-NL study. *Diabetes Care,* 33**,** 43-48.

Sluijs, I., van der Schouw, Y.T., van der A, D.L., Spijkerman, A.M., Hu, F.B., Grobbee, D.E. & Beulens, J.W., 2010b. Carbohydrate quantity and quality and risk of type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition-Netherlands (EPIC-NL) study. *American Journal of Clinical Nutrition,* 92**,** 905-911.

Smith, L.G., 2001. Plant cell division: Building walls in the right places. *Nature Reviews Molecular Cell Biology,* 2**,** 33-39.

Sonestedt, E., Øverby, N.C., Laaksonen, D.E. & Birgisdottir, B.E., 2012. Does high sugar consumption exacerbate cardiometabolic risk factors and increase the risk of type 2 diabetes and cardiovascular disease? *Food & Nutrition Research,* 56.

Song, Y., He, K., Levitan, E., Manson, J. & Liu, S., 2006. Effects of oral magnesium supplementation on glycaemic control in Type 2 diabetes: a meta‐analysis of randomized double‐blind controlled trials. *Diabetic Medicine,* 23**,** 1050-1056.

Southgate, D., 1978. Dietary fiber: analysis and food sources. *American Journal of Clinical Nutrition,* 31**,** S107-S110.

Southgate, D.A.T., 1969. Determination of Carbohydrates in Foods II._Unavailable carbohydrates *Journal of Agricultural and Food Chemistry,* 20**,** 331-335.

Southgate, D.A.T., 2001. Food Components associated with dietary fiber *In:* Spiller (ed.) *CRC Handbook of Dietary Fibre in Human Nutrition.* third edition ed. USA: CRC PressI Llc.

Spence, M., Cade, J.E., Burley, V.J. & Greenwood, D.C., Ability of the UK Women's Cohort Study food frequency questionnaire to rank dietary intakes: a preliminary validation study. Proceedings-Nutrition Society of London 2002. CABI Publishing; 1999, 117A.

Stampfer, M.J., Colditz, G.A., Willett, W.C., Manson, J.E., Arky, R.A., Hennekens, C.H. & Speizer, F.E., 1988. A prospective study of moderate alcohol

drinking and risk of diabetes in women. *American Journal of Epidemiology,* 128**,** 549-558.

Stephen, A.M., 1991. Starch and dietary fibre: their physiological and epidemiological interrelationships. *Canadian Journal of Physiology and Pharmacology,* 69**,** 116-120.

Stevens, J., Ahn, K., Juhaeri, Houston, D., Steffan, L. & Couper, D., 2002. Dietary fiber intake and glycemic index and incidence of diabetes in African-American and white adults. *Diabetes Care,* 25**,** 1715-1721.

Storlien, L.H., Higgins, J., Thomas, T., Brown, M.A., Wang, H., Huang, X.-F. & Else, P., 2000. Diet composition and insulin action in animal models. *British Journal of Nutrition,* 83**,** S85-S90.

Szmuilowicz, E.D., Stuenkel, C.A. & Seely, E.W., 2009. Influence of menopause on diabetes and diabetes risk. *Nature Reviews Endocrinology,* 5**,** 553-558.

Taylor, E., Burley, V., Greenwood, D. & Cade, J., 2007. Meat consumption and risk of breast cancer in the UK Women's Cohort Study. *British Journal of Cancer,* 96**,** 1139-1146.

The Institute of Food Science & Technology. 2007. Institute of Food Science and Technology Information Statement: Dietary Fibre.

Thompson, L., Yoon, J., Jenkins, D., Wolever, T. & Jenkins, A., 1984. Relationship between polyphenol intake and blood glucose response of normal and diabetic individuals. *American Journal of Clinical Nutrition,* 39**,** 745-751.

Threapleton, D., Greenwood, D., Evans, C., Cleghorn, C., Nykjaer, C., Woodhead, C., Cade, J., Gale, C. & Burley, V., 2012. Dietary fibre intake and risk of stroke: A systematic review and meta-analysis of prospective studies. *Proceedings of the Nutrition Society,* 71**,** E248.

Tierney, A., McMonagle, J., Shaw, D., Gulseth, H., Helal, O., Saris, W., JA Paniagua, I.G. & lstrok. 2011. Effects of dietary fat modification on insulin sensitivity and on other risk factors of the metabolic syndrome—LIPGENE: a European randomized dietary intervention study. *International Journal of Obesity,* 35**,** 800-809.

Tiwari, U. & Cummins, E., 2011. 5 - Functional and physicochemical properties of legume fibers. *In:* Brijesh, Aoife, Brian McKennaA2 - Brijesh K. Tiwari & Brian (eds.) *Pulse Foods.* San Diego: Academic Press.

Tonstad, S., 2009. Cigarette smoking, smoking cessation, and diabetes. *Diabetes Research and Clinical Practice,* 85**,** 4-13.

Trowell, H., 1972. Ischemic-disease and dietary fiber *American Journal of Clinical Nutrition,* 25**,** 926-932.

Trowell, H., Southgate, D., Wolever, T., Leeds, A., Gassull, M. & DJ., J., 1976. Letter: Dietary fibre redefined. *Lancet,* 1**,** 967.

Tu, Y.K. & Greenwood, D.C., 2012. *Modern methods for epidemiology,* Springer.

Tungland, B.C. & Meyer, D., 2002. Nondigestible Oligo- and Polysaccharides (Dietary Fiber): Their Physiology and Role in Human Health and Food. *Comprehensive Reviews in Food Science and Food Safety,* 1**,** 90-109.

Tuomilehto, J., Lindstrom, J., Eriksson, J.G., Valle, T.T., Hamalainen, H., Ilanne-Parikka, P., Keinanen-Kiukaanniemi, S., Laakso, M., Louheranta, A., Rastas, M., Salminen, V., Uusitupa, M., Aunola, S., Cepaitis, Z., Moltchanov, V., Hakumaki, M., Mannelin, M., Martikkala, V., Sundvall, J. & Finnish Diabet Prevention Study, G., 2001. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *New England Journal of Medicine,* 344**,** 1343-1350.

U.S Department of Agriculture, A.R.S., 2010. USDA National Nutritent Database for Standard Reference, Release 23. 2010 ed.

van Dam, R.M., 2006. Coffee and type 2 diabetes: From beans to beta-cells. *Nutrition, Metabolism and Cardiovascular Diseases,* 16**,** 69-77.

Van Dam, R.M. & Hu, F.B., 2005. Coffee consumption and risk of type 2 diabetes. *JAMA: Journal of the American Medical Association,* 294**,** 97-104.

Van Dam, R.M., Stampfer, M., Willett, W.C., Hu, F.B. & Rimm, E.B., 2002. Dietary fat and meat intake in relation to risk of type 2 diabetes in men. *Diabetes Care,* 25**,** 417-424.

Vazquez, G., Duval, S., Jacobs, D.R. & Silventoinen, K., 2007. Comparison of body mass index, waist circumference, and waist/hip ratio in predicting incident diabetes: a meta-analysis. *Epidemiologic Reviews,* 29**,** 115-128.

Veena, A., Urooj, A. & Puttaraj, S., 1995. Effect of processing on the composition of dietary fiber and starch in some legumes *Nahrung-Food,* 39**,** 132-138.

Venn, B.J. & Mann, J.I., 2004. Cereal grains, legumes and diabetes. *European Journal of Clinical Nutrition,* 58**,** 1443-1461.

Viera, A.J. & Garrett, J.M., 2005. Understanding interobserver agreement: the kappa statistic. *Family Medicine* 37**,** 360-363.

Villegas, R., Gao, Y.T., Yang, G., Li, H.L., Elasy, T.A., Zheng, W. & Shu, X.O., 2008. Legume and soy food intake and the incidence of type 2 diabetes in the Shanghai Women's Health Study. *JAMA: Journal of the American Medical Association,* 87**,** 162-167.

Walley, T. & Mantgani, A., 1997. The UK General Practice Research Database. *Lancet,* 350**,** 1097-1099.

Walters, S.J., 2009. Consultants' forum: should post hoc sample size calculations be done? *Pharmaceutical Statistics,* 8**,** 163-169.

Wannamethee, S., Papacosta, O., Whincup, P., Carson, C., Thomas, M., Lawlor, D., Ebrahim, S. & Sattar, N., 2010. Assessing prediction of diabetes in older adults using different adiposity measures: a 7 year prospective study in 6,923 older men and women. *Diabetologia,* 53**,** 890-898.

Wannamethee, S.G., Thomas, M.C., Whincup, P.H. & Sattar, N., 2009. Associations Between Dietary Fiber and Inflammation, Hepatic Function, and Risk of Type 2 Diabetes in Older Men Potential mechanisms for the benefits of fiber on diabetes risk. *Diabetes Care,* 32**,** 1823-1825.

Ward, H., Keogh, R., Lentjes, M., Luben, R., Wareham, N. & Khaw, K., 2011. 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.

Weickert, M.O., Möhlig, M., Schöfl, C., Arafat, A.M., Otto, B., Viehoff, H., Koebnick, C., Kohl, A., Spranger, J. & Pfeiffer, A.F.H., 2006. Cereal Fiber Improves Whole-Body Insulin Sensitivity in Overweight and Obese Women. *Diabetes Care,* 29**,** 775-780.

Weickert, M.O. & Pfeiffer, A.F., 2008. Metabolic effects of dietary fiber consumption and prevention of diabetes. *Journal of Nutrition,* 138**,** 439-442.

Weng, L.C., Lee, N.J., Yeh, W.T., Ho, L.T. & Pan, W.H., 2012. Lower intake of magnesium and dietary fiber increases the incidence of type 2 diabetes in Taiwanese. *Journal of the Formosan Medical Association,* 111**,** 651-659.

Westenbrink, S., Brunt, K. & van der Kamp, J.-W., 2013. Dietary fibre: Challenges in production and use of food composition data. *Food Chemistry,* 140**,** 562-567.

Whiting, D.R., Guariguata, L., Weil, C. & Shaw, J., 2011. IDF Diabetes Atlas: Global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Research and Clinical Practice,* 94**,** 311-321.

WHO/FAO, F.a.A.O.W.H.O., 1997. Carbohydrates in human nutrition: Report of a Joint FAO/WHO Expert Consiltation, Rome, 14-18 April 1997. FAO.

Wild, S., Roglic, G., Green, A., Sicree, R. & King, H., 2004. Global prevalence of diabetes - Estimates for the year 2000 and projections for 2030. *Diabetes Care,* 27**,** 1047-1053.

Will, J.C., Galuska, D.A., Ford, E.S., Mokdad, A. & Calle, E.E., 2001. Cigarette smoking and diabetes mellitus: evidence of a positive association from a large prospective cohort study. *International Journal of Epidemiology,* 30**,** 540-546.

Willett, W., 1998. *Nutritional Epidemiology,* New York, Oxford Univeristy Press

Willett, W., 2012. *Nutritional epidemiology,* Oxford University Press.

Willi, C., Bodenmann, P., Ghali, W.A., Faris, P.D. & Cornuz, J., 2007. Active smoking and the risk of type 2 diabetes: A systematic review and meta-analysis. *JAMA: Journal of the American Medical Association,* 298**,** 2654-2664.

Williamson, G., 2013. Possible effects of dietary polyphenols on sugar absorption and digestion. *Molecular Nutrition & Food Research,* 57**,** 48-57.

Wolters, M.G.E., Verbeek, C., Vanwesterop, J.J.M., Hermus, R.J.J. & Voragen, A.G.J., 1992. Comparison of different methods for determination of dietary fiber. *Journal of AOAC International,* 75**,** 626-634.

Wood, P.J., 2007. Cereal β-glucans in diet and health. *Journal of Cereal Science,* 46**,** 230-238.

Wood, P.J., 2010. Review: Oat and rye β-glucan: properties and function. *Cereal Chemistry,* 87**,** 315-330.

World Health Organization. 1999. *Definition, diagnosis, and classification of diabetes mellitus and its complications. Report of a WHO consultation* [Online]. Available: http://whqlibdoc.who.int/hq/1999/who_ncd_ncs_99.2.pdf [Accessed 5 May 2011].

World Health Organization. 2003. Diet, Nutrition and the Prevention of Chronic Diseases. *Joint WHO/FAO Expert Consultation.* Geneva, Switzerland.

World Health Organization. 2006. *Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation* [Online]. Available: http://www.who.int/diabetes/publications/Definition%20and%20diagnosis%20of%20diabetes_new.pdf [Accessed 5 May 2011].

World Health Organization. 2011. Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus. *Geneva (Switzerland): The Organization*.

World Health Organization. 2013. *Obesity and overweight* [Online]. Available: http://www.who.int/mediacentre/factsheets/fs311/en/ [Accessed 5 March 2013 2013].

Würsch, P. & Pi-Sunyer, F.X., 1997. The role of viscous soluble fiber in the metabolic control of diabetes: a review with special emphasis on cereals rich in β-glucan. *Diabetes Care,* 20**,** 1774-1780.

Wyness, L., 2009. Understanding the role of diet in type 2 diabetes prevention. *British Journal of Community Nursing,* 14**,** 374-379.

Xu, Q., Park, Y., Hollenbeck, A., Schatzkin, A., Chen, H. & Song, Y., 2011. Multivitamins, Individual Vitamin and Mineral Supplements, and Risk of Diabetes Among Older US Adults.

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**,** 1304-1313.

Yeh, H.-C., Duncan, B.B., Schmidt MIs, W.N. & Brancati, F., 2010. Smoking, smoking cessation, and risk for type 2 diabetes mellitus. *Annals of Internal Medicine,* 152**,** 10-17.

Appendix A: Description of the main dietary fibre fractions captured by current definitions

Fibre component	Description
Cellulose	Polysaccharides comprising up to 10,000 closely packed glucose units, arranged linearly, making cellulose very insoluble
Hemicelluloses	Polysaccharides containing sugars other than glucose. Present in both water soluble and insoluble forms
Pectins	Polysaccharides comprising galacturonic acid and a variety of sugars; soluble in hot water and forms gels on cooling.
B-glucan	Glucose polymers that, unlike cellulose, have a branched structure enabling them to form viscous solutions.
Resistant starch	Starch and starch degradation products that are not absorbed in the small intestine. Four classes have been identified: physically inaccessible starch, RS1; native starch granules, RS2; retrograded starch, RS3; and chemically modified starch, RS4.
Non-digestible oligosaccharides (NDOs)	NDOs comprising 3-9 sugar units occur naturally in plants. also be made chemically or enzymatically from mono- or disaccharides or by enzyme hydrolysis of polysaccharides. Fermentable and have prebiotic properties
Other synthetic carbohydrate compounds	Synthetic derivatives of cellulose (for example, methyl cellulose and hydroxypropylmethyl cellulose) are non-digestible and are soluble. But hardly fermented by microflora. Polydextrose has an average degree of polymerisation of 12 and is synthesised from glucose and sorbitol. It is partially fermented in the colon (~50% in humans) and has bulking and prebiotic properties.
Gums and mucilages	Gums are hydrocolloids derived from plant exudates. Mucilages are present in the cells of the outer layers of seeds of the plantain family, for example pysillium. Both are used as gelling agents, thickeners, stabilisers and emulsifying agents.
Lignin	Not a polysaccharide but chemically bound to hemicelluloses in plant cell walls.
Other minor components	Phytic acid (inositol hexaphosphate) is associated with fibre in some foods, especially cereal grains. May reduce mineral absorption in the small intestine as it binds strongly. Other compounds associated with fibre include tannins, cutins and phytosterols.

adopted from (Buttriss and Stokes, 2008)

Appendix B: Search key words for dietary fibre and type 2 diabetes mellitus

1. exp dietary fib$/
2. exp pectin/
3. non-starch polysaccharides.tw
4. dietary fibre.tw
5. dietary fiber.tw
6. non-starch polysaccharides.mp
7. complex carbohydrates.tw
8. dietary carbohydrates.tw
9. NSP.mp
10. or/1-9
11. limit 10 to (yr="1990 – Current")
12. diabetes mellitus.tw
13. diabetes.tw
14. diabetes mellitus/or exp insulin dependent diabetes mellitus/ or exp maturity onset diabetes mellitus/ or exp non-insulin dependent diabetes mellitus/
15. NIDDM.mp
16. exp diabetes prevention/
17. exp diabetes mellitus, Type 2/
18. diabetes mellitus type 2.mp
19. or/12-18
20. limit 19 to (yr="1990-Current")
21. cohort stud$.mp
22. exp prospective stud$.mp
23. incidence.mp or exp incidence/
24. relative risk.mp.
25. risk.mp
26. epidemiological stud$.mp
27. hazard ratio.mp
28. odds.ratio.mp
29. or/21-28 28
30. limit 19 to (yr="1990-Current")

Appendix C: Forest plots of dietary fibre intake, insoluble fibre intake and fibre sources with the risk of T2DM[16]

Figure 1.c Forest plot of total dietary fibre intake and risk of T2DM in USA studies

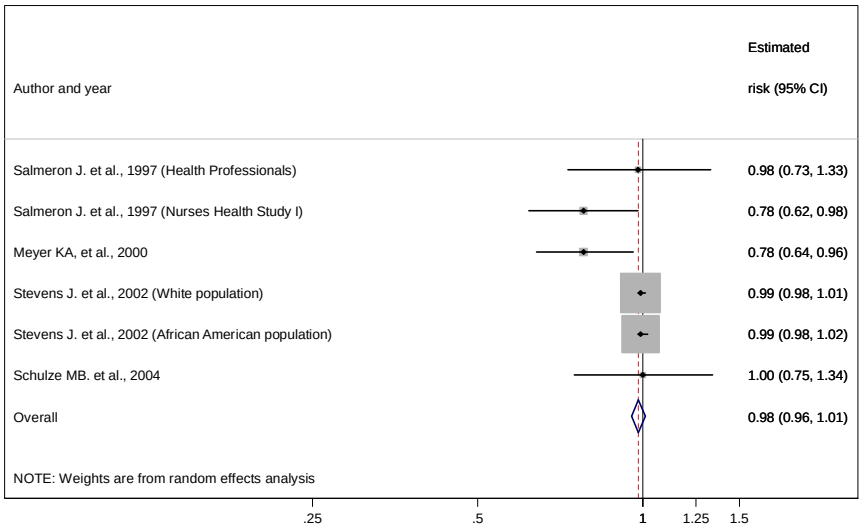
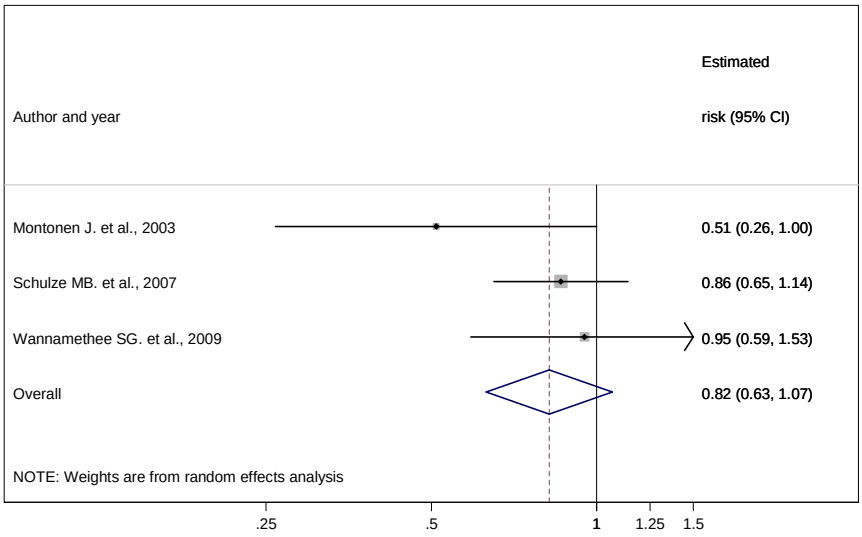


Figure 2.c Forest plot of total dietary fibre intake and risk of T2DM in European studies



[16] In all forest plots, the size of square is proportional to the weight of each study in the overall effect estimate and horizontal lines represent the 95%CI. Pooled estimate and its 95%CI are indicated by diamond figure.

Figure 3.c Forest plot of total dietary fibre and risk of T2DM of cohort studies with less than 7 years follow-up duration

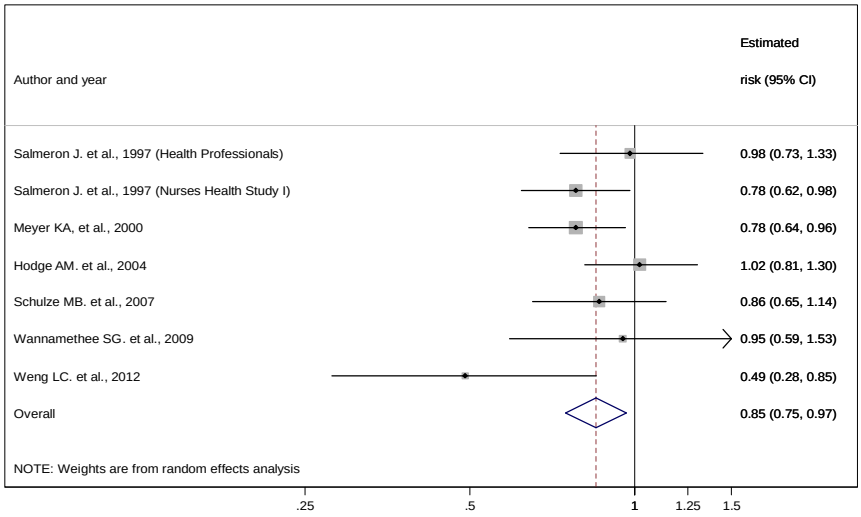


Figure 4.c Forest plot of total dietary fibre intake and risk of T2DM of studies with more than 7 years follow-up duration

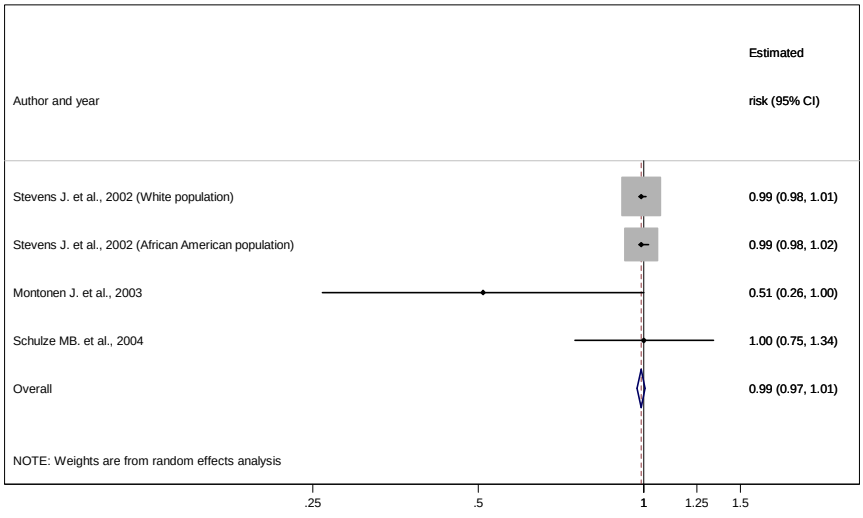


Figure 5.c Forest plot of insoluble dietary fibre intake and risk of T2DM

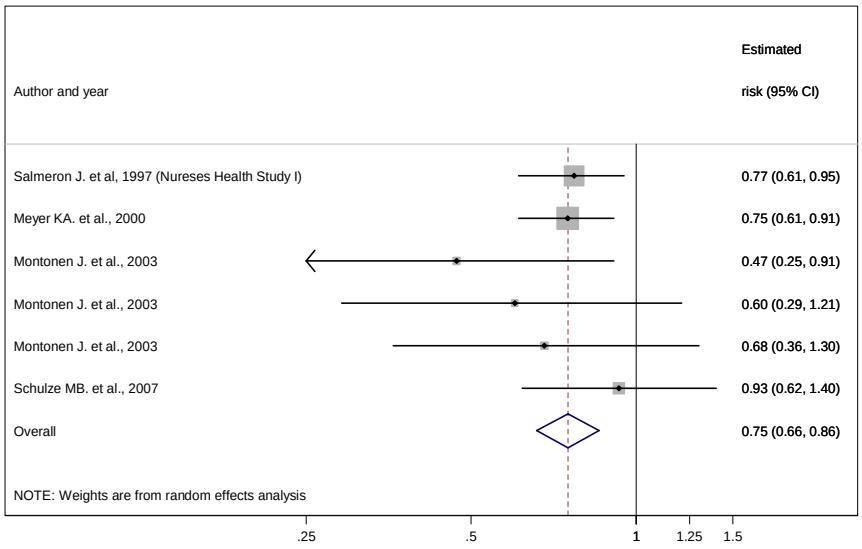


Figure 6.c Forest plot of cereal fibre intake (expressed as continuous variable) and risk of T2DM

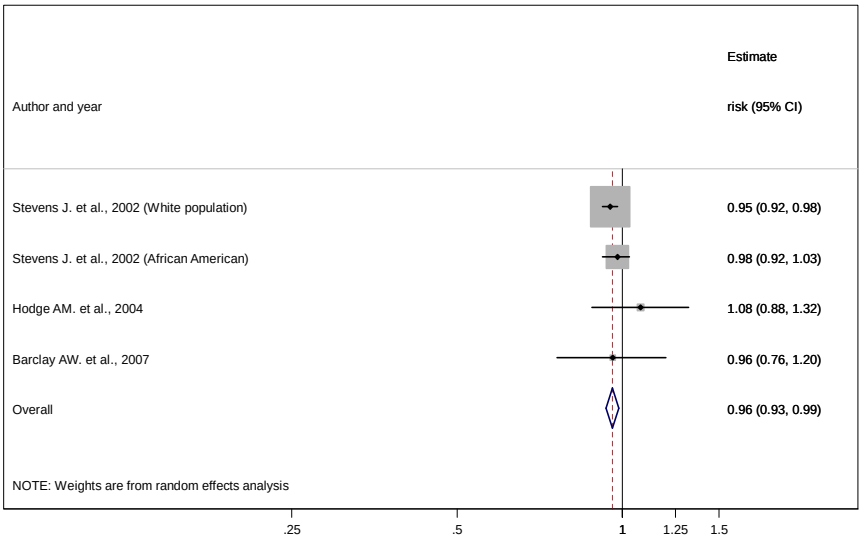


Figure 7.c Forest plot of vegetable fibre intake and risk of T2DM

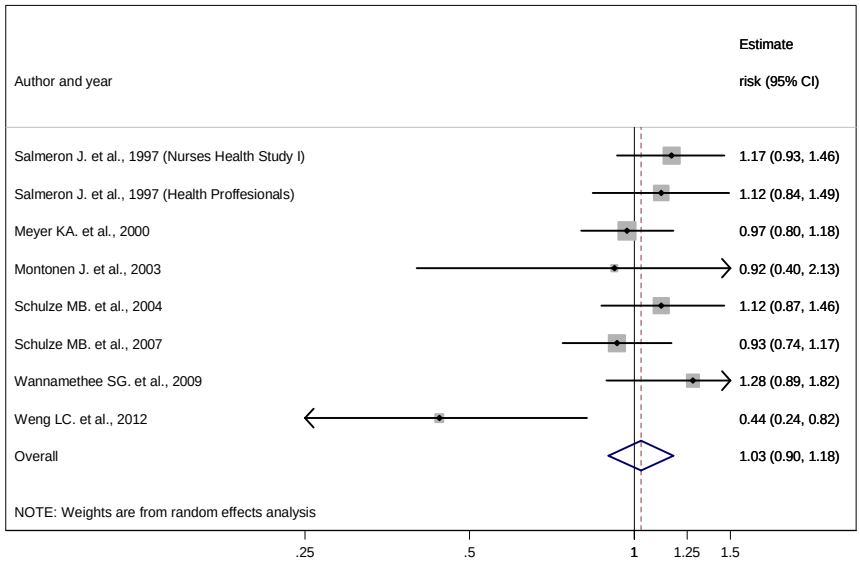


Figure 8.c Forest plot of fruit fibre intake and risk of T2DM

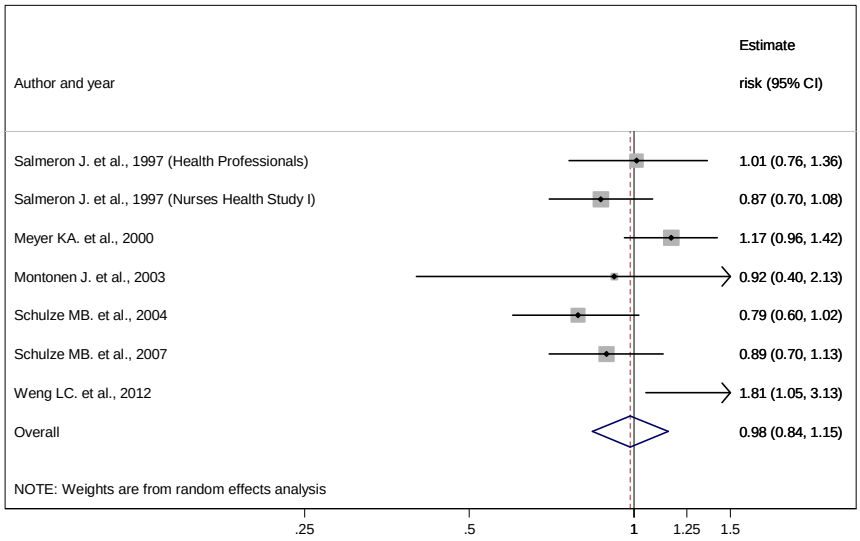


Figure 9.c Forest plot of estimated risk and 95% confidence interval for greater intake of soluble dietary fibre and incidence of T2DM

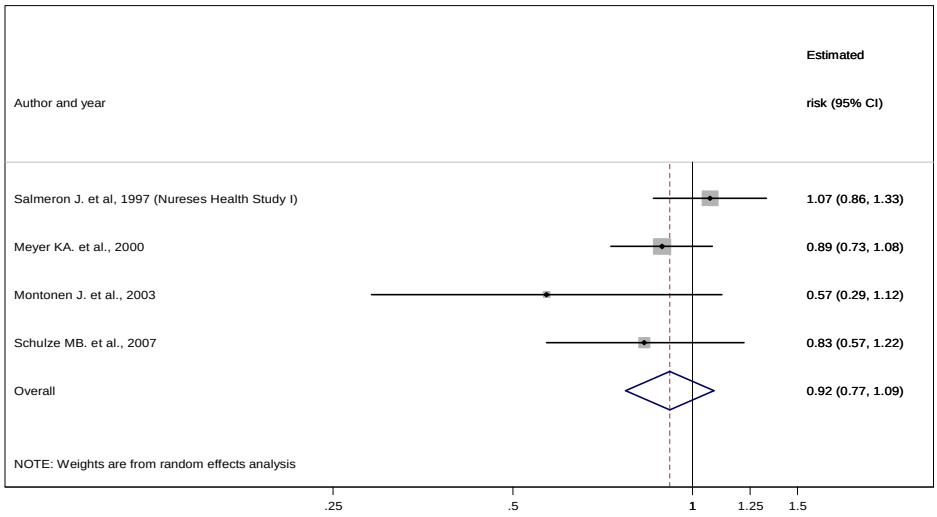
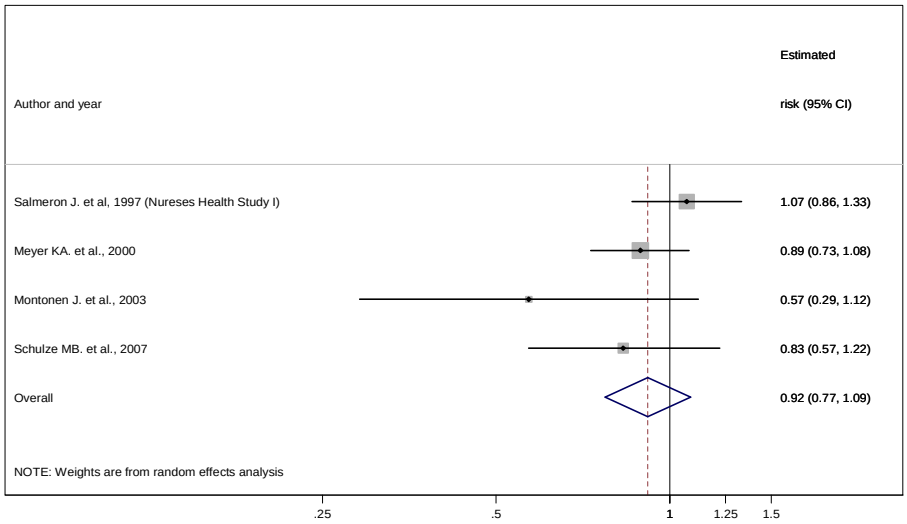


Figure 10.c Forest plot of estimated risk and 95% confidence interval for greater intake of dietary fibre and incidence of T2DM in women



Appendix D: Kjeldahl method

Sample

Weigh sample between 0.5 to 1 g of dry sample into a Kjeldahl flask

⇩

Digestion

Add two catalyst tablets with15mL of concentrated sulphuric acid then place the tubes in a heating block at 440°C for 30 min, or until bright green

⇩

Distillation

25mL of 4% boric acid solution into a 250 mL conical flask and add 2-3 drops of the screened methyl red indicator. NaOH will be dispensed into the tube, and the NH_3 gas distilled into the 250mL conical flask Sodium hydroxide is used to react with the ammonium sulphate, producing ammonia (gas) that is distilled off and bubbled into a solution of boric acid to produce ammonium borate. This is then titrated against hydrochloric acid to determine the level of nitrogen present.

⇩

Titration

Titrate the solutions against 0.05 M HCL (in burette) and record the end point to the nearest 0.1 mL

The carbon present is converted to carbon dioxide, the hydrogen to water, and the nitrogen to ammonium sulphate.

⇩

Calculation

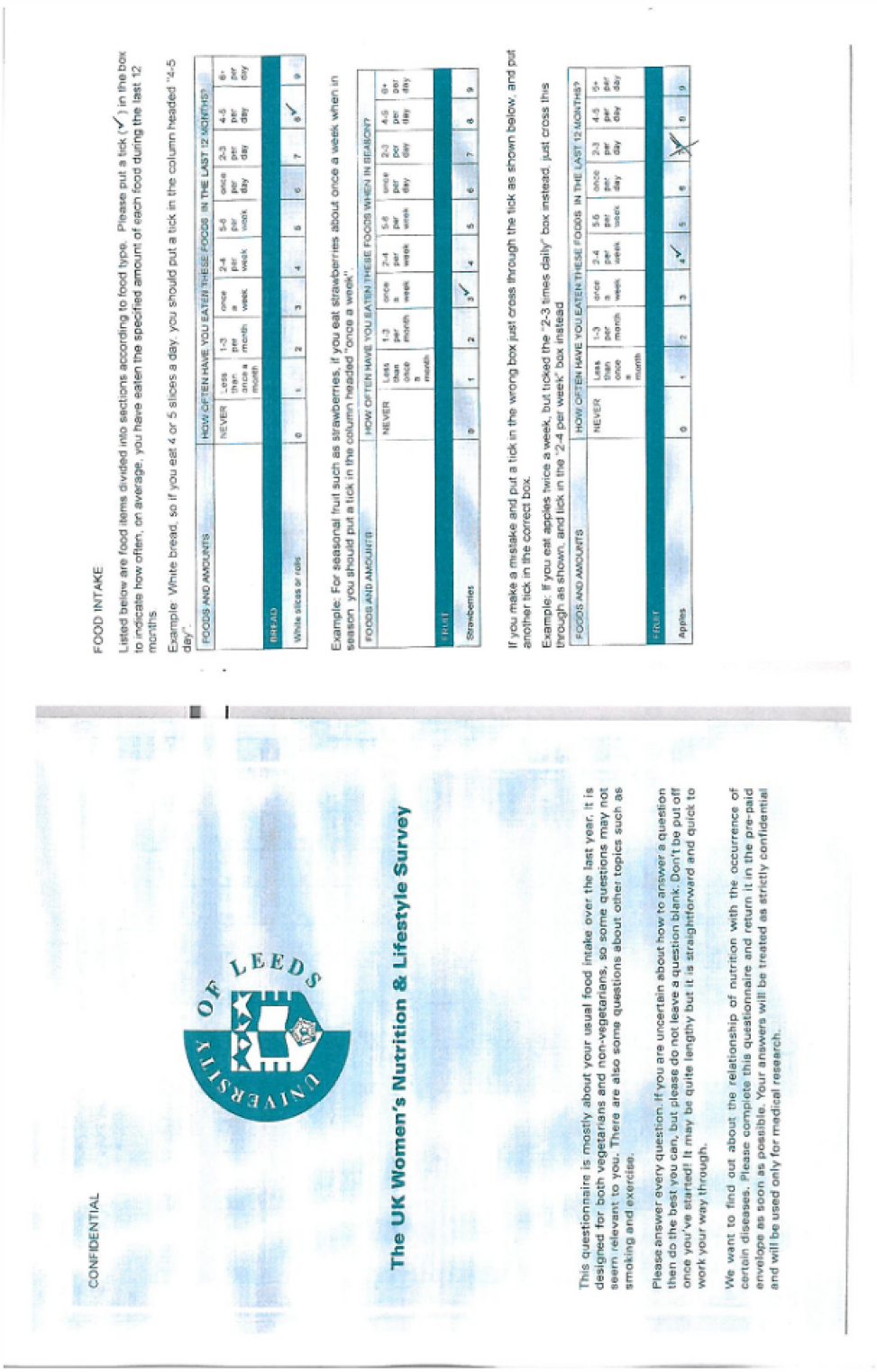
1 moles of HCl = 1 mole of NH_3 1mL of 0.05M HCl is equivalent to 0.0007g N N (g) = titre (mL) x 0.0007

$$\%N = \frac{N\ (g)}{\text{Dry sample weight (g)}} \times 100$$

; % protein = %N x 6.25

(Jones Jr, 1991)

Appendix E: Baseline semi-quantitative FFQ completed by UKWCS participants



CONFIDENTIAL

UNIVERSITY OF LEEDS

The UK Women's Nutrition & Lifestyle Survey

This questionnaire is mostly about your usual food intake over the last year. It is designed for both vegetarians and non-vegetarians, so some questions may not seem relevant to you. There are also some questions about other topics such as smoking and exercise.

Please answer every question. If you are uncertain about how to answer a question then do the best you can, but please do not leave a question blank. Don't be put off once you've started! It may be quite lengthy but it is straightforward and quick to work your way through.

We want to find out about the relationship of nutrition with the occurrence of certain diseases. Please complete this questionnaire and return it in the pre-paid envelope as soon as possible. Your answers will be treated as strictly confidential and will be used only for medical research.

FOOD INTAKE

Listed below are food items divided into sections according to food type. Please put a tick (✓) in the box to indicate how often, on average, you have eaten the specified amount of each food during the last 12 months.

Example: White bread, so if you eat 4 or 5 slices a day, you should put a tick in the column headed "4-5 day".

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										
White slices or rolls	0	1	2	3	4	5	6	7	8✓	9

Example: For seasonal fruit such as strawberries, if you eat strawberries about once a week when in season you should put a tick in the column headed "once a week".

FOODS AND AMOUNTS	HOW OFTEN HAVE YOU EATEN THESE FOODS WHEN IN SEASON?									
	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
FRUIT										
Strawberries	0	1	2	3✓	4	5	6	7	8	9

If you make a mistake and put a tick in the wrong box just cross through the tick as shown below, and put another tick in the correct box.

Example: If you eat apples twice a week, but ticked the "2-3 times daily" box instead, just cross this through as shown, and tick in the "2-4 per week" box instead

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
FRUIT										
Apples	0	1	2	3	4✓	5	6	✗	8	9

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
Papadums	0	1	2	3	4	5	6	7	8	9
Tortillas	0	1	2	3	4	5	6	7	8	9
Pitta Bread	0	1	2	3	4	5	6	7	8	9
Crispbread e.g. Ryvita	0	1	2	3	4	5	6	7	8	9
Cream crackers, cheese biscuits	0	1	2	3	4	5	6	7	8	9
BREAKFAST CEREALS										
Porridge, Readybrek	0	1	2	3	4	5	6	7	8	9
Sugar coated cereals e.g. Sugar Puffs	0	1	2	3	4	5	6	7	8	9
Non-sugar coated cereals e.g. Cornflakes, Rice Krispies	0	1	2	3	4	5	6	7	8	9
Muesli	0	1	2	3	4	5	6	7	8	9
All Bran, Bran Flakes	0	1	2	3	4	5	6	7	8	9
Weetabix, Shredded Wheat	0	1	2	3	4	5	6	7	8	9
POTATOES, RICE & PASTA										
Potatoes e.g. boiled, mashed	0	1	2	3	4	5	6	7	8	9
Chips	0	1	2	3	4	5	6	7	8	9
Jacket Potato	0	1	2	3	4	5	6	7	8	9
Roast Potatoes	0	1	2	3	4	5	6	7	8	9
Potato Salad	0	1	2	3	4	5	6	7	8	9
White Pasta e.g. Spaghetti, Green Pasta, Red Pasta, Noodles	0	1	2	3	4	5	6	7	8	9
Wholemeal Pasta, Brown Spaghetti	0	1	2	3	4	5	6	7	8	9
White Rice	0	1	2	3	4	5	6	7	8	9
Brown Rice	0	1	2	3	4	5	6	7	8	9
Wild Rice	0	1	2	3	4	5	6	7	8	9
Macaroni Cheese	0	1	2	3	4	5	6	7	8	9

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
DAIRY & NON-DAIRY PRODUCTS										
Thick & Creamy Yoghurt (125g carton)	0	1	2	3	4	5	6	7	8	9
Low fat Yoghurt (125g carton)	0	1	2	3	4	5	6	7	8	9
Diet Yoghurt (125g carton)	0	1	2	3	4	5	6	7	8	9
Greek Yoghurt (125g carton)	0	1	2	3	4	5	6	7	8	9
Fromage Frais/Creme Fraiche (125g carton)	0	1	2	3	4	5	6	7	8	9
Dairy Dessects (125g carton)	0	1	2	3	4	5	6	7	8	9
Single/Sour Cream (tablespoon)	0	1	2	3	4	5	6	7	8	9
Double/Clotted Cream (tablespoon)	0	1	2	3	4	5	6	7	8	9
Icecream	0	1	2	3	4	5	6	7	8	9
Milk Puddings	0	1	2	3	4	5	6	7	8	9
Low-fat Cheese	0	1	2	3	4	5	6	7	8	9
Cheese e.g. Cheddar, Brie, Edam	0	1	2	3	4	5	6	7	8	9
Cottage Cheese	0	1	2	3	4	5	6	7	8	9
Cheese and Onion Pastie	0	1	2	3	4	5	6	7	8	9
Soya Cheese	0	1	2	3	4	5	6	7	8	9
Soya Yoghurt	0	1	2	3	4	5	6	7	8	9
MARGARINES/BUTTERS & SPREADS										
Butter (enough for 1 slice of bread)	0	1	2	3	4	5	6	7	8	9
Block Margarine e.g. Stork, Krona, NOT in tub (enough for 1 slice of bread)	0	1	2	3	4	5	6	7	8	9
Polyunsaturated Margarine e.g. Flora, Sunflower, Granose, in tub (enough for 1 slice of bread)	0	1	2	3	4	5	6	7	8	9
Other soft Margarine Dairy spreads e.g. Blue Band, Clover in tub (enough for 1 slice of bread)	0	1	2	3	4	5	6	7	8	9
Low fat spread e.g. Outline, Gold, Flora Lite, in tub (enough for 1 slice of bread)	0	1	2	3	4	5	6	7	8	9
Very low fat spread, in tub e.g. St Ivel Lowest Fat Spread (enough for 1 slice of bread)	0	1	2	3	4	5	6	7	8	9
Monounsaturated Margarine eg. Meno Olivio (enough for 1 slice of bread)	0	1	2	3	4	5	6	7	8	9

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
SPREADS										
Marmite/Bovril/Vegemite	0	1	2	3	4	5	6	7	8	9
Peanut Butter	0	1	2	3	4	5	6	7	8	9
Chocolate/Chocolate & Nut Spread	0	1	2	3	4	5	6	7	8	9
Jam/Marmalade	0	1	2	3	4	5	6	7	8	9
Honey	0	1	2	3	4	5	6	7	8	9
Vegetable pâté	0	1	2	3	4	5	6	7	8	9
Nut Pâté	0	1	2	3	4	5	6	7	8	9
SAUCES & SOUPS										
Low Calorie Salad Cream (tablespoon)	0	1	2	3	4	5	6	7	8	9
Mayonnaise, Salad Cream Type Dressing (tablespoon)	0	1	2	3	4	5	6	7	8	9
French Type Dressing (tablespoon)	0	1	2	3	4	5	6	7	8	9
Sauces e.g. white/cheese/Cook In/curry	0	1	2	3	4	5	6	7	8	9
Tomato Ketchup (tablespoon)	0	1	2	3	4	5	6	7	8	9
Pickles/Chutney/Pesto sauce	0	1	2	3	4	5	6	7	8	9
Packet Soups - Meat & Veg (Bowl)	0	1	2	3	4	5	6	7	8	9
Other - Vegetable Soups (Bowl)	0	1	2	3	4	5	6	7	8	9
Other - Meat Soups (Bowl)	0	1	2	3	4	5	6	7	8	9
Low Calorie Soups (Bowl)	0	1	2	3	4	5	6	7	8	9
GRAINS (Medium serving)										
Barley	0	1	2	3	4	5	6	7	8	9
Oats	0	1	2	3	4	5	6	7	8	9
Bulgar Wheat	0	1	2	3	4	5	6	7	8	9
Wheat Germ (tablespoon)	0	1	2	3	4	5	6	7	8	9
Cous-cous	0	1	2	3	4	5	6	7	8	9
White Rice	0	1	2	3	4	5	6	7	8	9
Brown Rice	0	1	2	3	4	5	6	7	8	9
NUTS & SEEDS										
Peanuts/Pistachio Nuts	0	1	2	3	4	5	6	7	8	9
Cashew Nuts & Almonds	0	1	2	3	4	5	6	7	8	9
Pecan Nuts/Walnuts	0	1	2	3	4	5	6	7	8	9
Sunflower Seeds/ Sesame Seeds	0	1	2	3	4	5	6	7	8	9

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
PULSES (include when used in recipes)										
Lentils dals	0	1	2	3	4	5	6	7	8	9
Chick Peas, Chanas	0	1	2	3	4	5	6	7	8	9
Hummus	0	1	2	3	4	5	6	7	8	9
Baked beans	0	1	2	3	4	5	6	7	8	9
Mung Beans & Red Kidney Beans	0	1	2	3	4	5	6	7	8	9
Bean Sprouts	0	1	2	3	4	5	6	7	8	9
Black Eyed Beans	0	1	2	3	4	5	6	7	8	9
Butter Beans/Broad Beans	0	1	2	3	4	5	6	7	8	9
EGGS/EGG DISHES										
Boiled/ Poached egg	0	1	2	3	4	5	6	7	8	9
Omelette Scrambled egg	0	1	2	3	4	5	6	7	8	9
Fried egg	0	1	2	3	4	5	6	7	8	9
Quiche	0	1	2	3	4	5	6	7	8	9
VEGETABLE DISHES										
Quorn	0	1	2	3	4	5	6	7	8	9
Textured vegetable protein/ Sosmix/burger mix/soya sausages	0	1	2	3	4	5	6	7	8	9
Vegetarian Chili/Vegetable Curry	0	1	2	3	4	5	6	7	8	9
Mixed Bean Casserole/Ratatouille	0	1	2	3	4	5	6	7	8	9
Stir-fry vegetables	0	1	2	3	4	5	6	7	8	9
Vegetable - Lasagne/Moussaka/Ravioli/ filled pasta with sauce	0	1	2	3	4	5	6	7	8	9
Vegetable Pizza	0	1	2	3	4	5	6	7	8	9
MEAT										
Beef e.g. roast, steak	0	1	2	3	4	5	6	7	8	9
Beef Stew/Casserole/Mince/Curry	0	1	2	3	4	5	6	7	8	9
Beefburger/Hamburger	0	1	2	3	4	5	6	7	8	9
Pork e.g. Roast, Chops, Slices	0	1	2	3	4	5	6	7	8	9
Pork Stew/Casserole	0	1	2	3	4	5	6	7	8	9
Lamb e.g. Roast, Chops	0	1	2	3	4	5	6	7	8	9
Lamb Stew/Casserole	0	1	2	3	4	5	6	7	8	9

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
OTHER MEATS										
Chicken/Turkey roast, slices	0	1	2	3	4	5	6	7	8	9
Breadcrumbed e.g chicken nuggets/kievs	0	1	2	3	4	5	6	7	8	9
Chicken/Turkey in creamy sauce, curry	0	1	2	3	4	5	6	7	8	9
Bacon	0	1	2	3	4	5	6	7	8	9
Ham	0	1	2	3	4	5	6	7	8	9
Corned Beef, Spam, Luncheon Meats	0	1	2	3	4	5	6	7	8	9
Sausages e.g. Beef Pork	0	1	2	3	4	5	6	7	8	9
Pies/Pasties/Sausage Rolls	0	1	2	3	4	5	6	7	8	9
Offal e.g. Liver, Kidney	0	1	2	3	4	5	6	7	8	9
Liver Pâté/Sausage, Salami	0	1	2	3	4	5	6	7	8	9
Meat - Lasagne/Moussaka/Ravioli/ filled pasta with sauce	0	1	2	3	4	5	6	7	8	9
Meat Pizza	0	1	2	3	4	5	6	7	8	9
FISH										
Fish fingers/cakes	0	1	2	3	4	5	6	7	8	9
Fried fish in batter (as in fish and chips)	0	1	2	3	4	5	6	7	8	9
White fish e.g. Cod, Haddock, Plaice, Sole, Halibut (fresh or frozen)	0	1	2	3	4	5	6	7	8	9
Oily fish e.g. Mackerel, Kippers, Tuna, Salmon, Sardines, Herring	0	1	2	3	4	5	6	7	8	9
Shellfish e.g. Crab, Prawns, Mussels	0	1	2	3	4	5	6	7	8	9
Fish Roe, Taramasalata	0	1	2	3	4	5	6	7	8	9
Fish Pie/Fish Lasagne	0	1	2	3	4	5	6	7	8	9
VEGETABLES										
Beetroot	0	1	2	3	4	5	6	7	8	9
Broccoli, Spring Greens, Kale	0	1	2	3	4	5	6	7	8	9
Brussel Sprouts	0	1	2	3	4	5	6	7	8	9
Cabbage	0	1	2	3	4	5	6	7	8	9
Carrots	0	1	2	3	4	5	6	7	8	9
Cauliflower	0	1	2	3	4	5	6	7	8	9
Celery	0	1	2	3	4	5	6	7	8	9

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
VEGETABLES (continued)										
Coleslaw	0	1	2	3	4	5	6	7	8	9
Low-calorie Coleslaw	0	1	2	3	4	5	6	7	8	9
Courgettes, Marrow, Squash	0	1	2	3	4	5	6	7	8	9
Cucumber	0	1	2	3	4	5	6	7	8	9
Garlic	0	1	2	3	4	5	6	7	8	9
Green Beans, Runner Beans	0	1	2	3	4	5	6	7	8	9
Leeks	0	1	2	3	4	5	6	7	8	9
Lettuce	0	1	2	3	4	5	6	7	8	9
Mushrooms	0	1	2	3	4	5	6	7	8	9
Aubergine, Okra/Ladies Finger	0	1	2	3	4	5	6	7	8	9
Olives	0	1	2	3	4	5	6	7	8	9
Parsnips	0	1	2	3	4	5	6	7	8	9
Peas, Mushy peas, Mange-tout	0	1	2	3	4	5	6	7	8	9
Peppers - Red, Green, Yellow, Black etc	0	1	2	3	4	5	6	7	8	9
Swedes	0	1	2	3	4	5	6	7	8	9
Sweetcorn	0	1	2	3	4	5	6	7	8	9
Tomatoes - raw/canned/sauce	0	1	2	3	4	5	6	7	8	9
Turnip	0	1	2	3	4	5	6	7	8	9
Watercress, Mustard & Cress	0	1	2	3	4	5	6	7	8	9
FRUIT										
Apples	0	1	2	3	4	5	6	7	8	9
Avocado	0	1	2	3	4	5	6	7	8	9
Bananas	0	1	2	3	4	5	6	7	8	9
Grapes	0	1	2	3	4	5	6	7	8	9
Kiwi Fruit	0	1	2	3	4	5	6	7	8	9
Mangoes	0	1	2	3	4	5	6	7	8	9
Oranges, Satsumas, Grapefruit, etc	0	1	2	3	4	5	6	7	8	9
Papaya	0	1	2	3	4	5	6	7	8	9
Pears	0	1	2	3	4	5	6	7	8	9
Pineapple	0	1	2	3	4	5	6	7	8	9

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
SEASONAL FRUIT										
How often have you eaten these fruits, when they are in season?										
Apricots	0	1	2	3	4	5	6	7	8	9
Melon	0	1	2	3	4	5	6	7	8	9
Nectarines	0	1	2	3	4	5	6	7	8	9
Peaches	0	1	2	3	4	5	6	7	8	9
Plums	0	1	2	3	4	5	6	7	8	9
Raspberries	0	1	2	3	4	5	6	7	8	9
Red currants/Black currants	0	1	2	3	4	5	6	7	8	9
Rhubarb	0	1	2	3	4	5	6	7	8	9
Strawberries	0	1	2	3	4	5	6	7	8	9
DRIED FRUIT										
Dates	0	1	2	3	4	5	6	7	8	9
Figs	0	1	2	3	4	5	6	7	8	9
Prunes	0	1	2	3	4	5	6	7	8	9
Mixed Dried Fruit e.g. Apricots, Apples, Pears, Mangoes	0	1	2	3	4	5	6	7	8	9
Currants, Raisins, Sultanas	0	1	2	3	4	5	6	7	8	9
SWEET SNACKS										
Cereal Bars/Flapjacks (one)	0	1	2	3	4	5	6	7	8	9
Fruit bars (one) eg Apricot, Date	0	1	2	3	4	5	6	7	8	9
Chocolate Snack Bars e.g. Mars, Crunchie (1 bar)	0	1	2	3	4	5	6	7	8	9
Mini chocolate snack bars, Chocolates - singles or squares (1)	0	1	2	3	4	5	6	7	8	9
Boiled Sweets, Toffees, Mints	0	1	2	3	4	5	6	7	8	9
SAVOURY SNACKS										
Crisps (1 bag)	0	1	2	3	4	5	6	7	8	9
Other fried snacks e.g. Wotsits (1 bag)	0	1	2	3	4	5	6	7	8	9
Low fat or baked snacks e.g. Low-fat Crisps (1 bag)	0	1	2	3	4	5	6	7	8	9
Bombay Mix (small handful)	0	1	2	3	4	5	6	7	8	9
Peanuts/Pistachio Nuts (small handful)	0	1	2	3	4	5	6	7	8	9
Mixed Nuts and Raisins (small handful)	0	1	2	3	4	5	6	7	8	9

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
BEVERAGES										
Tea (cup)	0	1	2	3	4	5	6	7	8	9
Herbal Tea (cup)	0	1	2	3	4	5	6	7	8	9
Coffee - instant/ground (cup)	0	1	2	3	4	5	6	7	8	9
Coffee - decaffeinated (cup)	0	1	2	3	4	5	6	7	8	9
Coffee substitute e.g Caro/Bambu (cup)	0	1	2	3	4	5	6	7	8	9
Coffee whitener (teaspoon)	0	1	2	3	4	5	6	7	8	9
Cocoa, Hot Chocolate (cup)	0	1	2	3	4	5	6	7	8	9
Horlicks, Ovaltine (cup)	0	1	2	3	4	5	6	7	8	9
Low Calorie/Low-fat Horlicks, Ovaltine, Hot Chocolate (cup)	0	1	2	3	4	5	6	7	8	9
Orange Juice (Pure Fruit) (glass)	0	1	2	3	4	5	6	7	8	9
Other (100%) Pure fruit juices (glass)	0	1	2	3	4	5	6	7	8	9
Fruit Squash/Cordial - diluted (glass)	0	1	2	3	4	5	6	7	8	9
Fizzy soft drinks e.g. Coke, Lemonade (glass/can)	0	1	2	3	4	5	6	7	8	9
Low Calorie/Diet Soft Drinks (glass/can)	0	1	2	3	4	5	6	7	8	9
ALCOHOLIC BEVERAGES										
Wines (one glass/full)	0	1	2	3	4	5	6	7	8	9
Beer, Lager (half pint)	0	1	2	3	4	5	6	7	8	9
Cider (half pint)	0	1	2	3	4	5	6	7	8	9
Port, Sherry, Liqueurs (glass)	0	1	2	3	4	5	6	7	8	9
Spirits e.g. Whisky, Gin, Vodka, Brandy (single/1 measure)	0	1	2	3	4	5	6	7	8	9
BISCUITS, SWEETS & PUDDINGS										
Plain Biscuits e.g. Marie, Nice, Digestive (one)	0	1	2	3	4	5	6	7	8	9
Chocolate Biscuits (one)	0	1	2	3	4	5	6	7	8	9
Sandwich/Cream Biscuits (one)	0	1	2	3	4	5	6	7	8	9
Fruitcake (1 slice)	0	1	2	3	4	5	6	7	8	9
Sponge cakes (1 slice)	0	1	2	3	4	5	6	7	8	9
Buns/Pastries e.g. Croissants, Doughnuts, Tray Bakes, (one)	0	1	2	3	4	5	6	7	8	9
Scones/Pancakes/Muffins/Crumpets (1)	0	1	2	3	4	5	6	7	8	9
Fruit Pies, Tarts, Crumbles, (1 slice)	0	1	2	3	4	5	6	7	8	9
Sponge Puddings (1 serving)	0	1	2	3	4	5	6	7	8	9

1: **Other Foods**

Are there any other foods which you eat more than once a week? Yes [ ] 1 No [ ] 2
If yes, please list below

Food	Usual serving size	Number of times eaten each week

2: Would you describe yourself as a vegetarian? Yes [ ] 1 No [ ] 2
If yes, how long have you been vegetarian? [ ][ ] years.

Would you describe yourself as a vegan? Yes [ ] 1 No [ ] 2
If yes, how long have you been vegan? [ ][ ] years.

3: Do you use herbs and spices at least once per week when cooking food? Yes [ ] 1 No [ ] 2
Which fresh herbs and spices would you use at least once a week? Please list here

Which dried herbs and spices would you use at least once a week? Please list here

PORTION SIZE:

4: Compared to other people would you describe your typical average portion size of foods as?
Small? [ ] 1 Medium? [ ] 2 Large? [ ] 3

PULSES:

5: Do you eat pulses e.g. beans, peas, lentils etc. Yes [ ] 1 No [ ] 2
If no, please go to question 7.

6: Can you please indicate how much of the Pulses you eat are Fresh, Frozen, Canned or Dried. Please tick the appropriate boxes, e.g. 1/4 Dried, 3/4 Frozen.

	Proportion				
	Never	1/4	1/2	3/4	All
Fresh	1	2	3	4	5
Frozen	1	2	3	4	5
Dried	1	2	3	4	5
Canned	1	2	3	4	5

How do you usually cook pulses? Tick all applicable.

Steaming/Boiling/Pressure Cooking	[ ] 1	Stewing/Casseroling/Baking	[ ] 4
Microwaving	[ ] 2	Stir Frying/Frying	[ ] 5
Roasting	[ ] 3	Raw/soaked/Raw-sprouted	[ ] 6

VEGETABLES:

7: How many servings of vegetables or vegetable containing dishes, (excluding potatoes) do you usually eat each week? [ ][ ]

8: Can you please indicate how much of the vegetables you eat are Fresh, Frozen, Canned or Dried.
Please tick the appropriate boxes, e.g. 1/4 Dried. 3/4 Frozen.

	Proportion				
	Never	1/4	1/2	3/4	All
Fresh	1	2	3	4	5
Frozen	1	2	3	4	5
Dried	1	2	3	4	5
Canned	1	2	3	4	5

9: Do you ever eat raw vegetables apart from salad vegetables? Yes [ ] 1 No [ ] 2

10: How do you usually cook your vegetables.? (Excluding potatoes). Tick more than one box if necessary.

Boiling	[ ]
Steaming	[ ]
Grilling/Barbecuing/Baking/Roasting (Cooked dry or using a small amount of oil)	[ ]
Stir Frying/Frying/Sauté	[ ]
Microwaving	[ ]
Deep frying - including in batter	[ ]
Casseroling//Baking in sauce	[ ]
Other	[ ]

Please describe

FRUIT:

11: How many servings of fruit or fruit containing dishes do you usually eat each week? [ ][ ]
Can you please indicate how much of the fruit you eat is Fresh, Canned, Dried or Stewed.
Please tick the appropriate boxes e.g. 1/4 Fresh, 3/4 Canned

	Proportion				
	Never	1/4	1/2	3/4	All
Fresh	1	2	3	4	5
Stewed	1	2	3	4	5
Dried	1	2	3	4	5
Canned	1	2	3	4	5

12: Do you ever cook the fruit you eat? Yes [ ] 1 No [ ] 2

13: If so, how do you usually cook your fruit?

Stewing	[ ] 1	Poaching/Steaming	[ ] 2
Baking	[ ] 3	Microwaving	[ ] 4
Other	[ ] 5		

Please describe

MEAT: (If you **never** eat meat please go to question 16)

14: How many servings of meat or meat containing dishes do you usually eat each week? ☐☐

What do you do with the visible fat on your meat?
- Eat all/most of the fat ☐ 1
- Eat some of the fat ☐ 2
- Eat as little as possible/none ☐ 3

15: How do you usually cook meat? Tick more than one box if necessary.
- Grilling/Barbecuing/Baking/Roasting (Cooked dry or using a small amount of oil) ☐ 1
- Stir Frying/Frying ☐ 2
- Microwaving ☐ 3
- Deep frying - including in batter ☐ 4
- Casseroling/Baking in sauce ☐ 5
- Other ☐ 6

Please describe

16: How many servings of fish or fish containing dishes do you usually eat each week? ☐☐

How do you usually cook fish. Tick more than one box if necessary.
- Boiling ☐ 1
- Steaming ☐ 2
- Grilling/Barbecuing/Baking/Roasting (Cooked dry or using a small amount of oil) ☐ 3
- Stir Frying/Frying ☐ 4
- Microwaving ☐ 5
- Deep frying - including in batter ☐ 6
- Casseroling/Baking in sauce ☐ 7
- Other ☐ 8

Please describe

MILK:

17. What type of milk do you use most often? Select one only
- Full cream (Silver Top) ☐ 1
- Semi-skimmed (Red/White Top) ☐ 2
- Skimmed/fat free ☐ 3
- Channel Islands (Gold Top) ☐ 4
- Dried Milk ☐ 5
- Soya ☐ 6
- Sterilised ☐ 7
- None ☐ 8
- Other ☐ 9 Specify

If you used soya milk, please describe brand and type

18: How much milk do you drink each day, including milk with tea, coffee, milky drinks, cereals etc?
- None ☐ 1
- ¼ Pint ☐ 2
- ½ Pint ☐ 3
- ¾ Pint ☐ 4
- 1 Pint ☐ 5
- More than 1 Pint ☐ 6

BREAKFAST:

19: Are there any breakfast cereals that you normally eat that were not mentioned earlier? Yes ☐ 1 No ☐ 2

If yes, which brand and type of breakfast cereal, do you usually eat?

List the types most often used

Brand	Type

20: Do you usually take sugar on your breakfast cereal? Yes ☐ 1 No ☐ 2

If yes, how many teaspoons? ☐☐ teaspoons

21: Do you usually take sugar/honey in tea, herbal tea, coffee or coffee substitute? Yes ☐ 1 No ☐ 2

If Yes, please write the number of teaspoons per cup.
- Sugar/honey in tea ☐☐ teaspoons
- Sugar/honey in herbal tea ☐☐ teaspoons
- Sugar/honey in coffee ☐☐ teaspoons
- Sugar/honey in coffee substitute ☐☐ teaspoons

Do you use sweeteners instead of sugar or honey. Yes ☐ 1 No ☐ 2

Which brand of sweetener do you use, please specify

If yes how many tablets per day, or how many teaspoons of powder per day? ☐☐

22: On days when you eat bread, how many slices of bread or rolls do you eat ? ☐☐ slices/rolls per day

USE OF FATS:

23: Do you usually spread butter/margarine on your bread?
Yes ☐ 1 No ☐ 2 Sometimes ☐ 3

How many slices of bread/rolls/crackers do you have with spread each day? ☐☐

How much spread do you use?
- Just a scrape/thinly spread ☐ 1
- medium ☐ 2
- Thickly spread ☐ 3

24: What kind of fat do you most often use for frying, roasting, grilling etc?

Tick more than one if applicable
- Butter ☐ 1
- Lard/Dripping ☐ 2
- Vegetable Oil ☐ 3
- Solid White Vegetable Fat ☐ 4
- Margarine ☐ 5
- None ☐ 6

If you used vegetable oil, or margarine, please give type e.g. corn, sunflower

25: What kind of fat do you most often use for baking cakes etc.? Tick more than one if applicable

Butter ☐ 1 — Solid White Vegetable Fat ☐ 2
Lard/Dripping ☐ 3 — Margarine ☐ 4
Vegetable Oil ☐ 5 — None ☐ 6

If you use margarine, please give Brand e.g. Flora, Stork

USE OF SALT:

26: How often do you add salt to food while cooking?

Always ☐ 1 — Usually ☐ 2
Sometimes ☐ 3 — Rarely ☐ 4
Never ☐ 5

27: How often do you add salt to any food at the table?

Always ☐ 1 — Usually ☐ 2
Sometimes ☐ 3 — Rarely ☐ 4
Never ☐ 5

28: Do you regularly use a salt substitute (e.g. Losalt)? Yes ☐ 1 No ☐ 2

If yes, which brand?

USE OF SUPPLEMENTS:

29: Do you take any vitamins, minerals, fish oils, fibre or other food supplements? Yes ☐ 1 No ☐ 2

If yes, please fill in details below.

Name and Brand of Supplements	How much do you take at a time	How often do you take these? Daily	Weekly	Monthly	Less often
		1	2	3	4
		1	2	3	4
		1	2	3	4

SPECIAL DIETS:

30: I) Have you changed your diet over the last 12 months? Yes ☐ 1 No ☐ 2

If yes, please indicate if the change was for any of the reasons listed below?

Tick more than one box if applicable

High Blood Pressure ☐ 8 — Stomach problems (e.g. ulcer or gastritis) ☐ 1
Bowel problems (e.g. irritable bowel or diverticulitis) ☐ 2 — Concern over eating a healthy Diet ☐ 3
Concern over a family history of illness ☐ 4 — High Blood Cholesterol/Lipids ☐ 5
Overweight/Obesity ☐ 6 — Diabetes ☐ 7
Allergies (e.g. skin rash) ☐ 9 — Other ☐ 0

Specify

II) Describe below how your diet has changed

Do you currently follow any of these diets? Tick more than one box if neccessary.

Low fat ☐ 1 — Low salt ☐ 2 — Diabetic ☐ 3
Slimming ☐ 4 — Gluten free ☐ 5 — High fibre ☐ 6
Other ☐ 7

Please give details

CONSUMPTION OF ALCOHOL:

31: How often, if ever do you drink alcohol?

More than once a week ☐ 1 — Once a week ☐ 2
Less than once a week ☐ 3 — Never drink alcohol ☐ 4

32: In a typical week, how much do you drink?

Beer or cider ☐☐ pints each week
Wine ☐☐ glasses each week
Sherry/Fortified Wines ☐☐ glasses each week
Spirits ☐☐ glasses (singles) each week

33: Five years ago, how many alcoholic drinks did you have each week?

Beer or cider ☐☐ pints each week
Wine ☐☐ glasses each week
Sherry/Fortified Wines ☐☐ glasses each week
Spirits ☐☐ glasses (singles) each week

SMOKING:

34: Which one of the following best describes you?

I smoke every day ☐ 1 — I smoke occasionally, but not every day ☐ 2
I used to smoke every day, but do not smoke at all now ☐ 3 — I have never smoked ☐ 4

If you have never smoked, please go to question 37.

35: Do/did you smoke?

Cigarettes ☐ 1
Cigars ☐ 2
A combination of the above ☐ 3

If you currently smoke or used to smoke cigarettes how many do/did you smoke each day?

☐☐ cigarettes

If you currently smoke or used to smoke cigarettes which brand of cigarettes do/did you usually smoke?

36: If you have stopped smoking for what period of time have you been a non-smoker?

1 year or less ☐ 1 2-5 years ☐ 2

6-10 years ☐ 3 Over 10 years ☐ 4

SIZE:

37: Approximately how much did you weigh when you were born?

☐☐ lbs or ☐☐ Kg or Don't Know ☐

38: Approximately how much did you weigh when you were 20 years old?

☐☐ stones ☐☐ pounds or ☐☐☐ Kg or Don't Know ☐

39: Approximately how much do you weigh at present?

☐☐ stones ☐☐ pounds or ☐☐☐ Kg or Don't Know ☐

40: Have you lost more than half a stone in the last year? Yes ☐ 1 No ☐ 2

Have you gained more than half a stone in the last year? (Please ignore weight gained during pregnancy.) Yes ☐ 1 No ☐ 2

41: What is your present waist size? ☐☐ inches or ☐☐☐ centimetres or Don't Know ☐ 1

42: What is your present hip size? ☐☐ inches or ☐☐☐ centimetres or Don't Know ☐ 1

43: What is your present height? ☐ ft ☐☐ inches or ☐☐☐ centimetres or Don't Know ☐ 1

44: What size of blouse do you wear? Size ☐☐

45: What size of skirt do you wear? Size ☐☐

PHYSICAL ACTIVITY

46: In a typical week during the last 12 months, how many hours did you spend on each of the following activities? Put "0" if none

Housework, such as cleaning, washing, cooking, child care		☐☐ hours	☐☐ minutes per week
Do-it-Yourself		☐☐ hours	☐☐ minutes per week
Gardening	In Summer	☐☐ hours	☐☐ minutes per week
	In Winter	☐☐ hours	☐☐ minutes per week
Walking, including to work, shopping & leisure	In Summer	☐☐ hours	☐☐ minutes per week
	In Winter	☐☐ hours	☐☐ minutes per week
Cycling, including to work & leisure	In Summer	☐☐ hours	☐☐ minutes per week
	In Winter	☐☐ hours	☐☐ minutes per week
Other physical exercise, such as keep-fit, aerobics, jogging, tennis, swimming	In Summer	☐☐ hours	☐☐ minutes per week
	In Winter	☐☐ hours	☐☐ minutes per week

47: In a normal week, do you do any of these activities vigorously enough to cause sweating or a faster heartbeat? Yes ☐ 1 No ☐ 2

If yes, for how long each week do you do such vigorous physical activity? ☐☐ hours ☐☐ minutes per week

ILLNESS:

48: Have you ever been told by a doctor that you have, or had, any of the following conditions? Please tick all which apply and give the age at which each condition was first diagnosed.

Heart attack, coronary thrombosis, myocardial infarction	Yes ☐ 1	No ☐ 2	at age ☐☐	yrs old
Angina	Yes ☐ 1	No ☐ 2	at age ☐☐	yrs old
Stroke	Yes ☐ 1	No ☐ 2	at age ☐☐	yrs old
High Blood Pressure (Hypertension)	Yes ☐ 1	No ☐ 2	at age ☐☐	yrs old
High Blood Cholesterol, Hyperlipidaemia	Yes ☐ 1	No ☐ 2	at age ☐☐	yrs old
Diabetes	Yes ☐ 1	No ☐ 2	at age ☐☐	yrs old
Gallstones	Yes ☐ 1	No ☐ 2	at age ☐☐	yrs old
Polyps in the large intestine	Yes ☐ 1	No ☐ 2	at age ☐☐	yrs old
Cancer	Yes ☐ 1	No ☐ 2	at age ☐☐	yrs old

If yes, what type of cancer? _____

Any other illnesses or operations?
Do not include hysterectomy or breast surgery. These are covered later in the questionnaire.

Condition/ operation / disease	Age first diagnosed
	☐☐ yrs old
	☐☐ yrs old
	☐☐ yrs old
	☐☐ yrs old

49: Are you currently receiving long-term treatment for any illness or condition? Yes ☐ 1 No ☐ 2

If yes, please give details of treatment. If no please go to question 50:

Illness or condition	Treatment	Dose	Frequency

50: Have your mother and/or father ever suffered from cancer or heart attack/heart disease?

Yes ☐ 1 No ☐ 2 Don't Know ☐ 3

If yes, please give details

51: If you have brothers and/or sisters, have they ever suffered from cancer or heart attack/heart disease?

Yes ☐ 1 No ☐ 2 Don't Know ☐ 3

If yes, please describe details

EDUCATION:

52: How old were you when you finished your full time education? ☐☐ yrs old

53: Do you have any of the following qualifications? Tick all applicable

CSE	☐ 1	"A" Level, Highers	☐ 4
GCE "O" Level	☐ 2	Teaching diploma, HNC	☐ 5
City & Guilds	☐ 3	Degree	☐ 6
Other ☐ 7 describe		None of these	☐ 8

EMPLOYMENT:

54: Have you ever had a paid job? Yes ☐ 1 No ☐ 2

If yes, please answer for your current or most recent job

What is/was your job title?

What do/did you do in your job?

What does/did the organisation you work for make or do?

Are/were you a Manager? ☐ 1 Foreman/woman? ☐ 2

Supervisor? ☐ 3 None of these? ☐ 4

Are/were you self-employed? Yes ☐ 1 No ☐ 2

Do you have a paid job at present? Yes ☐ 1 No ☐ 2

If no, how would you describe yourself?

Housewife ☐ 1 Unemployed ☐ 2

Retired ☐ 3 Student ☐ 4

Other ☐ 5 describe

When did you last have paid employment 19☐☐ (year) or Never ☐ 1

55: What is your marital status?

Married or living as married ☐ 1 Divorced ☐ 3

Widowed ☐ 3 Single ☐ 4

Separated ☐ 5

If you are not married or living as married, please go to question 57.

PARTNER'S EMPLOYMENT:

56: If married or living as married, has your partner ever had a paid job? Yes ☐ 1 No ☐ 2

If yes, please answer for your partner's current or most recent job.

What is/was your partner's job title?

What does/did the organisation your partner works for make or do?

Is/was your partner a Manager? ☐ 1 Foreman/woman? ☐ 2

Supervisor? ☐ 3 None of these? ☐ 4

Is/was your partner self-employed? Yes ☐ 1 No ☐ 2

Does your partner have a paid job at present? Yes ☐ 1 No ☐ 2

If no, how would you describe your partner?

House-husband ☐ 1 Student ☐ 2

Unemployed ☐ 3 Retired ☐ 4

Other ☐ 5 describe

57: Which of these groups would you consider you belong to?

White ☐ 1 Bangladeshi ☐ 2

Indian ☐ 3 Chinese ☐ 4

Pakistani ☐ 5 Black - Caribbean ☐ 6

Black - other ☐ 7

Other ☐ 8

MENSTRUAL & OBSTETRIC HISTORY:

58: How old were you when you had your first menstrual period ☐☐ years old

59: What is the usual length of your menstrual cycle?

(i.e. from the first day of one period to the first day of the next period e.g. 26 days)? ☐☐ days.

60: Have you ever been pregnant? Yes ☐ 1 No ☐ 2

Are you pregnant at the moment? Yes ☐ 1 No ☐ 2

How many times have you been pregnant? ☐☐

Have you ever had a miscarriage/still birth? Yes ☐ 1 No ☐ 2

If you have had children, please go to question 61. If not please go to question 63.

61: Have you had any children? Yes ☐ 1 No ☐ 2

If yes, how old were you when your first child was born ☐☐ years

If yes, how many children have you had? ☐☐ children

If none please go to question 63.

Please can you write in each child's sex and approximate birthweight.

Child	Sex of Child	Approximate Birthweight	Child's D.O.B
CHILD 1:			19☐☐
CHILD 2:			19☐☐
CHILD 3:			19☐☐
CHILD 4:			19☐☐
CHILD 5:			19☐☐

62: Did you ever breast feed any of your children? Yes ☐ 1 No ☐ 2

If yes, for those children you breast-fed, please describe how long you continued breast feeding after each birth, (even only occasional breast feeding). Tick the appropriate box.

	1–6 days	1–4 weeks	1–3 months	4–6 months	6+ months	12+ months
CHILD 1:	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
CHILD 2:	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
CHILD 3:	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
CHILD 4:	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
CHILD 5:	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

63: Have you ever seen a doctor because of fertility problems? Yes ☐ 1 No ☐ 2

If yes, has a doctor ever told you that you were infertile? Yes ☐ 1 No ☐ 2

64: Have you ever used oral contraceptives (the pill)? Yes ☐ 1 No ☐ 2

If yes, how old were you when you first started to use the pill? ☐☐ years old

For how long altogether did you use the pill? ☐☐ years

Are you currently using the pill? Yes ☐ 1 No ☐ 2

If no, how old were you when you last used it? ☐☐ years old

65: Have you ever used a coil or intra-uterine device (IUD)? Yes ☐ 1 No ☐ 2

If yes, do you have a coil or IUD at present? Yes ☐ 1 No ☐ 2

66: How many "natural" menstrual periods have you had in the last 12 months?

Do not count bleeding while using the pill or HRT (Hormone Replacement Therapy)

None ☐ 1

1 to 3 ☐ 2

4 to 5 ☐ 3

6 to 9 ☐ 4

10 or more ☐ 5

Not applicable because using the Pill or HRT or currently pregnant ☐ 6

67: When did you last have a "natural" menstrual period? Do not count bleeding while using the pill or HRT (Hormone Replacement Therapy). Record as fully as possible

Date: ☐ or age ☐☐ years old Don't know ☐

68: Have you ever used HRT (Hormone Replacement Therapy) for menopause? Yes ☐ 1 No ☐ 2

If yes, how old were you when you first used HRT? ☐☐ years old

For how long altogether have you used HRT? years ☐☐ and months ☐☐

Are you currently using HRT? Yes ☐ 1 No ☐ 2

If no, how old were you when you last used HRT? ☐☐ years old

69: Have you had a hysterectomy ? If no please go to question 71.

Yes ☐ 1 Age at time of operation ☐☐ years old No ☐ 2 Don't know ☐ 3

70: Have you had an operation to remove one or both your ovaries?

Yes ☐ 1 No ☐ 2 Don't know ☐ 3

If yes, how old were you? ☐☐ years old

Were one or both ovaries removed? One ☐ 1 Both ☐ 2 Don't know ☐ 3

71: Have you ever had a breast biopsy (minor surgery to remove tissue from your breast for diagnostic purposes)? Yes ☐ 1 No ☐ 2 Don't know ☐ 3

If yes, how old were you (first occurrence)? ☐☐ years old

Appendix F: Distribution of information resources to extract AOAC-fibre values (g/100g) based on food groups in the FFQ

No	Food groups	FSA	UK review	EuroFIR databases	USDA database	Food labelling	Recipe calculation	Total number of food items
1	Bread-savoury biscuits	9	1	3	0	2	0	15
2	Breakfast cereals	11	5	2	2	4	0	24
3	Potatoes, rice and pasta	1	5	10	15	3	0	34
8	Grains	0	3	2	2	0	0	7
23	Biscuits , sweets and pudding	6	4	1	16	7	4	38
15	Vegetables	1	25	9	27	5	0	67
12	Vegetable dishes	1	2	6	8	7	5	29
16	Fruit	0	8	2	5	0	0	15
17	Seasonal fruit	0	6	6	4	0	0	16
18	Dried fruit	0	3	3	7	2	0	15
9	Nuts and seeds	0	6	2	2	0	0	10
10	Pulses	0	0	1	16	1	1	19
4	Dairy and non dairy products	2	0	1	37	7	2	49
13	Meat	3	0	4	39	25	4	75
14	Fish	0	0	1	23	8	3	35
11	Egg/egg dishes	0	0	1	5	1	4	11
5	Margarines/butters and spreads	0	0	0	10	0	0	10
6	Spread	0	0	2	4	6	1	13
19	Sweet snacks	0	0	3	5	7	0	15
20	Savoury snacks	0	0	1	8	1	1	11
7	Sauces and soups	0	0	0	7	13	0	20
21	Beverages	0	0	2	10	2	0	14
22	Alcohol	0	0	0	5	0	0	5
24	Other	0	0	0	1	0	0	1
	Total number of food items	36	68	62	258	101	25	545

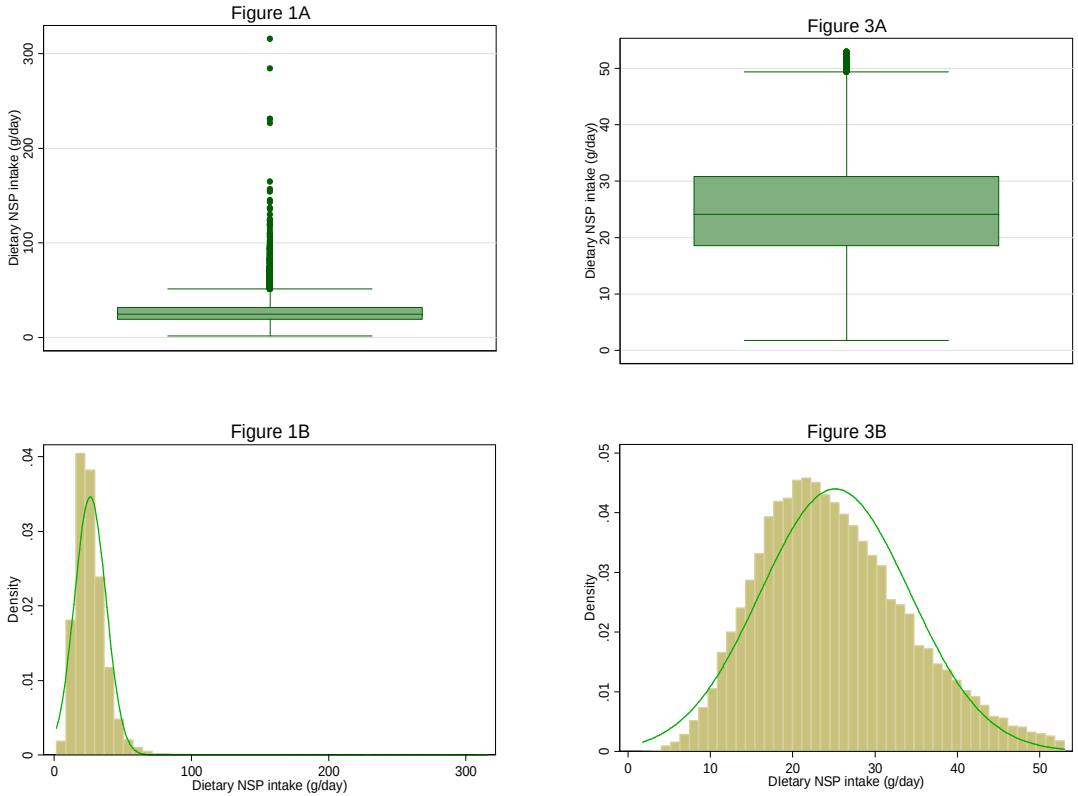
Appendix G: Baseline FFQ food items list in each food group

Food groups	FFQ food items numbers	Food items
Cereals high in fibre	3,5,8,13,14,15, 68,70	wholemeal bread average, papadums, crispbread, muesli, all bran, Weetabix, oat, wheat germ
Cereals source of fibre	2,4,6,9,10,11,2 2,69,173,204,2 06	brown bread average, chapattis, tortillas, cream crackers, porridge, sugar coated cereals , wholemeal pasta ,burger, cereal bar, chocolate biscuits, fruitcake, buns, scones, fruit pie
Other cereals	1,7,12,21,2 ,26 , 67, 71, 203,205, 207,208,211	white bread, pitta, white pasta, white rice, brown rice, macaroni cheese, couscous, non-sugar cereal, wild rice, plain biscuits, sandwich biscuits sponge cakes, sponge puddings
Potatoes	16-20	boiled potato, jacket potato, potato salad chips, roasted potato barley
Fruit	53, 149-172 174	apple, avocado, bananas, grapes, kiwi, mango, orange, papaya, pears, pineapple, apricot, melon, nectarines, peaches, plums, raspberries, currants, rhubarb, strawberries, dates, fig, prunes, mixed dried fruits, currants & sultanas
Vegetables	55, 90-96, 123-134, 136-148	Quron, Textured vegetable protein, veg chilli, mixed beans casserole, stir fry veg, veg dishes, veg pizza, beetroot, broccoli, Brussels, cabbage, carrots, cauliflower, celery, coleslaw, low cal coleslaw, courgettes, cucumber, garlic, lettuce, leeks mushrooms, olives, peppers, Swede, sweet corn, tomatoes, turnip, watercress
Legumes	78-85, 135,142	lentils, chickpeas, hummus, baked beans, red kidney beans, bean sprout, blacked eyed beans, butter beans, green beans, peas
Nuts	51, 56, 74-77,181-183	Peanut butter, nut pate, peanuts, cashews, pecans, sunflower seeds, Bombay mix, peanuts/pistachio, mix nuts
Milk and milk products	27-42 , 212-218	thick & creamy yogurt, Greek yogurt, fromage frais, dairy desserts, sour cream, soya yogurt, double cream, ice cream, milk pudding, cheese, cottage cheese, cheese and onion, soya cheese, milk whole, milk channel island, soy milk, sterilized milk
Meat and meat products	97-115	beef, beef stew, pork, pork dishes, lamb, lamb stew, pies, pastries, offal, pate, meat dishes, meat pizza beef burger, bacon, ham, canned meat, sausages chicken & turkey roast, breaded chicken, chicken & turkey sauce
Fish and fish dishes	116-122	Fish fingers, battered fish, white fish, oily fish, shellfish, fish Roe, fish pie.
Non-alcoholic beverages	184-197	tea, herb tea, coffee, decaf coffee, coffee substitute, coffee whitener, coca, Horlicks, low calorie hot chocolate, orange juice, other fruit juice, fruit squash, fizzy soft drinks, low diet soft drinks
Alcoholic beverages	198-202	wines, beer, cider, port, spirits
Eggs and eggs dishes	86-89	boiled egg, omelettes scrambled, fried egg, quiche
Fat spread	43-49	Butter, block margarine, polyunsaturated margarine, other soft margarine, low fat spreads, very low fat spread, monounsaturated margarine
Savoury snacks	178-180	crisps, fried snacks, low fat crisps
miscellaneous foods	57-66	Low calorie salad cream, mayonnaise, French dressing, white sauces, tomato ketchup, pickles, packet soup, vegetable soup, meat soup low calorie soup.
Sugar and confectionary	50, 52-54, 175, 176, 177,219	Marmite, chocolate spread, jam, honey, chocolate bars(Mars), mini chocolate bars, boiled sweets, table sugar

Appendix H: Degree of agreement express as Kappa (K) and weight Kappa (K_w) with percentage between NSP and AOAC-fibre intake in different fibre sources

No	Dietary fibre intake from food groups	Degree of agreement	
		K (%)	K_w(%)
1	Cereals and cereals products	0.8(83%)	0.9(95%)
1.a	Cereals high in fibre	0.9(90%)	0.9(97%)
1.b	Cereals source of fibre	0.7(76%)	0.8(94%)
1.c	Other cereal products	0.8(80%)	0.9(95%)
2	Vegetables	0.8(80%)	0.9(94%)
3	Fruit	0.9(90%)	0.9(97%)
4	Potatoes	0.5(61%)	0.8(90%)
5	Legumes	0.8(84%)	0.9(95%)
6	Nuts	0.9(96%)	0.9(99%)
7	Sugars, preserves and snacks confectionary	0.9(98%)	0.9(99%)
8	Sauces, soups and miscellaneous foods	0.5(58%)	0.7(86%)
9	Meat and meat products	0.8(90%)	0.9(97%)
10	Milk and milk products	0.7(76%)	0.8(94%)
12	Fish and fish dishes	0.8(81%)	0.9(95%)

Appendix I: Distributions of NSP intake by box plots and histograms with and without energy restriction



All figures shows dietary fibre intake (g/day). Fig.1A & 1B Box plot and histogram for all participants, Fig. 2A & 2B Box plot and histogram for participants with total energy intake restricted between 500 – 4000 kcal/day.

Appendix J: List of food items for TDF calculated from recipies

No	Food item name	TDF content g/100g
1	Pizza, tomato, whole meal	3.5
2	Fish pie	0.9
3	Fish cake, cod, homemade	1.0
4	Fish cakes, salmon, homemade	1.0
5	Moussaka	1.5
6	Beef stew	0.8
7	Pork and apple casserole	0.7
8	Quiche, cheese, egg, wholemeal	2.9
9	Quiche lorraine, wholemeal	3.2
10	Quiche cheese and egg	0.7
11	Quiche, cheese, egg, mushroom, wholemeal	2.9
12	Scone, wholemeal, sultana	7.2
13	Pork, stir-fried with vegetables	1.1
14	Apple pie, wholemeal, pastry top and bottom	5.0
15	Crumble, fruit, wholemeal	3.3
16	Apple pie, wholemeal, one crust	3.6
17	Quorn Korma	5.2
18	Lasagne, vegetable	3.5
19	Chilli, vegetable , retail	3.1
20	Dal Dhokari	3.8
21	Fruit cake, wholemeal	3
22	Nut and vegetable roast, with egg	4.8

UNIVERSITY OF LEEDS

UK Women's Cohort Study

FOOD AND ACTIVITY DIARY

We would be grateful if you could record all your food and drink for 4 consecutive days and your physical activity on day 3. You can find full instructions in the diary.

Please complete and return this diary at your convenience, but preferably within one month. We appreciate that completing this diary will take some time and we wish to take this opportunity to thank you for your contribution to our research.

If you have any queries, please contact a member of the team on 0113 343 7452 or email cohorteam@leeds.ac.uk

Thank you very much for your help.

INSTRUCTIONS ON HOW TO COMPLETE THIS DIARY

This diary is designed for you to record everything you eat and drink for four days, plus one day for physical activity.

To give us an accurate picture, please fill out the diary in as much detail as possible. It is very important that you do not change what you eat and drink, or the physical activities that you carry out just because you are keeping a record.

FOOD DIARY

- Please date your record when you start your diary.
- Please record the time you had something to eat or drink in the left-hand column marked **"Time of food/drink"**.
- In the column marked **"Description of food or drink consumed"**, please give a full record of the food/drink and how it was prepared (cooking method). If possible, please record each individual food and drink item separately **(see example on page 4)**.
- In the last column, please record the **amount of food or drink you consumed** by giving the weight if on the packet or carton e.g. 150g pot of yoghurt, 56g bar of milk chocolate. For other foods we would like you to weigh the foods you consume. If you do not have scales at home, or if you are eating food away form the home, then describe the food you eat using household measures e.g. tablespoons, cups, large glass etc. Please document what you ate & drank in as much detail as possible.

- To establish that your weighing scales are accurate, please weigh ***at least*** one of the following foods in (grams) and record your results in the boxes below :

1 raw egg (still in the shell) □□ g
1 large tin of baked beans (420g) □□□ g
1 full bag of flour □□□□ g

- At the end of the third page, there is space to write down all recipes and if more than one serving, how much was consumed by you. Also use this space to record details of any foods/drinks eaten away from home and to record the brand of any manufactured products.

- If you eat ready made foods that have the nutritional information on the packet, then please could you write down this information in the space provided on the fourth page. It is important to state if the information is for either per serving, or per 100 grams. If it is for a serving, then please write down the serving size that you had.

- At the end of the diary section are a few questions about your diet in general over the four days you have recorded. Please remember to complete these before returning your food diary.

PHYSICAL ACTIVITY DIARY

We would like to find about your physical activity for one full day. **Please fill in the physical activity diary on day three.**

Please refer to the specific instructions, which are situated just before the physical activity diary.

EXAMPLE DAY - UP TO LUNCH

Date: 14 October 2005	Day of the week Friday	
Time of food or drink	**Description of food or drink consumed (Include brand name where possible)**	**Amount**
7.15 am	Filter Coffee	1 cup (200m1)
	semi-skimmed milk	3 tablespoons
7.30 am	Sainsbury's orange juice, un-sweetened	1 glass (150ml)
	Sainsbury's Bran flakes	40g
	semi-skimmed milk	180ml
10.30am	Plain chocolate digestives (large biscuits)	2
	Earl Grey tea (weak) no milk	1 cup
11 am	banana (medium sized)	95g
11.30 am	London herb company Lemon Zester tea	1
12.10 pm	Local bakery's wholemeal bread un-sliced loaf (cut thickly)	1 slice 47g
	Tesco sunflower margarine	thinly spread
	home made mushroom risotto (see recipe)	About $^1/_3$ of recipe
	green seedless grapes	32g
	Cox's Orange Pippin apple (medium)	82g
	Sainsbury's wholemilk fruit yogurt (150g)	1 pot
	London herb company sweet berry tea	1
2pm	Warburton's Carrot cake - with cream cheese topping (see nutritional information)	1 slice - 75g (on packet)

EXAMPLE - CONTINUED

Recipes/description of foods eaten away from home/ any other comments

Home made mushroom risotto
2 Tablespoons of olive oil
3 oz Butter
1 Clove of garlic
1lb Button Mushrooms
10oz Rice (long grain rice)
1 large onion
1 ½ Pints of vegetable stock
Pinch of salt & pepper
Pinch of rosemary, thyme & chilli powder.

Fry onion & garlic in butter & olive oil until soft. Add chilli & mushrooms, cook gently for approx 5 mins. Add all of the rice, then add stock ladle by ladle until completely absorbed, this should take around 20 mins.

Usually approx 3 servings from this of equal proportion

EXAMPLE - CONTINUED

Nutritional Information for Ready Made / Packaged foods
Please state, if the information is for a serving or per 100g

Food	Serving size	per 100g	Energy kcal	Protein	Carbohydrate	Fat	Fibre
Carrot cake		✓	490	4.5	52.4	30.6	0.7

Any extras not already recorded, e.g. sweets/snacks/drinks?

2 polo mints during morning at work, & pack of Salt & Vinegar crisps (Walkers 28g) on way home

END OF EXAMPLE

Appendix L of Stata code

Chapter 4: Pooled estimate and forest plot of the meta-analysis

```
metan lnrror1 lnlci1 lnuci1, lcols(authornameyear country) boxsca (100) null(1) effect
(Estimated risk) xlabel (0.25, 0.5, 1.0, 1.25,1.50) force eform randomi nohet nowt texts(180)
/* Subgroup analysis for potential sources of heterogeneity*/
/* gender*/
metan lnrror1 lnlci1 lnuci1 if gender==1, lcols(authornameyear country) boxsca (100) null(1)
effect (Estimated risk) xlabel (0.25, 0.5, 1.0, 1.25,1.50) force eform randomi nohet nowt
texts(180)
metan lnrror1 lnlci1 lnuci1 if gender==2, lcols(authornameyear country) boxsca (100) null(1)
effect (Estimated risk) xlabel (0.25, 0.5, 1.0, 1.25,1.50) force eform randomi nohet nowt
texts(180)
/* Origin of study*/
metan lnrror1 lnlci1 lnuci1 if origin==1, lcols(authornameyear ) boxsca (100) null(1) effect
(Estimated risk) xlabel (0.25, 0.5, 1.0, 1.25,1.50) force eform randomi nohet nowt texts(180)
metan lnrror1 lnlci1 lnuci1 if origin==2, lcols(authornameyear ) boxsca (100) null(1) effect
(Estimated risk) xlabel (0.25, 0.5, 1.0, 1.25,1.50) force eform randomi nohet nowt texts(180)
/* Duration of follow-up*/
metan lnrror1 lnlci1 lnuci1 if period==0, lcols(authornameyear) boxsca (100) null(1) effect
(Estimated risk) xlabel (0.25, 0.5, 1.0, 1.25,1.50) force eform randomi nohet nowt texts(180)
metan lnrror1 lnlci1 lnuci1 if period==1, lcols(authornameyear ) boxsca (100) null(1) effect
(Estimated risk) xlabel (0.25, 0.5, 1.0, 1.25,1.50) force eform randomi nohet nowt texts(250)
```

Chapter 7: Kappa statistics [AOAC-fibre and NSP intakes]

```
/*check normality of DF variables by graphs*/
histogram aoacout if kcal>=500 & kcal<=4000, normal
histogram engfibout if kcal>=500 & kcal<=4000, normal
graph box aoacout if kcal>=500 & kcal<=4000
sum engfibout aoacout if kcal>=500 & kcal<=4000 , detail
xtile aoacoutq = aoacout , n(5)
xtile engfiboutq = engfibout , n(5)
gen NSP1000kcal = (engfibout*1000)/kcal
gen DF1000kcal = (aoacout*1000)/kcal
xtile DFg1000kcalq = DFg1000kcal , n(5)
xtile NSPg1000kcalq = NSPg1000kcal , n(5)
by aoacoutq, sort : summarize aoacout if kcal>=500 & kcal<=4000, normal
by DFg1000kcalq , sort : summarize DFg1000kcal if kcal>=500 & kcal <=4000
by engfiboutq, sort : summarize engfibout if kcal>=500 & kcal<=4000, normal
by NSPg1000kcalq , sort : summarize NSPg1000kcal if kcal>=500 & kcal <=4000
summarize NSPg1000kcal if kcal>=500 & kcal<=4000, detail
summarize DFg1000kcal if kcal>=500 & kcal<=4000, detail
tabulate aoacoutq engfiboutq if kcal>=500 & kcal<=5000, row col
/* assess degree of agreement*/
kap engfiboutq aoacoutq if kcal>=500 & kcal<=4000, tab
kap engfiboutq aoacoutq if kcal>=500 & kcal<=4000, tab w(w)
```

kap DFg1000kcalq NSPg1000kcalq if kcal>=500 & kcal<=4000, tab w(w)
tabulate DFg1000kcalq NSPg1000kcalq if kcal>=500 & kcal<=4000, row
regress aoacout engfibout if kcal>=500 & kcal<=4000/*regression equation*/
predict resids, residuals/* the residuals are the diff between the observed value and fitted value or the estimated expected value*/
histogram resids, normal/* distribution to meet condition*/
predict fitted, xb
scatter resids fitted
scatter aoacout engfibout if kcal>=500 & kcal<=4000
/*high NSP and AOAC-fibre consumers*/
/*testing normality of baseline variables [age as example] */
by engfiboutq, sort : summarize age if kcal>=500 & kcal<=4000
hist age if kcal>=500 & kcal<=4000, by (engfiboutq) normal
oneway age engfiboutq if kcal>=500 & kcal<=4000, tabulate
by engfiboutq, sort : summarize NSP1000kcal if kcal>=500 & kcal<=4000
hist NSP1000kcal if kcal>=500 & kcal<=4000, by (engfiboutq) normal
/* chi square test used for categorical variables of interest for descriptive tables */
tab newclass engfiboutq if kcal>=500 & kcal<=4000, row col chi
testparm i.newclass /* overall p value is recorded in the table*/
regress engfibout age bmi vitamin smoke vitc alkcal iron i.newclass pamins i.foodgrp3 if kcal>=500 & kcal<=4000
predict r3, rstudent
hist r3, normal
predict f3, xb
rvfplot
regress aoacout age bmi vitamin smoke vitc alkcal iron i.newclass pamins i.foodgrp3 if kcal>=500 & kcal<=4000
predict r4, rstudent
hist r4, normal
predict f4, xb
rvfplot

Chapter 8: Stata Code for logistic regression analyses [AOAC-fibre intake as example]

/*For continuous OR*/
/*unadjusted OR model*/
xi: logistic DM aoacout if kcal>=500 & kcal<=5000, or
/*model 1 age adjustment */
xi: logistic DM aoacout age if kcal>=500 & kcal<=5000, or
/*model 2 adjust age(con) weight change(con) kcal(con) smoke(2cat)*/
xi: logistic DM aoacout age wtchangep smoke if kcal>=500 & kcal<=5000
xtile ethanol3 = ethanol2007, n(3)
/*model 3 adjust age(con), kcal(cont), SES(3cat) weight change(con) alcohol intake(3 groups) smoke(2cat) METs(con)*/

xi: logistic DM aoacout age i.newclass wtchangep i.ethanol3 smoke total_mets if kcal>=500 & kcal<=5000

/* Model 4 add family history of diabetes*/

xi: logistic DM aoacout age MG i.newclass wtchangep ethanol2 kcal smoke total_mets famhisdiab if kcal>=500 & kcal<=5000, or

/* categorical OR*/

xtile aoacoutq = aoacout, n(5)

/* age adjusted model*/

xi: logistic DM i.aoacoutq age if kcal>=500 & kcal<=5000, or

test _Iaoacoutq_2 _Iaoacoutq_3 _Iaoacoutq_4 _Iaoacoutq_5

/*model 2 adjust age(con) weight change(con) kcal(con) smoke(2cat)*/

xi: logistic DM i.aoacoutq age kcal smoke wtchangep if kcal>=500 & kcal<=5000, or

test _Iaoacoutq_2 _Iaoacoutq_3 _Iaoacoutq_4 _Iaoacoutq_5

/*model 3 adjust age(con), kcal(cont), SES(3cat) weight change(con) alcohol intake(3 groups) smoke(2cat) METs(con)*/

xi: logistic DM aoacoutq age i.newclass wtchangep i.ethanol3 smoke total_mets if kcal>=500 & kcal<=5000

test _Iaoacoutq_2 _Iaoacoutq_3 _Iaoacoutq_4 _Iaoacoutq_5

/* Model 4 add family history of diabetes*/

xi: logistic DM aoacoutq age MG i.newclass wtchangep ethanol2 kcal smoke total_mets famhisdiab if kcal>=500 & kcal<=5000, or

test _Iaoacoutq_2 _Iaoacoutq_3 _Iaoacoutq_4 _Iaoacoutq_5

/*subgroup analysis*/

gen BMI2 = .

replace BMI2=0 if bmi<25

replace BMI2=1 if bmi>=25

label define BMI2 0 "Normal BMI " 1 "Overweight and obese"

label values BMI2 BMI2

tab DM BMI2

sort BMI2

by BMI2: logit DM aoacout age if kcal>=500 & kcal<=5000, or

sort BMI2

by BMI2: logit DM i.aoacoutq age if kcal>=500 & kcal<=5000, or

sort BMI2

by BMI2: logit DM engfibout age if kcal>=500 & kcal<=5000, or

sort BMI2

by BMI2: logit DM i.engfiboutq age if kcal>=500 & kcal<=5000, or

Chapter 9: Stata code for logistic regression analysis [legumes intake and risk of T2DM]

/* legumes intake UKWCS dataset */

gen DL = gpd78+ gpd79+ gpd80+ gpd81+ gpd82+ gpd83+ gpd84+ gpd85

label variable DL "dried legumes g per day"

gen FL = gpd135+ gpd142

label variable DL "fresh legumes g per day"

gen legumes = DL+FL
label variable legumes " total legumes intake g/day"
sum legumes DL FL, if kcal>=500 & kcal<=5000
tabstat legumes DL FL, statistics(mean sd iqr p50 max min p25 p75 p99), if kcal>=500 & kcal<=5000
/*unadjusted model dose response relationship*/
logistic DM DL if kcal>=500 & kcal<=5000
logistic DM FL if kcal>=500 & kcal<=5000
logistic DM legumes if kcal>=500 & kcal<=5000
/*age adjusted model */
logistic DM DL age if kcal>=500 & kcal<=5000
logistic DM FL age if kcal>=500 & kcal<=5000
logistic DM legumes age if kcal>=500 & kcal<=5000
/*fully adjusted model */
xi: logistic DM DL age kcal i.smoke wtchangep i.newclass i.ethanol3 total_mets if kcal>=500 & kcal<=5000, or
xi: logistic DM FL age kcal i.smoke wtchangep i.newclass i.ethanol3 total_mets if kcal>=500 & kcal<=5000, or
xi: logistic DM legumes age kcal i.smoke wtchangep i.newclass i.ethanol3 total_mets if kcal>=500 & kcal<=5000, or
/* increment per half portion daily with risk of diabetes */
gen DL40 = DL/40
label variable DL40 "dried legumes 40g intake per day equivalent to 1/2 recommanded portion"
gen FL40 = FL/40
label variable FL40 "fresh legumes 40g per day 30g equivalent to 1/2 recommanded portion"
gen legumes40 = legumes/40
label variable FL40 "total legumes 40g per day 30g equivalent to 1/2 recommanded portion "
/*unadjusted model*/
logistic DM DL40 if kcal>=500 & kcal<=5000
logistic DM FL40 if kcal>=500 & kcal<=5000
logistic DM legumes40 if kcal>=500 & kcal<=5000
/*age adjusted model*/
logistic DM DL40 age if kcal>=500 & kcal<=5000
logistic DM FL40 age if kcal>=500 & kcal<=5000
logistic DM legumes40 age if kcal>=500 & kcal<=5000
/*fully adjusted model */
xi: logistic DM DL40 age kcal i.smoke wtchangep i.newclass i.ethanol3 total_mets if kcal>=500 & kcal<=5000, or
xi: logistic DM FL40 age kcal i.smoke wtchangep i.newclass i.ethanol3 total_mets if kcal>=500 & kcal<=5000, or
xi: logistic DM age kcal i.smoke wtchangep i.newclass i.ethanol3 total_mets if kcal>=500 & kcal<=5000, or
/* legumes intake is categorized into quartiles to examine the trend of increasing intake with risk of diabetes*/

```
xtile DLq = DL , n(3)
xtile FLq = FL , n(3)
xtile legumesq = legumes , n(3)
/*unadjusted model */
xi:logistic DM i.DLq if kcal>=500 & kcal<=5000
test _IDLq_2 _IDLq_3
xi:logistic DM i.FLq if kcal>=500 & kcal<=5000
test _IFLq_2 _IFLq_3
xi:logistic DM i.legumesq if kcal>=500 & kcal<=5000
test _Ilegumesq_2 _Ilegumesq_3
/* age model 1*/
xi:logistic DM i.DLq age if kcal>=500 & kcal<=5000
test _IDLq_2 _IDLq_3
xi:logistic DM i.FLq age if kcal>=500 & kcal<=5000
test _IFLq_2 _IFLq_3
xi:logistic DM i.legumesq age if kcal>=500 & kcal<=5000
test _Ilegumesq_2 _Ilegumesq_3
/*fully adjusted model*/
xi: logistic  DM i.DLq age kcal i.smoke wtchangep i.newclass i.ethanol3 total_mets if
kcal>=500 & kcal<=5000, or
test _IDLq_2 _IDLq_3
xi: logistic  DM i.FLq age kcal i.smoke wtchangep i.newclass i.ethanol3 total_mets if
kcal>=500 & kcal<=5000, or
test _IFLq_2 _IFLq_3
xi: logistic  DM i.legumesq age kcal i.smoke wtchangep i.newclass i.ethanol3 total_mets if
kcal>=500 & kcal<=5000, or
test _Ilegumesq_2 _Ilegumesq_3
```

Chapter 10 Comparison between baseline FFQ-fibre, repeated FFQ-fibre and diary-fibre

```
xtile engfiboutq = engfibout, n(5)
xtile rfibreq = rfibre, n(5)
xtile nsp4ddq = nsp4dd, n(5)
kap engfiboutq rfibreq  if rfibre!=.
kap engfiboutq rfibreq if rfibre!=., w(w)
kap engfiboutq nsp4ddq if nsp4dd!=. & MaxOfDay>=3
kap engfiboutq nsp4ddq if nsp4dd!=. & MaxOfDay>=3, w(w)
kap rfibreq nsp4ddq if rfibre!=.
kap rfibreq nsp4ddq if rfibre!=., w(w)
tabulate engfiboutq rfibreq , chi row col
tabulate engfiboutq nsp4ddq if MaxOfDay>=3 , chi row col
tabulate rfibreq nsp4ddq if rfibre!=., chi row col
/* fibre density intake*/
xtile dd4q = dd4, n(5)
```

```
xtile ffqq = ffq, n(5)
xtile rffqq = rffq, n(5)
kap ffqq rffqq  if rfibre!=.
kap ffqq rffqq if rfibre!=., w(w)
kap ffqq dd4q if nsp4dd!=. & MaxOfDay>=3
kap ffqq dd4q if nsp4dd!=. & MaxOfDay>=3, w(w)
kap rffqq dd4q if rfibre!=.
kap rffqq dd4q if rfibre!=., w(w)
tabulate ffqq rffqq , chi row col
tabulate ffqq dd4q if MaxOfDay>=3 , chi row col
tabulate rffqq dd4q if rfibre!=., chi row col
/* normal distributed  Pearson correlation used and if not normally distributed then spearman
correlation used */
pwcorr nsp4dddensity nsp4dd _Englyst IDF4dd SDF4dd engfibout NSPdensity insolfibout
solfibout if MaxOfDay>=3, sig
/*not normal distributed*/
spearman Cereal4dd Fruit4dd Vegetables4dd Nut4dd cnspffqbl fnsoffqbl vnspffqbl nnspffqbl if
MaxOfDay>=3
ci2 Cereal4dd cnspffqbl if MaxOfDay>=3, spearman
ci2 Fruit4dd fnsoffqbl if MaxOfDay>=3, spearman
ci2 Vegetables4dd vnspffqbl if MaxOfDay>=3, spearman
ci2 Nut4dd nnspffqbl if MaxOfDay>=3, spearman
/* Bland-Altman method for continuous variables between 2 methods [diary-fibre and
baseline FFQ-fibre as example]*/
baplot nsp4dd engfibout if nsp4dd!=. & MaxOfDay>=3
baplot nsp4dddensity NSPdensity if nsp4dd!=. & MaxOfDay>=3
/* Bland-Altman method for main fibre sources [ cereal fibre as example]*/
baplot cnspffqbl Cereal4dd  if nsp4dd!=. & MaxOfDay>=3
/* Bland-Altman method for main fibre sources [fibre density]*/
gen cffqk = cnspffqbl*1000/kcal
baplot cffqk c4ddk  if nsp4dd!=. & MaxOfDay>=3
/*Kappa Statistic [diary-fibre and baseline FFQ-fibre as example]*/
xtile nsp4ddq = nsp4dd, n(5)
xtile engfiboutq = engfibout, n(5)
kap engfiboutq nsp4ddq if nsp4dd!=. & MaxOfDay>=3
kap engfiboutq nsp4ddq if nsp4dd!=. & MaxOfDay>=3 , w(w)
tabulate engfiboutq nsp4ddq if nsp4dd!=. & MaxOfDay>=3 , chi row col
```

WARRINGTON LREC

141294

NCH
NORTH CHESHIRE HEALTH
Lister Road, Astmoor West Estate
Runcorn, Cheshire WA7 1TW
Tel: 01928 593000 Fax: 01928 569532
Tel: 0928 592200 Fax: 0928 569532

Our Ref: W94/33 (Filename JCade)
Please quote Ref No on all correspondence

Direct Line 0928 593050

6 December, 1994

Dr Janet Cade
Principal Investigator
Nuffield Institute for Health
71-75 Clarendon Road
Leeds
LS2 9PL

Dear Dr Cade

The UK Women's Nutrition and Lifestyle Survey - Reg Ref: W94/33

Thank you for your application to the local Research Ethical Committee to carry out the above study. I am pleased to tell you that at the meeting on 1st December, 1994 ethical approval was given for the study to proceed.

I would point out the requirement to refer proposed amendments to the protocol to the LREC for further review and to obtain LREC approval prior to implementation (except only in cases of emergency were the welfare of the subject is paramount). If applicable the LREC should be informed should the research be discontinued or any subject withdrawn.

Investigators should furnish the LREC with details of the progress of the research periodically (eg 6 monthly).

Yours sincerely

Dr D Pearson
Chairman Warrington LREC

WORLD AIDS DAY
1 DECEMBER
An initiative by North Cheshire Health in support of WADD94

Incorporating
North Cheshire Health Authority and the North Cheshire Division of Cheshire Family Health Services Authority

North Worcestershire Health Authority

LOCAL RESEARCH ETHICS COMMITTEE

Chairman:
Michael A Spittle

The Croft
Sutton Park Road
Kidderminster
Worcs., DY11 6LJ
Telephone: (01562) 824711
Fax: (01562) 822695

KG

180995

14th September 1995

Amanda Woodhouse
Division of Public Health
The Nuffield Institute for Health
71-75 Clarendon Road
Leeds, LS2 9PL

Dear Ms Woodhouse

Protocol for Research:
LREC: 3/95 (February 1995) - The UK Women's Nutrition and Lifestyle Survey

Following the meeting of the Local Research Ethics Committee on the 12th September 1995, we write to confirm that, with the revised questionnaire now received, the Committee has given approval to the above project. Concern was, however, expressed regarding the length of the questionnaire and the time it would take for completion.

In keeping with the Committee's protocol, would you please inform us of the results of the study when it is completed.

Yours sincerely

Kath Garrad
Secretary to the Local Research Ethics Committee