UNIVERSITY OF SOUTHAMPTON

Sociodemographic and lifestyle predictors of young women's diets

Sarah Rachael Crozier

Submitted for PhD

Medical Research Council
Environmental Epidemiology Unit

May 2004

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES

MEDICAL RESEARCH COUNCIL ENVIRONMENTAL EPIDEMIOLOGY UNIT

Doctor of Philosophy

SOCIODEMOGRAPHIC AND LIFESTYLE PREDICTORS OF YOUNG WOMEN'S DIETS

by Sarah Rachael Crozier

The diet of a woman both before and during pregnancy is an important determinant of her own health, and of the growth and development of her child. Impaired fetal growth is known to be linked to higher rates of adult disease in later life. Eating patterns vary with sociodemographic and lifestyle information, and thus women likely to have inadequate diets need to be identified so that public health advice can be targeted most appropriately.

These issues were explored in two datasets: the Princess Anne Hospital (PAH) study of 620 pregnant women and the Southampton Women's Survey (SWS) of 6,129 non-pregnant women aged 20-34. Sociodemographic and lifestyle measures were defined in detail. In particular, occupation-based measurement of women's social position is often problematic because relatively few women work full-time, and therefore methods using her own occupation, her partner's occupation, or a combination were explored.

Diet was assessed using a food frequency questionnaire (FFQ), and principal component analysis (PCA) was employed to generate overall dietary scores. The first two scores were stable across different methods of food grouping; the first score described a 'prudent' diet, and the second a 'high-energy' diet. Principal component analysis of four-day food diaries kept by the PAH women in early pregnancy resulted in prudent and high-energy scores that were positively correlated with those found using the FFQ. Weighting responses to the FFQ using fruit and vegetable cross-check questions from the SWS data made negligible difference to the scores. Longitudinal elements of the two studies revealed good stability of the prudent diet score over time, whilst that of the high-energy diet score was considered moderate.

A higher prudent diet score was independently associated with older age, higher qualifications, living with fewer children and not smoking. In the PAH study there were additional effects of a higher Cambridge score, and a lower Townsend index in late pregnancy, whereas SWS women from ethnic minorities, who took strenuous exercise, watched less television or were currently dieting were more likely to eat a prudent diet.

A higher-energy diet score was independently associated with a lower Cambridge score living with more children, not dieting and a lower BMI in both early and late pregnancy in the PAH study. In early pregnancy there were additional effects of younger age, smoking and eating more food since becoming pregnant, whereas in the SWS only living with more children was associated with a higher-energy diet score.

Analyses in this thesis have demonstrated that PCA is a valuable means of assessing diet. The results provide information about appropriate target groups for public health initiatives aiming to improve the diets of young women, and hence to influence the growth and development of their children. Further research is required to examine other characteristics, such as knowledge and attitudes, through which the associations observed are mediated.

Contents

List of Tables

List of Figures

Acknowledgements

I am very grateful to my supervisors who have always been available to answer questions and offer guidance. Hazel Inskip has been consistently supportive and helpful, and I am especially thankful for her willingness to be involved in all aspects of my thesis, from data cleaning through to commenting on analyses. I have very much appreciated Siân Robinson's thoughtful input on nutritional matters, and contribution of astute ideas. I would also like to thank Ann Berrington for her enthusiasm and encouragement, particularly as I explored the social statistical issues in the data.

I am indebted to both the Princess Anne Hospital study and Southampton Women's Survey researchers who collected and processed the data analysed in this thesis. In particular I would like to thank the teams of nurses who worked diligently to interview the many thousands of participants. Vanessa Cox and Hugh Darrah have been the careful custodians of the vast quantity of information collected and have cheerfully provided me with reliable datasets to analyse. Sincere thanks also go to the women of Southampton who have given their time to participate in the two studies.

On a more personal note, I am very grateful to Holly Syddall, Isabel Reading, Samantha Leary, Brian Yuen and Julia Saperia for their support and contribution towards my enjoyment of life at the MRC. I would also like to thank Louise O'Sullivan and Sharon Dirckx for their very valuable, enduring friendships.

My final thanks are to Mum and Dad whose unconditional love has made me who I am, and to my husband, Philip Crozier, whose love and friendship are the greatest gifts I have.

Author's contribution

The data collection for the PAH study was performed by S. Crofts, V. Davill, J. Hammond, L. Greenaway, S. Mitcham and S. White, whilst that for the SWS was completed by a large team of research nurses. Data checking and cleaning of the PAH study was performed by K. Godfrey, S. Robinson and V. Cox. The author and S. Borland checked and cleaned the SWS dietary data, whilst the rest of the SWS data were checked and cleaned by the author in conjunction with H. Inskip. The author performed all the analyses presented in this thesis, and the work was done mainly whilst in registered postgraduate candidature.

Abbreviations

BMI	Body Mass Index
BMR	Basal Metabolic Rate
CI	Confidence Interval
DETR	Department of the Environment, Transport and the Regions
DNSBA	Dietary and Nutritional Survey of British Adults
DRV	Dietary Reference Value
EAR	Estimated Average Requirement
ED	Enumeration District
FFQ	Food Frequency Questionnaire
GHS	General Household Survey
GP	General Practitioner
IHD	Ischaemic Heart Disease
IPF	Iterated Principal Factors
JSA	Jobseeker's Allowance
ML	Maximum Likelihood
MRC	Medical Research Council
NS-SEC	National Statistics Socio-Economic Classification
NTF	Nutrition Task Force
ONS	Office for National Statistics
OUG	Occupational Unit Group
PAH	Princess Anne Hospital
PC	Principal Component
PCA	Principal Component Analysis
PCF	Principal Component Factors
PFA	Principal Factor Analysis
RDA	Recommended Daily Amount
r	Pearson's correlation coefficient
r_S	Spearman's correlation coefficient
SC	Social Class based on occupation
SEG	Socio-Economic Group
SOC	Standard Occupational Coding
SWS	Southampton Women's Survey
USDA	US Department of Agriculture

1 Introduction

The diet of a woman both before and during pregnancy is an important determinant of her own health, and of the growth and development of her child. Impaired fetal growth is known to be linked to higher rates of adult disease in later life. Section 1.1 describes why young women's diets are considered critical.

Since nutrition is recognised as an important public health issue, guidelines have been drawn up to advise on healthy eating. Specific advice is available for pregnancy, and current guidelines are summarised in Section 1.2.

The importance of young women's diets suggests that appropriate methods of characterising them must be found. Section 1.3 describes three methods of dietary pattern analysis that have been used previously: the related techniques of principal component analysis and factor analysis, cluster analysis and diet quality indices.

Diet is known to vary with sociodemographic and lifestyle factors, and thus it may be possible to identify groups of women who are likely to have inadequate diets, to whom public health advice could be targeted most appropriately. Section 1.4 summarises the relationships observed in the literature.

1.1 Why are young women's diets important?

1.1.1 Women's diets influence their own health

Diet is an important determinant of health. James Lind conducted the first trial of dietary intervention amongst sailors with scurvy in 1747[1]. Those sailors who ate daily citrus fruit were rapidly cured, and able to return to sea within 6 days. It is now known that the sailors were lacking dietary vitamin C during long voyages. The first half of the 20th century, when all of the vitamins known today were discovered, is often referred to as the Golden Age of Nutrition[2]. Indeed, in the 1940s Oxford University decided that the establishment of a nutrition department was unnecessary because it was thought that there were no remaining unsolved problems in human nutrition[3].

However, although deficiency disease such as that discovered by Lind is relatively rare in the UK population today, the effect of nutrition on health is still apparent. For example, in observational studies increased meat or red meat consumption has been seen to be associated with a greater risk of colorectal cancer[4], and increased fruit and vegetable consumption with decreased cardiovascular disease, particularly stroke[5]. In a systematic review of randomised controlled trials of reduction in dietary fat intake, a decrease in cardiovascular risk was seen[6]. Dietary guidelines have been formulated to recommend increased fruit and vegetable consumption, decreased red meat and fat intake, and other healthy dietary behaviours as described in Section 1.2.

Thus young women's diets are likely to be important in determining their own health. Furthermore there is some evidence that young women's dietary habits persist into pregnancy[7,8], and diet during pregnancy is known to be an influence on pregnancy outcome (Section 1.1.3). Indeed pre-pregnant diet may itself be an independent influence on pregnancy outcome through the effects on nutrient status and body composition prior to conception, although limited research exists in this area (Section 1.1.3).

1.1.2 Diet may be a mediator of inequalities in health

Differences in mortality across socioeconomic group have been recognised in the UK from as early as 1845[9]. Marmot et al.'s[10] important study classified 17,530 Whitehall civil servants according to employment grade between 1967 and 1969, and found a steep inverse relation between grade and 10-year mortality. Socioeconomic differentials in morbidity are also evident; for example, Mackenbach et al.[11] found that across 11 western European countries, self-reported morbidity was higher in lower socioeconomic groups, and patterns were generally similar across countries.

Researchers are keen to discover reasons for such inequalities in health. James et al.[12] note that diet was not originally considered a likely explanation since most of the classic nutritional deficiencies are rare in Britain. However, modern nutritional research has presented a new perspective: diet is known to affect health further than through the simple elimination of deficiency disease. James et al. propose possible pathways throughout the lifecycle along which diet might cause social class differences in health, and conclude that 'diet affects the health of socially disadvantaged people from the cradle to the grave. The social and economic reasons are complex, but the potential for health gain through improved diet is enormous'.

Dubois and Girard[13] investigated nutrient gradients across four different social position indicators in Canada and the USA between 1988 and 1994. They note that the associations indicate 'a possible intermediate role for nutrition in the distribution of social health inequalities in North America'. Davey Smith and Brunner[9] observed that fat intakes do not differ greatly across social classes, but that large disparities exist in micronutrient density. They hence speculate that micronutrient and antioxidant intakes are the most likely nutritional influences on health inequalities. Bolton-Smith et al.'s[14] results from 10,000 Scottish men and women concur with this theory.

In 1998 the Independent Inquiry into Inequalities in Health[15] recognised that members of lower socioeconomic groups tend to eat less fruit and vegetables, and less dietary fibre, and these factors may be a cause of the higher mortality and greater morbidity seen amongst the more deprived. Recommendations to reduce health inequalities include policy changes to increase the availability of an adequate and affordable diet, especially amongst the disadvantaged. The Independent Inquiry also noted the importance of a mother's nutrition, both before and during pregnancy, in terms of effects on the health of the next generation. Policies that improve the nutrition of women of childbearing age are specifically recommended. Indeed, one of the two main recommendations of the report was that 'a high priority is given to policies aimed at improving health and reducing health inequalities in women of childbearing age, expectant mothers and young children'.

1.1.3 Women's diets affect the early growth and development of their children

Since birthweight is an indicator of fetal undernutrition, it is thought that maternal diet may be an important determinant of pregnancy outcome. de Onis et al.[16] provide a useful systematic review of randomised controlled trials of nutritional interventions to prevent intrauterine growth retardation. Twelve interventions were considered, and one was found to be borderline statistically insignificant in reducing the incidence of small-for-gestational age birthweights: a balanced protein-energy supplement, in which protein accounted for < 25% of the total energy content of the supplement. The treatment effect odds ratio was 0.77 with 95% confidence interval (CI) from 0.58 to 1.01. Other interventions indicating some beneficial effect that the authors felt merited further research were zinc, folate and magnesium.

Although randomised controlled trials are considered the most rigorous method available for assessing the benefit of health care interventions, observational research can also provide useful information, especially for hypothesis generation. For example, a study of 538 mothers in Southampton assessed diet in early and late pregnancy[17]. High carbohydrate intake in early pregnancy and low animal protein intake in late pregnancy were both associated with lower placental and birthweights. In a further analysis of the same cohort[18] mothers who had a high intake of carbohydrate in early pregnancy and a low intake of dairy protein in late pregnancy were more likely to have infants who were thin at birth. Ramakrishnan et al.[19] have reviewed the role of micronutrients in observational as well as randomised controlled trials. They concluded that there is strong evidence that zinc, calcium and magnesium supplementation could improve birthweight, prematurity and hypertension, particularly in high-risk groups.

Assessment of longer-term outcomes after pregnancy interventions has also been possible. Belizán et al.[20] showed a dietary influence on childhood blood pressure at a mean age of 7 years in a randomised controlled trial. The 298 women in the experimental group receiving a supplement of 2g of calcium per day during the second half of pregnancy had offspring with systolic blood pressure on average 1.4 mmHg lower than those in the placebo group. This effect was predominantly seen amongst children with body mass index greater than the median.

Retrospective observational research into outcomes of maternal nutrition in adults has been feasible amongst those whose mothers' diets were recorded during pregnancy. Between 1948 and 1954 women attending Aberdeen Maternity Hospital took part in a survey of diet over one week during the seventh month of pregnancy[21]. It was possible to trace and study 253 of the offspring, now aged around 40 years. Amongst those mothers whose animal protein intake was less than 50g per day, higher carbohydrate intake was associated with increased systolic blood pressure in the offspring. Conversely, amongst mothers whose animal protein intake was greater than 50g per day, carbohydrate intake was negatively related to systolic blood pressure.

The findings in Aberdeen were supported by a further study in Motherwell[22]. Here pregnant women attending the maternity hospital between 1952 and 1976 were advised to eat a high-meat, low-carbohydrate diet, in an attempt to prevent preeclampsia. Six hundred and twenty-six of their offspring were followed up at age 27 to 30. Those whose mothers reported high consumption of meat and fish in the second half of pregnancy had higher systolic blood pressures.

Evidence from births at times of famine in Europe during the Second World War indicates that severe food shortages can reduce birthweight. The German siege of Leningrad lasted from August 1941 to January 1943, and daily rations of bread dropped as low as 300g for manual workers, and 150g for non-manual workers[23]. Such acute famine conditions led to a decrease in mean birthweight amongst babies carried to term of 629g for boys, and 542g for girls.

The Dutch 'hunger winter' lasted between December 1944 and May 1945, and during this period average daily calorie intake fell below 1,000[24]. Obstetric data indicate that the effects of food shortages may have been most acute in the second half of pregnancy; babies exposed to the famine during the first half of gestation, but not during the second, had normal birthweights. However, those exposed during the second half of gestation had notably lower birthweights. Exposure during the entirety of pregnancy did not appear to have had a more detrimental affect than exposure during the second half of pregnancy only.

It is thought that pregnancy outcome may be affected by diet before pregnancy, through effects on body composition and nutrient storage. There is evidence that the impact of diet on birthweight may differ according to the nutritional status of the mother. In a study of dietary supplementation to increase energy intake during pregnancy in Gambian women[25], babies born to supplemented mothers were found to be 101g heavier than those born to unsupplemented mothers. However, a more detailed analysis revealed that the supplement had no effect in the dry season, but increased birthweight by 186g in the wet season when the women are usually in marked negative energy balance because food is scarce and the agricultural workload is high.

Related evidence has been found in a study of Asian mothers at Sorrento Maternity Hospital in Birmingham, England[26]. Women were defined as nutritionally at risk if their increase in triceps skinfold thickness during the second trimester was considered inadequate ($< 20\mu m$ per week). Babies born to at risk mothers who were randomised to a protein energy supplement during the third trimester had an increased birthweight. There was no effect of supplementation amongst mothers not defined to be at risk.

More recently, the relationship between maternal nutrition and birth size was examined in an observational study in Pune, India[27]. Birth size was strongly associated with green leafy vegetable and fruit intake at 28 weeks of pregnancy. These relationships were stronger amongst those mothers with a pre-pregnancy weight of less than 40kg. The authors suggest that micronutrients found in these foods may be the most important limiting nutrients in undernourished women.

More conclusive research into the impact of pre-pregnancy diet on birth outcomes has been used to inform dietary guidelines concerning folate and vitamin A consumption (Section 1.2.2); folic acid supplementation before pregnancy is now routinely recommended to prevent neural tube defects[28,29], and high intakes of vitamin A around the time of conception have been related to congenital abnormalities[28].

Jackson and Robinson[30] have recently reviewed current evidence for dietary guidelines in pregnancy. They note several limitations of research to date, which generally uses birthweight as a marker of the relative success of pregnancy. Birthweight is a crude indicator of fetal growth and development, and furthermore birth is only part way through the 'fetal-infant' stage of development. Additionally, the majority of studies have been conducted in mid to late pregnancy, when it may be too late for nutrient effects to be seen, since the fetal growth trajectory is established early in gestation. A further argument presented is that consideration of single nutrients in isolation does not take account of the confounding of other nutrients, or interactions between nutrients. Jackson and Robinson conclude that 'no clear relationship between maternal nutrition and pregnancy outcome emerges, and there are very few nutrients for which definitive guidelines can be established'. The authors therefore consider that 'dietary guidelines during pregnancy should emphasize the need to consume a varied, balanced diet'.

1.1.4 Early growth and development influence adult disease

Impaired fetal growth, resulting in low birthweight or thinness or shortness at birth, is known to be linked to higher rates of cardiovascular disease and type II diabetes in later life[31]. The theory behind this observation is that adaptations to a poor intrauterine environment permanently alter the structure and function of a body's developing organs in ways that lead to disease in adulthood. This is often termed the 'fetal origins of adult disease'.

Evidence for early life origins of adult disease was first described by Rose in 1964[32]. He compared the health of the families of patients with ischaemic heart disease (IHD) with that of the families of hospital controls. Relatives of the IHD patients were found to suffer excessive mortality, most evident in early life. In particular the siblings of male cases had stillbirth and infant mortality rates twice that of the controls. This disparity led Rose to suggest that 'it may be that ischaemic heart disease tends to occur in individuals who come from a constitutionally weaker stock'.

Later, Forsdahl[33] used county-level data from Norway and found that mortality rates from arteriosclerotic heart disease were correlated with infant mortality rates in the same county from the period when the cohort were children or adolescents. Forsdahl concluded that childhood poverty was a risk factor for arteriosclerotic heart disease.

A similar geographical association was observed by Barker and Osmond[34] using data from England and Wales. At the level of local authority areas they found a close geographical association between current IHD mortality rates and infant mortality rates in 1921 - 1925. The authors therefore suggested that influences on adult health may operate even earlier than suggested by Forsdahl, and could have origins in prenatal and early postnatal life.

Since increased infant mortality is strongly associated with low birthweight, Barker and colleagues conjectured that low birthweight babies who survived childhood might be at greater risk of IHD in later life. To explore this further, birth records from 5,654 men born from 1911-1930 in Hertfordshire were located, and the men were traced via the National Health Service Central Register. Using information on weight at birth and at 1 year, the combination of poor prenatal and postnatal growth was found to be associated with the highest death rates from IHD[35]. Further analyses amongst an extended cohort of 10,141 men and 5,585 women[36] revealed the same patterns amongst men, and a negative relationship between birthweight only and cardiovascular disease rates in women.

Similar findings have been made in other countries. For example, Forsén et al.[37] observed that amongst 3,202 men born in Finland between 1924 and 1933 high death rates from coronary heart disease were found for those who were thin at birth. Higher death rates were also seen amongst those of low birthweight, although this trend was not statistically significant.

Early studies were limited to the use of mortality as an outcome, since subjects were traced through national mortality records. Subsequently researchers have been able to locate subjects and assess morbidity. A wide range of adult disorders have been related to poor fetal growth and development, including impaired glucose tolerance[38], hypertension[39] and poorer adult lung function[40]. Recently work has shown that men in Finland who were thin at birth were more vulnerable to the effects of poor living standards in later life on coronary heart disease[41].

Since fetal growth is a predictor of health in later life, it is essential to determine the influences on fetal development. Although genes may have an influence, Brooks et al.[42] studied 62 ovum donations and found that the birthweight of the baby was positively correlated with the weight of the recipient mother, but not with that of the donor. Carr-Hill et al.[43] observed that the correlations between mothers' birthweights and those of their children were between 0.14 and 0.18, after standardisation for sex, maternal height and gestational age, indicating that genetic factors play only a small part in determining birthweight. Further evidence is available from sibling studies, such as that by Nance et al.[44] who analysed birthweights of offspring of twins. The correlation between full-sibling pairs was 0.48, and that between maternal half-siblings (i.e. children whose mothers were monozygotic twins) was 0.31. However, the correlation between paternal half-siblings was –0.03. These studies indicate that environmental influences are more important than genetic influences on birthweight.

In conclusion, diet is an important influence on health, and may be a significant mediator of health inequalities. Young women's diets are of particular interest in light of dietary effects on fetal growth and development, which have implications for adult disease.

1.2 What is a 'healthy' diet?

1.2.1 Current dietary guidelines

It is important to understand current dietary guidelines for two reasons; firstly, to be familiar with the dietary advice that young women are exposed to, and secondly to assess whether diets observed are in line with the recommendations. Dietary guidelines from the UK are considered here. However, Margetts and Nelson[45] note that:

> 'dietary guidelines to reduce the risks of chronic diseases have been prepared by many organizations in many countries; there is remarkable agreement between these guidelines, perhaps not surprisingly, because they are all based on a review and development of consensus from the same body of evidence'.

Many researchers have studied various aspects of diet and their influences on health and disease. Public health policy makers aim to synthesise the results of this research in order to promote health through dissemination of current scientific consensus on diet. However, an optimal diet is difficult to define, and is more than the avoidance of clinical features of deficiency disease. Instead such a diet should optimise health, including reducing the risk of chronic disease. Additional guidelines are required to support specific processes such as pregnancy, or recovery from illness.

In 1979 the Department of Health and Social Security published Recommended Daily Amounts (RDAs) for ten nutrients[46]. These nutrient goals were defined as 'the average amount of the nutrient which should be provided per head in a group of people if the needs of practically all members of the group are to be met'.

In 1991 the Dietary Reference Values for Food Energy and Nutrients for the United Kingdom[47] were published, covering a large number of macronutrients, vitamins and minerals. These revised nutrient goals (DRVs) were proposed to succeed RDAs. The single figure RDA was considered to be misleading since amounts referred to are averages for a group of people, and not quantities which an individual must consume. Instead DRVs assume that the requirement of a population for a particular nutrient is normally distributed, and encompass the following:

- Estimated Average Requirement (EAR) – mean requirement
- Lower Reference Nutrient Intake – two standard deviations below the EAR
- Reference Nutrient Intake – two standard deviations above the EAR

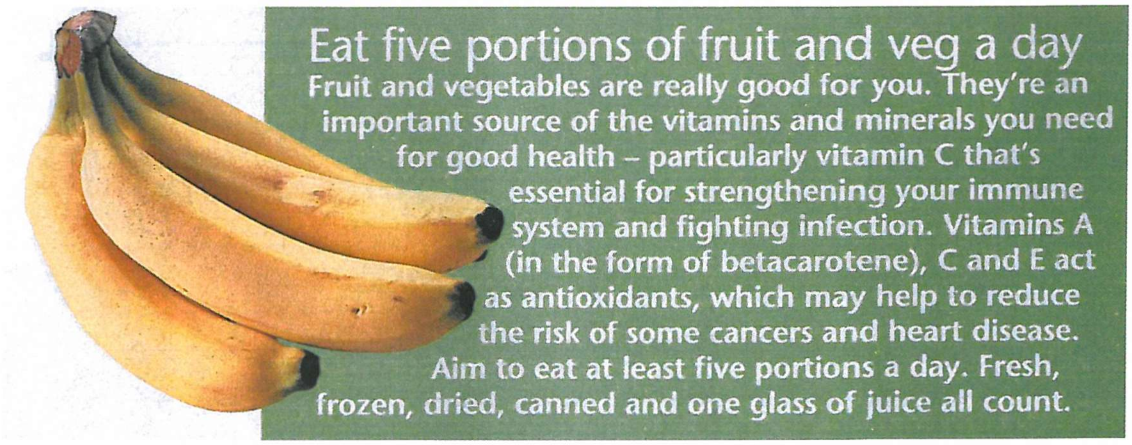
A further development from the RDAs was in the consideration of the criteria of adequacy. DRVs recognise that nutrient requirements may be greater than that necessary simply to prevent clinical signs of deficiency. No single criterion was found to define requirements for all nutrients. Instead, for each nutrient the parameters used to characterise adequacy are described individually. Recommendations are provided by age and sex, and the additional requirements of growth, pregnancy and lactation are considered.

In 1992 the Health of the Nation white paper[48] was published, with the stated aims of reducing coronary heart disease and stroke, cancers, mental illness, HIV/AIDS and sexual health, and accidents. Diet and nutrition targets were included as a means of achieving these objectives:

- To reduce the average percentage of food energy derived by the population from saturated fatty acids by at least 35% (to no more than 11% of food energy).
- To reduce the average percentage of food energy derived from total fat by the population by at least 12% (to no more than about 35% of total food energy).
- To reduce the proportion of men and women aged 16-64 who are obese by at least 25% and 33% respectively (to no more than 6% of men and 8% of women).
- To reduce the proportion of men drinking more than 21 units of alcohol per week and women drinking more than 14 units per week by 30% (to 18% of men and 7% of women).

The Nutrition Task Force (NTF)[49,50] was established in 1993 in order to develop a program of action to achieve these targets. One strategy of the NTF was to provide consumers with dietary guidelines regarding intake of fruit and vegetables, bread and cereals, and fish. Simple, practical messages giving an indication of amounts of foods to be consumed were preferred. Of particular note is the NTF advice to eat five portions of fruit and vegetables per day. This message was originally suggested by the Committee on Medical Aspects of Food Policy[51] through the recommendation that fruit and vegetable consumption in the UK should be increased by at least 50%. Williams[52] calculates that this is equivalent to a total intake of 400g, or 5 portions, per day. The message to 'Eat 5 a day' has been widely disseminated (Figure 1).

Figure 1 Dietary advice from Safeway magazine



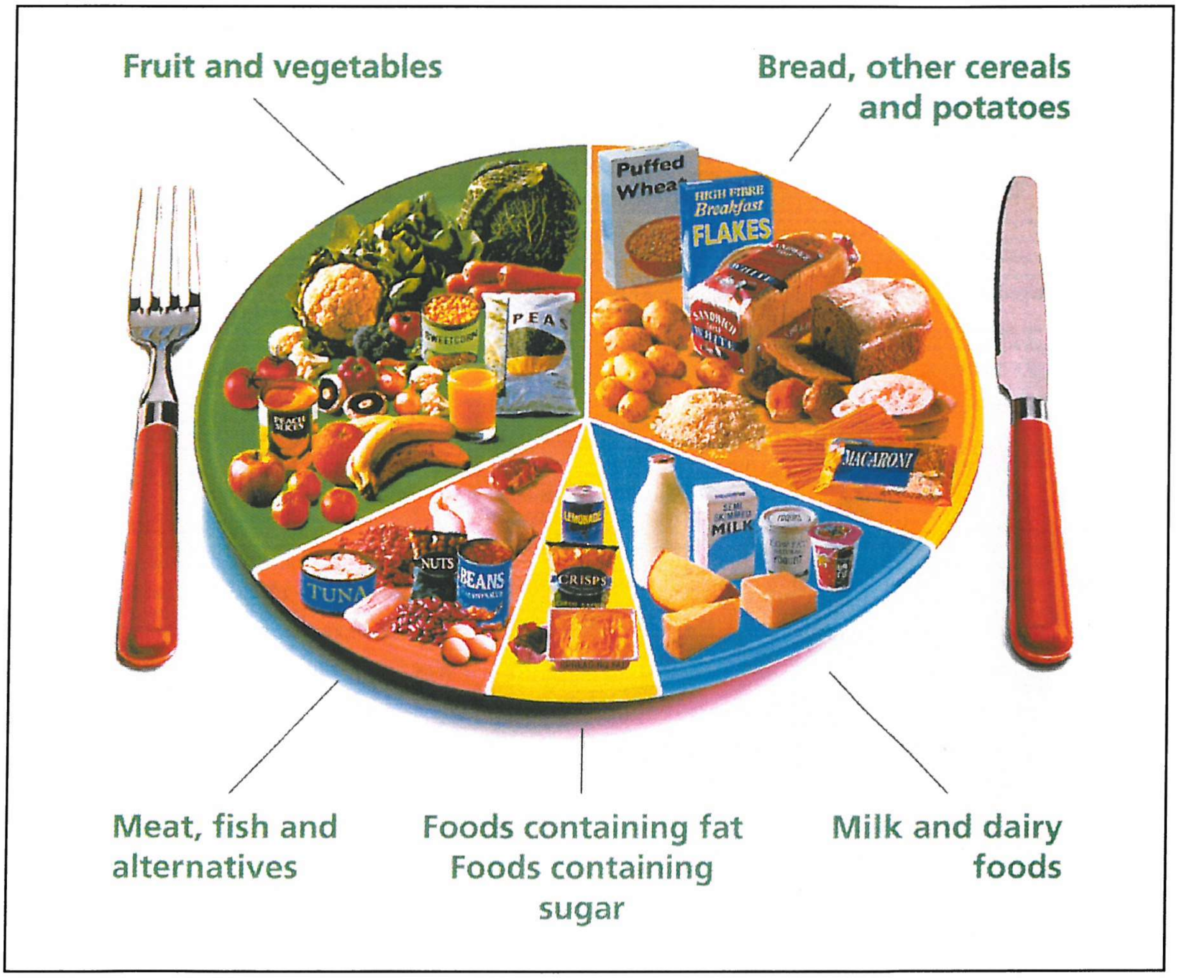
Source: Safeway magazine, January 2001

A further government initiative in the 1990s came from the Health Education Authority in the form of the Eight Guidelines for a Healthy Diet[53]. This advice aims to summarise the main nutrition messages of the DRVs:

- Enjoy your food
- Eat a variety of different foods
- Eat the right amount to be a healthy weight
- Eat plenty of foods rich in starch and fibre
- Eat plenty of fruit and vegetables
- Don't eat too many foods that contain a lot of fat
- Don't have sugary foods and drinks too often
- If you drink alcohol, drink sensibly

An important guide to aid public understanding of healthy eating is 'The Balance of Good Health' (Figure 2), which encourages consumption of foods from each group in proportion to the areas represented.

Figure 2 The Balance of Good Health



Source: Eight Guidelines for a Healthy Diet, Health Education Authority[53]

1.2.2 Current dietary guidelines in pregnancy

Dietary advice available to pregnant women, such as that offered by the Health Education Authority[54], aims to provide a food-based interpretation of current recommendations. Generally, a diet in agreement with 'The Balance of Good Health' (Figure 2) is advised. However, additional nutrients are mentioned to be of particular importance: calcium, iron, folic acid, and vitamins C and D.

Dairy products, bread, nuts, green vegetables and fish with edible bones are recommended to provide calcium during pregnancy. Green, leafy vegetables, lean meat, dried fruit and nuts are cited as good sources of iron. All women are advised to take a daily supplement of 400 micrograms of folic acid prior to conception, and to continue this during the first twelve weeks of pregnancy, to reduce the risk of neural tube defects such as spina bifida. Foods rich in folate including green leafy vegetables, and supplemented breakfast cereals and bread are also recommended. Citrus fruit, tomatoes, broccoli, blackcurrants and potatoes are good sources of Vitamin C, whereas Vitamin D is available from margarine and oily fish.

Additionally pregnant women are advised not to consume untreated milk, undercooked meats or eggs, pâté, soft and blue-veined cheeses, or unwashed fruit and vegetables. This guidance is intended to minimise the risk of listeriosis, toxoplasmosis, and salmonellosis. Liver and products made from liver should not be eaten due to the particularly large amounts of vitamin A found in some livers, that may be harmful to a developing baby. Large quantities of peanuts are to be avoided, to reduce the risk of the baby developing a peanut allergy. Alcohol should be limited to one or two units once or twice a week.

In conclusion, some dietary recommendations have been formulated for both the general population and pregnant women. It is nevertheless unclear how an overall assessment can be made of all aspects of diet. For example, how should a diet meeting recommendations regarding fruit and vegetables, but not regarding fat content, be judged? Dietary pattern analyses can be used to address this issue. Advantages of the use of summaries of dietary patterns are described in Section 2.4.1. Three methods of dietary pattern analysis have been employed by researchers to assess overall dietary quality[55]. These are detailed in Section 1.3.

1.3 Dietary pattern analysis

1.3.1 Principal component analysis and factor analysis

The two related statistical techniques, principal component analysis (PCA) and factor analysis, have been used in the dietary literature. These methods of analysis and the relationship between them are described in detail in Sections 2.4.2 and 2.4.3. Both aim to reduce the dimension of dietary data, by objectively producing scores as linear combinations of food intake.

Results from PCA and factor analysis can be interpreted as eating patterns and researchers have used these techniques to produce dietary scores. Recently an emerging theme in the literature has been the identification of two major eating patterns labelled 'Western' and 'prudent'. The Western diet indicates high intakes of meat, especially red meat, refined grains, chips, and high-fat dairy products, and low levels of fruit and vegetables. The prudent diet describes high consumption of fish, poultry, whole grains, fruit and vegetables. Slattery et al.[56] analysed food frequency questionnaire data from 2,410 American adults in 1998, and identified 6 patterns, the first two being labelled Western and prudent. Subsequently, Hu et al.[57] studied men aged 40-75 living in the Boston area and participating in the Health Professionals Follow-up Study. They recognised that the first two patterns identified were similar to those found by Slattery et al., and therefore labelled their factors in the same way. Hu et al. also noted that patterns labelled 'satiating capacity' and 'healthfulness' from a PCA of 939 Swiss men and women by Gex-Fabry et al.[58] in 1988 were similar to the Western and prudent dietary patterns.

Fung et al.[59] were able to reproduce the Western and prudent dietary patterns in US women participating in the Nurses Health Study. Osler et al.[60] found that results from a factor analysis of the dietary patterns of 5,872 Danish adults could be labelled Western and prudent. McCann et al.[61] used three different methods of grouping dietary data from the Western New York Diet Study. Factor analysis produced two dietary patterns in each instance, that were labelled 'healthy' and 'high-fat'. Descriptions of foods weighting highly on these variables revealed that they are very similar to the prudent and Western dietary patterns. Tseng and DeVellis[62] analysed food frequency questionnaire responses from 5,788 adults in the US National Health and Nutrition Examination Survey. The 'vegetable-fruit' and 'red meat-starch' factors were very close to the prudent and Western dietary patterns. Terry et al.[63] found factors labelled 'healthy' (similar to the prudent pattern) and Western amongst 61,463

Swedish women aged 40-74. Recently Sánchez-Villegas et al.[64] analysed data from 3,847 Spanish men and women and labelled their factors 'Western' and 'Spanish-Mediterranean', a pattern that reflected many prudent dietary characteristics.

Other researchers have described dietary patterns similar to the prudent pattern. Whichelow and Prevost[65] found four patterns in the 1984-5 Health and Lifestyle Survey. The first indicated frequent fruit, salad and vegetable consumption with infrequent intake of high-fat foods. This component seems to be much the same as the prudent dietary pattern. The second component described a diet weighted towards high-starch foods, most vegetables and meat. The third component indicated frequent consumption of high-fat foods, and the fourth denoted high intake of sweets, biscuits and cakes, with negative weightings for vegetables. None of these subsequent components appear to reflect the Western dietary pattern well.

Gregory et al.[66] found five dietary patterns in the 1986-7 Dietary and Nutritional Survey of British Adults (DNSBA), using factor analysis. The first was considered a health-conscious diet, and again clearly reflects the prudent dietary pattern. Subsequent patterns were described as a traditional British diet with meat, vegetables, white bread and potatoes being of importance, a traditional British diet emphasising puddings and cakes, a varied, less traditional diet with wines and spirits, and a fast foods and snack foods diet.

Williams et al.[67] described four components in a study of 40-65 year olds in Cambridgeshire, the first of which was considered to represent 'a conventionally accepted balanced healthy diet'. Beaudry et al.[68] found three patterns in a factor analysis of 24-hour food recall data from 2,104 adults from the Quebec Nutrition Survey. These were labelled 'high-energy density', 'traditional', and 'health-conscious'. This third pattern is similar to the prudent dietary pattern. Barker et al.[69] used PCA to analyse the diets of a random sample of 592 men and women in Northern Ireland. Four distinct dietary patterns were generated, described as 'traditional', 'cosmopolitan', 'convenience' and 'meat and two veg', none of which replicated the prudent or Western patterns. Similarly van Dam et al.'s[70] 'cosmopolitan', 'traditional' and 'refined-foods' patterns found amongst Dutch men and women aged 20-65 do not reflect the prudent or Western factors.

The literature described above indicates that the prudent diet is a major international dietary pattern, being found in all analyses except those by Barker et al.[69] and van Dam et al.[70]. The Western pattern is also an important pattern, but was not observed by Whichelow and

Prevost[65], Gregory et al.[66], Williams et al.[67], Beaudry et al.[68], Barker et al.[69] or van Dam et al.[70]. Variations in dietary patterns may be due to disparities in methods of data collection or analysis. Alternatively there may be differences across the populations studied, or the time period of data collection. Although it is impossible to determine which of these has the greatest influence, it is notable that all the research from the United States has revealed a Western dietary pattern, and therefore this pattern may be most important amongst the US population.

Hu et al.[57] have investigated the important issues of reproducibility and validity of dietary patterns defined by factor analysis in their sample of 127 men in the Health Professionals Follow-up Study. Cohort members completed the same food frequency questionnaire (FFQ) twice, one year apart, and also kept two one-week food diaries during this time. The average intake of each food from the two food diaries was mapped to an equivalent FFQ category. Factor analysis based on 40 food groups produced two major patterns of food intake labelled prudent and Western, similar for both FFQs and the food diary. The correlations between the two FFQs were 0.70 for the prudent pattern and 0.67 for the Western pattern, and the correlations between the two FFQs and the diaries ranged between 0.34 and 0.64 for the two patterns. The authors therefore concluded that the data 'indicate reasonable reproducibility and validity of the major dietary patterns defined by factor analysis'.

Further analysis of a larger sample from the Health Professionals Follow-up Study by Fung et al.[71] also assessed reproducibility of dietary patterns amongst 466 men. Food frequency questionnaires administered in 1986, 1990 and 1994 found reproducible patterns, and at each of three time points it was again considered appropriate to use the labels Western and prudent.

Dietary patterns defined by PCA or factor analysis have been associated with health outcomes. Osler et al.[60] compared their prudent and Western dietary patterns with all-cause mortality and concluded that overall dietary patterns can predict mortality. The dietary pattern associated with lowering risk was the one in line with current recommendations for a prudent diet. Hu et al.[72] examined the associations between the prudent and Western dietary patterns with coronary heart disease amongst 44,875 men aged 40-75 participating in the Health Professionals Follow-up Study. A more prudent diet was negatively associated with disease risk, and a more Western diet was positively associated. However, Osler et al.[73] found no association between the prudent or Western dietary scores and coronary heart disease risk in 7,316 Danish adults.

In the Health Professionals Follow-up Study van Dam et al.[74] observed a substantial association between high Western dietary scores and an increased risk of diabetes. Williams et al.[67] found an association between low prudent diet scores and an increased risk of diabetes. Both men and women in Slattery et al's[56] paper demonstrated a higher risk of colon cancer amongst those eating a Western diet, and a lower risk amongst those consuming a prudent diet. These associations were replicated recently amongst women in the Nurses Health Study, although that for the prudent dietary pattern was non-significant[75]. However, amongst Swedish women Terry et al.[63] found no clear associations between Western or prudent dietary patterns and colon cancer risk. Fung et al.[71] analysed biomarkers of obesity and cardiovascular disease risk from a subsample of 466 men from the Health Professionals Follow-up Study; a more favourable biomarker profile was associated with high prudent diet scores and low Western diet scores.

1.3.2 Cluster analysis

An alternative means of summarising dietary patterns is to use cluster analysis to identify mutually exclusive, homogeneous groups of subjects. Researchers have used cluster analysis methods in order to group subjects into clusters with similar diets, and to examine the characteristics of these clusters.

Hulshof et al.[76] performed cluster analysis to classify individuals into groups on the basis of similarity in nutrient intake in 3,781 Dutch adults. The authors chose seven nutrients for which dietary guidelines have been formulated in the Netherlands. Eight clusters were identified as the most adequate representation of the data, and details were presented for three where the greatest differences in food intake, sociodemographic and lifestyle variables were seen. The clusters were labelled based on their fat and alcohol content as high fat/low alcohol, moderate fat/low alcohol and high fat/high alcohol. The latter cluster was considered the poorest dietary intake since it was also associated with high cholesterol intake and low dietary fibre intake.

Data from a study of university employee volunteers, and two studies of weight-loss amongst moderately obese volunteers, were combined by Wirfält and Jeffery[77]. Six clusters were identified from data on percentage energy contributed by each food or food group. These clusters were labelled by the foods that contributed the highest percentage of total energy: 'soft drinks', 'pastry', 'skimmed milk', 'meat', 'meat-cheese' and 'white bread'.

Greenwood et al.[78] studied the dietary clusters resulting from the UK women's cohort study. These women, aged 35-69, were recruited from the World Cancer Research Fund mailing list; vegetarians were deliberately over-sampled. Seven clusters were found, labelled 'monotonous low-quantity omnivores', 'health conscious', 'traditional meat, chips and pudding eaters', 'higher-diversity traditional omnivores', 'conservative omnivores', 'low diversity vegetarians' and 'high diversity vegetarians'.

Recently Pryer et al.[79] performed cluster analysis on the DNSBA. Men and women were analysed separately and each formed four main dietary groups. The male clusters (categorising 93% of men in the sample) were labelled 'convenience food/beer', 'traditional British', 'mixed but sweet' and 'healthier'. The female clusters (categorising 86% of women) were labelled 'traditional British', 'healthier cosmopolitan', 'convenience food' and 'healthier but sweet'.

Cluster analysis has been found to be useful in more elderly populations to identify eating patterns related to cancer[80,81], coronary heart disease[81], the metabolic syndrome[82] and bone mineral density[83]. However, no consistent dietary patterns have been seen across different studies using cluster analysis, indicating that groupings revealed by the technique may be less robust than PCA and factor analysis, and making it more difficult to compare results across studies. No data are available on the reproducibility of the results of cluster analysis[55].

1.3.3 Diet quality indices

Many researchers have attempted to develop diet quality indices. Kant[84] provides a review of 54 indices of overall diet quality, dividing them into those that examine nutrient intake, food intake, or a combination of both. Several indices examining nutrients are based on adequacy: the intake of a nutrient relative to recommended levels. Alternatively, single-index nutrients can be chosen, using the idea that the intake of nutrients such as fat or a selected micronutrient may permit inference regarding the intake of other nutrients. Amongst indices examining food groups, scores are derived either as an overall dietary score, with higher scores for a larger number of positive dietary habits (and perhaps a reduction in scores for negative habits), or as a measure of dietary diversity.

The Healthy Eating Index is a popular dietary quality score, developed by the US Department of Agriculture (USDA)[85]. It is an aggregate of 10 components, based on both food and nutrient intake. The first 5 components indicate whether the subject is consuming an appropriate number of servings of grain, vegetables, fruits, milk and meat, according to the USDA's Food Guide Pyramid. Two components indicate the percentage energy from fat and from saturated fat. Two further components measure total cholesterol intake and total sodium intake. The final component is a measure of dietary variety. Each of the 10 components can be scored between 0 and 10, leading to a maximum score of 100. This score has been frequently utilised as a research tool, and has been validated against several plasma biomarkers[86].

Two other diet quality indices are of note because they have been used to examine sociodemographic and lifestyle predictors of diet; both have been developed in Finland. Prättälä et al.[87] defined four 'bad dietary habits': spreading butter on bread, drinking high-fat milk, taking more than two lumps of sugar per cup of coffee, and not eating vegetables at least once a week. A 'good' diet in their analyses was simply identified as a subject who had no more than one 'bad' habit. Roos et al.'s[88] diet quality index was based on the Finnish Nutrition Council guidelines and required at least five out of six dietary recommendations to be achieved. These were: low use of fat on bread, avoidance of high-fat milk, and daily use of bread, fruit, vegetables and potatoes.

In conclusion, principal component analysis, factor analysis, cluster analysis and diet quality indices have all been used by researchers as methods of summarising the many facets of diet. The advantages and disadvantages of each technique for use in this thesis are discussed in Section 2.4.1.

1.4 Predictors of diet

1.4.1 Sociodemographic and lifestyle predictors of diet in non-pregnant populations

Many researchers have investigated dietary variations associated with sociodemographic and lifestyle characteristics. Despite the range of settings, subjects, data collection and analysis techniques involved in these studies, it is possible to draw some common conclusions from their findings. Only one study[89] has concentrated on young women exclusively, however, and therefore this literature review summarises evidence from studies that include young women within a more widely defined sample.

1.4.1.1 Receipt of benefits

Information about receipt of benefits was collected as a measure of deprivation predicting dietary intake in the DNSBA. Those in receipt of benefits were found to consume less fruit and vegetables[90] and fewer vitamins and minerals (amongst women)[66], but also less fat[66]. A high proportion of subjects living in households receiving benefits were found in the DNSBA traditional and convenience clusters[79].

1.4.1.2 Education

Education has been used in many studies as a measure of social position, and appears to be consistently related to a diet resembling current dietary guidelines. In Danish[91] and Norwegian[92] adults increasing education is associated with greater fruit and vegetable consumption. For example, Norwegian women[92] receiving less than 13 years of education were likely to eat 381 g/10MJ of fruit and vegetables, whereas those with 13 or more years of education were likely to eat 448 g/10MJ. Amongst residents of a highly deprived area of Leeds low educational attainment was associated with reduced fruit and vegetable consumption[93]. Higher scores on Tseng and DeVellis's[62] 'vegetable-fruit' factor were found for those who had completed higher educational grades.

The Danish[91] and Norwegian studies[92] also showed a reduction in fat intake with increasing education; in a study of women living with children in the Netherlands and Belgium[89] such trends were only apparent for absolute fat intake, and not fat density. Greater intake of fibre amongst those with more education was seen in Norwegian adults[92], and Dutch and Belgian women living with children[89].

Roos et al.[88] and Prättälä et al.'s[87] indices of a healthy diet (Section 1.3.3) are associated with higher levels of education amongst Finnish adults aged 25 or older, whilst in another Finnish study Roos et al.[94] found that educational level was associated with increased intakes of carotenoids and vitamin C. Women in the 'higher' educational group had a 26% greater intake of vitamin C (g/10MJ) compared to the 'basic' educational group, and a 19% greater intake of carotenoids.

Roos et al.[95] performed a systematic review of 33 studies of fruit and vegetable consumption amongst European adults. Generally higher educational levels were associated with greater fruit and vegetable intake. However, exceptions to this trend were found in countries in southern and eastern Europe, where consumption of fruit and vegetables is higher on the whole.

1.4.1.3 Social position

Patterns of eating associated with non-manual social class are similar to those for subjects of higher education; across Europe, non-manual social classes eat more fruit[90,92], vegetables[90,92] and fibre[88], and less fat[92,96]. Cluster analysis of the DNSBA[69] revealed that non-manual social classes are likely to eat a healthy diet, but are less likely to eat a convenience diet, or a traditional British diet.

Principal component analysis of the DNSBA[66] similarly found that the non-manual social classes are less likely to eat a diet described as 'traditional British: meat, veg, white bread and potatoes'. Scores on the other principal components also varied by social class: non-manual social classes scored highly on the 'varied, less traditional with wines and spirits' component, and women of non-manual social class had low scores on the 'fast food and snack food' component. The 'health conscious' component was associated with non-manual social class, a pattern that is reflected by Whichelow and Prevost's[65] finding that non-manual social classes had greater scores on their component representing high consumption of fruit, vegetables and salad, and infrequent consumption of high-fat foods.

Several studies in Europe, the United States and Australia have used an alternative measure of social position, based either on a measure of occupational prestige, or a combination of education and occupational status. Amongst these studies adults with a higher social position ate more fruit[62,97], vegetables[62,97] and fibre[98,99,100], and less fat[98,99,101]. However, Hulshof et al.[99] found that participants with a higher social position consumed more saturated fat, relative to polyunsaturated fat. Increased intakes of vitamins and minerals were also associated with higher social position[98,99]. Total energy intake has been found to be lower in those with a higher social position[100], although Hulshof et al.[99] only recognised this pattern amongst men.

Income level can also be used as a measure of social position[102], and has been found to be positively associated with intakes of fruit and vegetables in Norwegian adults[103], and vitamin C and carotenoids in Finnish adults aged 25 or older[94].

1.4.1.4 Ethnicity

A significantly higher proportion of Black/ethnic minority women, but not men, in the DNSBA were found to comply with dietary fat recommendations from the Committee on Medical Aspects of Food Policy[96]. However, Black subjects consumed fewer vitamins and minerals amongst men and women aged 25 and over, in the United States Nationwide Food Consumption Survey[101]. Pryer et al.'s[79] 'convenience' clusters derived from the DNSBA had higher proportions from ethnic minorities.

1.4.1.5 Age

Although this thesis focuses on young women of childbearing age, within this relatively narrow range age may still be an important predictor of diet, and thus dietary associations with age are considered here. In the DNSBA[90] assessment of the diet of British men and women aged 16-64, increasing age was positively associated with fruit and vegetable consumption. This trend was confirmed in US adults[62], in two samples of Norwegian adults[92,103], and amongst residents of a deprived area of Leeds[93]. Also within the DNSBA[66] and in one of the samples of Norwegian men and women[92], fibre intake was found to increase with age. Consumption of vitamins and minerals was seen to rise significantly with age for both sexes in the DNSBA[66].

Whichelow and Prevost[65] found a striking relationship between increasing age and low scores on their component representing frequent consumption of high fat foods, using data from the Health and Lifestyle Survey of British adults aged 18 or over. Women's score on the 'health conscious' principal component of the DNSBA[66] was seen to increase with age. As might be expected, age is associated with more traditional diets in British[66] and Northern Irish[69] adults, and less convenience or snack foods in British[66,79] and Northern Irish[69] adults.

1.4.1.6 Household composition and marital status

Wandel[103] demonstrates that adults living in households of more than two persons in Norway tend to eat more fruit, vegetables and potatoes. Increasing size of household of participants in the Health and Lifestyle Survey[65] was associated with higher scores on the 'high starch, meat and veg' component, and low scores on the 'more sweets, biscuits and cakes, less vegetables' component. Furthermore, in Norway adults who live with children eat more fruit[103], and Finnish adults living with children scored higher on Roos et al's index of a healthy diet[88]. However, in Denmark women living in families with children had lower intakes of fruit and vegetables compared to single women or couples[91], and residents of a deprived area of Leeds with more children in the household ate less fruit and vegetables[93]. A higher proportion of married men and women were found in the 'healthy' clusters of the DNSBA[79], and married subjects also had higher scores on Roos et al.'s[88] index of a healthy diet.

1.4.1.7 Physical activity

In Johansson et al.'s[92] sample of Norwegian adults, frequency of leisure exercise was associated with increased intake of fruit, vegetables and fibre and decreased consumption of fat. Slattery et al.'s[56] prudent dietary pattern and Tseng and DeVellis's[62] 'vegetable-fruit' pattern were positively associated with more frequent physical activity. Participants classified as sedentary by Gillman et al.[104] were more likely to consume a diet considered suboptimal than subjects taking vigorous or moderate exercise.

1.4.1.8 Dieting

Information about dieting was collected in the DNSBA, where dieters scored higher on a principal component labelled 'health conscious'[66]. Similarly US adults who had attempted weight loss in the last 12 months had higher scores on a 'vegetable-fruit' factor[62]. In a cluster analysis of data from Dutch adults, slimming on the subject's own initiative was more likely in a cluster described as 'moderate fat/low alcohol', that was also noted to include a high intake of fruit and vegetables[76].

1.4.1.9 Smoking

In a 1998 meta-analysis of the relationship between smoking and nutrient intake, Dallongeville et al.[105] found that amongst 51 studies from 15 countries smokers declared significantly higher intakes of energy, total fat, saturated fat, cholesterol and alcohol than non-smokers, and lower intakes of polyunsaturated fat, fibre, vitamin C, vitamin E and β-carotene. Intakes of protein and carbohydrate did not differ between smokers and non-smokers. In food-based analyses there is strong evidence that fruit and vegetable intake is lower amongst smokers in British[90,93], Norwegian[92] and Canadian adults[106].

Amongst studies using principal component analysis or factor analysis, smokers scored highly on components labelled 'more sweets, biscuits and cakes, less vegetables'[65], 'high fat foods' (in women)[65], 'traditional British: meat, veg, white bread and potatoes'[66] and 'red meat-starch'[62].

Smokers had low scores on components labelled 'high starch, meat and veg'[65], 'more fruit, vegetables and salad, less fat'[65], 'cosmopolitan'[69], 'health conscious'[66] and 'vegetable-fruit'[62].

1.4.1.10 Obesity

Much of the research around the associations between obesity and food intake has focused on dietary fat. In their review of evidence from epidemiology, Lissner and Heitmann[107] assessed 13 cross-sectional studies and found a positive association between obesity and concentration of fat in the diet, with correlations ranging from 0.17 to 0.38. A useful review of the associations between measures of obesity and dietary patterns defined by diet quality indices, factor analysis and cluster analysis is provided by Togo et al.[108]. Relationships were generally found to be inconsistent, although it was observed that all significant associations between measures of obesity and diet quality indices were negative. It is important to note that causal relationships cannot be defined between measures of obesity and dietary intake in cross-sectional studies, and thus food consumption may influence obesity, as well as obesity affecting food consumption.

1.4.1.11 Summary

Although there are variations between the studies considered above, some general patterns of dietary intake are seen. Subjects receiving benefits appear to consume less healthy diets, and those with more education, or higher social position have healthier eating habits. Evidence exists that increasing age is associated with a diet more in line with current public health advice, and is also associated with more traditional eating patterns. Trends amongst ethnic groups are less clear, however, particularly since few studies have considered this characteristic as a predictor.

The literature indicates that married men and women tend to have healthier diets, and there is some evidence that those who are married or live with children, or who live in larger households, have eating patterns more in line with current dietary recommendations. As might be anticipated, subjects who engaged in health-seeking behaviour (exercising regularly, dieting and not smoking) are more inclined to have dietary habits in agreement with current guidelines. There is some evidence that obese subjects eat diets containing greater amounts of fat, and which have worse scores on diet quality indices, than subjects who are not obese.

It is important to remember that all dietary intake described above is merely reported, and thus any differences may not reflect variations in actual food consumed. It is wholly plausible, for example, that subjects with more education have a superior knowledge of current dietary advice, and therefore report diets in closer accord with those standards.

1.4.2 Sociodemographic and lifestyle predictors of diet in pregnant women

A small amount of research has investigated sociodemographic and lifestyle predictors of diet specifically within pregnant samples. Smithells et al.[109] found that pregnant women of lower social classes had lower intakes of both macronutrients and micronutrients. Haste et al. repeated this observation for macronutrients[110] and micronutrients[111] , whilst Schofield et al. observed the pattern for micronutrients[112], but not macronutrients[113].

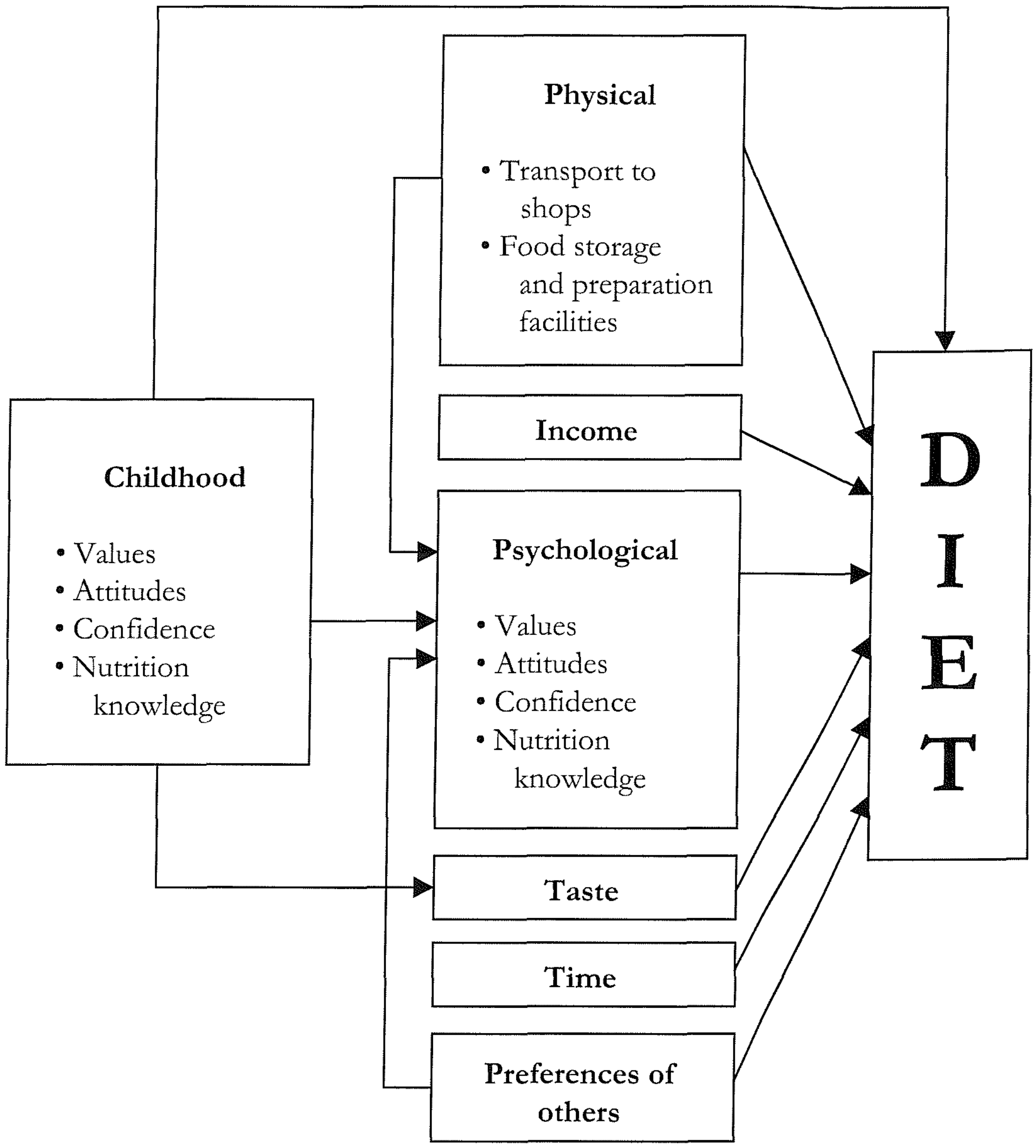
Pregnant women with less education are seen to consume fewer micronutrients[109,114], and younger women similarly have a lower micronutrient intake[110,114]. Amongst studies in the UK and Norway consistent patterns of lower micronutrient intake amongst smokers were seen[109,110,111,114,115]. In a study of 11,833 pregnant women in the South West of England, Rogers et al.[116] found that non-smokers consumed diets that might be described to be in closer accordance with current dietary recommendations.

In summary, the predictors of diet that have been investigated in pregnancy appear to have similar effects to those observed in non-pregnant populations.

1.4.3 Conceptual model of influences on diet

Although knowledge about sociodemographic and lifestyle predictors of diet is helpful, it is important to remember that these influences are often not direct, but are effected through other factors. For example, there is evidence that higher levels of education are associated with a healthier diet (Section 1.4.1.2), but this could be due to the impact of education on psychological factors such as values and confidence, or greater knowledge of public health guidelines. Alternatively, money, the physical environment a woman uses to store and prepare food, or influences in childhood may be associated with education. Since research to date is limited, many models of these factors could be constructed[117]. Figure 3 illustrates a model of potential pathways of proximal influences on diet.

Figure 3 Conceptual model of influences on diet



Habits formed in childhood are likely to have a strong effect on a woman's food choices. In a qualitative study of influences on fruit and vegetable intake[118], habit or tradition were mentioned repeatedly in focus groups. Subjects often stated that 'they were eating the way they were taught at home in the past and continued eating according to those habits when they left their parents to go and live on their own or started their own family'. These effects are likely to act directly and through values, attitudes, knowledge and cooking confidence acquired during childhood, as well as effects on taste preferences, as indicated by the arrows in Figure 3.

Physical barriers to consuming a healthy diet are outlined by Leather[119]; the weight of fruit and vegetables can be a obstacle to purchasing them, and may be a considerable difficulty for those who shop by bus, particularly when travelling with a buggy and small children. Income might also affect dietary choices since a limited income necessitates careful consideration of the price of food. In a study of perceived influences on food choice of adults in the European Union, Lennernäs et al.[120] report that price was regarded as an influence by 43% of respondents.

Lennernäs et al.[120] also note that 'trying to eat healthily' was an effect on the food choices of 32% of respondents, indicating that 'healthy eating' is an important value in this population. Confidence is a further psychological influence; Leather[119] describes how lack of confidence and skill in purchasing and preparation are barriers to higher fruit and vegetable consumption. The effect of nutrition knowledge on food choice was investigated by Wardle et al.[121] in a study amongst adults in England. They found that respondents in the highest quintile for knowledge were almost 25 times more likely to meet recommendations for fruit, vegetable and fat intake than those in the lowest quartile.

Pollard et al.[117] note that 'food is not just eaten for its nutrient value; for many people it is a source of pleasure, an enjoyable experience and even a comforting activity'. Therefore taste, along with other sensory properties such as texture, smell and appearance, play a role in food choice. Taste and lack of time were the main perceived barriers to healthy eating in a study of adults from the European Union[122], particularly amongst younger participants. Time available to go shopping, and time required for preparation are both pertinent issues, as well as the related concern of the perishable nature of fruit and vegetables, necessitating more frequent shopping trips.

'Taste preferences of family and friends' was selected as a barrier to healthy eating by 13% of the subjects in the European Union study by Kearney and McElhone[122]. Brug et al.[118] distinguish between influences of members of the same household (partner and other family members) and influences of others (friends etc.). It is likely that preferences of others will affect food consumption, particularly when sharing a meal, and that values and attitudes of others will influence those of the woman herself, as indicated by the arrow in Figure 3.

Since variables measuring the concepts in Figure 3 are not available in the PAH study or the SWS, associations between sociodemographic and lifestyle factors and diet will instead be explored. These may be influenced by aspects of childhood, but their effects on diet are likely to be somewhat mediated by factors in the second column of Figure 3. Therefore in the analysis of associations between sociodemographic and lifestyle factors and diet it should be remembered that some of the effects of the available factors might not be direct, but instead mediated through the more proximal influences in Figure 3. Frameworks of influences of data available in the PAH study and the SWS are developed in Section 5.2 and Section 7.2.

1.5 Objectives of this thesis

Young women's diets are important in terms of their own health and that of their children. Since diet is a complex, multi-faceted phenomenon, appropriate multivariate statistical analysis is essential to summarise aspects of dietary behaviour. Diets vary according to sociodemographic and other factors, and it is necessary to characterise women's social position appropriately to understand which sectors of society have the least favourable diets.

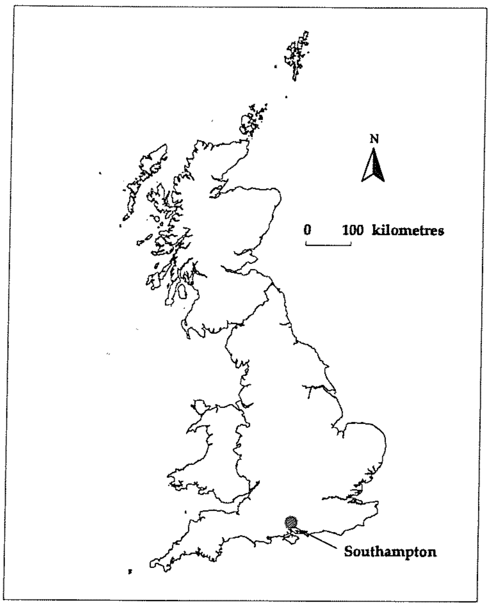
The aims of this thesis are therefore:

1) To examine how sociodemographic and lifestyle measures available in the data can be used to summarise the situations of young women in the UK.
2) To employ suitable statistical methods for characterising the diets of these women, using all aspects of food consumption.
3) To assess whether dietary information collected using diaries provides similar results to that collected using food frequency questionnaires.
4) To examine changes in women's diets, both between early and late pregnancy, and over a period of 2 years amongst non-pregnant women.
5) To study the effects of sociodemographic and lifestyle measures on women's diets, in both pregnant and non-pregnant young women.

In analyses of sociodemographic and lifestyle predictors of women's diets it is important to make explicit the hypothetical mechanisms of effect[123]. Therefore models of causal pathways of influence are proposed, and analyses proceed using this framework.

Two datasets are analysed: a pregnant and a non-pregnant sample of women from Southampton, UK (Figure 4).

Figure 4 Location of Southampton, UK



2 Data and methods

Two relevant datasets are available to investigate sociodemographic predictors of young women's diets in Southampton, both collected by the Medical Research Council (MRC) Environmental Epidemiology Unit. The Princess Anne Hospital study is described in Section 2.1 and the Southampton Women's Survey in Section 2.2. These studies used the same food frequency questionnaires to assess diet, and this method of data collection is discussed in Section 2.3. Summarising dietary patterns revealed by food frequency questionnaires is complex, and two statistical techniques are employed in this thesis: principal component analysis and factor analysis. Fruit and vegetable consumption is a simpler dietary summary pertinent to current public health advice. These methods are described in Section 2.4.

2.1 Princess Anne Hospital study

The Princess Anne Hospital (PAH) study was conducted at the Princess Anne Maternity Hospital, the only National Health Service maternity hospital in Southampton. The primary objective of the study was to assess how maternal nutrition in early and late pregnancy influences fetal growth and development. All white women aged 16 years or older, with singleton pregnancies at less than 17 weeks gestation, who registered under two consultants between October 1991 and October 1992, and attended the midwives' antenatal booking clinic were approached. Women were excluded if they had a past history of diabetes or treatment for infertility. Of the 655 eligible women, 620 (95%) were willing to take part and were visited by a trained research nurse at home shortly after recruitment (median gestation 15.3 weeks): the 'early pregnancy' interview. Of these 620 women, 27 subsequently miscarried, moved away or did not wish to continue with the study, resulting in 593 'late pregnancy' interviews (median gestation 32.7 weeks).

At each early and late pregnancy visit the nurse administered a questionnaire (Appendices 1 and 2). The women were asked about their physical activity, smoking habits, alcohol consumption, qualifications, medication, nausea and illnesses, as well as their own and the baby's father's work. Their menstrual and obstetric history was recorded, and the woman's and baby's father's birthweights. The woman was asked at the early pregnancy visit to recall her weight before she became pregnant, and her height was measured at the antenatal clinic. The woman's date of birth and information about number of previous live-born infants was available from her antenatal records. In addition, a large part of both the early and late pregnancy questionnaires consisted of a food frequency questionnaire (Section 2.3.3).

At the early pregnancy booking an aliquot of routine non-fasting blood sample was taken and used for estimation of serum vitamin C concentration by the ascorbate oxidase - orthophenylene diamine assay on a Cobas analyser. Serum samples were stabilised with 10% metaphosphoric acid within 6 hours of collection and stored at –40°C prior to analysis. Samples were available for 579 women.

Further to the FFQ, women were asked to keep a prospective food diary for four days (including one weekend day) following their early pregnancy visit. This was an open diary in which participants recorded food intake using measured dimensions, weight, number, or in terms of household measures. At the late pregnancy visit the nurse checked the diary, and the volume of household measures used was determined. Additionally the nurse subjectively recorded whether she considered the diary to be 'excellent', 'average', or 'very poor, probably incomplete'. Diary data for all four days are available for 588 women.

Subsequent measurements were made of the baby at birth, 9 months and 9 years, but these data are not considered in this thesis. Rather the pregnancy dietary data provide an opportunity to study sociodemographic predictors of diet amongst a representative sample of young, pregnant women in Southampton. A strength of the PAH study is that dietary changes between early and late pregnancy can be examined.

2.2 Southampton Women's Survey

The Southampton Women's Survey (SWS) was initiated to extend the Medical Research Council's research into intra-uterine origins of chronic disease. Previous studies recruiting women on presentation at antenatal clinic are limited because by this point (usually around 12 weeks of pregnancy) it is thought that much of the trajectory of growth may have been established[124]. The trajectory is an important influence on subsequent fetal nutritional requirements. Little is known about maternal influences that set the fetal growth trajectory, but animal experiments indicate that diet and body composition around the time of conception will be important[124]. Since diet and body composition change in early pregnancy it is difficult to study their effect retrospectively on women attending antenatal clinics. Instead, the SWS provides data with which to study these effects prospectively.

The aim of the SWS was to interview 12,000 women aged 20 to 34 resident in the city of Southampton. Participants were recruited via names and addresses obtained from all general practitioners (GPs) in the city. Also self-referrals were encouraged using local publicity in an attempt to recruit those amongst this mobile age group with out-of-date addresses on their GP register, or not registered with a GP at all. The recruitment procedure was thus designed to provide as representative a sample as possible. Approximately 75% of all women approached agreed to take part in the survey.

Women aged 20 to 34 are of peak child-bearing age, and if a woman in the SWS becomes pregnant then her pregnancy is monitored in detail. In this way, prospective information will be available to allow the study of pre-pregnant as well as pregnancy influences on fetal development. Inskip[125] gives further details of methods and ancillary studies related to the whole SWS. However, baseline information collected on the total non-pregnant population provides a unique dataset in which to study the sociodemographic predictors of diet amongst a contemporary, representative sample of young women in Southampton.

A trained nurse administered the baseline SWS questionnaire (Appendix 3). The woman was asked about her own and partner's occupation, physical activity, smoking, family background, education, ethnicity, housing, household composition, childcare arrangements, benefits, general health, and menstrual and obstetric history. Furthermore, a large part of the questionnaire consisted of an FFQ, virtually identical to that used in the PAH study (Section 2.3). Detailed body composition measurements were taken, including weight, height, waist and hip circumference, and skinfold measurements at four sites. The nurses were carefully trained and regular inter-observer variability studies performed to ensure measurements were as accurate as possible.

Data from 6,153 pre-pregnancy interviews are available for analysis in this thesis. These women were interviewed between April 1998 and June 2000. Twenty-four of the 6,153 interviews were found to be outside the valid age range, and were thus excluded, leaving an analysis sample of 6,129.

Women who were willing were asked to give a blood sample during the second half of her menstrual cycle. Immediately after collection the samples had a full blood count performed and red cell folate was assessed by microparticle enzyme immunoassay. Measurements were available from 3,823 women who were willing to provide a blood sample.

An additional component of the SWS consisted of a 'stability study'. This involved re-interviewing 94 women approximately two years after their initial interview using a questionnaire virtually identical to that in Appendix 3. None of the 94 women had been pregnant during the two year interval. These data about the stability of their body composition and lifestyle are of particular interest to the study of the intra-uterine origins of chronic disease because women are unlikely to conceive directly after their initial interview; thus an assessment can be made of the stability of parameters that might be most important around the time of conception. However, the data are also of interest for this thesis, in order to assess the stability of dietary patterns over time, as well as sociodemographic and lifestyle predictors of changes that may occur.

2.3 Food frequency questionnaires

2.3.1 Choice of FFQs

Diet is very difficult to measure accurately, and several different methods are used in epidemiological research. Food diaries are a prospective method and require a participant to record their entire food and drink consumption over a number of days. Conventionally a pair of scales is provided with which the participant weighs all intake. An alternative prospective method is the duplicate diet, where the participant is asked to reserve an identical portion of every food and drink consumed for researchers to analyse.

Dietary recall requires the participant to recall retrospectively all food and drink consumption over a recent period, typically the previous 24 hours. Food frequency questionnaires are another retrospective technique which enable a participant to describe usual frequency of food consumption, rather than needing to remember foods eaten at any specific meal in the past. Thus a participant is presented with a list of food and food groups, and asked to choose the frequency with which they consume each, from several categories.

The diet history method is a combination of retrospective and prospective methods. It consists of a three-part assessment: a detailed face-to-face interview with a trained interviewer, with the aim of assessing usual consumption of a wide variety of foods, a cross-check FFQ, and a three-day food diary[126].

FFQs were chosen for use in the Southampton studies because they require minimal coding and are less demanding than other methods to complete, so data may therefore be available for a more representative sample from the population, including the illiterate and semi-literate. This property is particularly attractive in the context of this thesis since a response bias could clearly be related to socio-demographic factors. A further advantage of FFQs over food diaries, dietary recall and duplicate diet methods is that they refer to a longer time frame, and thus may give a better approximation to usual diet.

Interviewer-administered FFQs are completed far more fully than participant-completed dietary responses. This is an important feature in multivariate analyses using all aspects of diet, since just one missing item means that the participant is excluded from the entire analysis, unless imputation based on other variables is used. The Southampton MRC FFQs were administered by trained nurses who have ensured virtually non-existent missing data, and therefore imputation is not necessary. In the PAH early pregnancy data only three participants (out of 620) are excluded with missing teaspoons of sugar per day, and in the late pregnancy data just two participants (out of 593) have missing responses (to the question about white bread intake). In the SWS data four women (out of 6,129) have not reported their wine intake, and two of these also omitted information about teaspoons of sugar.

Interviewer-administered FFQs may have the disadvantage that a participant could be less honest in answering a question posed by an interviewer than if they completed the questionnaire individually. Also, all FFQs are subject to recall bias due to errors in conceptualisation or memory of food consumption. A further disadvantage of FFQs is that although they have been seen to rank individuals meaningfully within distributions of nutrient intake[127], they are not considered such an accurate measure of an individual's diet.

However, food frequency questionnaires are regarded by many to be the method of dietary assessment best suited to most epidemiological applications, and have become the primary method of measuring dietary intake in epidemiological studies[128]. When FFQs are employed it is important that they are made relevant to the population under consideration, by tailoring to typical food consumption patterns; FFQs are not necessarily transferable between populations.

Four-day food diaries were used additionally in the PAH study to enable validation of the FFQ[127]. The diary data from those who completed all four diary days are utilised in this thesis to further explore dietary patterns amongst the PAH participants. Diary entries were converted to FFQ categories to enable the same analyses on both types of data.

2.3.2 Reproducibility and validity of FFQs

Both reproducibility and validity are important criteria for a useful FFQ to fulfil. 'Reproducibility' in this thesis (following Willett[128]) refers to the consistency of responses from an FFQ when administered to the same subject on different occasions, recognising that conditions can never be exactly replicated. 'Validity' refers to the consistency of FFQ responses with an external, reference measure, such as a food diary.

Much research has addressed the question of reproducibility of FFQs. In a summary of reproducibility studies by Willett[128], correlations for nutrient intakes in a wide variety of settings were generally seen to range from 0.5 to 0.7. Correlations for foods were considered somewhat more variable, ranging between approximately 0.3 and 0.7. Willett comments that although levels of reproducibility for nutrients may not seem very high, they are comparable with those of many other biological measurements on free-living subjects such as blood pressure and serum cholesterol.

Reproducibility is an important criterion, but a completely reproducible questionnaire may nevertheless fail to provide a measure of the true intake of a particular food, and thus not be valid. The validity of FFQs has also been assessed in many studies. In a summary of research into the validity of both nutrient and food intakes by Willett[128] in a wide variety of populations, validity of FFQs was considered 'generally encouraging'; just one study was discrepant, perhaps because it was conducted on lactating women. Amongst the more recent studies, average correlations for energy-adjusted nutrients were generally between 0.6 and 0.7.

It is often noted that no method of dietary assessment is perfect; Margetts and Thompson[129] suggest that validation studies should be considered instead as 'calibration' studies. Thus in a study comparing FFQs with food diaries, error is associated with the diary as well as the FFQ itself. Margetts et al.[130] attempt to model this variation, as well as variation that might be expected from other sources, in a comparison of an FFQ with a 24-hour food diary. The sources of variation considered were within-person variation, diet record measurement error, trend in diet, and food frequency measurement error. They found that the variation between the FFQ and the food diary was of a very similar magnitude to that expected with realistic estimates of error in their model, and concluded that their FFQ was able to categorise individuals by nutrient intake in a similar way to a 24-hour food diary.

The Southampton MRC FFQ (Section 2.3.3) nutrient intakes have been validated in the PAH study against the four-day food diaries and serum vitamin C[127]. Spearman rank correlation coefficients between the FFQ and diaries for macronutrients ranged from 0.27 (protein and starch) to 0.37 (fat). Using serum vitamin C as an independent biomarker of intake, the FFQ classified 34% of women to the correct quarter, and 8% of women to the opposite quarter. The FFQ was found to give an average ratio of energy intake to basal metabolic rate (BMR) consistent with that expected, whereas energy intakes determined by the food diaries appeared to be low. However, there was also some over-reporting of energy intake on the FFQ since 5% of the population had a ratio of energy intake to BMR ratios above 2.8 (the maximum value obtained from the food diaries).

Alongside the evidence for the reproducibility and validity of FFQs, there are known biases in FFQ data. In particular, studies have demonstrated that FFQs typically overestimate consumption of fruit and vegetables[131,132]. It is thought this may be due to repeated reporting in related, but distinct food categories. A contrasting bias is the tendency for obese participants to under-report food intake when assessed using FFQs as well as when using diet history, food diaries or 24-hour recall methods[126,133].

In conclusion, FFQs are seen to be adequately reproducible and valid. The FFQ is the most appropriate tool for use within the Southampton studies, although limitations of the technique must be considered when results are interpreted.

2.3.3 The Southampton MRC FFQ

The Southampton MRC FFQ is administered by a trained interviewer. At the start of the FFQ the interviewer ensures that the participant is thinking about the previous three-month period in her response to the questions. The participant is then asked to specify her frequency of consumption of each of 100 foods and food groups, using categories ranging from 'never' to 'more than once a day'. If the frequency is more than once a day, the participant is asked how many times per day. Several additional questions at the end of the FFQ provide supplementary dietary information.

The Southampton MRC FFQ was developed to ascertain frequency of consumption of foods that contribute at least 90% of the macronutrients, iron and vitamin C to the diet[127]. This emphasis on nutrients means that foods with similar nutrient compositions have been grouped together. If these foods are consumed in different ways then analyses of eating patterns based on foods, rather than nutrients, may be less clear. A notable example is the grouping of blackcurrant drinks with blackcurrants (FFQ category (46)). Also chips are combined with roast potatoes in FFQ category (18); although these foods have a similar nutrient composition, chips are often eaten as a snack bought outside the home, whereas roast potatoes are usually eaten as part of a more formal meal at home or in a restaurant.

There were slight differences between the FFQs used in the PAH study and the SWS (Appendices 1 and 3). Notably FFQ category (86) (Meat stock cubes/Oxo/Knorr/ Bovril) in the PAH questionnaire is split into two separate categories in the SWS questionnaire ((86A) Gravy granule and powders, and (86B) Stock cubes and Marmite). This permitted a more accurate assessment of iron intake. Also, frequencies of red wine and white wine/cider (FFQ categories (91) and (92)) in the PAH questionnaire are not included in the SWS FFQ. Instead the SWS interview includes more detailed information about alcohol consumption in Section 7. Here frequency of wine/sherry/Martini/Cinzano is one category, and information on number of glasses normally consumed is also available. The SWS questionnaire incorporates additional questions about extra foods that the participant has eaten or drunk once a week or more, but are not included in the main FFQ.

Prompt cards are used with the Southampton MRC FFQ; the participant is shown the relevant prompt cards as she responds to the questions. The prompt cards indicate all the foods relevant to each question, and are used to standardise prompting across interviewers. As an example, the 'fruit-based puddings' PAH prompt card is shown in Figure 5. Minor alterations to the foods included on the prompt cards were made between the PAH study and the SWS to ensure foods mentioned were relevant to contemporary diets.

Figure 5 Example of PAH FFQ prompt card

FRUIT BASED PUDDINGS

e.g. Apple crumble/apple pie
Lemon meringue pie
Trifle
Apple strudel
Summer pudding

2.3.4 Preparation of dietary data

2.3.4.1 Bread and potatoes

In the Southampton FFQ each participant specifies the number of slices/rolls of bread eaten at a typical meal, as well as the number of potatoes eaten at a typical meal. To incorporate this additional quantitative information, frequency of white bread consumption was multiplied by the number of white slices/rolls to give a frequency of consumption of one slice of white bread. For example, two slices of white bread once a fortnight was converted to one slice of white bread once a week. Brown and wholemeal bread was modified in the same manner, and 'boiled and jacket potatoes' and 'roast potatoes and chips' were converted to frequency of consumption of one potato, where a large baked potato counts as 3, and a new potato counts as 0.5. The PAH diary data was coded only as frequency, and not quantity, of bread and potatoes.

2.3.4.2 Wine

In the PAH study, information about wine intake is available in the FFQ. In the SWS the average number of glasses of wine consumed is additionally supplied. Similarly to the bread and potatoes categories, this information can be used to gain a more accurate measurement of wine consumption. Thus frequency of wine intake was multiplied by number of glasses typically consumed, to give a frequency of drinking one glass of wine per week. Since only frequency of wine, and no other alcoholic drink, is included in the PAH study, only the response to the wine question is used in analysis of the SWS.

2.3.4.3 Fats

Intake of fats was collected in FFQ categories (93) – (100), and coded into brand information, and then into equivalent McCance and Widdowson codes[134]. Since a woman's choice between spreading fats might be an important aspect of her dietary pattern, FFQ categories (93) - (95) were subsequently classified as either full-fat spreads, or reduced-fat spreads (including low-fat spreads and very low-fat spreads). Frying fats and oils and other vegetable oils, particularly those used in salad dressings (FFQ categories (96) to (100)) were grouped together and labelled as cooking fats and salad oils. Thus, by summing the relevant FFQ categories, a frequency of full-fat spreads, reduced-fat spreads, and cooking fats and salad oils was available for each woman.

2.3.4.4 *Milk*

The additional dietary questions following the FFQ included information about milk intake. Participants were asked about types and pints of milk used per day. Most milks were liquid, but some women specified powdered milks. In this case the number of grams used was recorded in the PAH study, and the number of teaspoons reported in the SWS. In the SWS the number of teaspoons was converted to grams using the conversion 1 teaspoon = 5 g. In many analyses it was necessary to combine liquid and powdered milk intake. Examination of several powdered milk cartons indicated that approximately 50g of powdered milk are needed to make up 1 pint. Therefore in all analyses powdered milk intake in grams was divided by 50 and analysed as a pint equivalent. Information on pints of liquid milk and grams of powdered milk consumed was available from the PAH food diaries, and the data was prepared in the same way.

In the PAH study all milks were coded using McCance and Widdowson codes[134]. In the SWS respondents classified their milk as whole, semi-skimmed, skimmed or 'other'. The 'other' codes were converted to McCance and Widdowson codes. Subsequently all milks were coded as either full-fat or reduced-fat, reduced-fat milks including both skimmed and semi-skimmed. Therefore a frequency of pints of full-fat liquid milk, reduced-fat liquid milk, full-fat powdered milk and reduced-fat powdered milk was available for each participant. Each of these daily frequencies was multiplied by seven to give a frequency of pints of milk per week.

2.3.4.5 *Teaspoons of sugar*

A further additional dietary question referred to the number of teaspoons of sugar each participant added to breakfast cereals, tea and coffee, puddings etc. per day. This figure was multiplied by seven for each woman to give the number of teaspoons of sugar per week. Information on teaspoons of sugar consumed was available from the PAH food diaries, and the data was prepared in the same way.

2.3.4.6 *Final categories*

In the PAH study the 100 FFQ categories were reduced to 95 by summarising the eight fat responses using three variables. However, the additional four milk categories, along with frequency of teaspoons of sugar intake meant that 100 PAH food variables were analysed in total. In the SWS there was an additional variable created by the splitting of gravy and stock cubes, but only one variable for wine, as compared to the two in the PAH study. Therefore there were also 100 SWS food variables in total. All FFQ categories were converted to an average frequency per week using the factors in Table 1.

Table 1 Conversion of FFQ frequencies

FFQ category	Frequency per week
Never	0
Once every 2-3 months	0.1
Once a month	0.25
Once a fortnight	0.5
1-2 times per week	1.5
3-6 times per week	4.5
Once a day	7
More than once a day	Number of times a week specified by the participant

2.3.4.7 Food nutrient intakes

In addition to frequency of food consumption, food nutrient intakes were available for analysis in the PAH study. These were calculated by converting the FFQ categories and additional dietary information into standard portion sizes[135], and then multiplying by McCance and Widdowson's food composition data[134,136,137,138]. Note that nutrient intake from supplements is not included in analyses in this thesis.

2.4 Statistical methods of dietary analysis

2.4.1 Multivariate techniques

Traditionally overall diet has been assessed by the measurement of nutrients, but this approach has many limitations. Such a description of diet is hardly adequate in reducing the number of dimensions under consideration: there are still many nutrient outcomes. Also, as Jacques and Tucker[139] remark 'We don't eat nutrients, we eat foods'. Reducing dietary patterns to summaries of nutrients is unlikely to be very relevant to the choices people make every day with regard to their diet; when investigating sociodemographic and lifestyle predictors of diet it is clearly important that the outcomes chosen are as closely linked to behaviour in the population as possible. Furthermore, foods consist of substances other than nutrients, and thus food intake may not be adequately represented by nutrient intake.

The vast range of food available creates a need for summaries of dietary patterns, if analyses using foods are to be useful. In addition, since foods are not generally eaten in isolation, overall dietary patterns may have more of an influence on health than individual foods. Combinations of food intakes might be more important than individual intakes. Hu et al.[57] note that iron absorption is enhanced in the presence of vitamin C, and thus the effect of a food high in iron might be greater when eaten in proximity to a food high in vitamin C. Furthermore, the intakes of some foods are so highly correlated that it becomes difficult to examine their effects separately. For example, Jacques and Tucker[139] note that whole-grain consumption is positively associated with fruit, vegetable and fish consumption. Therefore a disease inversely associated with whole-grain consumption might actually be caused by low fruit and vegetable or fish intake. However, the analysis of patterns of diet within a population uses such collinearity of foods to advantage.

With the availability of powerful computing capabilities it is now possible to use complex multivariate statistical techniques to seek to produce multidimensional descriptions of diet. The main disadvantage of such dietary pattern analyses is that they are uninformative about specific nutrient-disease relationships. However, the ethos of dietary pattern analysis is summarised by Hu et al.[57]: 'Examination of patterns would more closely parallel real-world conditions, under which dietary intakes consist of nutrients that occur together in common foods'.

Principal component analysis (Section 2.4.2) and factor analysis (Section 2.4.3) are techniques of dietary pattern analysis commonly used in the literature (Section 1.3.1). Both aim to reduce the dimension of food data by producing a smaller number of derived variables. Cluster analysis is a further technique of dietary assessment often used (Section 1.3.2). An important difference between the two methods is that PCA results in a continuous score for each individual, whereas cluster analysis assigns each participant to one of several groups. Since PCA therefore has the potential to characterise relatively small differences between individuals it was considered more suitable for the purposes of summarising diet in this thesis than cluster analysis. Greater consistency has also been found amongst principal component and factor analysis patterns in the literature, than cluster analysis patterns, and no data are available on the reproducibility of cluster analysis results (Sections 1.3.1 and 1.3.2). Furthermore, it was felt that characterising an individual's diet using a continuous score might be a more pragmatic reflection of dietary choice than assignment to one of several discrete categories.

Diet quality indices are frequently used as a measure of dietary assessment in the literature (Section 1.3.3). The attraction of scoring a participant on a scale designed to describe an aspect of diet is clear. However, available scores usually have the specific aim of avoidance of one or more chronic diseases, and are therefore limited in their application. Such scores are also limited by current knowledge[55], and are only based on a few aspects of diet, rather than involving information about all foods or nutrients consumed. Furthermore, there is no guarantee that the axis represented by the score is one along which participants will vary markedly. Since PCA and factor analysis are based on the whole range of foods within the diet, and PCA provides a score with the property of maximum variance amongst subjects by definition, these techniques are employed, rather than diet quality indices.

All data analyses presented in this thesis were performed using Intercooled Stata 7.0 for Windows[140].

2.4.2 Principal component analysis

One hundred food variables were available for analysis from the Southampton MRC FFQs. Rather than analysing each food separately, it is natural to aim to reduce the number of variables under consideration. The simplest reduction technique would be to focus on consumption of a small number of foods or food groups. If an overall description of diet is required, however, such a choice of foods is arbitrary and may conceal a significant proportion of the variation in the data.

Principal component analysis is a dimension-reduction technique, which aims to construct a small number of derived variables from a large number of original variables. In the context of dietary data the intention is to construct new variables that can be understood as recognisable aspects of diet. The new variables are linear combinations of the original variables. They are chosen to retain the maximum amount of the variation in the original data and to be independent of each other.

2.4.2.1 Theory

Using the notation of Joliffe and Morgan[141], PCA aims to construct m derived variables $z_1, z_2, \ldots, z_m$ from p measured variables $x_1, x_2, \ldots, x_p$, where m is much less than p. The variables $z_1, z_2, \ldots, z_m$ are defined to be linear functions of $x_1, x_2, \ldots, x_p$.

The first component z_1 has maximum possible variance amongst all possible linear functions

$$z = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_p x_p = \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{x}$$

However, the variance is unbounded since if any potential solution $\alpha_1, \alpha_2, \ldots, \alpha_p$ is multiplied by 10, for example, the variance increases by a factor of 100. Therefore a constraint is imposed, typically $\boldsymbol{\alpha}^T\boldsymbol{\alpha} = 1$.

Thus the aim is to find the $\boldsymbol{\alpha}$ that maximises

$$\mathrm{Var}(z) = \mathrm{Var}(\boldsymbol{\alpha}^T\boldsymbol{x}) = \boldsymbol{\alpha}^T(\mathrm{Var}(\boldsymbol{x}))\boldsymbol{\alpha} = \boldsymbol{\alpha}^T\mathbf{S}\boldsymbol{\alpha}$$

where $\mathbf{S}$ is the covariance matrix of $\boldsymbol{x}$, subject to $\boldsymbol{\alpha}^T\boldsymbol{\alpha} = 1$.

A Lagrange multiplier λ is used:

$$Q = \boldsymbol{\alpha}^T\mathbf{S}\boldsymbol{\alpha} - \lambda(\boldsymbol{\alpha}^T\boldsymbol{\alpha} - 1)$$

$$\frac{\partial Q}{\partial \alpha} = 2\mathbf{S}\boldsymbol{\alpha} - 2\lambda\boldsymbol{\alpha}, \text{ since } \mathbf{S} \text{ is symmetric.}$$

Setting this equal to $\mathbf{0}$ gives $\mathbf{S}\boldsymbol{\alpha} = \lambda\boldsymbol{\alpha}$.

This is an eigenequation, so that λ is an eigenvalue of $\mathbf{S}$, and $\boldsymbol{\alpha}$ is an eigenvector of $\mathbf{S}$.

The question of which eigenvalue to choose is answered by multiplying on the left by $\boldsymbol{\alpha}^T$.

$$\boldsymbol{\alpha}^T\mathbf{S}\boldsymbol{\alpha} = \boldsymbol{\alpha}^T\lambda\boldsymbol{\alpha} = \lambda, \text{ since } \boldsymbol{\alpha}^T\boldsymbol{\alpha} = 1.$$

Therefore the variance of z is maximised by choosing $\mathrm{Var}(z_1) = \lambda_1$, the largest eigenvalue. The first principal component of $\boldsymbol{x}$ is $\boldsymbol{\alpha}_1$, the corresponding eigenvector, with first principal component score $z_1 = \boldsymbol{\alpha}_1^T\boldsymbol{x}$.

The second principal component is chosen as the vector $\boldsymbol{\alpha}$ which maximises $\mathrm{Var}(z)$ where $z = \boldsymbol{\alpha}^T\boldsymbol{x}$, again subject to $\boldsymbol{\alpha}^T\boldsymbol{\alpha} = 1$. Additionally, however, the second principal component is chosen to be orthogonal to the first, so $\boldsymbol{\alpha}^T\boldsymbol{\alpha}_1 = 0$.

Again Lagrange multipliers are employed:

$$Q = \boldsymbol{\alpha}^T\mathbf{S}\boldsymbol{\alpha} - \lambda(\boldsymbol{\alpha}^T\boldsymbol{\alpha} - 1) - \mu(\boldsymbol{\alpha}^T\boldsymbol{\alpha}_1)$$

$$\frac{\partial Q}{\partial \alpha} = 2\mathbf{S}\boldsymbol{\alpha} - 2\lambda\boldsymbol{\alpha} - \mu\boldsymbol{\alpha}_1, \text{ since } \mathbf{S} \text{ is symmetric.} \quad (*)$$

Setting this equation equal to 0 and premultiplying by $\boldsymbol{\alpha}_1^T$ gives

$$2\boldsymbol{\alpha}_1^T\mathbf{S}\boldsymbol{\alpha} - 2\lambda\boldsymbol{\alpha}_1^T\boldsymbol{\alpha} - \mu\boldsymbol{\alpha}_1^T\boldsymbol{\alpha}_1 = 0$$

$$2\boldsymbol{\alpha}_1^T\mathbf{S}\boldsymbol{\alpha} - \mu = 0, \text{ since } \boldsymbol{\alpha}_1^T\boldsymbol{\alpha} = 0 \text{ and } \boldsymbol{\alpha}_1^T\boldsymbol{\alpha}_1 = 1$$

Recall $\mathbf{S}\boldsymbol{\alpha}_1 = \lambda_1\boldsymbol{\alpha}_1 \Rightarrow \boldsymbol{\alpha}_1^T\mathbf{S} = \lambda_1\boldsymbol{\alpha}_1^T$. Substituting into the equation above:

$$2\lambda_1\boldsymbol{\alpha}_1^T\boldsymbol{\alpha} - \mu = 0$$

Since $\boldsymbol{\alpha}_1^T\boldsymbol{\alpha} = 0$, $\mu = 0$. Therefore from (*):

$$2\mathbf{S}\boldsymbol{\alpha} - 2\lambda\boldsymbol{\alpha} = \mathbf{0}$$

$$\mathbf{S}\boldsymbol{\alpha} = \lambda\boldsymbol{\alpha}$$

The variance of z is maximised by choosing $\mathrm{Var}(z_2) = \lambda_2$, the second largest eigenvalue. The second principal component of $\boldsymbol{x}$ is $\boldsymbol{\alpha}_2$, the corresponding eigenvector, with second principal component score $z_2 = \boldsymbol{\alpha}_2^T\boldsymbol{x}$.

A similar argument for all subsequent principal components yields the kth principal component $\boldsymbol{\alpha}_k$ to be the kth eigenvector, where $z_k = \boldsymbol{\alpha}_k^T\mathbf{x}$, with $\mathrm{Var}(z_k) = \lambda_k$ the kth largest eigenvalue.

The eigenvalues are crucial in determining the appropriate number of principal components to retain. λ_k is the variance of the kth principal component. Therefore the ratio $\sum_{k=1}^{m}\lambda_k \Big/ \sum_{k=1}^{p}\lambda_k$ describes the proportion of the total variation in the sample accounted for by the first m principal components. A 'scree' plot is often used to visualise the proportion of variance accounted for by each component; λ_k is plotted against k and the change from 'large' to 'small' eigenvalues noted.

The above theory is formulated in terms of **S**, the covariance matrix. However, in practice the covariance matrix is only suitable if all the variables are measured in the same units and have a similar magnitude[141]. Otherwise the correlation matrix is more appropriate. This is equivalent to using the covariance matrix once the original variables have been standardised to unit variance and zero mean. The standardisation to zero mean is not actually necessary, but results in principal components (PCs) that themselves all have zero mean. In the case of food data, intakes are very disparate across food categories: for example white bread is on average consumed far more frequently than dried fruit. Thus the correlation matrix is most appropriate, and is used in all analyses presented.

2.4.2.2 Application

Principal component analysis generates a matrix of coefficients ($\alpha_1, \alpha_2, \ldots, \alpha_p$), each column relating to one principal component. For each participant's diet, multiplying that individual's food intake by the first column of coefficients generates the first principal component score. Multiplying the participant's food intake by the n^{th} column of coefficients generates the n^{th} component score for that individual.

The first few principal components are retained based on the scree plot, and the utility of the outcome. It is hoped that the derived components will result in coefficients relating to a recognisable dietary pattern. For example the prudent dietary pattern referred to in Section 1.3.1 has high coefficients for fruit, vegetables, whole grains, fish and poultry, and low coefficients for high-sugar foods. Thus a subject consuming prudent foods frequently and imprudent foods infrequently will have a high score on the prudent diet scale derived using PCA.

A criticism of PCA is the degree of subjectivity involved in the decisions during analysis. The first decision is the selection of foods to be considered and how similar foods might be grouped, followed by the number of components, to be extracted. After the analysis is completed the labelling of the dietary patterns is also subjective. Despite these limitations, PCA is a very useful tool in dietary analyses. The impact of different food groupings and numbers of components extracted is assessed in the analyses of dietary data from the PAH study (Chapter 4).

2.4.3 Factor analysis

Factor analysis is an alternative dimension-reduction technique to PCA. Although similar in execution, the philosophy behind each is different. Factor analysis is based on the assumption that underlying the variables observed $x_1, x_2, \ldots, x_p$ are a smaller number of unobserved variables $f_1, f_2, \ldots, f_m$ known as 'common factors'. The observed variables are linear combinations of the common factors, plus a unique factor. These unique factors are assumed to be independent of each other.

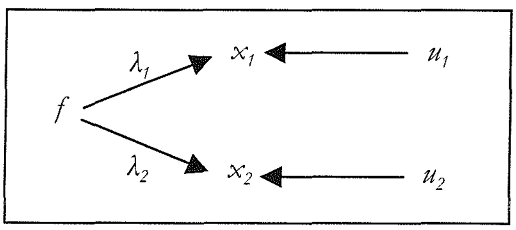
The distinguishing feature of factor analysis is the assumption that correlation between observed variables is due to the common factors. In the context of dietary data, factor analysis is used to discover whether there are common factors that explain the correlation in the food variables measured.

Factor analysis has two forms: confirmatory and exploratory. Confirmatory factor analysis begins with some prior knowledge of the relationships between the variables that are included; such a structure can be estimated and tested. However, in the analysis of the PAH and SWS dietary data no such knowledge exists, and instead exploratory factor analysis is used, where no preconceived ideas about the structure of the data are required. Exploratory factor analysis is henceforth simply referred to as 'factor analysis'.

2.4.3.1 Theory

A path causal diagram can be helpful in illustrating simple factor models, as utilised by Kim and Mueller[142]. A two-variable, one common factor model is illustrated in Figure 6.

Figure 6 Path causal diagram for a two-variable, one common factor model



Adapted from Kim and Mueller[142]

The model described by this diagram indicates that x_1 is a linear combination of f and u_1, with f weighted by λ_1. Similarly x_2 is a linear combination of f and u_2, with f weighted by λ_2. Since f is common to both x_1 and x_2 it is known as a common factor. λ_1 and λ_2 are termed 'factor loadings'. u_1 and u_2 are unique to each observed variable, and are therefore known as 'unique factors'. In algebraic form the model may be specified using the following three equations:

$$x_1 = \lambda_1 f + u_1$$

$$x_2 = \lambda_2 f + u_2$$

$$\mathrm{cov}(f, u_1) = \mathrm{cov}(f, u_2) = \mathrm{cov}(u_1, u_2) = 0$$

In more general terms, for p observed variables and m common factors, a factor analytic model can be described by:

$$x_j = \sum_{k=1}^{m} \lambda_{jk} f_k + u_j$$

$$\mathrm{cov}(f_k, u_j) = 0 \text{ for all } j,k$$

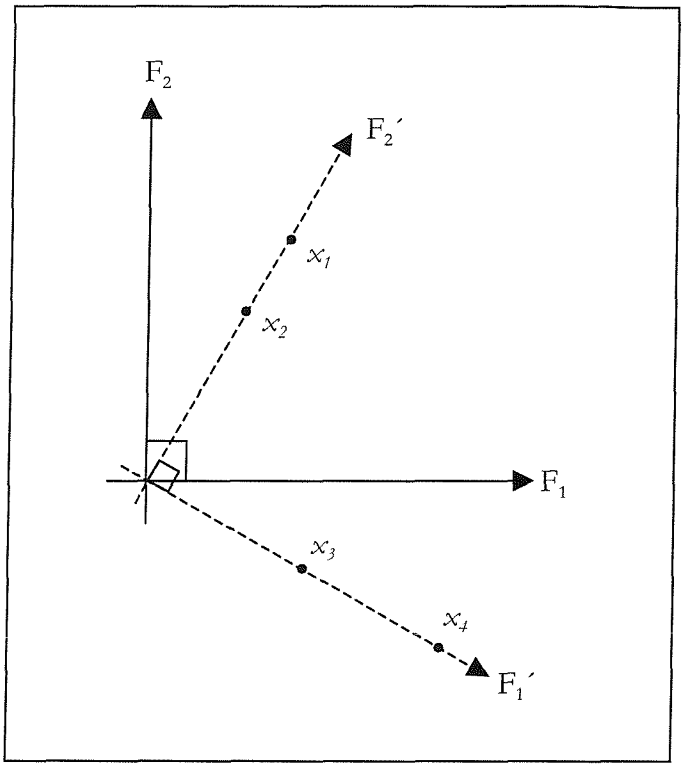
$$\mathrm{cov}(u_j, u_k) = 0 \text{ for all } j \neq k$$

An additional concept in factor analysis is the proportion of the variation in x_j that is common to the other variables, through the common factors. This is known as the 'communality' of x_j. If $\mathrm{var}(u_j) = \omega_j$, then the communality of x_j is $1 - \omega_j$.

The aim of factor analysis is to find loadings λ_{jk} and communalities ω_j which give a good fit to the model specified above. Any such solution is not unique, since that solution will be just as valid if rotated. This property can be used to the advantage of the researcher, by enabling the 'simplest' loadings to be chosen.

The idea of rotation to 'simplest' loadings is illustrated in Figure 7. Here the two co-ordinates for each variable x represent the loadings on two factors. The unique factors are all zero for the purposes of illustration. The initial solution with loadings on F_1 and F_2 is improved by rotating the reference axes to F_1', F_2', equivalent to multiplying the initial solution by a rotation matrix. The loadings on F_1', F_2' are easier to interpret, and in general the 'simplest' solution amongst all equally valid solutions is the one where as many as possible of the loadings are close to 0, -1 or +1.

Figure 7 Illustration of rotation to 'simple' factor loadings



Adapted from Kim and Mueller[142]

Factor analysis thus consists of three steps: extracting the initial factors, deciding upon the number of factors to retain, and rotation to a terminal solution. Additional choices are therefore necessary to PCA: the choice of extraction method, and the rotation technique to be employed.

There are several methods of extracting the initial factors. Principal component analysis is the method most commonly used in the dietary literature. Joliffe and Morgan[141] note that using PCs as initial factors frequently gives similar results to other extraction methods, such as those described below. When PCA is used as an initial extraction method in factor analysis, the technique is sometimes known as 'principal component factors' (PCF), and this terminology is used in this thesis for clarity.

A simple modification to the use of PCF as initial estimates is known as 'principal factor analysis' (PFA). Here the unit diagonal values of the correlation matrix are replaced by estimates of the communalities of each variable. In Stata the estimates of communalities are the squared multiple correlations of each variable with all other variables. The remainder of the PCF procedure is then performed on this matrix. Iterated principal factors (IPF) extends the idea of principal factor analysis further, by replacing the diagonal values of the correlation matrix with the new communalities found, and iterating to a stable solution. Maximum likelihood (ML) estimation determines the loadings on each variable that have the greatest likelihood of producing the observed correlation matrix.

Once the initial extraction method has been chosen, the number of factors retained is an important consideration because, unlike in PCA, a change from m to $(m + 1)$ factors can completely change the nature of each factor. Joliffe and Morgan[141] consider the choice of the number of factors to be more important than the selection of rotation method in determining the factors. The decision is affected by several criteria, and consideration of all relevant issues is advised[143]. Inspection of the scree plot (see Section 2.4.2) can be used to decide which initial subset of eigenvalues are substantially larger than those remaining. Consideration of the communalities is also important, and the interpretability of the solution is an essential criterion.

Once the first m factors have been chosen factor analysis permits rotation of these results. Numerous methods of rotation are available. Techniques are either orthogonal, as illustrated in Figure 7, or oblique where the resulting factors are no longer independent. Probably the most popular technique is known as 'varimax', an orthogonal rotation where the sum over factors of the within-factor variances of the squared loadings is maximised. This has the effect of forcing the loading towards –1, 0 or 1. Varimax retains the original assumption that the factors are orthogonal and have unit variance. Joliffe and Morgan[141] have found that for fixed m and initial solution, different rotation criteria lead to similar final solutions. This was confirmed using the PAH dietary data, and therefore only the varimax rotation technique is considered in this thesis.

2.4.3.2 Scoring

The loadings in factor analysis λ_{jk} are ostensibly analogous to coefficients in PCA α_{jk}, and indeed have a similar interpretation. Both factor loadings and PCA coefficients measure the relationship between a variable and a factor or component. However, in PCA the components are a linear combination of the original variables, whereas in factor analysis the original variables are linear functions of the factors, plus a unique factor. Therefore factors in factor analysis cannot be expressed as linear functions of the original variables in the same straightforward way that principal component scores can be generated. Rather factor analysis scores must be estimated. Two methods are available to estimate factor scores within Stata: Bartlett's and the regression method. Bartlett's method is unbiased, but regression scores may be more precise[140]. Kim and Mueller[143] suggest there is usually a very high correlation between the regression and Bartlett's methods and thus the choice between the two is often academic. This was confirmed using the PAH dietary data, and only the regression method is used here.

2.4.4 Fruit and vegetables

As described in Section 1.2.1 the consumption of five portions of fruit and vegetables per day is an important government health message. There is strong epidemiological evidence for the protective effect of fruit and vegetable consumption in studies of important diseases such as cancer, and coronary heart disease[144]. High fruit and vegetable intake is recognised by the general public as being associated with a healthy diet[145], and is a simple measure to calculate and understand. For these reasons fruit and vegetable consumption has been frequently used as a measure of diet in epidemiological research[90,91,97,146]. McCann et al.[147] found that fruit and vegetable intake performed slightly better than either of their first two principal components in predicting endometrial cancer cases. It is considered in this thesis as an alternative measure of one aspect of diet since it provides a simple indication of women who endeavour to eat in accordance with public health advice.

Williams[52] clarifies, and provides a rationale for, the calculation of fruit and vegetable intake. In particular baked beans and other pulses, dried fruit, frozen or tinned fruit and vegetables, and fruit and vegetables used as a main ingredient in recipes or composite foods are included. Nuts, potatoes and fruit squashes and cordials are excluded. It is advised that fruit juice should only count once per day towards the 'at least five a day' recommendation, since fruit juice provides most of the vitamins and minerals of fresh fruit, but the majority of fibre is lost. Unfortunately the Southampton MRC FFQ collects information about fruit juice consumption across three categories, two of which include fresh fruit: (44) Fresh oranges and orange juice, and (45) Grapefruit and grapefruit juice. It is therefore impossible to distinguish adequately between fresh fruit and fruit juice consumption, and thus every portion of fruit juice contributed to a fruit and vegetable score.

Williams also provides quantification of amounts that should constitute a 'portion' of commonly eaten fruit and vegetables. For example, one large slice of melon, one banana or two serving spoonfuls of carrots are all considered one 'portion'. However, the Southampton MRC FFQ only provides information on frequency and not on quantity. Therefore to calculate fruit and vegetable intake it is necessary to assume that each time a subject consumes a particular fruit or vegetable it is a 'portion' size on average.

In the light of the above considerations, fruit and vegetable intake is calculated in this thesis by summing frequency of intake of FFQ categories (13), (22) – (38), (40) – (45), (47) and (51) – (54). Note that examination of diary data from the PAH study revealed that of the 111 times that an item mapping to FFQ category (46) (Blackcurrants and blackcurrant drinks) was consumed, 108 were blackcurrant drinks, and only 3 were blackcurrant fruit. Therefore this category is best classified as a soft drink.

The SWS included two additional fruit and vegetable 'cross-check' questions. The participant was asked about the number of servings eaten in the past week of (i) vegetables and vegetable-containing dishes (excluding potatoes), and (ii) fruit and pure fruit juice. These questions endeavour to provide an assessment of over-estimation from the FFQ (Section 2.3.2), whilst acknowledging that servings in the past week might not necessarily be representative of average frequency over the past three months. The sum of these two 'cross-check' questions is used as an additional assessment of fruit and vegetable consumption in the SWS analysis.

3 Sociodemographic and lifestyle measures

An understanding of sociodemographic differentials in diet may provide insights into sociodemographic inequalities in health. It is therefore necessary to define sociodemographic measures with care. A discussion of sociodemographic measures encompasses both sociological and demographic characteristics. The sociological characteristics relevant here are measures of deprivation and social position. Both area-based and individual-based assessments of deprivation are discussed in Section 3.1, and the use of various measures of social position is considered in Section 3.2. Occupation-based measurement of women's social position is often problematic because relatively few women work full-time; therefore Section 3.3 examines alternative occupation-based measures of women's social position.

Demographic and lifestyle predictors are discussed in Section 3.4, and the representativeness of the PAH and SWS samples is assessed in Section 3.5. A summary of all sociodemographic and lifestyle measures available in the PAH study and the SWS is provided in Section 3.6.

3.1 Deprivation

3.1.1 What is deprivation?

Deprivation takes many forms, and has been defined in a variety of ways. Those who are deprived are in a general sense thought to be living without life's necessities. However, the definition of what is 'necessary' is not easy. A distinction is sometimes made between absolute and relative deprivation. In an early study on deprivation, B.S. Rowntree[148], working at the turn of the century, defined absolute deprivation (which he refers to as 'primary poverty') to be an income insufficient 'to provide the minimum of food, clothing and shelter needful for the maintenance of merely physical health'. He graphically illustrates the lifestyle of a family who would be considered to be in primary poverty:

> A family must... never spend a penny on railway fare or omnibus. They must never go into the country unless they walk. They must never purchase a halfpenny newspaper or spend a penny to buy a ticket for a popular concert. They must write no letters to absent children, for they cannot afford to pay the postage... The mother must never buy any pretty clothes for herself or for her children, the character of the family wardrobe, as for the family diet being governed by the regulation "Nothing must be bought but that which is absolutely

> necessary for the maintenance of physical health, and what is bought must be of the plainest and most economical description."

Such a stringent definition is nevertheless ambiguous. Poor health results from living standards far less desperate than those described above. Thus acceptable levels of susceptibility to disease and life expectancy cannot be defined in an absolute sense, but are determined by the norms of the society within which a person lives. Moreover, the standards of many societies go beyond physical needs to social and emotional needs. An illustration is given by Mack and Lansley[149] in a report on the London Weekend Television survey of lives of the poor in Britain in the 1980s. Mary is a single parent with a 5-year old son:

> "Now he is at school and tells me about other children's bikes, and the toys they take, and holidays, and days out with parents, and it breaks my heart for there is nothing for him; if he has food and clothes he can have nothing else".

Since Mary can afford food and clothes for her son, she might not be described as 'absolutely deprived', but most would consider Mary to be deprived by the standards of British society in the 1980s.

Although absolute deprivation certainly exists, an understanding of deprivation can only be gained relative to the society that a person is living in. Townsend[150] defines deprivation to be 'a state of observable and demonstrable disadvantage relative to the local community or the wider society or nation to which an individual, family or group belongs'. However, such a relative definition of deprivation has been criticised since it does not allow comparisons of deprivation across time or between communities.

Deprivation has many facets, and is certainly more complex than a simple measure of income. Townsend et al.[151] make a helpful distinction between material and social deprivation. Material deprivation 'entails the lack of the goods, services, resources, amenities and physical environment which are customary, or at least widely approved in the society under consideration'. Social deprivation is 'non-participation in the roles, relationships, customs, functions, rights and responsibilities implied by membership of a society and its sub-groups'.

Within the PAH study and the SWS no direct measure of deprivation was made. However, area-based measures of deprivation can be used. Also, receipt of benefits, housing tenure and housing conditions can be used as measures of components of deprivation. These are discussed in turn.

3.1.2 Area-based measures of deprivation

The increasingly recognised need to direct resources for infrastructure and healthcare to deprived areas has lead to the development of several area-based measures of deprivation. The areas for which relevant statistics are generally available are those used for census administration: wards or enumeration districts (EDs). Wards encompassed an average of 2,000 households in the 1991 census. EDs are the smallest unit for which census statistics are available in England and Wales, with an average of 200 households in the 1991 census. Within the city of Southampton there are 15 wards and 417 EDs.

Area-based scores developed to measure deprivation include the Townsend Material Deprivation score[151], the Jarman Underprivileged Area score[152,153], the Carstairs and Morris Scottish Deprivation score[154,155,156] and the Department of the Environment, Transport and the Regions (DETR) Index of Deprivation 2000[157]. Such measures can be used in individual-based research by assigning a score to each participant according to their area of residence. Since the city of Southampton comprises only 15 wards, a ward-level measure would be a relatively crude discriminator between participants. The DETR Index of Deprivation is such a measure, and thus is not a candidate for use in this thesis.

Morris and Carstairs[158] compare the Townsend, Jarman, and Carstairs scores, along with the 1983 Department of Environment score, a predecessor of the DETR Index of Deprivation 2000. Their work involved the correlation of a range of area health measures with the four deprivation scores, and concluded that the Carstairs and Townsend scores are found to explain most variation. These scores are very similar; both include information on unemployment, car ownership and overcrowding. The Townsend score also incorporates a measure of home ownership, whereas the Carstairs score includes information on the social class of residents within an area. The Carstairs score was developed on Scottish data where the use of home ownership data is less appropriate since a much higher proportion of housing in Scotland is in the public sector.

Jarman, Carstairs and Townsend[159] themselves argue that their three indices are highly correlated and 'more effort should be put into using the indices for the purposes for which they were developed and less into dissecting out the fairly small differences among them'.

Therefore, since the Carstairs score is a Scottish-based score, whereas the Townsend score was developed for use throughout Britain, the Townsend score was chosen as an area-based measure of deprivation in this thesis.

The Townsend score constitutes the following four variables:

- Unemployment — % of unemployed economically active residents aged 16-59/64
- Car ownership — % of private households who do not possess a car
- Home ownership — % of private households not owner-occupied
- Overcrowding — % of private households with more than one person per room

Reasons are given for the inclusion of each of the variables. Unemployment reflects the lack of income and facilities of employment, as well as a lack of material resources, and the insecurity to which this gives rise. Not possessing a car is considered a surrogate for current income, and non-owner occupation reflects lack of wealth as well as income. Overcrowding is included as a more general measure of living circumstances and housing conditions.

The Townsend index is calculated by log transforming the unemployment and overcrowding variables to reduce their skew. The scores of each variable are then standardised and summed to produce the final Townsend score. Therefore a high Townsend score indicates a high level of deprivation.

Area-based measures of deprivation such as the Townsend score are useful because they measure neighbourhood attributes which may contribute to deprivation but which cannot be captured by the participant's characteristics alone. A further important advantage of area-based measures is that they can be assigned to all participants who provide a postcode. Therefore the difficulty of assigning occupational-based measures of deprivation to those who don't work is avoided. The Townsend score has specifically been developed to measure material deprivation, as distinct from social deprivation (see Section 3.1.1), and therefore has a defined theoretical basis. In an analysis of wards in the north of England, Townsend et al.[151] found that their deprivation index explained more variance in an area-based measure of overall health than was explained by social class.

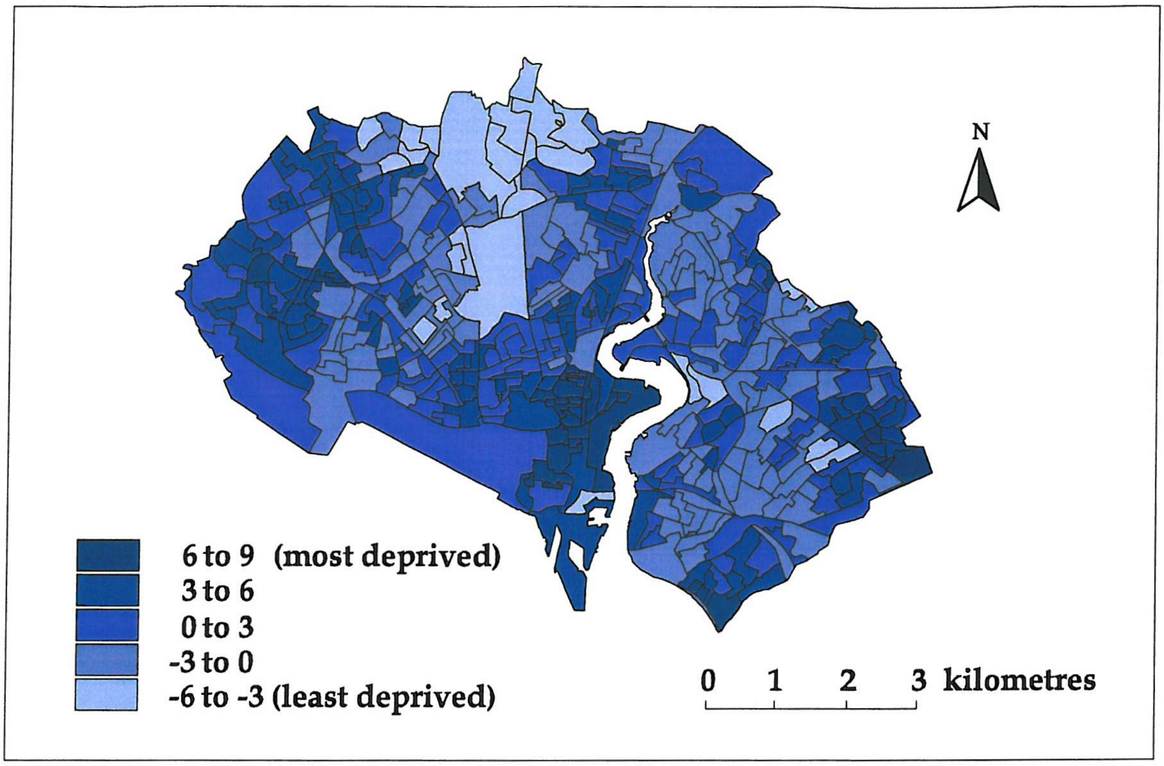
Area-based measures are often criticised since not all deprived people live in deprived areas. Equally, not everyone living in an area defined as deprived is deprived. This is known as the 'ecological fallacy'. However, Carstairs and Morris[154] point out that social class might be considered to suffer from the same difficulty. Social classes contain a diversity of occupational groups, but homogeneity of social position is assumed within each social class in the same way that homogeneity of deprivation is assumed within each geographical area for an area-based measure.

A further criticism of area-based measures of deprivation is that boundaries have often been drawn to define an appropriate number of voters, rather than to identify a 'natural' community. For example, within Southampton the relatively wealthy area of Bassett is immediately adjacent to the comparatively deprived area of the 'Flower's Estate'. One ED straddles the boundary of the two areas and has a Townsend score of 1.25, a figure that poorly reflects levels of deprivation in either area. For comparison, the adjacent ED to the west is wholly contained within the wealthier area of Bassett and has a Townsend score of –2.98, whereas the adjacent ED to the east lies entirely within the 'Flower's Estate' and has a Townsend score of 4.49.

The Townsend score itself has been criticised. In particular greater overcrowding in an area leads to a higher Townsend score, but in some communities it is thought that this phenomenon may in fact indicate greater levels of social support. Also, the inclusion of car ownership within the score is considered imprudent; in rural areas car ownership can be a necessity, rather than a luxury[160]. In such a situation the expense of car ownership might in fact lead to greater deprivation. For this reason, the use of the Townsend score is seen by some to be most appropriate when characterising urban areas. The PAH study recruited women referred to the PAH maternity hospital, the great majority of whom were from urban Southampton. The SWS recruited women from an entirely urban area. The geographical distribution of Townsend scores for EDs within the city of Southampton is illustrated in Figure 8.

531 (86%) of the 620 women completing PAH early pregnancy interviews have provided a valid postcode, as have 6,067 (99%) of the 6,129 women in the SWS analysis sample. The proportion of women providing a postcode in the SWS is higher because for most women who did not provide a postcode during the interview it was possible to extract their postcode from the information provided by their GP. Files on the Manchester Information and Associated Services (MIMAS) system (source: The 1991 Census, Crown Copyright. ESRC purchase) have been used to map each postcode to an ED and subsequently to a Townsend score.

Figure 8 Townsend scores of Southampton enumeration districts



3.1.3 Benefits

The SWS asked each participant whether she or her husband/partner were receiving income support, jobseeker's allowance, working families' tax credit, or housing benefit. The responses to these questions can be seen as a measure of financial deprivation and possibly the social deprivation resulting from unemployment, as described below.

3.1.3.1 Income Support

Income Support can be claimed by adults on a low income with savings of less than £8,000, who do not work, or are working less than 16 hours per week. It therefore reflects a level of financial deprivation, as well as the possible social deprivation of unemployment or limited employment.

3.1.3.2 Jobseeker's Allowance

Jobseeker's Allowance (JSA) is paid to people of working age who are capable of working, available for work and actively seeking work. JSA can be income-based or contribution-based. Income-based JSA is available to the unemployed with savings of less than £8,000 and a low income. It therefore indicates financial deprivation as well as the social deprivation of unemployment. Contribution-based JSA is available if a minimum number of National Insurance contributions have been paid. It is not affected by savings or income, and is paid for up to 6 months. Contribution-based JSA does not indicate financial deprivation, but does reflect the social disadvantages associated with unemployment.

3.1.3.3 Working Families' Tax Credit

Working Families' Tax Credit replaced Family credit in October 1999. Parents who work 16 hours a week or more and are bringing up children can claim Working Families' Tax Credit. Recipients must also have savings of less than £8,000 and may therefore be experiencing a level of financial deprivation.

3.1.3.4 Housing Benefit

Housing Benefit is available to those who pay rent and have a low income, with savings of less than £16,000. Since recipients have a limited income and savings, and do not own their own residence, they may be suffering some financial deprivation.

In summary, receipt of benefits is largely a measure of financial deprivation, and is therefore a somewhat limited assessment of the many facets of deprivation. However, receipt of benefits is information that can easily be collected, although it might be argued that such data are not reported as reliably as some other variables, due to perceived social stigma. Table 2 displays proportions in the SWS receiving each benefit.

Table 2 Benefits received by SWS participants or their partners

Receipt of benefit	Income support	Jobseeker's allowance	Working families' tax credit	Housing benefit
No	5461 (89%)	5980 (98%)	5771 (94%)	5310 (87%)
Yes	642 (10%)	122 (2%)	335 (5%)	792 (13%)
Missing	26 (0.4%)	27 (0.4%)	23 (0.4%)	25 (0.4%)
Total	6129 (100%)	6129 (100%)	6129 (100%)	6129 (100%)

Overall 1,177 (19%) of the 6,118 SWS participants who answered the benefit questions reported receiving at least one of these benefits. Questions about benefits were not included in the PAH study.

3.1.4 Housing tenure and conditions

3.1.4.1 Housing tenure

Housing tenure is commonly used in research as a measure of social circumstances since it is easy to collect and available for all members of the population, regardless of household composition. Typically participants are categorised as living in households which are owner-occupied, privately rented, or socially rented (from the local authority or housing association). In the UK, Right to Buy sales have led to an increase in the proportion of owner-occupied housing and proportionate decrease in socially rented accommodation.

Arber[161] describes the privately rented sector as 'increasingly becoming an intermediate sector, occupied at transitional life-cycle stages', and the council sector as 'increasingly becoming a residual sector for those whose structural circumstances make it impossible for them to be owner-occupiers'. In her analysis of the 1981-82 General Household Survey, Arber found that local authority tenants report limiting long-standing illness more frequently than owner-occupiers, with those in privately rented accommodation occupying an intermediate position. However, the health status of younger women (aged 20-29) in privately rented accommodation was much closer to that of owner-occupiers than local authority tenants. Young women in Southampton living in local authority housing might therefore be considered to be experiencing a level of deprivation greater than those in owner-occupied or privately rented accommodation.

Macintyre et al.[162] attempt to explain the reasons for the well-known associations between housing tenure and health[161,163]. The physical features of the home and area in which it is located are seen to be direct links between tenure and health. Such characteristics include damp, temperature, ventilation, crowding, access to amenities and stressors in the dwelling or its environs. Housing tenure is thus a measure of material and social deprivation. The authors consider that 'housing and area influences on health seem likely to be more proximate and direct… than the more generalized effects of income and social class'.

Responses to the SWS question about housing tenure are shown in Table 3. To simplify the categories some responses to the 'other' category were combined into living with parents, partner or friend. A question about housing tenure was not included in the PAH study.

Table 3 Housing tenure responses from the SWS

Tenure		
Own outright or buying with mortgage	2638	(43%)
Rent from private landlord	1357	(22%)
Rent from council or housing association	1263	(21%)
Lives with parents, partner or friend	821	(13%)
Missing	50	(1%)
Total	6129	(100%)

3.1.4.2 Housing conditions

Housing conditions were assessed in the SWS by asking participants whether they have a big problem, small problem or no problem with condensation, rising or penetrating damp, difficulty in keeping home warm, leaking roof, rot in window frames, timbers or floorboards, or not enough space (Appendix 3). Like housing tenure, this information is easy to collect and potentially available for all members of the population, regardless of household composition. It is a direct measure of housing-related material deprivation.

Questions about housing conditions were not included in the PAH study. Responses to the SWS question about housing conditions are shown in Table 4. Overall 4,100 (67%) of women had at least one small or big problem. 1,766 (29%) of participants had at least one big problem. Since deprivation is considered to be a lack of customary material or social resources (Section 3.1.1), 'having at least one big housing problem' is used as the most appropriate summary measure of housing conditions deprivation.

Table 4 Housing conditions responses from the SWS

	Big problem	**Small problem**	**Not a problem**	**Missing**
Condensation	500 (8%)	1474 (24%)	4147 (68%)	8 (0.1%)
Rising or penetrating damp	360 (6%)	633 (10%)	5128 (84%)	8 (0.1%)
Difficulty in keeping home warm	664 (11%)	913 (15%)	4544 (74%)	8 (0.1%)
Leaking roof	131 (2%)	243 (4%)	5747 (94%)	8 (0.1%)
Rot in window frames etc.	251 (4%)	733 (12%)	5135 (84%)	10 (0.2%)
Not enough space	836 (14%)	1283 (21%)	4001 (65%)	9 (0.1%)

3.2 Social position

3.2.1 What is social position?

Egalitarian society remains a dream, and all human societies have some form of social inequality. This is realised through the unequal distribution of power and prestige between individuals and social groups. In British society, as in many others, there are also marked differences in the division of wealth. An individual's social position is a measure of their location within the social structure, however that structure is perceived.

Social position is traditionally considered as a hierarchy of groups, within which individuals share similar lifestyles and common interests. Social Class based on occupation (SC) is the conventional measure of social position in British research and subscribes to the concept of a hierarchy of social groups[164].

An alternative sociological standpoint contends that a more appropriate model of British society consists of social groups, but that these groups are not hierarchical. The theoretical strengths of such a view have encouraged the ONS to recently produce a new measure of social position, the National Statistics Socio-Economic Classification (NS-SEC)[165].

A further idea proposed by sociologists is that models of social classes are no longer appropriate in Western industrial society. They suggest that social strata have been replaced by a continuous hierarchy of social positions. In such a model there is a continuum of occupational statuses representing a hierarchy of individuals. The Cambridge scale[166] is a measure of social position within a social structure thus perceived.

Bartley et al.[167] suggest that social stratification has three important dimensions: social status or prestige, material resources and employment relations and conditions. Although social class has an inadequate conceptual basis it is most akin to a measure of social status or prestige. The Cambridge scale is a more theoretically robust gauge of social status or prestige, while the NS-SEC assesses employment relations and conditions. Social stratification by material resources is considered in this thesis through measures of deprivation (Section 3.1).

Social class, NS-SEC and Cambridge scores are all based on occupational information. This information was collected for the woman, her partner and her father (at the time of her birth) in both the PAH study and the SWS. The job descriptions were coded using the Office for

National Statistics (ONS) Standard Occupational Coding[164] (SOC) into Occupational Unit Groups (OUGs) and these OUGs mapped to a class under the SC or NS-SEC, or to a Cambridge score.

Note that the participant's and her partner's occupation in the SWS is coded by current full-time occupation, or if that is unavailable, by last occupation, or if neither of these are available, by current part-time occupation. It is questionable whether coding by last occupation is superior to current part-time occupation, especially since it is possible that the participant's last occupation may have itself been part-time. However, partners are likely to have had a previous full-time occupation, and in these relatively modern cohorts most women themselves will have had a previous full-time job at some point. Since occupational downgrading to part-time jobs is common, especially amongst women after childbirth, it is reasonable to use part-time jobs as an indicator of occupation only if no other information is available. Only responses about full-time work or last job were collected in the PAH study.

Conventionally where full-time students are included in occupation-based analyses of social position they are given their position of origin[165]. In this thesis partner's social position is also considered (Section 3.3). However, within the SWS analysis sample, 405 (7%) women are full-time students and are not living with a partner. For these women their father's occupation at their birth is used to assign their social position. In the PAH study no information was collected on whether the woman was studying full-time. Nevertheless, participants within this pregnant cohort are relatively unlikely to be students, and more likely to be living with a partner, and thus it would probably have rarely been necessary to substitute the women's fathers' social position.

The occupation-based measures of social position used in this thesis are discussed in Sections 3.2.2 - 3.2.4. Rose and O'Reilly[165] note that 'occupation-based classifications will continue to remain vital tools for scientific and policy analyses for the foreseeable future. This is because, pragmatically, they are based on routinely and widely-collected data and, theoretically, because it remains the case that a person's employment situation is a key determinant of life chances'. However, education is a significant non-occupation-based measure of social position, and is discussed in Section 3.2.5.

An alternative ONS classification, Socio-Economic Group (SEG), is not considered since it is conceptually similar to the NS-SEC in terms of measuring employment relations and conditions, and has been superseded by it. The NS-SEC is also designed to have replaced SC, but the differing conceptual bases, wide recognition and ordinal nature of SC merit its inclusion here.

3.2.2 Social Class based on occupation

'Social grades' were first introduced in 1911 in the Registrar General's Annual Report on the census taken that year. THC Stevenson, a medical statistician, was responsible for devising the system as an eight-fold classification of both occupational and industrial groups. In 1921 there was a major revision into roughly the form of the social class scheme that exists today. The main change subsequently has been the division of Class III into manual and non-manual classes. Other more minor changes have often involved the allocation of occupations to different classes. Between 1921 and 1971 the scale was officially described as an ordinal classification of occupations according to their reputed 'standing in the community'. In 1980 social class was equated more with occupational skill. In 1990 this new emphasis became more explicit, so that the classification of OUGs was directly related to the competence required to perform in an occupation, and the scheme was renamed 'Social Class based on occupation'. The categories now in use are listed in Table 5.

Table 5 Social Class based on occupation

I	Professional, etc occupations
II	Managerial and technical occupations
IIIN	Skilled occupations (non-manual)
IIIM	Skilled occupations (manual)
IV	Partly skilled occupations
V	Unskilled occupations

The SC scheme has been heavily criticised as a measure of social position. A major objection is that occupations are assigned to social classes 'on the basis of judgements made by ONS staff and various other experts whom they consult, and not in accordance with any coherent body of social theory'[168]. It is commonly agreed that the scale lacks a valid conceptual basis.

In their vehement criticism of the use of SC, Jones and Cameron[169] argue that THC Stevenson allocated occupations to his social grades to achieve smooth class gradients in mortality. Thus differentials in health that are commonly cited as evidence for the value of SC are actually a direct result of the way in which the classes themselves were generated.

Other criticisms of SC include the lack of homogeneity within social classes, and the fact that changes in allocation over time mean that comparisons with historical data are crude. Importantly for the classifications of women's occupations in this thesis, the SC scheme is based on a male occupational structure, and thus 'allows little discrimination between 'female' occupations'[170].

Although these criticisms of SC are often acknowledged, it is still widely used, largely because it is well established and familiar. Although it may be a crude gauge of social position, the facilities to code SC are readily available and results can be compared to official figures and the results of other research. While the continuity of SC is not absolute, the historical nature of the measure does allow for some comparison of results over time. Krieger et al.[102] recognise the limitations of SC but consider its importance on the international stage to be such that it serves 'as a foundation for contemporary public health knowledge about patterns and trends in social inequalities in health'.

Social class distributions were very similar in the PAH study and the SWS (Table 6). A much higher proportion of women than men were classified as skilled non-manual (IIIN), and conversely many more men than women were classified as skilled manual (IIIM).

Table 6 Social class of PAH and SWS participants and partners

Social class	**PAH participant**	**PAH partner**	**SWS participant**	**SWS partner**
Professional (I)	14 (2%)	65 (10%)	294 (5%)	289 (8%)
Management and technical (II)	141 (23%)	123 (20%)	1706 (28%)	910 (25%)
Skilled non-manual (IIIN)	251 (40%)	60 (10%)	2122 (35%)	382 (11%)
Skilled manual (IIIM)	54 (9%)	211 (34%)	482 (8%)	1068 (29%)
Partly skilled (IV)	112 (18%)	87 (14%)	807 (13%)	546 (15%)
Unskilled (V)	15 (2%)	38 (6%)	208 (3%)	203 (6%)
Missing	33 (5%)	36 (6%)	510 (8%)	235 (6%)
Total	620 (100%)	620 (100%)	6129 (100%)	3633 (100%)

3.2.3 National Statistics Socio-Economic Classification

The National Statistics Socio-Economic Classification was developed in the 1990s as a response to criticism of the SC and SEC scales. It is said to follow a 'well-defined sociological position that employment relations and conditions are central to delineating the structure of socio-economic positions in modern societies'[168].

Rose and O'Reilly[165] presented the NS-SEC in 1998. Table 7 describes the NS-SEC and the collapsed version of the NS-SEC.

Table 7 The NS-SEC and collapsed NS-SEC

NS-SEC		Collapsed NS-SEC	
L1 L2 L3	Employers (25+ employees) Managers (25+ employees) Professionals	Class 1	Higher managerial and professional occupations
L4 L5 L6	Associate professionals Managers (<25 employees) Higher supervisors	Class 2	Lower managerial and professional occupations
L7	Intermediate occupations	Class 3	Intermediate occupations
L8 L9	Employers (<25 employees) Own account workers	Class 4	Small employers and own account workers
L10 L11	Lower supervisors Craft and related occupations	Class 5	Lower supervisory, craft and related occupations
L12	Semi-routine occupations	Class 6	Semi-routine occupations
L13	Routine occupations	Class 7	Routine occupations

The NS-SEC has been developed with close reference to Goldthorpe's employment relations theory[165]. Notably the classes are not to be considered ordinal, but rather nominal. The notion of skill and the manual/non-manual divide have no part in the conception of the scheme, but rather the categories of the NS-SEC have been developed as described below.

In Goldthorpe's employment relations theory a primary distinction is made between employers, who buy the labour of others, the self-employed, who neither buy labour nor sell their own, and employees, who sell their labour to employers. Employers are further divided into 'small' employers (L8) who are themselves responsible for managerial functions of their business, and 'large' employers (L1) where some of these functions are devolved to other staff. A pragmatic rule for differentiating these two types of employers uses the cut-off of 25 employees.

A similar distinction is made between 'large' and 'small' managers, but for different reasons. Employment relations theory discriminates between 'service relationships' and 'labour contracts'. The latter is typified by a relatively short-term and specific exchange of money for effort, whereas the former is a longer-term more diffuse exchange of compensation both in the present and the future. Assessment of the nature of employment relations was made using data on employment status characteristics of each OUG in terms of:

- Mode of payment
- Pay scale increments
- Notice required
- Promotion prospects
- Autonomy

A university teaching professional has a 'service relationship', whereas a car park attendant (without any supervisory responsibility) has a 'labour contract'. Intermediate forms of employment are considered to exist between these two classifications; for example a routine laboratory tester has an attenuated service relationship. 'Large' managers (L2) have a service relationship, whereas 'small' managers (L5) have an attenuated form of the relationship.

Professionals (L3) have a 'pure' service relationship, whereas associate professionals (L4) have a reduced form of the service relationship. Intermediate occupations with respect to the service relationship and labour contract are in clerical, administrative, service and lower technical positions (L7). Higher supervisors (L6) are described as having an attenuated form of the service relationship, and are identified as being in positions (other than managerial) involving the supervision of those in intermediate occupations. Own account workers (L9) are self-employed, with no employees other than family workers.

Employees in routine occupations (L13) have a basic labour contract, and those in semi-routine occupations (L12) have an only slightly modified labour contract. In the case of craft and related occupations (L11) the labour contract is considered modified, and lower supervisors (L10) also have a modified labour contract, but additionally are supervisors of those in routine, semi-routine, and craft and related occupations.

The NS-SEC has two further categories. L14 is 'never worked and long-term unemployed' and L15 is 'full-time students'. However, since no information in the Southampton studies has been collected on length of unemployment, analyses follow Rose and O'Reilly's[165] suggestion that 'some might not want to implement L14 at all and thus will exclude the never worked from the classification, and classify all unemployed persons in respect of their last main job'. Since full-time student participants without partners are classified by their fathers' NS-SEC as described in Section 3.2.1, and many others are classified by their last or part-time employment, there are too few subjects to merit the use of L15.

The strength of the NS-SEC is that it allows the possibility of explaining associations found, since the basis for occupational classification is far more explicit than for SC. Also, the NS-SEC is considered to reflect women's positions in employment more adequately than SC because women often perform non-manual jobs, such as sales assistants, which are less secure, less autonomous, and have fewer prospects than some manual jobs. Thus sales assistants are more appropriately represented by L12 (Semi-routine occupations) than SC IIIN (Skilled non-manual). Women have a more even distribution across the NS-SEC categories than SC or grouped SEG classes[165].

In an analysis of the General Household Survey (GHS)[165], a comparison of the NS-SEC with SC and collapsed SEG indicates that the NS-SEC produces a greater change in log-likelihood ratio than either of the previous measures, with respect to self-assessed health and limiting long-standing illness.

NS-SEC classes are allocated by OUGs. Unfortunately, coding for the NS-SEC is only available for 1990 OUGs. Since occupational information in the PAH study used 1980 OUGs it is not possible to generate NS-SEC classifications for the PAH study.

A refinement of the NS-SEC exists which differentiates between employment status within each OUG. Thus, for example, a 'large' self-employed electrician would be allocated to L1, whereas an electrician employee would be allocated to L11. Unfortunately, however, there is too little information in the SWS to use even the simplest such refinement, and thus the most basic version of the NS-SEC was used. This was found to correctly allocate 75% of the Labour Force Survey cases for those in paid employment, as compared to the most sophisticated version[165].

It is important to remember that the NS-SEC is a nominal measure. It is a class scheme, designed to measure the relational, rather than the distributive aspects of society. Rose and O'Reilly[165] elaborate further: 'Class schemata do not attempt to describe societies on a layer-cake model, but via more subtle, relational concepts'. The collapsed version (Table 7) of the NS-SEC is recommended for analyses[165]. The NS-SEC classifications of SWS participants and partners are described in Table 8.

Table 8 NS-SEC of SWS participants and partners

NS-SEC	SWS participant	SWS partner
Higher managerial and professional occupations	1223 (20%)	853 (23%)
Lower managerial and professional occupations	788 (13%)	437 (12%)
Intermediate occupations	1760 (29%)	312 (9%)
Small employers and own account workers	60 (1%)	300 (8%)
Lower supervisory, craft and related occupations	81 (1%)	385 (11%)
Semi-routine occupations	1248 (20%)	893 (25%)
Routine occupations	480 (8%)	238 (7%)
Missing	489 (8%)	215 (6%)
Total	6129 (100%)	3633 (100%)

3.2.4 Cambridge Scale

The Cambridge Scale is a continuous, rather than categorical, measure of social position that developers Prandy et al. believe can be understood as 'social order'; that is, a measure of general, hierarchical material and social advantage. A major benefit of the Cambridge Scale is that, unlike SC or NS-SEC, it is does not rely on any prior notion of the concept it is seeking to measure. Other approaches to the assessment of social position involve the categorisation or ranking of occupations by either 'expert' judges, samples of the population, or occupational characteristics. The Cambridge Scale, on the other hand, does not require any such assumptions. Furthermore, it does not even impose the belief that social classes or a quantity such as social standing indeed exist. The results of the empirical investigation determine the theoretical interpretation.

The Cambridge Scale is based on patterns of social interaction and assumes that close social relationships, such as friendship or marriage, occur in situations of social similarity. Thus two people who are friends or married are likely to be close to each other in the social order. The developers of the scale do not deny that friendships and marriages occur between individuals with a wide gulf in their social backgrounds, but rather that, put simply, it is far more likely that 'like befriends like' or 'like marries like'. Those in occupations close in the social order are more likely to have similar standards of lifestyle and to interact regularly both at work and in the community.

The Cambridge Scale was developed[166,171,172] by asking questions within existing surveys about details of occupations of 'people with whom you are friendly'. Additionally spouse's occupation was collected in some surveys, since marriage is argued to be equivalent to, and the culmination of a particular form of, friendship. Responses to these questions were analysed using Multi-Dimensional Scaling, and subsequently a simpler iterative technique, to produce a one-dimensional scale. This scale was well related to traditional stratification measures and inspection of the scores for each occupation suggests it is a measure of material and social advantage. Prandy[166] proposed that the Cambridge score could be considered a measure of 'generalised advantage' or 'lifestyle'.

The Cambridge Scale has been verified by Prandy[166]. The scale was compared to SC and was found to have relationships at least as strong as, or more often stronger than, SC with income, education, friendship, marriage, intergenerational comparisons and voting. Of particular relevance to this thesis, Prandy compared the performance of the Cambridge Scale to SC in the field of health inequalities[173]. He concluded that 'in terms of mortality, the population is not divided into discrete homogeneous categories'. Rather, there is evidence from Prandy's work that social order is a finely graded hierarchy, as exemplified by the Cambridge scale.

Cambridge Scale scores have been made available by OUG coding. Like the NS-SEC (Section 3.2.3) a refinement by employment status exists to produce a more accurate scoring. However, the PAH study did not enquire about employment status, and there was insufficient information in the SWS to use this refinement satisfactorily. Thus the simplest form of the Cambridge scale is utilised here.

Cambridge scores are available for men and women separately. Indeed this is seen as a significant advantage of the Cambridge Scale over SC; female scores are not based on a scheme designed initially to categorise male occupations, but rather the scale was developed separately for men and women. As Arber[174] advocates, 'There is a need to analyse the pattern for women in their own right, rather than holding it up to the male standard and treating any differences between the patterns for women's and men's occupations as an anomaly'.

The advantages of the Cambridge Scale over SC are summarised by Prandy[173]: 'the Cambridge Scale is to be preferred on both theoretical and empirical grounds: being constructed on a much sounder basis and superior in an explanatory sense'. The distribution of the Cambridge scores for the PAH study are shown Figure 9 and for the SWS in Figure 10.

Figure 9 Distribution of Cambridge scores for PAH participants and partners

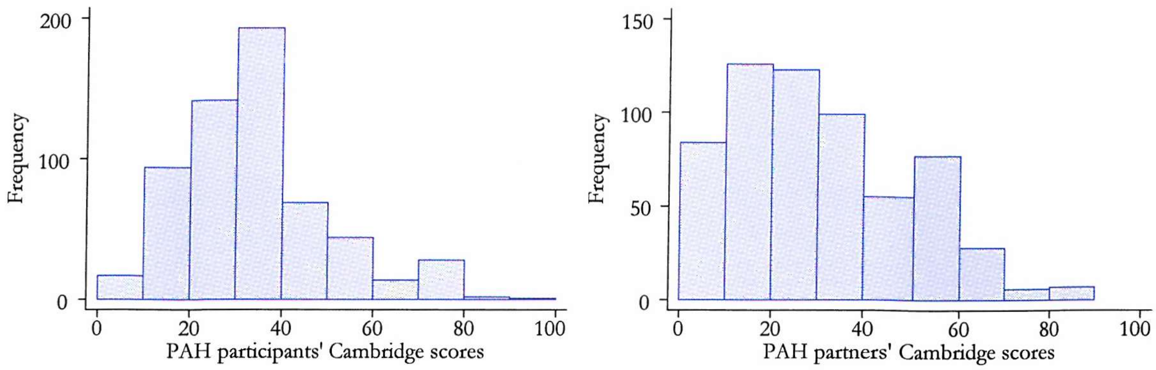
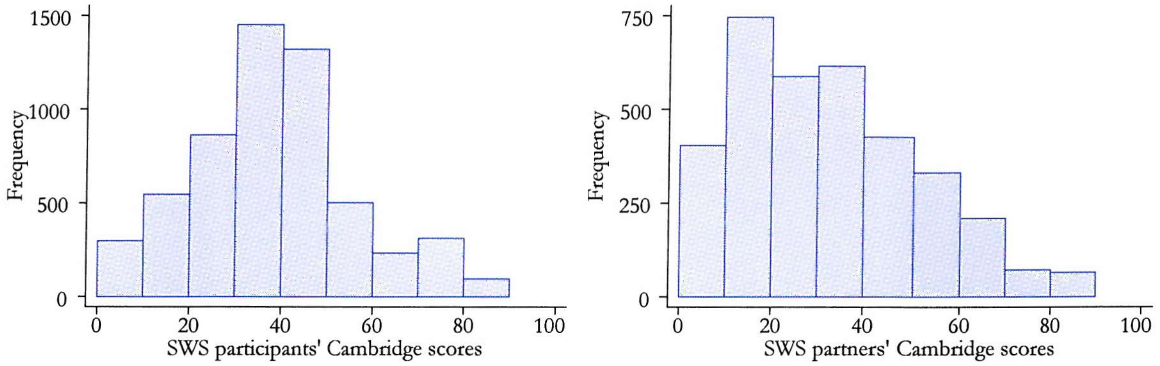


Figure 10 Distribution of Cambridge scores for SWS participants and partners



3.2.5 Education

Participants in the PAH and SWS were asked their highest qualification level; SWS participants were additionally asked the age at which they left full-time education. The advantage of qualification information is that it is available for all adults, irrespective of whether they are employed, and is relatively constant throughout the adult life-course. A disadvantage commonly cited is that the expansion of education has led to wide variations between educational attainment in different age cohorts. However, since the PAH and SWS only include young women, this concern is less relevant. Similarly, the relative youth of PAH and SWS participants is a reasonable defence against two other criticisms of the use of qualification information: that a high proportion of the population have no qualifications and that recall of qualifications is likely to be inaccurate.

The age at which participants left full-time education is not used as a measure of social position in the SWS since Arber[175] states that qualification attainment provides 'a better measure of labour market potential or human capital than completed years of full-time education'. Krieger et al.[102] agree that education is better measured in terms of credentials than years because a one year increase at a critical point of an individual's educational experience (such as the year A-levels are completed) will have more important implications for that individual's employment prospects than a one year increase at less critical times. The age of completion of full-time education is particularly useful in allowing international comparisons, but the Southampton studies are set entirely within the UK, and any international qualifications reported have been recoded as their English equivalent.

The SWS participants had somewhat higher educational attainment than PAH participants, probably because they are a more recent, non-pregnant cohort (Table 9).

Table 9 Qualification level of PAH and SWS participants

Highest qualification achieved	PAH	SWS
None	59 (10%)	424 (7%)
CSE/GCSE grade D or lower	114 (18%)	724 (12%)
O level/GCSE grade A,B,C	204 (33%)	1568 (26%)
A level/BTech	138 (22%)	1674 (27%)
HND	36 (6%)	387 (6%)
Degree	51 (8%)	1328 (22%)
Missing	18 (3%)	24 (0.4%)
Total	620 (100%)	6129 (100%)

Qualification levels amongst SWS participants were slightly negatively related to age (Spearman's correlation coefficient (r_S) = -0.13), reflecting improved access to higher education in more recent years. When education was examined by age in completed years it was notable that those aged 20 and 21 had slightly lower qualification levels than would be indicated by the overall negative trend, because women in these age groups had not yet had opportunity to complete qualifications such as degrees. Amongst PAH participants there was a positive correlation between age and educational level (r_S = 0.26), reflecting the fact that women with more education tend to have children later in life.

3.3 Occupation-based measures of women's social position

3.3.1 Background

The allocation of women's social position by occupation is a considerable difficulty in research into inequalities in women's health and lifestyles. Social classifications such as SC, NS-SEC or the Cambridge Scale require women to be characterised by occupation, but far fewer women than men currently work full-time. Conversely, a greater proportion of women than men currently work part-time (Table 10). The immediately intuitive classification of women by their own social position, the 'individual approach', therefore illustrates the current full-time occupation of a relatively small number of women, and has other limitations, as described in Section 3.3.2.2.

Table 10 Employment status of men and women in the UK in 1998 (millions)

	Men (aged 16-64)	Women (aged 16-59)
Full-time employees	11.3 (**60%**)	6.1 (**36%**)
Part-time employees	0.9 (**5%**)	4.6 (**27%**)
Self-employed	2.3 (**12%**)	0.8 (**5%**)
Others in employment	0.1 (**0.5%**)	0.1 (**6%**)
Unemployed	1.1 (**6%**)	0.7 (**4%**)
Economically inactive	3.0 (**16%**)	4.9 (**28%**)
Population of working age	18.7 (**100%**)	17.1 (**100%**)

Social Trends, ONS, 1999 Table 4.2

The traditional solution to this difficulty has been to classify women by their own occupation until marriage, and from then on to use their husband's occupation; the 'conventional approach'. The conventional approach has also been criticised (Section 3.3.2.3), and alternative measures to the individual or conventional allocation of social position have been suggested and evaluated. Measures other than the individual approach assume that a woman's social position is dependent on others in her household, specifically her husband or partner, if she has one. They are therefore called household-based measures.

3.3.2 Approaches to occupation-based measurement of women's social position

3.3.2.1 Summary

Single women are always classified by their own social position. There are however a variety of ways in which a married woman's social position can be defined. Table 11 summarises the alternative approaches, which are discussed in more detail below.

Table 11 Approaches to measurement of married women's occupation-based social position

		Woman's social position	**Husband's social position**
Individual		✓	
Conventional			✓
Dominance		Dominant of two	
Mixture:	not employed		✓
	employed	✓	

Note that a further variation on all the household-based measures is to use the occupation of the woman's partner for cohabiting women, rather than restricting such measures to married women. This adaptation has been used increasingly in recent research, as cohabitation has become more prevalent. For simplicity of description the following approaches are explained in terms of marriage.

3.3.2.2 The individual approach

The individual approach involves the classification of all women by their own occupation. It has the advantage that it is conceptually simple and is useful for investigating the direct effects of employment on women's health. However, there are several difficulties associated with using the woman's own occupation as a measure of social position. Firstly, since a relatively small proportion of women are currently employed full-time, the classification is frequently based on previous employment, or part-time employment, which are often an inadequate characterisation of the woman's current social position. More generally, considering women's lifetime employment experiences, their attachment to the labour force has been argued to be less stable than that of men[176]. Secondly, women's occupations are often taken to fit around childcare and other domestic responsibilities, and may therefore not reflect their skills, training or full potential[177,178]. This situation may be compounded by sex discrimination in the labour market. Finally, a small but significant proportion of women have never been in paid employment, and therefore cannot be assigned to an occupation.

3.3.2.3 The conventional approach

The conventional approach classifies single women according to their own occupation, but married women by their husband's occupation. The conventional approach is based on the assumption that a married woman is either a housewife or a secondary wage earner in the household.

The conventional approach is defended by Goldthorpe[176] who claims that husbands have a greater commitment to the workforce, since the vast majority of working wives withdraw from paid work at least once during their working life. Goldthorpe also argues that there are very few 'cross-class families', where the wife has a higher social class than her husband. Although a number of families may appear to be cross-class, this is very often because most female non-manual workers have routine or unskilled white-collar jobs. Goldthorpe claims that these women have a less favourable situation than their male non-manual counterparts in terms of pay, fringe benefits and job security, and therefore the husband's occupation is still the most appropriate measure of social position.

The conventional approach has many critics. In response to Goldthorpe, Stanworth[179] and Heath and Britten[180] cite evidence that the conventional approach is inappropriate in terms of its disregard for the contribution of women to the circumstances of the household. Arber[161] argues from the feminist viewpoint that 'it is no longer socially and politically acceptable for women to be considered an appendage of men'. Elsewhere Arber[175] reasons that increases in married women's employment alongside greater fluidity in marital status may mean that the conventional approach is found to be less predictive in empirical research. Pugh and Moser[181] add that if the conventional approach is used 'meaningful comparisons between women of different marital statuses are impossible' and in addition that this method 'does not provide a direct means of examining the effect of a married woman's own occupation on her health'.

3.3.2.4 The dominance approach

The dominance approach was originally suggested by Erikson[182] and uses a woman's occupation if she is unmarried, but the dominant of the wife and husband's occupation if she is married and employed. This assumes that the social advantage to the household is more dependent on one of the pair than the other assuming that one is dominant, and that it is irrelevant which of the two has the dominant position. The measure is therefore not gender-biased. It has the further advantage of allowing comparisons across different marital status groups on a more equal basis. The dominance approach clearly requires a dominance order in the measure of social position being considered. There is an obvious dominance order within SC and the Cambridge Scale, but not the NS-SEC.

3.3.2.5 The mixture approach

The mixture approach uses the woman's occupation if she is in employment, but the husband's for women who are economically dependent, that is, married and economically inactive. This is ostensibly a logical approach to the classification of the social position of women who are not currently working but whose social circumstances are arguably more closely related to her husband's current occupation than the woman's previous job. Arber[161] calls this method the 'employment-based' approach, but highlights a serious flaw in the measure. Categorising women from the 1981-2 GHS using the mixture approach reveals 'skilled manual' women to have an unusually high rate of limiting long-term illness. However, this is an artefact of the following two points. Firstly, few women, but a high proportion of men, work in skilled manual occupations, and secondly, women out of the labour force have been shown to report poorer health than those in current employment. Therefore those women classified as 'skilled manual' include a disproportionate number of women out of the labour force, many of whom have been selected out due to poor health. Conversely the proportion of women working in unskilled occupations is relatively high and therefore the rate of limiting long-term illness in this group appears very low. These problems of interpretation lead Arber to conclude that the mixture approach to classification is invalid, and thus it is not considered in this thesis.

3.3.3 Published empirical research

Empirical research into the most appropriate measure of women's social position has utilised diverse outcomes such as women's subjective class identification, earnings, voting patterns, hours spent performing household labour, household food purchasing patterns and childbearing patterns[183]. Epidemiological research provides another arena for investigations in this area, with particular pertinence to the current inequalities in health debate. Gradients in various measures of health are seen across categories of social position, and the comparison of different methods of allocation amongst women may help illuminate the causal pathways involved. If women's own occupations influence their health more than household-based measures then occupation itself might generate a substantial proportion of women's ill health. If the contrary is found then the major forces are probably located in household conditions.

Previously published epidemiological research comparing individual and household measures of social position in terms of class inequalities in women's health is summarised in Table 12.

3.3.4 Conclusions

Results from Table 12 are difficult to generalise since the studies differ so widely across all the attributes listed. However, in all but one of the studies that consider the dominance approach, this procedure is found to best describe the variation in outcomes considered. Indeed, the study that considers aspects of women's health and behaviours most relevant to childbearing (Krieger et al.[183]) concludes that the dominance approach is superior. Furthermore Arber's[161] analysis of the GHS indicates that the strength of social position gradients in women's health using the conventional method largely reflects the poor health of single and previously married women in manual occupations, whereas the gradient for married women is much weaker.

Table 12 Epidemiological research comparing estimates of class inequalities in women's health detected with individual versus household measures of social position (in order of years of study)

Authors	Years of study	No. of particip-ants	Age of particip-ants	Location of study	Measure of social position	Predictors	Outcomes	Variation in outcomes best described by…
Moser KA, Pugh HS & Goldblatt PO[184]	1976-1981	---†	15-59	England and Wales*	SC (manual/non-manual)	Individual Conventional (H)	Mortality	Conventional
Arber S[161]	1981-1982	14044	20-59	Britain**	Collapsed SEG	Individual Conventional (H) Dominance (H) Mixture (H)	Limiting long-standing illness	Dominance [Conventional better than individual, mixture dismissed on theoretical grounds]
Martikainen P[185]	1981-1985	---†	35-64	Finland	Education Finnish Socio-Economic Classification	Individual Conventional (H)	Mortality (total and five main causes)	No difference
Vågerö D[186]	1981-1986	---†	20-64	Sweden	Swedish Socio-Economic Classification	Individual Dominance (P)	Mortality	Dominance
Dahl E[187]	1983	1272	26-66	Norway	Norwegian Socio-Economic Classification	Individual Conventional (P) Dominance (P)	Self-assessed health	Conventional [Dominance better than individual]
Arber S & Ginn J[188]	1985	2083	65+	Britain**	Collapsed SEG	Individual Conventional (H)	Guttman's functional disability	No difference
Arber S[189]	1985-1986	12810	20-59	Britain**	Collapsed SEG	Individual Conventional (H)	Limited long-standing illness	Conventional
Arber S & Ginn J[188]	1985-1987	6486	65+	Britain**	Collapsed SEG	Individual Conventional (H)	Self-assessed health	Conventional
Krieger N[190]	1987	101	20-80	California	Wright's Typology (working class/not working class)	Individual Dominance (HoH)	Reproductive history	Dominance

*ONS Longitudinal Study
**General Household Survey
***Health and Lifestyle Survey

HoH = Head of household
H = Husband
P = Partner

†Number of participants not specified

Table 12 (continued) Epidemiological research comparing estimates of class inequalities in women's health detected with individual versus household measures of social position (in order of years of study)

Authors	Years of study	No. of particip-ants	Age of particip-ants	Location of study	Measure of social position	Predictors	Outcomes	Variation in outcomes best described by…
Krieger N, Chen JT & Selby JV[183]	1989-1990	684	30-91	California	Wright's Typology (working class/not working class)	Individual Conventional (P) Dominance (P)	Anthropometry Metabolic measures Self-assessed health Age at first pregnancy Smoking history Physical activity	Dominance [Conventional better than individual]
Chandola T[191]	1984-1996	4338	18+	England, Wales and Scotland***	SC Erikson-Goldthorpe Schema Cambridge Scale	Individual Conventional (P)	Coronary Heart Disease	Conventional
Sacker A, Firth D, Fitzpatrick R, Lynch K & Bartley M[192]	1986-1996	90327	16-60	England and Wales*	Cambridge Scale NS-SEC	Camb: Dominance (P) NS-SEC: Individual	Mortality	Dominance
Arber S[175]	1991-1992	12266	20-59	Britain**	Collapsed SEG	Individual Conventional (P)	Limiting long-standing illness Self-assessed health	LLI: Individual Self-assessed health: Conventional

*ONS Longitudinal Study
**General Household Survey
***Health and Lifestyle Survey

HoH = Head of household
H = Husband
P = Partner

†Number of participants not specified

The majority of the studies listed in Table 12 that consider the dominance method use the social position of the woman if she is single, and the dominant of her and her partner (rather than husband or head of household) if she is not. Cohabitation may resemble marriage in terms of living arrangements and emotional commitment for an increasing number of women (Section 3.4.2.1). Therefore the dominance method, considering the woman's partner if she is living with a partner, is employed in analyses using SC and the Cambridge Scale in this thesis. Bartley et al.[193] concur in their analyses using the Cambridge Scale that it is most appropriate to link attitudes and behaviours (such as food choice) to the highest level of 'general advantage' within a household.

The NS-SEC has nominal, rather than ordinal, categories and thus the dominance concept cannot be applied. The conventional approach may describe more variation in health outcomes than the individual approach (Table 12). However, it is important to consider the causal links between NS-SEC, as a measure of employment relations and conditions, and lifestyle. Bartley et al.[123] used the individual approach in the classification of women by the Erikson-Goldthorpe schema (the classification upon which the NS-SEC is based) because they argue that there is little plausibility in an aetiological pathway between an individual's behaviour and their partner's employment conditions. The NS-SEC is a measure of the employment relations and conditions dimension of social position; individual NS-SEC is considered the most logical measure when investigating links between these conditions and a woman's lifestyle, and is thus used in subsequent analyses.

3.4 Other measures

3.4.1 Demographic measures

Gender, age, and ethnicity are regarded as demographic measures[194]. Since only women's diets are examined in this thesis, gender is not relevant for consideration as a demographic predictor of diet here, but ethnicity and age are discussed below.

3.4.1.1 Ethnicity

Macintyre and Anderson[195] clarify the meaning of 'ethnicity' as it is used in this context:

> The term 'ethnicity' should be distinguished from 'race' in epidemiological and social research. 'Ethnicity' signifies shared and inherited cultural, linguistic, and

religious traditions. It is not necessarily the same as nationality, migrant status, or skin colour.

The PAH data was collected only from white women, and thus it is not possible to look at ethnicity as a predictor of women's diets in this study. However, the SWS was not restricted to women of any one ethnic group. Women in the survey were asked which ethnic group they considered they belonged to. Since numbers in some of the groups were very small, these were recoded using the groupings of the GHS[196] (Table 13).

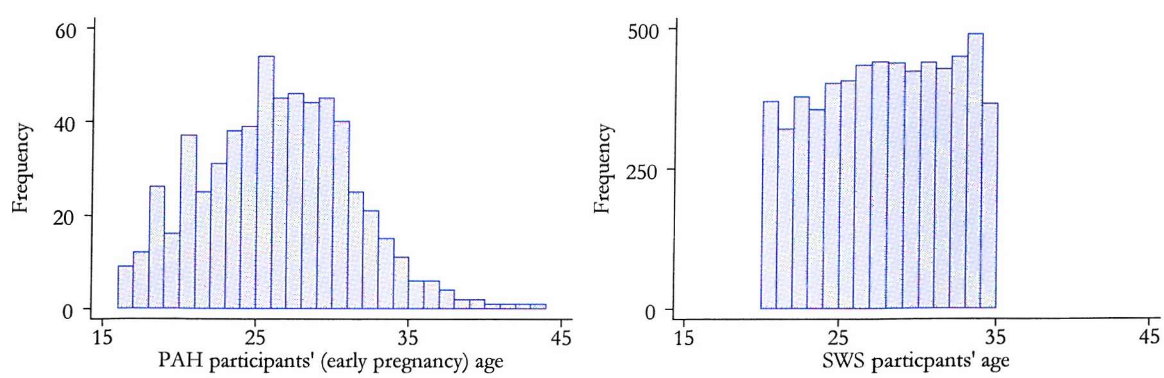
Table 13 Ethnicity of SWS participants

Ethnic group		
White	5691	(93%)
Black	84	(1.4%)
Indian	150	(2.5%)
Pakistani/Bangladeshi	69	(1.1%)
Other	118	(1.9%)
Missing	17	(0.3%)
Total	6129	(100%)

3.4.1.2 Age

Participants in the PAH study were all aged 16 or older at the early pregnancy interview, whereas those in the SWS were chosen to be aged 20-34. The SWS had a more even distribution of ages than the PAH study (Figure 11). This is to be expected, since the age distribution of PAH participants has a peak at the most common child-bearing ages, whereas the distribution of SWS respondents reflects the even age distribution of women aged 20-34 living in Southampton.

Figure 11 Age of PAH and SWS participants



Macintyre and Anderson[195] note that any apparently simple effect of age on dietary outcomes may itself have three different mechanisms:

- Biological states and processes (e.g. growth and maturation)
- Historical processes (e.g. access to higher education may have been more difficult for older women)
- Psychological and social processes (e.g. younger women may feel more pressure to be thin)

3.4.2 Household composition

In the PAH study only information about whether a woman is sharing her home with a partner is available, but the SWS recorded all members of a participant's household. A household was defined as a group of people who share a living room or eat together for at least one meal a day, although people living in hostels or halls of residence were classified as living alone (Appendix 3). A woman's household composition is thus clearly likely to be related to her dietary patterns.

3.4.2.1 Living with a partner

Macintyre and Anderson[195] comment that legal marital status 'is a civil status which may increasingly bear little relationship to actual living arrangements or emotional commitments, so it may be less valid than it once was for observers to infer other characteristics from legal marital status'. Certainly cohabitation is becoming more common[197] and therefore 'living with a partner' against 'living as a single person' is used in this thesis as an indicator of this aspect of a woman's household composition.

Information on living with a partner is available in both the PAH and SWS datasets. Substantially more women within the pregnant PAH cohort were living with a partner (Table 14).

Table 14 PAH and SWS participants living with a partner

Living with a partner	**PAH**	**SWS**
No	50 (**8%**)	2496 (**41%**)
Yes	542 (**87%**)	3633 (**59%**)
Missing	28 (**5%**)	0 (**0%**)
Total	620 (**100%**)	6129 (**100%**)

3.4.2.2 Number of children in household

Further information about others living in a participant's household is available from the SWS. Each participant was asked about her relationship to all members of her household, as well as how many days per week each member lived in the household on average. This enabled the number of children a woman lived with to be calculated. A 'child' is defined as either the woman's own child, a stepchild, an adopted child, a foster child or her cohabitee's child. Children who lived with the woman for only a few days per week are less likely to have a major influence on her diet. A child was thus considered a member of the household if they lived there for at least 5 days per week. The number of children a woman lived with is a highly skewed variable; very few women lived with five or more children. Therefore this variable was categorised as 0, 1, 2 or 3+ children (Table 15).

Information about living with children is not available in the PAH study. However, the woman's number of previous live births was extracted from her birth records. Clearly some live-born children may not have lived with their mother, or the PAH participant may have lived with children who are not her own. However, this variable was considered adequate to use as a proxy measure of the number of children that the woman was living with.

Table 15 Number of children living with PAH and SWS participants

Number of children	PAH	SWS
0	324 (**54%**)	3402 (**56%**)
1	192 (**32%**)	1019 (**17%**)
2	55 (**9%**)	1081 (**18%**)
3+	32 (**5%**)	627 (**10%**)
Missing	17 (**3%**)	0 (**0%**)
Total	620 (**100%**)	6129 (**100%**)

3.4.3 Lifestyle measures

3.4.3.1 Smoking

Smoking is related to food intake (Section 1.4.1.9). It is an important aspect of lifestyle and is known to create an altered pattern of demand for specific nutrients[198]. Smoking habits at time of interview are described in Table 16. As might be expected, a lower proportion of pregnant than non-pregnant women smoked.

Table 16 Smoking habits of PAH and SWS participants

Smoker	PAH (early)	PAH (late)	SWS
No	449 (72%)	459 (77%)	4076 (67%)
Yes	154 (25%)	133 (22%)	2049 (33%)
Missing	17 (3%)	1 (0.2%)	4 (0.1%)
Total	620 (100%)	593 (100%)	6129 (100%)

3.4.3.2 Physical activity

Physical activity is a very important predictor of energy requirements[128]. Information about physical activity in the PAH study was collected by asking how many times the participant had engaged in any sporting activity long enough to work up a sweat or get out of breath over the previous week (Appendices 1 and 2). This variable was dichotomised into women that had and hadn't take part in any exercise, since the variable was highly skewed (Table 17).

More detailed physical activity information was collected in the SWS using an adaptation of the Godin exercise questionnaire[199]. The Godin questionnaire asks about frequency of strenuous, moderate and gentle exercise per week. Strenuous activity is defined in the SWS as that 'which normally makes your heart beat rapidly and leaves you breathless' (Appendix 3). Strenuous (rather than moderate or gentle) exercise is therefore most comparable with physical activity information from the PAH study. Since Godin[199] also found frequency of strenuous exercise to be most predictive of two measures of fitness, strenuous exercise is used as a measure of physical activity in analyses of the SWS. Again the frequency of strenuous exercise was highly skewed so the variable was dichotomised into those that had and hadn't taken strenuous exercise. In the SWS exercise frequency was available up to once every two to three months, so the greater frequency of strenuous exercise amongst SWS participants (Table 17) may be partly because some took strenuous exercise less frequently than once per week, in addition to the fact that the participants are non-pregnant.

Table 17 Strenuous exercise of PAH and SWS participants

Strenuous exercise	PAH (early)	PAH (late)	SWS
No	496 (80%)	474 (80%)	2363 (39%)
Yes	106 (17%)	118 (20%)	3706 (60%)
Missing	18 (3%)	1 (0.2%)	60 (1%)
Total	620 (100%)	593 (100%)	6129 (100%)

3.4.3.3 Watching television

SWS participants were asked about the number of hours of television they watched per day. This variable reflects some aspects of physical activity, as well as providing information about lifestyle choice. The number of hours SWS participants spent watching television is described in Table 18, along with the numeric frequency equivalents used in data analysis.

Table 18 Hours spent watching television by SWS participants

Hours spent watching television	Frequency per week equivalent	n (%)
None	0	83 (1%)
Less than one hour	0.5	579 (9%)
1-2 hours	1.5	1594 (26%)
2-3 hours	2.5	1713 (28%)
3-4 hours	3.5	1191 (19%)
4-5 hours	4.5	677 (11%)
More than 5 hours	6	292 (5%)
Missing	Missing	0 (0%)
Total		6129 (100%)

3.4.3.4 Dieting

The decision to diet is clearly an influence on women's food choices. Ideally questions about dieting would refer to the 3-month period relevant to the FFQ. However, in the PAH study women were asked at their early pregnancy interview whether they had tried to lose weight by dieting in the past year. In the SWS women were asked whether they were currently trying to lose weight by dieting (Table 19).

Table 19 PAH and SWS participants trying to lose weight by dieting

Currently dieting	PAH	SWS
No	358 (58%)	4720 (77%)
Yes	245 (40%)	1402 (23%)
Missing	17 (3%)	7 (0.1%)
Total	620 (100%)	6129 (100%)

3.4.3.5 Currently employed

Whether a woman is currently working is an important indicator of her lifestyle. PAH participants were asked whether they were currently employed in paid work at both the early and late pregnancy interviews. SWS participants were asked whether they were in paid employment in the week ending the previous Sunday. This variable gives some information about the woman's lifestyle at the time of interview. Table 20 illustrates that fewer of the pregnant PAH sample than the non-pregnant SWS sample were working, particularly in late pregnancy.

Table 20 PAH and SWS participants currently employed

Currently employed	PAH (early)	PAH (late)	SWS
No	258 (**42%**)	472 (**80%**)	1668 (**27%**)
Yes	362 (**58%**)	121 (**20%**)	4461 (**73%**)
Missing	0 (**0%**)	0 (**0%**)	0 (**0%**)
Total	620 (**100%**)	593 (**100%**)	6129 (**100%**)

3.4.4 Body mass index

Body size is an important predictor of energy requirements[128]. Body mass index (BMI) is a measurement of body size calculated as:

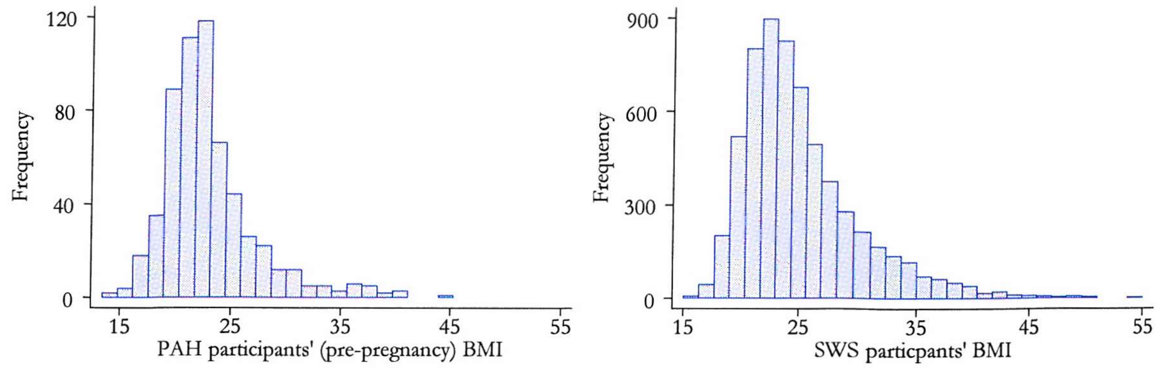
$$BMI = \frac{Weight}{Height^2}$$

It is highly correlated with weight, but aims to adjust for height, so that it has the same meaning for tall and short people. BMI should therefore be independent of height. Spearman's correlation between BMI and height is -0.02 in the PAH study, and -0.06 in the SWS. These coefficients compare favourably with the extensive data analysed by Garn et al[200].

BMI has been criticised as reflecting the weight of both lean and fat tissues when considered as a measure of adiposity[200]. However, in this context energy requirements are related to lean mass as much, if not more than, fat mass[128]. Therefore BMI is used as a measure of overall body size, rather than adiposity.

Participants in the PAH study were asked to recall their pre-pregnancy weight, and since this value does not include the weight of their baby, pre-pregnant BMI was used as a measure of body size. BMI is not normally distributed (Figure 12) and is therefore logged (using the natural logarithm) before analysis. However, since the coefficients for a logged predictor variable in a regression model are difficult to interpret, logged BMI was then standardised to a mean of zero and a standard deviation of one for presentation.

Figure 12 BMI of PAH and SWS participants



3.4.5 Pregnancy effects

3.4.5.1 Nausea

Nausea is commonly experienced during pregnancy and affects a woman's food intake. Women in the PAH study were asked in both early and late pregnancy whether they had experienced nausea in the previous three months, and if so whether this had been mild (nausea only), moderate (sometimes sick) or severe (regularly sick and unable to retain meals). Table 21 describes the nausea levels reported in early and late pregnancy. Nausea was worse in the first trimester of pregnancy, as would be expected.

Table 21 PAH participants experience of nausea

Nausea	PAH (early)	PAH (late)
None	108 (**17%**)	386 (**65%**)
Mild	235 (**38%**)	115 (**19%**)
Moderate	198 (**32%**)	83 (**14%**)
Severe	79 (**13%**)	8 (**1%**)
Missing	0 (**0%**)	1 (**0.2%**)
Total	620 (**100%**)	593 (**100%**)

3.4.5.2 Change in amount of food

Women may also experience a noticeable change of appetite during pregnancy, and the PAH study asked women at early pregnancy whether they were eating less, the same or more in amount since they became pregnant. The women's responses were very evenly spread between the three categories (Table 22).

Table 22 Change in amount of food since becoming pregnant in PAH early pregnancy interviews

Change in amount of food	
Less	202 (33%)
Same	197 (32%)
More	204 (33%)
Missing	17 (3%)
Total	620 (100%)

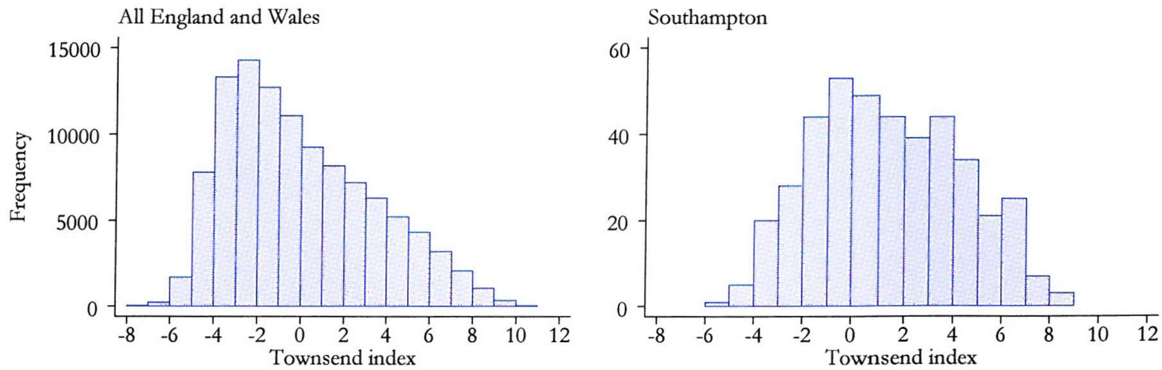
3.5 Representativeness

Information from sociodemographic and lifestyle measures available in the PAH study and the SWS provides an opportunity to assess the representativeness of these study samples in terms of the UK population.

3.5.1.1 Townsend index

Figure 13 illustrates the comparability of deprivation in Southampton with all of England and Wales, as measured by the Townsend score.

Figure 13 Distribution of Townsend scores for all England and Wales and Southampton enumeration districts



Southampton city EDs have a similar range of Townsend index scores to England and Wales EDs, although the very highest and very lowest scores are not represented in Southampton; only 0.9% of EDs in England and Wales have scores that are not represented in Southampton. The distribution of Townsend index scores indicates that Southampton contains fewer areas of less deprivation relative to areas of more deprivation, compared to all of England and Wales, perhaps because it is an urban area. The distribution of Townsend index scores of the PAH and SWS participants were almost identical to that of all enumeration districts in Southampton, implying that these samples are representative of general deprivation levels in Southampton.

3.5.1.2 Education

Qualification levels for women aged 20-29 are available for the 1992 GHS[201] and most recently from the 1996 GHS[202]. Table 23 shows that the qualification levels of PAH participants aged 20-29 are representative of the British population, although women with no qualifications are under-represented. SWS participants aged 20-29 are slightly better qualified than the 1996 GHS respondents. However, the GHS data is somewhat earlier than the SWS interviews, and more modern cohorts would be expected to have higher qualifications. Again women with no qualifications are under-represented. The higher proportion of women with a degree in the SWS is likely to be due to the presence of a large university in Southampton. The data analysed in this thesis includes women living around the main university campus, and may therefore have included a large number of postgraduate students, as well as women who have settled in Southampton after attending the university.

Table 23 Qualification level of PAH participants, GHS 1992 female respondents, SWS participants and GHS 1996 female respondents, aged 20-29

Highest qualification achieved	PAH	1992 GHS	SWS	1996 GHS
None	6%	14%	5%	14%
CSE/GCSE grade D or lower	21%	13%	9%	11%
O level/GCSE grade A,B,C	36%	41%	25%	33%
A level/BTech	25%	15%	30%	19%
HND	6%	8%	6%	7%
Degree	6%	8%	24%	14%
Missing	0.3%	1%	0.4%	2%

3.5.1.3 Ethnicity

Ethnicity of a representative sample of households in Great Britain is available from the 1998 GHS[196]. Table 24 shows that the ethnicity of SWS participants is highly representative of that of participants aged 25-44 in the GHS.

Table 24 Ethnicity of SWS participants and GHS subjects aged 25-44

Ethnic group	SWS	GHS
White	93%	93%
Indian	2%	2%
Pakistani/Bangladeshi	1%	1%
Black	1%	2%
Other	2%	2%

3.5.1.4 Smoking

The smoking rate in the 1992 GHS[201] was 34% amongst women aged 25-34. The rates for PAH participants aged 25-34 were quite a bit lower in early and late pregnancy (23% and 20%) as would be expected since women often try to give up smoking before or during pregnancy. The rate of smoking amongst SWS women aged 25-34 is 32%, highly comparable with the 33% of female respondents aged 25-34 to the 1998 GHS[196] who smoked.

3.6 Summary

- No overall measure of deprivation is available. Instead four measures indicate different facets of deprivation. The Townsend index has been developed as an indication of material deprivation, and is an area, rather than individual-based measurement. Benefits are largely an assessment of financial deprivation, although they certainly also reflect an element of social deprivation. Housing tenure indicates both material and social deprivation, and housing conditions are a more specific measure of accommodation-related material deprivation.

- Several means of assessing social position are available. Although SC has an insecure conceptual basis, it is useful since it is a widely used ordinal measure, and is analysed using the dominant household social class. The recently developed NS-SEC is a more valid measure based on employment relations theory. Since it is not an ordinal measure, each woman's NS-SEC is assigned using her own individual present or previous employment. The Cambridge Scale has been developed as an alternative measure that the authors consider represents 'generalised advantage' or 'lifestyle'. Each occupation has a score on a continuous scale, and the woman is assigned a score using the dominant household Cambridge score. Educational attainment is an alternative non-occupation-based measure of social position.

- Other relevant predictors of diet considered are ethnicity, age, household composition, smoking, physical activity, watching television, dieting, currently working, BMI and pregnancy effects. Table 25 provides a summary of all sociodemographic and lifestyle measures available in the PAH study and the SWS.

Table 25 Sociodemographic and lifestyle measures available for analysis

Sociodemographic or lifestyle measure	PAH	SWS
Measures of deprivation		
Townsend index	✔	✔
Benefits	✘	✔
Housing tenure	✘	✔
Housing conditions	✘	✔
Measures of social position		
Social Class (dominant)	✔	✔
NS-SEC (individual)	✘	✔
Cambridge score (dominant)	✔	✔
Education	✔	✔
Demographic measures		
Ethnicity	✘ †	✔
Age	✔	✔
Household composition		
Living with a partner	✔	✔
Number of children in household	✔	✔
Lifestyle measures		
Smoking	✔	✔
Physical activity	✔	✔
Watching television	✘	✔
Dieting	✔	✔
Currently employed	✔	✔
Body mass index		
Body mass index	✔	✔
Pregnancy effects		
Nausea	✔	✘ ‡
Change in amount of food	✔	✘ ‡

†All women in the PAH study are white, and therefore ethnicity is not relevant as a predictor of diet.
‡No women in the SWS are pregnant, and therefore pregnancy effects are not relevant as predictors of diet

- Women in the PAH study and the SWS are representative of the UK population, although those with no qualifications are under-represented, and they come from areas with slightly greater deprivation as measured by the Townsend index.

4 Dietary assessment in the PAH study

This chapter presents analyses of dietary data from the PAH study. Principal component analysis is found to be a valuable means of summarising FFQ data. Results from early and late pregnancy are interpreted and compared. The issue of robustness of patterns when food categories are collapsed is addressed, initially by comparison with two methods of grouping from the literature, and also by using a grouping designed to be most relevant to the Southampton FFQ. Useful principal component scores are developed, and these are used to assess whether women's diets change between early and late pregnancy.

Section 4.2 compares results obtained using factor analysis with those from PCA, and Section 4.3 relates the results of PCA on early pregnancy food diaries with results from the early pregnancy FFQ data. Finally, fruit and vegetable intake is considered in Section 4.4 as an alternative indication of one aspect of diet.

4.1 Principal component analysis

4.1.1 Early pregnancy PCA

The first five principal components from a PCA of data from 617 subjects using the 100 foods in the early pregnancy PAH data are shown in Table 26. Coefficients are coloured using the following scheme, chosen to highlight particularly high and low values:

≥ 0.16
0.06 to 0.15
-0.05 to 0.05
-0.15 to -0.06
≤ -0.16

Table 26 PCA coefficients from 100 foods in early pregnancy PAH data

Food	Component				
	1	2	3	4	5
White bread	-0.16	0.11	**0.02**	-0.10	**0.04**
Wholemeal bread	0.21	**0.03**	0.10	0.14	-0.06
'Bran' breakfast cereals	0.14	-0.06	**0.02**	0.09	-0.06
Other breakfast cereals	**-0.05**	0.10	**0.02**	0.09	**0.00**
Crackers	**0.05**	0.12	0.06	0.15	**-0.05**
Wholemeal crackers	0.07	0.07	**0.00**	0.20	**-0.02**
Pasta	0.11	0.10	0.09	**-0.03**	**-0.04**
Rice	0.20	**0.03**	**0.03**	-0.12	0.08
Quiche and pizza	0.06	0.07	0.13	**-0.02**	-0.11
Yorkshire pudding and savoury pancakes	-0.06	0.20	**-0.02**	0.19	**-0.03**
Cakes	**-0.03**	0.15	0.06	**0.01**	0.22
Buns	**0.03**	0.13	**-0.01**	0.09	0.23
Pastries	**0.02**	0.16	0.12	**0.04**	**0.04**
Biscuits	**-0.03**	0.16	0.07	**0.00**	0.18

Table 26 (continued) PCA coefficients from 100 foods in early pregnancy PAH data

Food	Component 1	Component 2	Component 3	Component 4	Component 5
Other biscuits	0.01	0.09	0.08	0.03	0.12
Fruit based puddings	0.04	0.12	0.04	0.12	0.06
Milk based puddings	0.04	0.15	0.01	0.09	-0.02
Other puddings	-0.01	0.14	0.02	0.05	0.04
Ice cream	-0.01	0.15	0.15	0.09	0.05
Full-fat liquid milk (pints)	-0.11	0.14	0.05	0.01	0.01
Full-fat powdered milk (pints)	-0.04	0.03	0.02	-0.01	-0.04
Reduced-fat liquid milk (pints)	0.11	-0.06	-0.06	0.00	-0.01
Reduced-fat powdered milk (pints)	0.02	-0.01	0.05	0.05	-0.01
Cream	0.08	0.13	0.10	0.01	0.10
Cheese	0.16	0.02	0.09	-0.08	0.03
Cottage cheese	0.10	0.01	0.09	0.05	-0.07
Yoghurt	0.18	0.04	0.06	0.11	-0.09
Eggs	0.00	0.08	-0.01	-0.04	-0.01
Omelette	-0.01	0.16	0.08	-0.08	-0.01
Full-fat spread	-0.02	0.14	0.03	-0.01	0.27
Reduced-fat spread	0.02	-0.01	0.01	0.03	-0.27
Cooking fats and salad oils	0.01	0.10	0.12	-0.21	-0.03
Chicken and turkey	0.10	0.11	-0.16	-0.09	-0.03
Pork	0.00	0.20	-0.22	-0.15	-0.07
Lamb	0.03	0.19	-0.23	-0.13	-0.03
Beef	-0.01	0.15	-0.16	-0.19	-0.10
Minced meat dishes	0.07	0.14	-0.16	-0.15	-0.02
Bacon and gammon	-0.02	0.20	0.00	-0.09	-0.04
Meat pies	-0.07	0.14	-0.08	0.02	-0.11
Sausage	-0.07	0.19	0.05	-0.11	-0.06
Ham and canned meats	0.02	0.12	-0.13	-0.03	-0.06
Liver and kidney	0.00	0.11	-0.10	-0.08	0.03
Pate and liver sausage	-0.01	0.07	-0.05	0.01	-0.01
Black pudding and faggots	-0.07	0.09	-0.10	0.00	-0.13
White fish	0.13	0.04	-0.03	0.13	-0.07
Fish pie	0.03	0.09	0.07	0.17	-0.14
Oily fish	0.16	0.00	-0.02	0.03	-0.12
Shell fish	0.08	0.00	-0.02	-0.07	0.03
Salad	0.21	0.02	0.13	-0.08	-0.08
Coleslaw	0.08	0.07	-0.03	0.05	-0.18
Tomatoes	0.17	0.05	0.00	-0.19	-0.06
Green pepper and watercress	0.25	-0.03	0.10	-0.14	0.02
Spinach	0.17	0.00	0.10	0.06	0.00
Sprouts and broccoli	0.13	0.10	-0.21	-0.01	0.12
Cabbage and cauliflower	0.08	0.12	-0.21	0.04	0.14
Peas	0.09	0.06	-0.24	0.07	-0.04
Courgettes, marrow & leeks	0.20	-0.02	0.03	-0.11	0.10
Carrots	0.14	0.07	-0.18	0.03	0.05
Parsnips, swedes and turnips	0.06	0.08	-0.19	0.01	0.09
Sweetcorn	0.12	0.12	-0.04	0.09	-0.09
Onion	0.17	0.05	0.06	-0.28	-0.06
Mushrooms	0.18	0.00	0.05	-0.11	-0.01
Beans and pulses	0.03	0.09	0.08	0.09	-0.17
Vegetable dishes	0.20	-0.03	0.23	-0.05	0.05
Vegetarian foods	0.13	-0.04	0.22	0.04	0.05
Tinned vegetables	-0.13	0.13	0.06	0.16	-0.13
Boiled potatoes	0.00	0.15	0.03	0.14	-0.05
Roast potatoes and chips	-0.10	0.17	0.14	0.07	-0.16
Crisps	-0.12	0.11	0.17	-0.05	0.07
Oranges and orange juice	0.13	0.05	0.01	0.04	0.01
Grapefruit and grapefruit juice	0.09	-0.04	-0.07	0.10	-0.02
Apples and pears	0.12	-0.01	-0.05	0.12	0.08
Banana	0.09	0.02	-0.04	0.10	0.09
Peaches, plums, apricots, cherries and grapes	0.12	0.05	-0.03	0.05	0.02

Table 26 (continued) PCA coefficients from 100 foods in early pregnancy PAH data

Food	Component 1	2	3	4	5
Strawberry and raspberry	0.08	0.10	0.02	0.02	-0.02
Melon, pineapple, mango and kiwi	0.05	0.03	-0.05	0.14	0.28
Other fruit juices	0.08	0.02	-0.03	0.17	0.26
Nuts	0.09	0.08	0.11	-0.20	0.02
Dried fruit	0.15	-0.01	0.04	0.06	0.01
Cooked fruit	0.10	0.07	-0.05	0.11	-0.06
Tinned fruit	0.06	0.11	-0.02	0.18	-0.07
Chocolate	-0.09	0.15	0.12	-0.05	0.11
Other confectionery	-0.05	0.11	0.14	0.03	0.19
Sweet spreads	0.09	0.08	0.00	-0.03	0.06
Added sugar (teaspoons)	-0.16	0.17	0.13	-0.03	-0.01
Blackcurrants and ribena	0.01	0.04	0.05	0.01	-0.17
Fizzy drinks and squashes	-0.05	0.05	0.08	0.00	0.06
Coke	-0.09	0.09	0.12	-0.13	0.02
Diet coke	-0.01	0.01	0.02	-0.08	0.04
Tea	-0.07	0.12	0.02	-0.06	0.08
Coffee	-0.06	0.05	0.07	-0.07	-0.12
Decaffeinated tea and coffee	0.08	-0.02	-0.05	0.02	0.07
Drinking chocolate	0.00	0.08	0.00	0.05	-0.06
Red wine	0.07	-0.03	0.03	-0.17	0.08
White wine and cider	0.03	-0.02	0.06	-0.15	0.06
Soup	0.06	0.04	0.03	0.05	0.00
Meat stock cubes	0.00	0.12	-0.24	-0.03	0.01
Mayonnaise and salad cream	0.10	0.10	0.07	-0.01	-0.19
Pickles, chutney and ketchup	-0.01	0.12	0.08	-0.08	-0.09
Added bran to foods	0.06	-0.02	0.01	0.06	0.07
Proportion of variation explained	6.1%	4.6%	2.6%	2.3%	2.2%

The first principal component has notably high coefficients for wholemeal bread, rice, cheese, yoghurt, oily fish, various vegetables and salad vegetables, and vegetable dishes. It has low coefficients for white bread and teaspoons of sugar. This diet could be thought of as one in line with current dietary recommendations (Section 1.2). A participant who scored highly on this component would have been more likely to choose wholemeal bread over white, reduced-fat over full-fat milk, and white meat over red, for example. Generally she would have tended to eat more of foods perceived as 'healthy' such as fruit, vegetables and fish, and fewer 'unhealthy' foods such as chips, coke, chocolate and added sugar.

Many researchers have used the term 'prudent' to describe a major dietary pattern (Section 1.3.1). Table 27 reproduces some of the descriptions of this diet that have been given. Analysis of the PAH data identified a pattern very similar to the descriptions in Table 27, and therefore the first principal component in Table 26 is labelled a 'prudent' pattern.

Table 27 Descriptions of a prudent diet from the literature

Authors	Description
Slattery ML, Boucher KM et al.[56]	'one in which all types of fruits and vegetables were consumed and fish and poultry were eaten more often than red meat or processed meat; this dietary pattern was inversely associated with high sugar foods'
Hu FB, Rimm E et al.[72]	'loaded heavily with the following foods or food groups: vegetables, legumes, whole grains, fruit, fish and poultry'
Osler M, Heitmann BL et al.[60]	'predominantly reflecting frequent intakes of wholemeal bread, vegetables, fruits and fish'

The second principal component has high coefficients for Yorkshire pudding, pastries, biscuits, omelettes, most meat and meat products, roast potatoes and chips, and added sugar. The lowest coefficient is for 'bran' breakfast cereals. A participant scoring highly on this component would have eaten high-fat products frequently, chosen meat regularly, and chosen full-fat, rather than reduced-fat spread.

Many researchers, particularly those studying populations in the United States, have used the term 'Western' to describe another major dietary pattern (Section 1.3.1). Table 28 reproduces some of the descriptions that have been given.

Table 28 Descriptions of a Western diet from the literature

Authors	Description
Slattery ML, Boucher KM et al.[56]	'high levels of red meat, processed meat, fast food, refined grains, and sugar-containing foods, and low levels of vegetables (other than potatoes) and fruits, with the predominant fruit being canned fruit'
Hu FB, Rimm E et al.[72]	'loaded heavily with red meat, processed meat, refined grains, sweets and dessert, French fries, and high-fat dairy products'
Osler M, Heitmann BL et al.[60]	'positively associated with frequent intakes of meat, sausages, potatoes, butter and white bread'

Analysis of the PAH data identified a pattern with some features of the descriptions in Table 28. However, it is noticeable that, unlike the Western diet, most of the coefficients in Table 26 are positive, indicating a generally high-intake diet. Correlating individual dietary scores on the second component with early pregnancy food energy intake gave $r_S = 0.82$ ($P < 0.0001$, $n = 600$). Similar patterns have been found elsewhere; McCann et al.[61] described their second factor as 'high-fat', and Beaudry et al.[68] labelled their first factor 'high-energy density'. Since fat and energy intake are highly correlated ($r_S = 0.92$ in the PAH data), either 'high-fat' or 'high-energy' might describe this diet. The effects of both were investigated using a linear regression model with scores on the second principal component as the outcome, and food fat intake and food energy intake as predictors (Willett[128]). Both were highly significant predictors ($t = 6.3$, $P < 0.0001$ and $t = 11.9$, $P < 0.0001$, respectively), but energy intake was the most statistically significant. In an alternative analysis considering percentage of energy from food fat alongside total food energy intake as predictors, food energy intake remained much more significant ($t = 46.9$ compared to $t = 6.3$). Therefore the second principal component in Table 26 is best described as a 'high-energy' pattern.

The third component in Table 26 loads highly on vegetable dishes, vegetarian foods and crisps. On the other hand, low coefficients are seen for meats and minced meat dishes and several vegetables. Women with particularly high scores on this component were therefore likely to be eating a vegetarian-style diet, characterised by a low intake of meat, and a high intake of alternatives such as vegetarian foods and vegetable dishes. There appears to be a further facet to this diet: the tendency to eat more snack-type foods, such as crisps, chips, and confectionery.

The fourth component has high coefficients for wholemeal crackers, Yorkshire pudding, fish pie, tinned vegetables, other fruit juices, and tinned fruit. Low coefficients are seen for cooking fats and salad oils, beef, tomatoes, onion, nuts and red wine. The fifth component has high coefficients for cakes, buns, biscuits, full-fat spread, melon, other fruit juices and other confectionery, and low coefficients for reduced-fat spread, beans and pulses, roast potatoes and chips, blackcurrants and ribena, and mayonnaise and salad cream. Since neither component four or five seem to have a useful interpretation in terms of recognisable dietary patterns, they are not considered any further.

The scree plot from the PCA of the 100 foods in the early pregnancy PAH data is shown in Figure 14.

Figure 14 Scree plot from PCA of 100 foods in early pregnancy PAH data

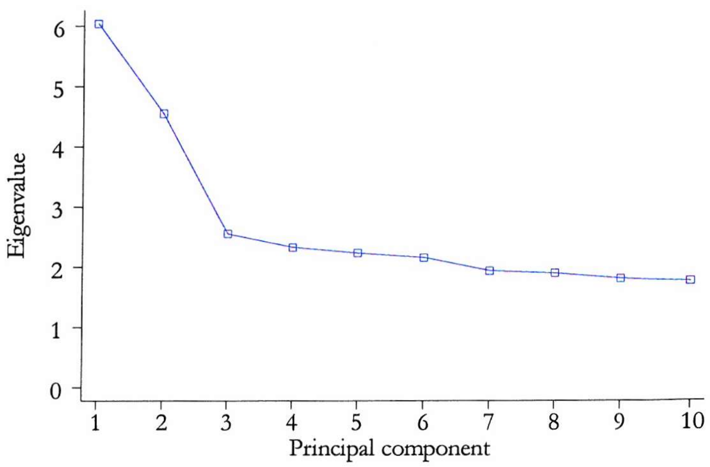


Figure 14 indicates that the first two components account for noticeably more variation in the data than the remaining components. The total variation retained by reducing the dimension of the data from 100 to 2 is 10.6%. However, as relevant as the amount of variation accounted for, is whether the coefficients describe a recognisable dietary pattern. Thus, in addition to the first two components indicated by the scree plot, the pattern revealed by the third principal component is of interest, and therefore this component is also investigated further.

4.1.2 Assessment of stability of patterns using late pregnancy PCA

Complete dietary information is available for 591 participants in late pregnancy. PCA was performed using this data to assess the stability of patterns between early and late pregnancy. For brevity and simplicity the table of coefficients resulting from the late pregnancy PCA is not presented here. Rather, since agreement between the early and late components is being assessed, Bland and Altman plots and 'limits of agreement'[203] are used to summarise the similarity between early and late pregnancy coefficients.

Bland and Altman plots display the difference between early and late pregnancy coefficients against the average of the two coefficients, for each component (Figure 15). A central line is plotted at the mean of these differences, indicating the average deviation in the differences from zero. Limits of agreement are plotted above and below the mean, enclosing 95% of the differences. Exact values for the limits of agreement are detailed in Table 29. For the first component, for example, the limits of agreement indicate that 95% of the coefficients in late pregnancy component 1 lie within –0.06 and 0.06 of their associated coefficients in early pregnancy.

Figure 15 Bland and Altman plots comparing PCA coefficients for early and late PAH data

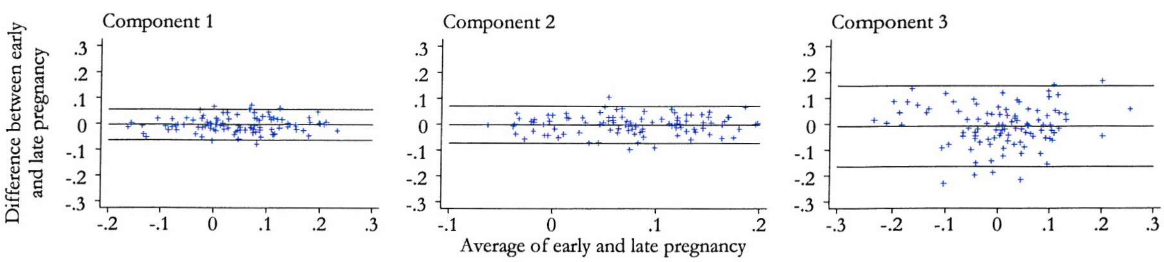


Table 29 Limits of agreement comparing PCA coefficients for early and late PAH data

Component	1	2	3
Limits of agreement	(-0.06, 0.06)	(-0.07, 0.07)	(-0.16, 0.15)

Limits of agreement for the first and second components are narrow, and thus indicate that prudent and high-energy dietary patterns for late pregnancy are strikingly similar to those for early pregnancy. The third component shows less agreement, and indeed inspection of the coefficients indicates that in late pregnancy, while there are similarly high coefficients for vegetarian foods and vegetable dishes, the negative loadings for meats are not as strong as in early pregnancy. Also, the third component in late pregnancy has an additional facet: a particularly low coefficient for reduced-fat spread, contrasting with very high coefficients for full-fat spread and oil use.

4.1.3 Grouping of food data

Table 26 is cumbersome, and since some foods, such as pork, lamb and beef, are eaten in similar ways and have comparable coefficient patterns, it would be helpful if such foods could be grouped together. Two approaches to food groupings were considered. Firstly an attempt was made to replicate two methods of grouping used in the literature, and secondly an independent, pragmatic reduction to 49 food groups was generated.

4.1.3.1 Replication of groupings used by Gregory et al.

Gregory et al.[66] have analysed the DNSBA, a detailed dietary survey from a representative sample of British adults. Fifty-four food groups were created from 84 original groups. Of these 54, 43 could be approximately recreated using the FFQ in the PAH study. Eleven of Gregory et al.'s categories could not be replicated, and in addition 5 food groups were used that did not correspond adequately with any of Gregory's groups. Thus 48 food groups were employed.

Participants' principal component scores were calculated for the first three components using all foods as described in Section 4.1, and using the 48 groups emulating Gregory et al.'s method. Spearman's correlation coefficients between participants' scores under the two methods (Table 30) indicate that little information is lost using Gregory's grouping in the first two components. The value of the third component using this grouping is less clear.

Table 30 Spearman's correlation coefficients for PAH PCA using 100 food groups compared with Gregory's 48 food groups

Component	Early pregnancy	Late pregnancy
1	0.90	0.96
2	0.87	0.97
3	0.60	0.62

4.1.3.2 Replication of groupings used by Hu et al.

Hu et al.[57,72,204] have promoted dietary pattern analysis using factor analysis. In particular they have found their dietary patterns to be reasonably reproducible and valid[57]. Their method condenses 131 foods into 40 food groups. Of these 40, 31 could be approximately recreated using the FFQ in the PAH study. Nine of Hu et al.'s categories could not be replicated, and in addition 8 food groups were used that did not correspond adequately with any of Hu's groups. Thus 39 food groups were employed, and again Table 31 indicates that relatively little information is lost using Hu's grouping in the first two components.

Table 31 Spearman's correlation coefficients for PAH PCA using 100 food groups compared with Hu's 39 food groups

Component	Early pregnancy	Late pregnancy
1	0.88	0.93
2	0.85	0.90
3	0.06	0.34

In early pregnancy the third component using Hu's grouping is completely unlike that using all 100 food groups (Table 31). However, since the components are ordered by the size of their eigenvalues, and the eigenvalues of the third and fourth components are 2.6 and 2.3 respectively (using 100 food groups), it could be anticipated that slightly differing analyses might not generate components in an identical order. Comparing the fourth early pregnancy component using Hu's grouping with the third component resulting from all 100 food groups gives a Spearman correlation coefficient (r_S) of 0.69, indicating a clear similarity between these components. In late pregnancy the 100-food third component appears less robust, not being picked up well by the third (Table 31), fourth (r_S= 0.42) or fifth (r_S= 0.38) component under Hu's grouping.

4.1.3.3 Pragmatic grouping

The 'pragmatic' grouping was developed as the most relevant means of combining foods from the Southampton FFQ of similar nutrient composition and comparable usage, as described in Table 32. A few foods were not condensed into a similar usage group since health or preparation considerations were likely to make the pattern of consumption of this food very different from the rest of the group. For example, 'tinned vegetables' were not combined with 'vegetables', and 'decaffeinated tea and coffee' were not combined with 'tea and coffee'. A miscellaneous category includes foods that are relatively unimportant in the diet in terms of frequency and quantity of intake, such as soup and condiments. Note that 'blackcurrants and ribena' were included as a 'high energy soft drink' because examination of the diary data from the PAH study indicated that women in this cohort consumed blackcurrant drinks far more frequently than blackcurrant fruit (Section 2.4.4).

Table 32 Pragmatic food grouping

All (n = 100)	Pragmatic (n = 49)
White bread	White bread
Wholemeal bread	Wholemeal bread
'Bran' breakfast cereals Other breakfast cereals	Breakfast cereals
Crackers Wholemeal crackers	Crackers
Pasta Rice	Rice and pasta
Quiche and pizza	Quiche and pizza
Yorkshire pudding and savoury pancakes	Yorkshire pudding and savoury pancakes
Cakes Buns Pastries Biscuits Other biscuits	Cakes and biscuits
Fruit based puddings Milk based puddings Other puddings Ice cream	Puddings
Full-fat liquid milk (pints) Full-fat powdered milk (pints)	Full-fat milk (pints)
Reduced-fat liquid milk (pints) Reduced-fat powdered milk (pints)	Reduced-fat milk (pints)
Cream	Cream
Cheese Cottage cheese	Cheese and cottage cheese
Yoghurt	Yoghurt
Eggs Omelette	Eggs and egg dishes
Full-fat spread	Full-fat spread
Reduced-fat spread	Reduced-fat spread
Cooking fats and salad oils	Cooking fats and salad oils
Chicken and turkey	Chicken and turkey

Table 32 (continued) Pragmatic food grouping

All (n = 100)	Pragmatic (n = 49)
Pork Lamb Beef Minced meat dishes	Red meat
Bacon and gammon Meat pies Sausage Ham and canned meats	Processed meat
Liver and kidney Pate and liver sausage Black pudding and faggots	Offal
White fish Fish pie Oily fish Shell fish	Fish and shell fish
Salad Coleslaw Tomatoes Green pepper & watercress	Salad vegetables
Spinach Sprouts and broccoli Cabbage and cauliflower Peas Courgettes, marrow & leeks	Green vegetables
Carrots Parsnips, swedes & turnips	Root vegetables
Sweetcorn Onion Mushrooms	Other vegetables
Beans and pulses	Beans and pulses
Vegetable dishes Vegetarian foods	Vegetable dishes and vegetarian foods
Tinned vegetables	Tinned vegetables
Boiled potatoes	Boiled potatoes
Roast potatoes and chips	Roast potatoes and chips
Crisps	Crisps
Oranges and orange juice Grapefruit & grapefruit juice	Citrus fruit and fruit juices
Apples and pears Banana Peaches, plums, apricots, cherries and grapes Strawberry and raspberry Melon, pineapple, mango and kiwi	Other fruit
Other fruit juices	Other fruit juices
Nuts	Nuts
Dried fruit	Dried fruit
Cooked fruit Tinned fruit	Cooked and tinned fruit
Chocolate Other confectionery	Confectionery
Sweet spreads	Sweet spreads
Added sugar (teaspoons)	Added sugar (teaspoons)
Blackcurrants and ribena Fizzy drinks and squashes Coke	High-energy soft drinks
Diet coke	Diet coke

Table 32 (continued) Pragmatic food grouping

All (n = 100)	Pragmatic (n = 49)
Tea Coffee	Tea and coffee
Decaffeinated tea and coffee	Decaffeinated tea and coffee
Drinking chocolate	Drinking chocolate
Red wine White wine and cider	Wine
Soup Meat stock cubes Mayonnaise and salad cream Pickles, chutney & ketchup Added bran to foods	Miscellaneous

Spearman's correlation coefficients between participants' scores under all foods and the pragmatic food grouping (Table 33) indicate that very little information is lost using the pragmatic grouping in the first two components. The stability of the third component using this grouping is less convincing.

Table 33 Spearman's correlation coefficients for PAH PCA using 100 food groups compared with 49 pragmatic food groups

Component	Early pregnancy	Late pregnancy
1	0.98	0.97
2	0.97	0.96
3	0.43	0.69

In conclusion, individual scores from the first two principal components appear stable under all the methods of grouping considered. This concurs with the findings from a similar analysis by McCann et al.[61] whose first two factors ('healthy' and 'high-fat') were replicated when two alternative methods of food-classification were considered. It is clear that the third component in the PAH is less stable, and such instability was also seen in other groupings investigated, but not reported here. McCann et al. similarly found less stability in their third and subsequent factors. Furthermore, the third component was seen to explain substantially less variation in food intake than the first two components (Figure 14), and therefore only these first two components are retained for use in further analysis.

The choice of grouping method remains. The pragmatic grouping was designed as an appropriate grouping of the FFQ under consideration. Furthermore, participants' scores using the pragmatic grouping produced notably high correlations with scores from all 100 foods (Table 33). Thus the pragmatic grouping is utilised in subsequent analyses.

Principal component analysis showed a high level of agreement between early and late pregnancy in prudent and high-energy coefficients using the pragmatic grouping (Table 34).

Table 34 Prudent and high-energy diet coefficients for pragmatic food grouping in early and late pregnancy PAH data

Food	Prudent diet		High-energy diet	
	Early pregnancy	Late pregnancy	Early pregnancy	Late pregnancy
White bread	-0.20	-0.21	0.14	0.19
Wholemeal bread	0.26	0.27	0.04	0.01
Breakfast cereals	0.06	0.06	0.09	0.08
Crackers	0.10	0.11	0.13	0.05
Rice and pasta	0.22	0.20	0.09	0.12
Quiche and pizza	0.08	0.05	0.08	0.10
Yorkshire pudding and savoury pancakes	-0.06	-0.11	0.25	0.17
Cakes and biscuits	-0.01	0.01	0.24	0.28
Puddings	0.05	0.07	0.27	0.19
Full-fat milk (pints)	-0.15	-0.22	0.19	0.13
Reduced-fat milk (pints)	0.15	0.19	-0.09	-0.08
Cream	0.09	0.13	0.14	0.14
Cheese and cottage cheese	0.21	0.22	0.03	0.12
Yoghurt	0.23	0.24	0.06	0.10
Eggs and egg dishes	0.00	0.00	0.18	0.15
Full-fat spread	-0.03	-0.05	0.18	0.22
Reduced-fat spread	0.03	0.06	-0.01	-0.06
Cooking fats and salad oils	0.01	0.01	0.13	0.18
Chicken and turkey	0.13	0.09	0.10	0.14
Red meat	0.05	0.01	0.23	0.25
Processed meat	-0.04	-0.07	0.28	0.28
Offal	-0.03	-0.07	0.15	0.13
Fish and shell fish	0.20	0.14	0.06	0.14
Salad vegetables	0.30	0.29	0.05	0.09
Green vegetables	0.21	0.17	0.09	0.15
Root vegetables	0.17	0.15	0.08	0.10
Other vegetables	0.27	0.23	0.10	0.11
Beans and pulses	0.04	0.02	0.12	0.09
Vegetable dishes and vegetarian foods	0.22	0.17	-0.02	0.00
Tinned vegetables	-0.16	-0.18	0.18	0.14
Boiled potatoes	0.01	-0.06	0.19	0.15
Roast potatoes and chips	-0.11	-0.19	0.22	0.19
Crisps	-0.15	-0.12	0.16	0.17
Citrus fruit and fruit juices	0.19	0.15	0.04	0.08
Other fruit	0.20	0.23	0.02	0.11
Other fruit juices	0.10	0.08	0.00	-0.01
Nuts	0.11	0.07	0.09	0.02
Dried fruit	0.18	0.18	0.00	0.01
Cooked and tinned fruit	0.13	0.06	0.15	0.15
Confectionery	-0.10	-0.07	0.22	0.22
Sweet spreads	0.12	0.10	0.09	0.10
Added sugar (teaspoons)	-0.20	-0.23	0.22	0.17
High-energy soft drinks	-0.09	-0.07	0.15	0.15
Diet coke	-0.01	-0.01	0.01	-0.03
Tea and coffee	-0.13	-0.13	0.17	0.18
Decaffeinated tea and coffee	0.10	0.12	-0.03	-0.07
Drinking chocolate	0.01	0.03	0.11	0.06
Wine	0.04	0.15	-0.03	-0.03
Miscellaneous	0.09	0.06	0.21	0.25
Proportion of variation explained	8.7%	9.0%	7.1%	6.1%

In order to better understand the dietary scores, and because nutrients ultimately drive metabolic processes, correlations between dietary scores and selected macro and micro-nutrients calculated from food intakes (Section 2.3.4.7) are shown in Table 35 and Table 36.

Table 35 Spearman's correlation coefficient (P-value) between nutrients and prudent diet scores for pragmatic food grouping in early and late pregnancy PAH data

	Prudent diet	
Nutrient	**Early pregnancy**	**Late pregnancy**
Calories	**0.02** (0.67)	**-0.04** (0.34)
Fat	**-0.02** (0.60)	**-0.12** (0.003)
Protein	**0.27** (<0.001)	**0.20** (<0.001)
Total carbohydrate	**-0.01** (0.76)	**-0.04** (0.34)
Carbohydrate: Starch	**-0.08** (0.06)	**-0.09** (0.03)
Carbohydrate: Sugars	**0.06** (0.11)	**0.01** (0.88)
Calcium	**0.24** (<0.001)	**0.15** (<0.001)
Beta-carotene	**0.59** (<0.001)	**0.53** (<0.001)
Total vitamin A	**0.31** (<0.001)	**0.20** (<0.001)
Thiamin	**0.38** (<0.001)	**0.33** (<0.001)
Riboflavin	**0.25** (<0.001)	**0.17** (<0.001)
Niacin	**0.37** (<0.001)	**0.36** (<0.001)
Vitamin C	**0.49** (<0.001)	**0.46** (<0.001)
Vitamin E	**0.49** (<0.001)	**0.48** (<0.001)
Iron	**0.31** (<0.001)	**0.32** (<0.001)
Folate	**0.54** (<0.001)	**0.46** (<0.001)
Zinc	**0.34** (<0.001)	**0.27** (<0.001)

Table 36 Spearman's correlation coefficients (P-value) between nutrients and high-energy diet scores for pragmatic food grouping in early and late pregnancy PAH data

	High-energy diet	
Nutrient	**Early pregnancy**	**Late pregnancy**
Calories	**0.87** (<0.001)	**0.87** (<0.001)
Fat	**0.85** (<0.001)	**0.84** (<0.001)
Protein	**0.73** (<0.001)	**0.77** (<0.001)
Total carbohydrate	**0.77** (<0.001)	**0.78** (<0.001)
Carbohydrate: Starch	**0.69** (<0.001)	**0.68** (<0.001)
Carbohydrate: Sugars	**0.66** (<0.001)	**0.66** (<0.001)
Calcium	**0.51** (<0.001)	**0.49** (<0.001)
Beta-carotene	**0.29** (<0.001)	**0.40** (<0.001)
Total vitamin A	**0.57** (<0.001)	**0.56** (<0.001)
Thiamin	**0.63** (<0.001)	**0.64** (<0.001)
Riboflavin	**0.53** (<0.001)	**0.49** (<0.001)
Niacin	**0.62** (<0.001)	**0.64** (<0.001)
Vitamin C	**0.28** (<0.001)	**0.35** (<0.001)
Vitamin E	**0.52** (<0.001)	**0.59** (<0.001)
Iron	**0.62** (<0.001)	**0.65** (<0.001)
Folate	**0.51** (<0.001)	**0.54** (<0.001)
Zinc	**0.70** (<0.001)	**0.72** (<0.001)

The prudent dietary pattern is unrelated to energy intake, but is strongly positively correlated with all micronutrients, particularly beta-carotene, vitamin C, vitamin E and folate. It shows a positive association with protein intake, a small negative association with fat and starch, and a small positive association with sugars. Associations between the prudent dietary pattern and nutrient densities were also explored, but these were very similar to the correlations with crude nutrient values because the prudent dietary pattern is itself unrelated to energy intake.

The high-energy dietary pattern is strikingly highly correlated with energy intake. In general the high-energy diet score is strongly positively associated with all macro and micronutrients, reflecting the large intake in all dimensions characterised by this dietary pattern. Correlations with nutrient densities were also calculated, and associations were much smaller than those for crude nutrient values and often non-significant. A few correlations, including those for calcium, were significantly negative.

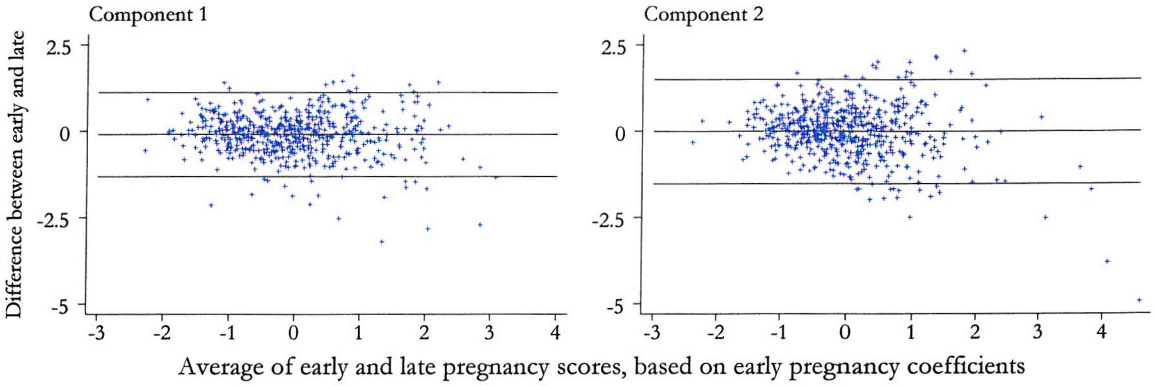
These results give credence to the use of the 'prudent' and 'high-energy' names. Fung et al.'s[71] finding that the prudent dietary pattern is associated with a biomarker profile described as favourable supports the associations with the prudent dietary pattern. It is clear that the dietary scores provide a useful summary of patterns of food intake that are related to nutrient consumption. Indeed the patterns may capture aspects of diet that are not apparent from nutrient-based analyses, such as the interactions between nutrients. For the prudent dietary pattern these factors contribute to the associations seen in the literature with reduced overall mortality[60], and reduced risk of coronary heart disease[72] and diabetes[67].

4.1.4 Do women's individual diets change from early to late pregnancy?

The similarity between early and late pregnancy coefficients in Table 34 indicates only that diets in early and late pregnancy varied along very similar axes. There remains the possibility that entirely different women were eating prudent and high-energy diets in early compared to late pregnancy. This scenario was investigated by comparing individual's scores at these two time points. However, it is only reasonable to make a direct comparison if these scores are generated based on the same set of coefficients. Therefore scores were calculated for each woman in early and late pregnancy, both based on the coefficients for early pregnancy in Table 34.

Since agreement of participants' scores between early and late pregnancy is being investigated, Bland and Altman plots are appropriate (Figure 16). Figure 16 differs from Figure 15 in that agreement of individual scores is being assessed, rather than agreement of PCA coefficients. The magnitude of scores generated from a PCA must be understood for an important difference in scores on the Bland and Altman plot to be judged. Since principal component scores are produced on an arbitrary scale, both early and late pregnancy scores were divided by the standard deviation of the relevant early pregnancy score before the creation of the plot. Thus an individual difference of –1 is interpreted as a decrease of one standard deviation from her early pregnancy score.

Figure 16 Bland and Altman plots comparing PAH participants' scores in early and late pregnancy, based on early PCA coefficients



The limits of agreement for component 1 in Figure 16 are –1.3 to 1.1; a participant's late pregnancy score (based on early pregnancy coefficients) is likely to be within approximately 1 standard deviation of her early pregnancy score (based on the same coefficients). Throughout this thesis limits of agreement on a standardised variable of ± 1 will be considered to be good, ± 1.5 to be moderate, and ± 2 to be weak. Thus there is a moderate-to-good degree of agreement between early and late pregnancy prudent diet scores based on early pregnancy coefficients. There is a slight tendency for participants' prudent diet scores to drop in late pregnancy. This drop is 0.10 on average and is statistically significant ($P < 0.001$).

The limits of agreement for component 2 in Figure 16 are –1.5 to 1.5. Thus there is a moderate degree of agreement between early and late pregnancy high-energy diet scores based on early pregnancy coefficients, with an individual's late pregnancy score likely to be within 1.5 standard deviations of her early pregnancy score. This agreement is slightly less than for prudent diet scores. The drop between early and late pregnancy for the high-energy score is 0.02 on average, and is not statistically significant ($P = 0.60$).

The converse analysis to the above, calculating both early and late pregnancy scores based on the late pregnancy coefficients in Table 34, is equally valid. As expected, it was found to yield very similar results.

4.2 Factor analysis

Having produced robust and interpretable dietary patterns using PCA, the additional level of sophistication provided by factor analysis, as commonly used in the literature, is considered here.

The most important consideration in factor analysis is the number of factors retained (Section 2.4.3.1). As previously described (Section 4.1.3.3) the first two components from the PCA were the most helpful in terms of the scree plot, stability, and interpretability. For these reasons two factors are considered here.

Spearman's correlation coefficients between scores from PCF, and varimax rotation following PCF, PFA, IPF and ML were calculated for the first two factors from the PAH data (Table 37). Participants' scores are correlated here, rather than loadings, because loadings in factor analysis, although superficially similar to coefficients in PCA, are not used to produce a linear function alone, but an error term is also included in the model (Section 2.4.3.2). Therefore there is no strict analogy between PCA coefficients and FA loadings, and scores for each individual are used instead.

Table 37 Spearman's correlation coefficients of PAH participants' scores using different methods of factor analysis in PAH data

	PCF (u)	PCF (v)	PFA (v)	IPF (v)	ML (v)
PCF (u)	1				
PCF (v)	**0.957** 0.945 **0.887** 0.872	1			
PFA (v)	**0.957** 0.935 **0.894** 0.850	**0.997** 0.996 **0.998** 0.991	1		
IPF (v)	**0.960** 0.943 **0.904** 0.867	**0.998** 0.999 **0.997** 0.998	**0.999** 0.996 **0.998** 0.993	1	
ML (v)	**0.959** 0.943 **0.900** 0.861	**0.997** 0.999 **0.997** 0.998	**0.998** 0.996 **0.998** 0.993	**0.999** 1.000 **1.000** 1.000	1

u = unrotated, v = varimax rotated
Early pregnancy factor 1, early pregnancy factor 2, **late pregnancy factor 1,** late pregnancy factor 2

It is clear from Table 37 that initial method of extraction before varimax rotation makes an imperceptible practical difference to the factor scores (compare PFA (v), IPF (v) and ML (v) with PCF (v)). This is the case for both factors 1 and 2, and for early and late pregnancy. Joliffe and Morgan's assertion[141] is confirmed: PCF gives very similar results to other extraction methods.

Table 37 indicates that scores from unrotated PCF (PCF (u)) are very similar to those for any of the rotated methods. Rotating seems to have more of an impact on the late pregnancy data, but nevertheless it is striking that all correlation coefficients are above 0.85. Thus there is a high level of association between unrotated PCF scores and rotated factor scores, and relatively little gain in introducing the additional complexity of rotation. Investigation revealed no improvement in interpretability of the factors resulting from rotation. Furthermore, unrotated PCs are to be preferred since they are designed to indicate the independent axes along which participants vary the most. The ability to differentiate between individuals is an attractive feature of a dietary score. Therefore the PCA (unrotated PCF) scores created in Section 4.1 are used as final dietary scores.

4.3 Principal component analysis using diary data

4.3.1 Do the diary data produce the same dietary patterns as the FFQ data?

Women completing the early pregnancy interview in the PAH study were asked to keep a four-day food diary following the visit. 588 women kept diaries for the full four days. This additional information permitted validation of the FFQ[127], and is utilised here to examine how dietary patterns resulting from PCA on the diary data compare to those from PCA on the FFQ data.

To allow equivalent analysis of the food diaries it was necessary for the diary information to be coded into FFQ categories. The food diary data in this format aimed to capture how the woman would have completed an FFQ covering the four day period of her diary. Therefore frequencies with which foods were mentioned in the food diaries were used, rather than amounts. The frequencies of food or food group use per four days were converted to the equivalent frequency per week.

The 49 pragmatic food groupings were utilised, and PCA performed on the diary data. The resulting coefficients from the first and second components are shown in Table 38, along with the equivalent coefficients from the early pregnancy FFQ. The pattern of coefficients in the first diary component is seen to be very similar to that for the first FFQ component, and is therefore labelled a 'prudent' diet. The pattern of coefficients in the second diary component is also similar to that for the second FFQ component and the correlation between the second diary component and calculated energy intake in the diaries is $r_S = 0.58$. Therefore the second diary component is also labelled a 'high-energy' diet.

Table 38 Prudent and high-energy diet coefficients for pragmatic food grouping in early pregnancy PAH data using FFQ and diaries

	Prudent diet		High-energy diet	
Food	**FFQ**	**Diary**	**FFQ**	**Diary**
White bread	-0.20	-0.16	0.14	0.23
Wholemeal bread	0.26	0.28	0.04	-0.10
Breakfast cereals	0.06	0.17	0.09	0.03
Crackers	0.10	0.09	0.13	0.12
Rice and pasta	0.22	0.12	0.09	0.00
Quiche and pizza	0.08	0.00	0.08	-0.01
Yorkshire pudding and savoury pancakes	-0.06	-0.05	0.25	0.21
Cakes and biscuits	-0.01	0.16	0.24	0.17
Puddings	0.05	0.11	0.27	0.09
Full-fat milk (pints)	-0.15	-0.13	0.19	0.24
Reduced-fat milk (pints)	0.15	0.18	-0.09	-0.12
Cream	0.09	0.17	0.14	0.12
Cheese and cottage cheese	0.21	0.21	0.03	0.14
Yoghurt	0.23	0.20	0.06	-0.04
Eggs and egg dishes	0.00	0.05	0.18	0.20
Full-fat spread	-0.03	0.04	0.18	0.37
Reduced-fat spread	0.03	0.07	-0.01	-0.21
Cooking fats and salad oils	0.01	-0.10	0.13	0.28
Chicken and turkey	0.13	0.06	0.10	0.03
Red meat	0.05	-0.04	0.23	0.15
Processed meat	-0.04	-0.05	0.28	0.09
Offal	-0.03	0.02	0.15	-0.08
Fish and shell fish	0.20	0.15	0.06	-0.05
Salad vegetables	0.30	0.22	0.05	0.03
Green vegetables	0.21	0.20	0.09	0.14
Root vegetables	0.17	0.17	0.08	0.02
Other vegetables	0.27	0.24	0.10	0.13
Beans and pulses	0.04	-0.01	0.12	0.00
Vegetable dishes and vegetarian foods	0.22	0.12	-0.02	-0.02
Tinned vegetables	-0.16	-0.16	0.18	0.03
Boiled potatoes	0.01	0.11	0.19	-0.05
Roast potatoes and chips	-0.11	-0.22	0.22	0.22
Crisps	-0.15	-0.05	0.16	0.22
Citrus fruit and fruit juices	0.19	0.23	0.04	0.03
Other fruit	0.20	0.25	0.02	0.01
Other fruit juices	0.10	0.09	0.00	-0.03
Nuts	0.11	0.15	0.09	0.07
Dried fruit	0.18	0.16	0.00	0.06
Cooked and tinned fruit	0.13	0.12	0.15	0.01
Confectionery	-0.10	0.11	0.22	0.24
Sweet spreads	0.12	0.18	0.09	0.08
Added sugar (teaspoons)	-0.20	-0.14	0.22	0.20
High-energy soft drinks	-0.09	0.01	0.15	0.16
Diet coke	-0.01	0.01	0.01	-0.04
Tea and coffee	-0.13	-0.03	0.17	0.21
Decaffeinated tea and coffee	0.10	0.13	-0.03	-0.11

Table 38 (continued) Prudent and high-energy diet coefficients for pragmatic food grouping in early pregnancy PAH data using FFQ and diaries

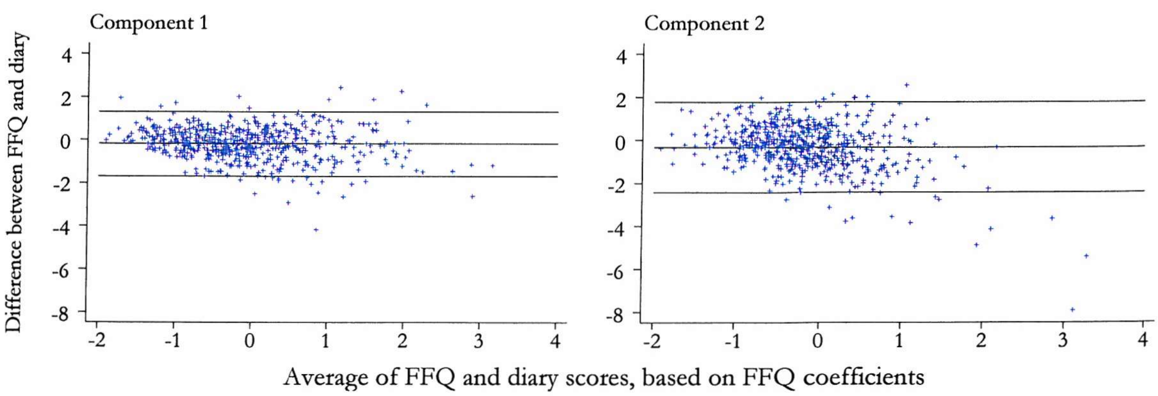
	Prudent diet		High-energy diet	
Food	**FFQ**	**Diary**	**FFQ**	**Diary**
Drinking chocolate	**0.01**	**0.05**	0.11	0.06
Wine	**0.04**	0.14	**-0.03**	**0.02**
Miscellaneous	0.09	0.17	0.21	0.24
Proportion of variation explained	8.7%	9.7%	7.1%	4.8%

Confectionery, and cakes and biscuits have notably higher coefficients on the first diary component. Full-fat spread has a noticeably higher coefficient, and processed meat, puddings, tinned vegetables, boiled potatoes, offal and reduced-fat spread have lower coefficients on the second diary component. These differences mean that the first two principal components have slightly different emphases between the diary and FFQ analyses. However, generally the agreement between the sets of coefficients is high, and the interpretations of prudent and high-energy dietary patterns are appropriate in both instances.

4.3.2 Do individual women's diets differ between FFQ and diary reports?

Although the prudent and high-energy dietary patterns in Table 38 are similar it is nevertheless possible that different women reported prudent and high-energy diets using the FFQ and the food diaries. Therefore Bland and Altman plots were used to assess agreement between individual scores using the FFQ and diaries. This analysis is equivalent to that done to compare early and late pregnancy scores from the FFQ (Section 4.1.4), and again all scores are calculated based on early pregnancy FFQ coefficients, and divided by the standard deviation of the relevant early pregnancy FFQ score.

Figure 17 Bland and Altman plots comparing PAH participants' scores from the early pregnancy FFQ and diary, based on FFQ coefficients



There is a tendency for scores on component 1 to be 0.18 standard deviations lower in the diaries than the FFQ, with Bland and Altman limits of agreement –1.69 to 1.32. The average drop from the FFQ to diaries for component 2 is -0.33 with limits of agreement –2.43 to 1.78. Therefore there is moderate agreement between diary and FFQ scores on the first component, but somewhat less agreement on the second component.

However, since the principal aim of an FFQ is to rank women according to nutrient intake[127], it is perhaps most pertinent to compare scores from the FFQ and diary analyses using Spearman's correlation coefficient. For the prudent diet $r_S = 0.68$, and for the high-energy diet $r_S = 0.24$. These correlations compare favourably with those found for nutrient intake, calculated using FFQ and diary data from the PAH study, which ranged from 0.26 to 0.44[127]. The correlations are also comparable with an equivalent analysis comparing FFQ and diary data by Hu et al.[57] who evaluated results from administration of two FFQs one year apart and two one-week diaries kept within this period. Intake from the two diaries was averaged. Correlations between the diary prudent pattern and the two FFQ scores were 0.34 and 0.41. Correlations between the diary Western pattern and the two FFQ scores were 0.51 and 0.64.

Discrepancies between FFQ and diary data may be a result of the period of reference. The FFQ data refers to the three months before the early pregnancy interview (median gestation 15.3 weeks), whereas the diary data refers to four days following the interview. In early pregnancy the issue of nausea is particularly pertinent. Nausea typically peaks between 8-12 weeks gestation, and declines thereafter[205], and may therefore be an important difference between the two periods. Women were asked to indicate at the early pregnancy interview if they had experienced any nausea since becoming pregnant, and if so whether this was mild (nausea only), moderate (sometimes sick) or severe (regularly sick, unable to retain meals).

Previous analyses of the PAH study have found that correlations between nutrients fell as the degree of nausea experienced rose[127]. The authors note that this is 'consistent with the suggestion that dietary changes caused by nausea may well have been less evident by the time our food diaries were kept'. Correlations between diary and FFQ prudent and high-energy scores were therefore calculated in each of the nausea groups (Table 39). However, there is no trend apparent across these groups, indicating that nausea has little effect on the maintenance of similar dietary patterns in the early stages of pregnancy.

Table 39 Spearman's correlations between PAH diary and FFQ scores for prudent and high-energy diets, by nausea status in early pregnancy

Nausea status	Prudent diet	High-energy diet	n
None	0.62	0.23	91
Mild	0.69	0.24	223
Moderate	0.68	0.24	193
Severe	0.68	0.32	78

Women in the PAH study were also asked in early pregnancy whether the amount they had eaten since becoming pregnant was more, the same or less than before pregnancy. This information may give an additional indication of women whose diets are less stable in the early stages of pregnancy. The correlations between prudent diet scores from the diaries and FFQs did not differ much across these groups (Table 40), although there is an indication that those claiming to eat less since becoming pregnant have a higher correlation between their diary and FFQ high-energy scores.

Table 40 Spearman's correlations between PAH diary and FFQ scores for prudent and high-energy diets, by changes in amount eaten since becoming pregnant

Amount	Prudent diet	High-energy diet	n
Less	0.70	0.34	196
Same	0.71	0.18	187
More	0.63	0.14	196

When the food diaries were collected the nurse recorded subjective information about diary quality. This information was used to investigate whether correlations between diary and FFQ data were dependent on diary quality. Table 41 indicates that the 29 women whose diaries were coded as 'very poor, probably incomplete' have notably lower correlations. Thus it appears that discrepancies between individual PC scores from FFQ and diary data may be partly the result of poor completion of the diaries by some women.

Table 41 Spearman's correlations between PAH diary and FFQ scores for prudent and high-energy diets, by diary quality

Diary quality	Prudent diet	High-energy diet	n
Excellent	0.64	0.26	148
Average	0.67	0.26	376
Very poor, probably incomplete	0.34	0.01	29

A further PCA was performed only on data from women whose diaries were coded as 'excellent' or 'average', without the 29 of 'very poor' quality, and the 32 who completed food diaries but for whom diary quality had not been recorded by the nurse. However, this made very little difference to the dietary patterns produced; no coefficient differed by more than 0.04 from the analysis on the whole sample.

4.3.3 Is the validity of the dietary scores affected by over or under-reporting?

Calculation of food patterns using both the FFQ and diary data provides an opportunity to consider how the validity of the dietary scores is affected by over and under-reporting. This is assessed using the ratio of energy intake to BMR, predicted from each woman's age and weight using the Schofield formulae[47]. An allowance of 0.2MJ/day was added to the pre-pregnancy BMR as recommended by Prentice et al.[206].

It is known that some women may over-report their intakes using the Southampton MRC FFQ (Section 2.3.2). Thus correlations between FFQ and diary scores by energy intake to BMR ratio calculated using FFQ energy intake are given in Table 42.

Table 42 Spearman's correlations between PAH diary and FFQ scores for prudent and high-energy diets, by FFQ energy intake to BMR ratio in early pregnancy

FFQ energy intake to BMR ratio	Prudent diet	High-energy diet	n
– 1.45	0.73	0.17	189
– 1.85	0.68	0.12	198
> 1.85	0.64	0.23	182

Intakes recorded using diaries in the PAH study are likely to be subject to some under-reporting (Section 2.3.2). Therefore an equivalent analysis was performed using the BMR ratio calculated according to diary intake (Table 43).

Table 43 Spearman's correlations between PAH diary and FFQ scores for prudent and high-energy diets, by diary energy intake to BMR ratio in early pregnancy

Diary energy intake to BMR ratio	Prudent diet	High-energy diet	n
– 1.15	0.70	0.12	179
– 1.50	0.66	0.11	200
> 1.50	0.70	0.26	190

These results indicate that there is little effect of over-reporting in the FFQ or under-reporting in the diaries on the correlations between the FFQ and diary scores.

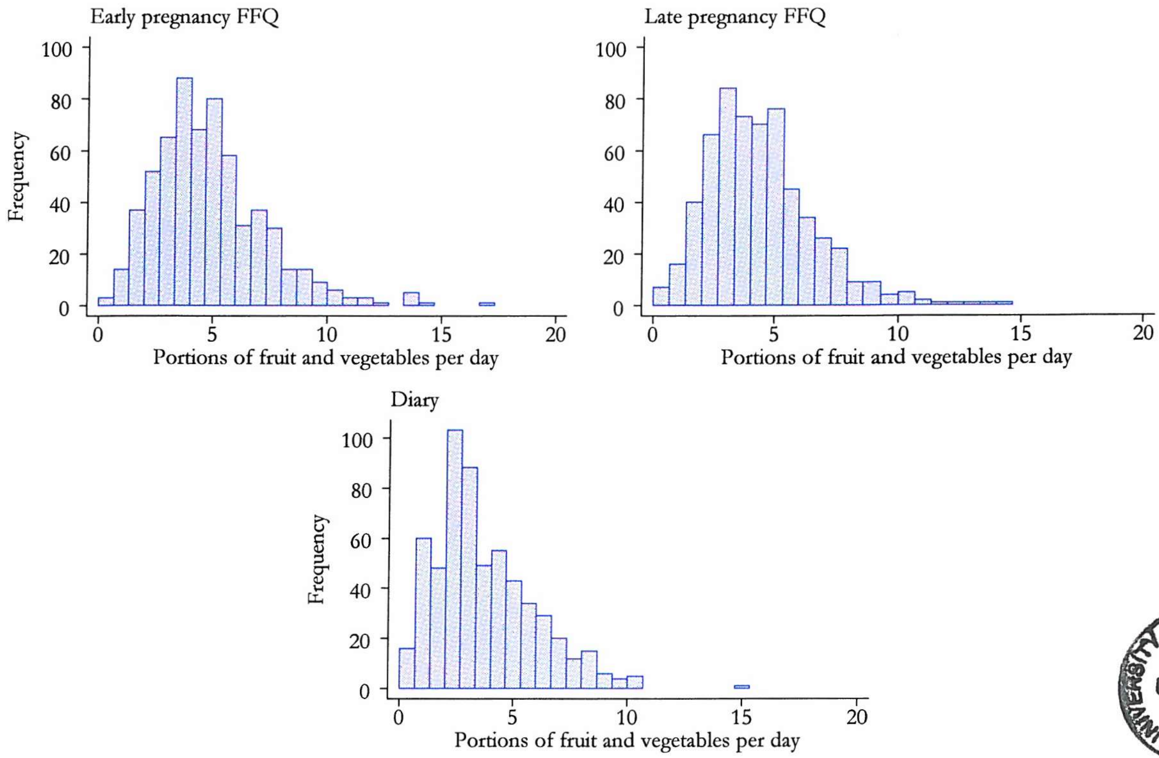
When considering the validity of the FFQ prudent diet score it is also of interest to note that the Spearman's correlation between the score in early pregnancy and serum vitamin C is 0.30. Although the prudent diet score is affected by intake of foods other than those containing vitamin C, and serum measurements cannot themselves be considered a gold standard[129], this correlation gives some indication that the FFQ prudent diet score is positively associated with an external measure of micronutrient intake. The serum vitamin C measurements are correlated with vitamin C intake in early pregnancy ($r_S = 0.21$).

4.4 Fruit and vegetables

Assessing fruit and vegetable intake provides an indication of women who endeavour to eat in accordance with public health advice, since this is an important contemporary dietary guideline (Section 1.2). Comparisons between early and late pregnancy FFQ, and diary consumption are of interest, as well as associations with the PCA dietary scores

Fruit and vegetable consumption was calculated from FFQ data in early and late pregnancy, and from diary data, using the criteria described in Section 2.4.4. The result was divided by seven to give portions per day, and the distributions of scores are shown in Figure 18.

Figure 18 Distribution of portions of fruit and vegetables per day in the PAH study



4.4.1 Early compared to late pregnancy FFQ

Between early and late pregnancy there was a tendency for a slight reduction in portions of fruit and vegetables, by 0.4 portions per day on average. There was reasonable agreement between portions per day: Bland and Altman limits of agreement are –4.1 to 3.2. Note that 'portions of fruit and vegetables' is not a standardised variable and therefore the limits of agreement do not have the same interpretation as described in Section 4.1.4. In early pregnancy 41% of the women claimed to achieve the recommended five portions of fruit and vegetables per day, compared to 34% of the women in late pregnancy.

4.4.2 Diaries compared to early pregnancy FFQ

There was a tendency for subjects to report more fruit and vegetable intake using the FFQ than when completing the food diaries. On average each participant reported 1.2 more portions per day using the FFQ than the food diary. Only 26% of the women recorded achieving the recommended five portions of fruit and vegetables per day in their diaries. Other researchers have also observed this phenomenon (Section 2.3.2), which is attributed to both over-reporting on FFQs (perhaps due to a high number of similar FFQ categories) and under-reporting in food diaries (by reducing intake during the four day period, or not recording all foods consumed). However, fruit and vegetable consumption is ranked similarly in early pregnancy using diaries and FFQs ($r_S = 0.50$).

4.4.3 Fruit and vegetable intake compared to prudent and high-energy dietary scores

It might be expected that, since fruit and vegetables contribute a number of high coefficients to the prudent diet score (Table 34), portions of fruit and vegetables and the prudent diet score might be correlated. There is indeed a strikingly high correlation (Table 44). The correlation between fruit and vegetable intake and high-energy diet score is positive, but weaker, particularly for the diary data (Table 44).

Table 44 Spearman's correlations between PAH fruit and vegetable intake and prudent and high-energy diet scores

Food data	Prudent diet	High-energy diet	n
Early pregnancy FFQ	0.77	0.20	617
Late pregnancy FFQ	0.72	0.37	591
Diaries	0.82	0.03	582

4.5 Summary

- Principal component analysis on FFQ data yielded prudent and high-energy dietary patterns in both early and late pregnancy. These patterns were found to be robust across three different methods of food grouping. The 'pragmatic' food grouping was chosen, and the same analysis in late pregnancy produced very similar coefficients, indicating stability of the patterns across these two time points. When individual's scores were compared between early pregnancy and late pregnancy a moderate level of agreement was found; a participant's late pregnancy prudent diet score is likely to be within approximately 1 standard deviation of her early pregnancy score, and her high-energy diet score is likely to be within 1.5 standard deviations. Results from various factor analysis techniques were found to be little different from PCA scores, and are therefore not considered further.

- Data from food diaries kept by 588 of the women in the PAH study were translated into an equivalent format to the FFQ data. Principal component analysis revealed prudent and high-energy patterns similar to those from the FFQ data. Women's scores from the diary and early pregnancy FFQ are positively correlated for both the prudent pattern ($r_S = 0.68$) and the high-energy pattern ($r_S = 0.24$). These correlations do not differ greatly according to nausea status, or whether amount consumed had changed since becoming pregnant, but are notably lower amongst those whose diary quality was considered very poor, probably incomplete.

- Fruit and vegetable intake is lower in late pregnancy than early pregnancy, using FFQ data. It is also reported to be lower using food diaries than FFQs in early pregnancy, although there is a fair amount of correlation between the two ($r_S = 0.50$). Both FFQ and diary data show very strong positive correlations between prudent diet scores, and fruit and vegetable intake.

- In summary, PCA prudent and high-energy dietary scores are informative summary measures of diet. Individual change in diet between early and late pregnancy is also of interest, based on early pregnancy PCA coefficients. Sociodemographic and lifestyle predictors of these dietary characteristics will be investigated in Chapter 5. Predictors of prudent and high-energy diet scores from the diary data are not considered further since the diary scores are positively correlated with those from the FFQ analysis; the diary data has the disadvantages that it is not available in late pregnancy, and that FFQs, rather than diaries, are regarded as the primary method of measuring dietary intake in epidemiological studies[128]. Furthermore, fruit and vegetable intake is not utilised as a separate dietary outcome as it is strikingly highly correlated with prudent diet score, and the prudent diet score is preferred since it is calculated using all foods measured by the FFQ, not just fruit and vegetables.

5 Sociodemographic and lifestyle predictors of diet in the PAH study

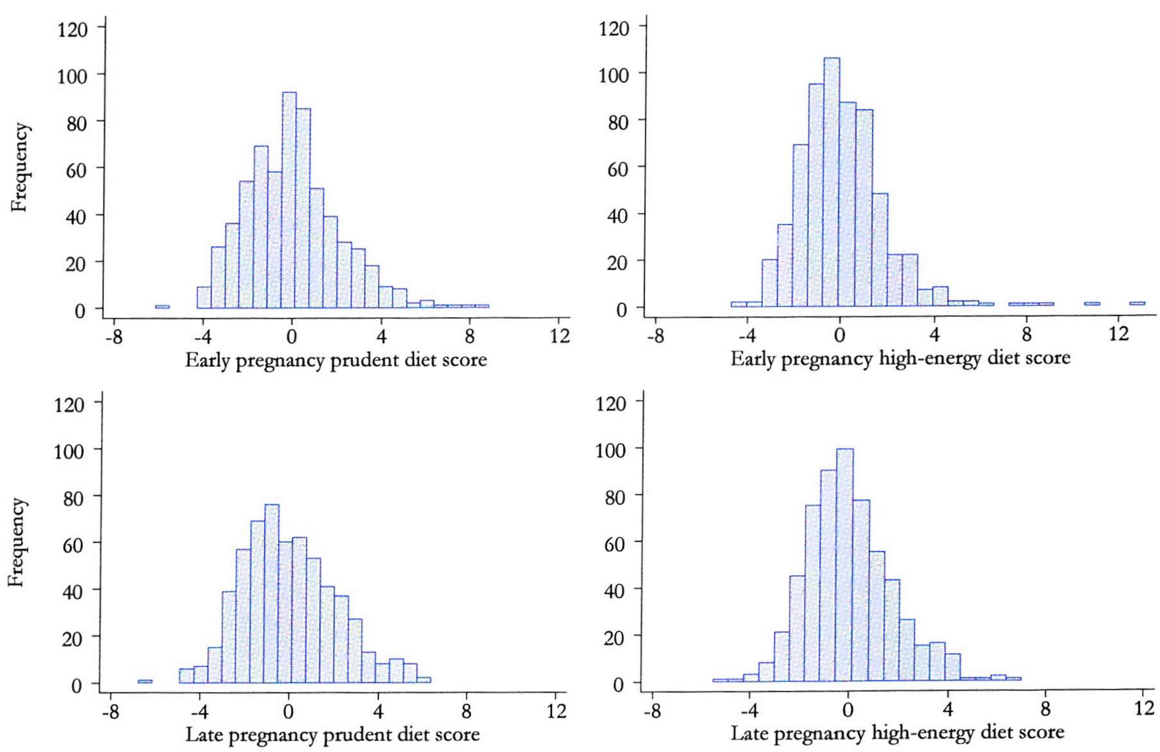
Four outcome variables are studied in this chapter as measures of diet in the PAH study. These are prudent and high-energy dietary scores on the early and late pregnancy FFQ data, defined using principal component analysis. Sociodemographic and lifestyle predictors of these scores are examined, as well as changes in scores between early and late pregnancy.

Details of the procedure for investigating the most important predictors of the prudent diet score defined using the early pregnancy data are presented comprehensively in Section 5.3. Subsequently only final models for other outcomes found using the same procedures are shown in Section 5.4. Predictors of change in prudent and high-energy scores between early and late pregnancy are considered in Section 5.5.

5.1 Data preparation

Early and late pregnancy prudent and high-energy diet scores are described in Figure 19.

Figure 19 Distributions of PAH early and late pregnancy prudent and high-energy scores



Multivariate linear regression with diet score as the outcome enables assessment of the most significant predictors of dietary patterns resulting from PCA. This statistical technique is based on the assumption that the outcome variable follows an approximately normal distribution. However, the dietary scores were positively skewed, particularly in early pregnancy (Figure 19). The log transformation was not found to be useful since some dietary scores were negative, and the addition of a constant to each value led to difficulty in interpretation of the results of regression models. Instead Fisher-Yates normal scores were utilised[207]. These rank each woman by the dietary score under consideration. An inverse-normal transformation is applied to the woman's rank divided by (n + 1), where n is the total number in the sample. This gives an equivalent observation from a standard normal distribution.

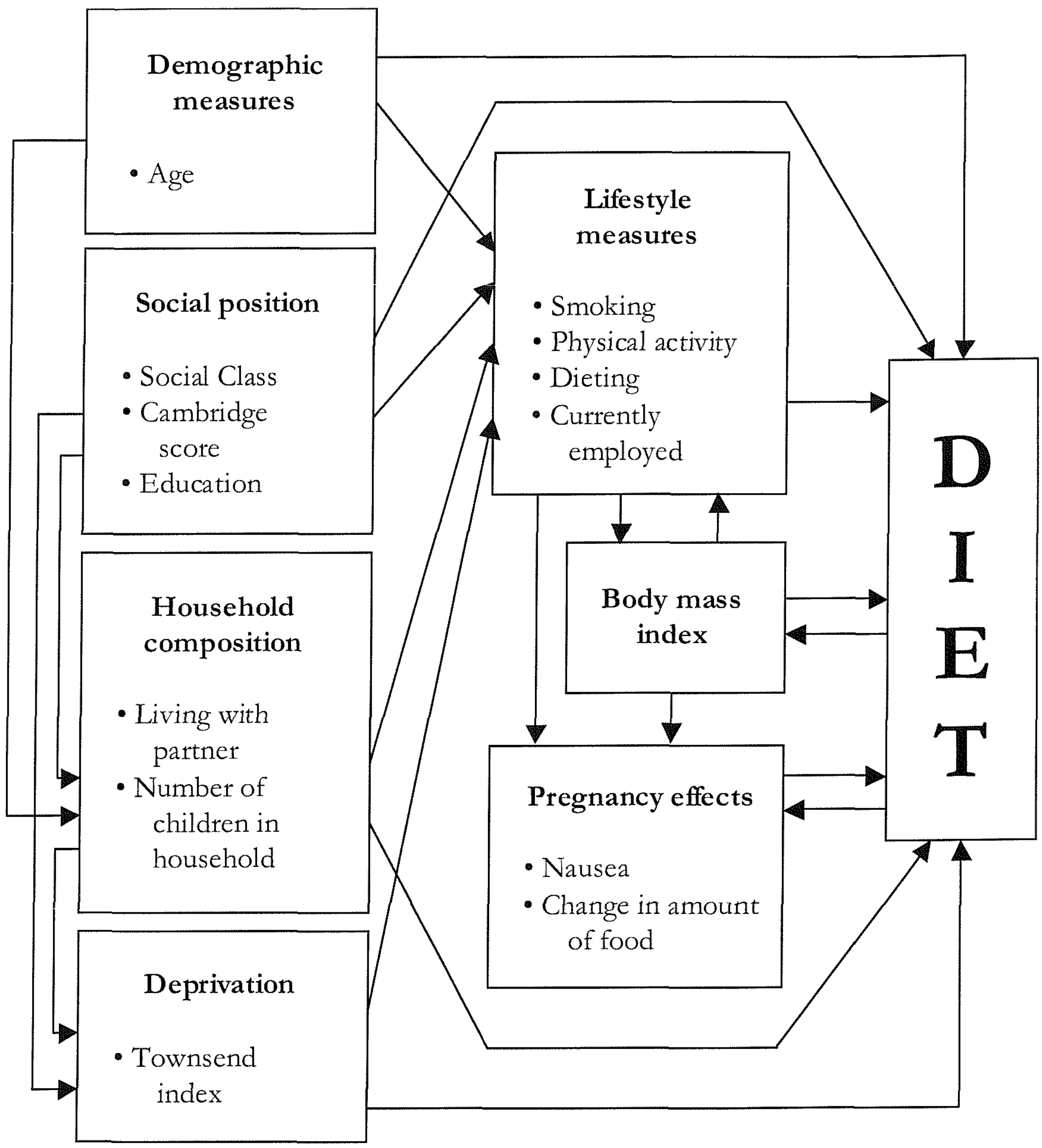
The Fisher-Yates-transformed dietary scores are thus by definition normally distributed, with a mean of 0 and a standard deviation of 1. Original scores that were already normally distributed are standardised, but otherwise unchanged. Original scores that were not normally distributed are transformed to normality; thus their relative magnitude is lost, but ranking is always preserved.

5.2 Framework of pathways of influence on diet

Figure 20 shows a proposed framework of the major pathways of influence of the dietary predictors available in the PAH study (Table 25). Older women are more likely to be living with a partner and to have children. Social position, as defined by Social Class, Cambridge score and education, also affects household composition; women of a higher social position more often delay marriage and childbearing[208]. Additionally, social position is clearly an influence on deprivation since women of a lower social position are likely to be exposed to greater deprivation. Household composition also affects deprivation; more members of a household cause a greater drain on the resources available. All of these factors have an influence on diet, directly and also partly through their effects on lifestyle measures, such as smoking, physical activity and current employment. Lifestyle measures in turn influence diet directly, and through BMI, which itself can affect lifestyle measures. BMI is a predictor of diet[128], and is also affected by diet. In the pregnant PAH population the effects of pregnancy are additional important dietary predictors. Nausea and change in amount of food consumed may be influenced by lifestyle measures and body mass index, and are closely associated with diet.

Analysis proceeds using the framework of causal pathways illustrated in Figure 20. This approach is considered more illuminating than simply building a model based on statistical significance. Instead models are firstly built for the combined effects of each group of variables on diet. Multivariate analysis then allows consideration of how much the effects shown to be important in these direct analyses are mediated through other factors thought to be on the causal pathway.

Figure 20 Framework of proposed major pathways of influence of dietary predictors in the PAH study



In the following analyses a cut-off of 0.3 is chosen to indicate significance because the more conventional cut-off of 0.05 is considered too stringent; it is possible that non-significant univariate predictors might have significant multivariate effects. A summary model is then built up using all predictors with P-values of less than 0.3. Once multivariate relationships are being considered within the summary model, the standard P-value cut-off of 0.05 is used to identify significant predictors. The proportion of variation in the outcome explained by predictors in a model is encapsulated by the R^2 statistic. All tables of regression models presented in this chapter are mutually adjusted.

5.3 Predictors of early pregnancy prudent diet score

5.3.1 Demographic measures

The effect of age on early pregnancy prudent diet score is displayed in Table 45. For clarity of presentation, the coefficients shown are the change in score expected from a 5 year increase in age. Thus older women in the study ate more prudently in early pregnancy; an increase in age of 5 years was associated with an increase in prudent diet score of 0.37 standard deviations, with a 95% confidence interval from 0.29 to 0.45. This effect is statistically significant, and age explains 13.3% of the variation in prudent diet score (adjusted for the number of variables in the model).

Table 45 Age as a predictor of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
Age†	0.37	(0.29, 0.45)	9.6	< 0.001

Adjusted R^2 = 13.3%, n = 600 †Per 5 year increase

5.3.2 Social position

The Cambridge score, SC and qualifications are all available in the PAH study as measures of social position (Table 25). Since they are all measuring similar concepts, it is important to understand how they are associated.

Figure 21 illustrates the strong correlation between the Cambridge score and SC (Pearson's correlation coefficient (r) = -0.85). This is perhaps not surprising since both are dominant household measures relating to 'social status' or 'prestige'.

The Cambridge score is slightly more strongly correlated with qualifications (r = 0.62) than SC is (r = -0.54). Note that correlations with SC are negative since being in a class of lower numeric value indicates a higher social position.

All three of the measures of social position are highly significant predictors of early pregnancy prudent diet score. However, Cambridge score is a more significant predictor than SC, and a linear regression model with both Cambridge score and SC as predictors indicates that Cambridge score is a significant predictor of prudent diet score once adjustment is made for SC, but SC has no effect independent of Cambridge score. Qualifications and Cambridge score are both positive independent predictors of early pregnancy prudent diet score (Table 46). Note that for presentation the Cambridge score has been scaled to have a standard deviation of one. Thus, at each qualification level, a one standard deviation increase in Cambridge score is associated with an increase of 0.30 standard deviations in the prudent diet score (95% CI 0.21 to 0.39).

Figure 21 Cambridge score and Social Class in the PAH study

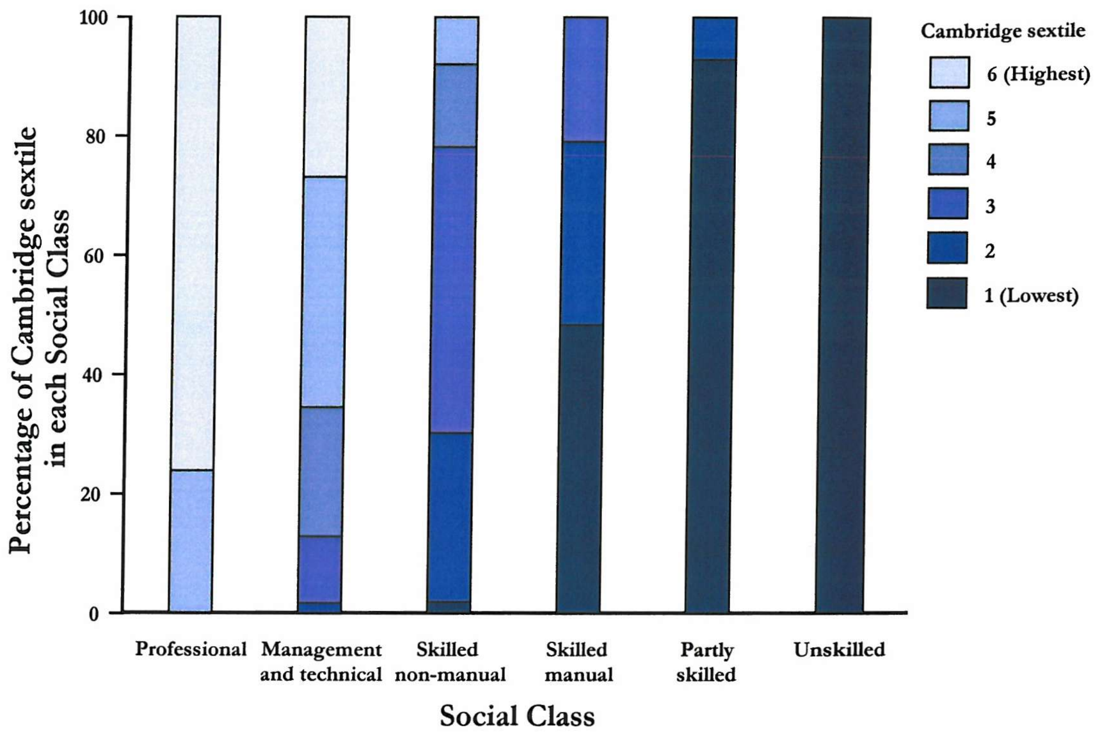


Table 46 Social position measures as predictors of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
Cambridge score*	0.30	(0.21, 0.39)	6.8	< 0.001
Qualifications	0.21	(0.14, 0.27)	6.2	< 0.001

Adjusted R^2 = 26.8%, n = 592 *Scaled to a standard deviation of 1

5.3.3 Household composition

There is a slight trend for women living with a partner to have more children (Table 47).

Table 47 Living with a partner and number of children in the PAH study

Live with partner	Number of children 0	1	2	3+	Total
No	35 (70%)	13 (26%)	1 (2%)	1 (2%)	50 (100%)
Yes	282 (52%)	177 (33%)	54 (10%)	29 (5%)	542 (100%)
Total	317 (54%)	190 (32%)	55 (9%)	30 (5%)	592 (100%)

Both aspects of household composition are seen to have independent effects on early pregnancy prudent diet score (Table 48). Women living with a partner and those with fewer children were somewhat more likely to eat prudently in early pregnancy. Investigation of the effect of living with children revealed a particularly marked effect amongst those living with two or more children, rather than a difference between simply living with children and not living with children.

Table 48 Household composition measures as predictors of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
Living with partner	0.35	(0.06, 0.63)	2.4	0.02
Number of children	-0.11	(-0.20, -0.01)	-2.2	0.03

Adjusted R^2 = 1.3%, n = 589

5.3.4 Deprivation

The Townsend index is an area-based measure of deprivation, and is the only assessment of deprivation available in the PAH study. The variable was scaled to have a standard deviation of one for presentation. Table 45 shows that an increase in Townsend index of one standard deviation leads to an average decrease in the early pregnancy prudent diet score of 0.29 standard deviations.

Table 49 Townsend index as a predictor of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
Townsend index*	-0.29	(-0.37, -0.21)	-7.1	< 0.001

Adjusted R^2 = 8.5%, n = 528 *Scaled to a standard deviation of 1

5.3.5 Lifestyle measures

Women in the PAH study who smoked were less likely to take strenuous exercise, and less likely to be currently employed (Table 50). However, there was no relationship between strenuous exercise and current employment, and no relationship between dieting and any of the other lifestyle measures.

Table 50 Smoking, strenuous exercise and current employment in the PAH study

Smoker	Percentage taking strenuous exercise	Percentage currently employed
No	20%	63%
Yes	12%	47%
Total	18%	59%

Smoking, strenuous exercise and current employment are all seen to have significant independent effects on early pregnancy prudent diet score (Table 51). Smokers were much less likely to eat a prudent diet. PAH participants who took strenuous exercise and were currently employed were more likely to eat a prudent diet.

Table 51 Lifestyle measures as predictors of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
Smoking	-0.65	(-0.82, -0.47)	-7.3	< 0.001
Strenuous exercise	0.29	(0.09, 0.49)	2.9	0.004
Currently employed	0.21	(0.06, 0.37)	2.7	0.007

Adjusted R^2 = 11.6%, n = 599

5.3.6 Body mass index

BMI is positively skewed (Figure 12), and was therefore logged before analyses. There was no linear association between a woman's pre-pregnant BMI and prudent diet score (Table 52). However, further investigation indicated that a quadratic relationship is present (Table 53). Since BMI is squared the effect is more difficult to interpret, and therefore Figure 22 shows the relationship between BMI and early pregnancy prudent diet score. Women with particularly low or high BMIs had lower prudent diet scores than those with average BMIs.

Table 52 Body mass index as a predictor of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
BMI‡	-0.01	(-0.09, 0.07)	-0.2	0.84

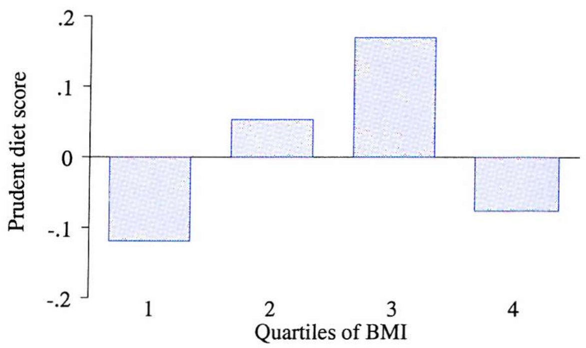
Adjusted R^2 = 0.0%, n = 586 ‡Logged and scaled to a standard deviation of 1

Table 53 Body mass index as a quadratic predictor of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
BMI‡	0.06	(-0.03, 0.15)	1.3	0.19
BMI‡ squared	-0.08	(-0.12, -0.03)	-3.1	0.002

Adjusted R^2 = 1.3%, n = 586 ‡Logged and scaled to a standard deviation of 1

Figure 22 Body mass index and PAH early pregnancy prudent diet score



5.3.7 Pregnancy effects

As might be expected, nausea and change in amount of food are negatively associated (Table 54).

Table 54 Nausea and change in amount of food in the PAH study

	Change in amount of food			
Nausea status	**Less**	**Same**	**More**	**Total**
None	13 (14%)	43 (47%)	35 (38%)	91 (100%)
Mild	63 (27%)	81 (34%)	91 (39%)	235 (100%)
Moderate	79 (40%)	56 (28%)	63 (32%)	198 (100%)
Severe	47 (59%)	17 (22%)	15 (19%)	79 (100%)
Total	202 (34%)	197 (33%)	204 (34%)	603 (100%)

Women experiencing nausea tended to eat somewhat less prudently, independent of change in amount of food consumed (Table 55). Women eating more food tended to eat slightly more prudently; although this effect is relatively small the P-value is less than the cut-off of 0.3 specified in Section 5.2 and therefore this variable is considered in the summary model.

Table 55 Pregnancy effects as predictors of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
Nausea	-0.10	(-0.19, -0.01)	2.2	0.03
Change in amount of food	0.05	(-0.04, 0.15)	1.1	0.28

Adjusted R^2 = 1.2%, n = 600

5.3.8 Summary model

A model of the joint effects of age and social position is fitted below. Household composition terms are added to this model, to assess whether the effects of age and social position are mediated through household composition (Figure 20). It may be that the effects of age, social position and household composition on diet act partially through deprivation, and therefore Townsend index is subsequently added to the model. All measures then significant in the model may have direct effects, or may be mediated via lifestyle measures, and so the significant lifestyle variables are fitted. The quadratic effect of BMI was significant (Table 52) and so this is next considered as a predictor in the multivariate model. Finally, the pregnancy effects variables are added to the model of hitherto significant predictors. Decisions about variables to retain are based solely on statistical significance, using the P-value cut-off of 0.05 described in Section 5.2.

Age and social position are seen to have independent effects on early pregnancy prudent diet score (Table 56).

Table 56 Age and social position measures as predictors of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
Age†	0.21	(0.13, 0.28)	5.6	< 0.001
Cambridge score*	0.24	(0.15, 0.33)	5.4	< 0.001
Qualifications	0.19	(0.13, 0.26)	5.9	< 0.001

Adjusted R^2 = 30.3%, n = 592 †Per 5 year increase *Scaled to a standard deviation of 1

Once these variables are controlled for, number of children also has an independent effect on diet, but living with a partner does not (Table 57). This is largely because living with a partner is strongly related to age; the average age of those living with a partner is 26.7, whereas the average age of those not living with a partner is 22.2. Thus living with a partner has very little effect on prudent diet score independent of age, and is dropped from the model. The coefficients of age, Cambridge score and qualifications are broadly unchanged by the addition of number of children to the model, indicating that all have independent effects on diet.

Table 57 Age, social position and household composition measures as predictors of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
Age†	0.26	(0.18, 0.34)	6.3	< 0.001
Cambridge score*	0.23	(0.14, 0.32)	5.0	< 0.001
Qualifications	0.17	(0.11, 0.24)	5.2	< 0.001
Living with partner	-0.18	(-0.44, 0.08)	-1.4	0.17
Number of children	-0.12	(-0.21, -0.03)	-2.6	0.008

Adjusted R^2 = 30.5%, n = 581 †Per 5 year increase *Scaled to a standard deviation of 1

Once age, social position and number of children are controlled for, Townsend index has no independent affect on diet (Table 58); the level of deprivation of the area the woman lives in is relatively unimportant in predicting her early pregnancy prudent diet score.

Table 58 Age, social position, number of children and Townsend index measures as predictors of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
Age†	0.22	(0.14, 0.31)	5.0	< 0.001
Cambridge score*	0.19	(0.09, 0.28)	3.9	< 0.001
Qualifications	0.21	(0.14, 0.28)	6.0	< 0.001
Number of children	-0.11	(-0.20, -0.03)	-2.5	0.01
Townsend index*	-0.06	(-0.14, 0.02)	-1.6	0.12

Adjusted R^2 = 33.4%, n = 508 †Per 5 year increase *Scaled to a standard deviation of 1

Exercise and current employment don't have any independent effects on early pregnancy prudent diet score (Table 59). Further investigation indicated that this is mainly due to the effect of social position. Once a woman's social position is controlled for there is very little additional independent effect of exercise or current employment. This is largely because a woman's frequency of exercise and chance of current employment are strongly related to her qualifications (Table 60).

Table 59 Age, social position, number of children and lifestyle measures as predictors of PAH early pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
Age†	0.26	(0.18, 0.34)	6.4	< 0.001
Cambridge score*	0.19	(0.10, 0.27)	4.2	< 0.001
Qualifications	0.16	(0.10, 0.23)	5.0	< 0.001
Number of children	-0.16	(-0.25, -0.06)	-3.2	0.001
Smoking	-0.37	(-0.52, -0.21)	-4.6	< 0.001
Strenuous exercise	0.10	(-0.08, 0.27)	1.1	0.28
Currently employed	-0.11	(-0.26, 0.04)	-1.5	0.15

Adjusted R^2 = 33.7%, n = 591 †Per 5 year increase *Scaled to a standard deviation of 1

Table 60 Qualifications, exercise and current employment in the PAH study

Qualifications	Percentage taking strenuous exercise	Percentage currently employed
None	5%	27%
CSE	10%	55%
O-levels	17%	58%
A-levels	23%	72%
HND	33%	61%
Degree	27%	71%

Smoking, however, does have a strong independent effect on diet, and the addition of this variable to the model does not reduce the effects of other terms already included. Nausea and change in amount of food do not have significant independent effects on early pregnancy prudent diet score. The reduction in significance of nausea is partly due to a negative association between nausea and age, as found by O'Brien and Zhou[209]. Quadratic BMI terms are also no longer significant predictors once the other factors in the model are accounted for. Further investigation indicated that this is due to a quadratic relationship between social position and BMI; women of higher social position (as measured by qualifications or Cambridge score) are less likely to have very high or very low BMIs.

The final model of all significant independent predictors of a prudent diet score is shown in Table 61. No two-way interactions between the predictors were statistically significant.

Table 61 All significant independent predictors of PAH early pregnancy prudent diet score

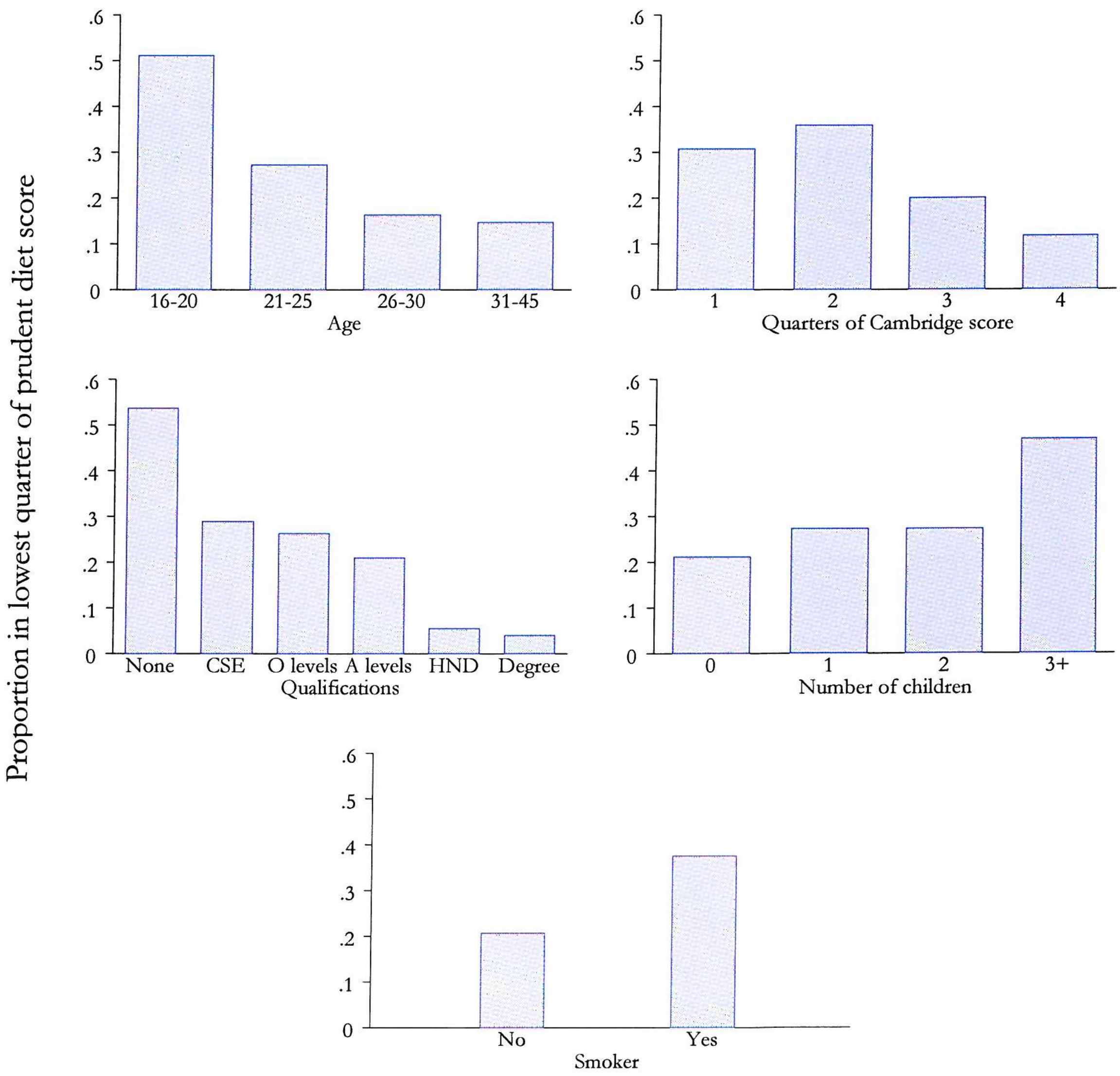
Predictor	β	95% CI for β	t	P
Age†	0.25	(0.17, 0.33)	6.2	< 0.001
Cambridge score*	0.18	(0.10, 0.27)	4.2	< 0.001
Qualifications	0.17	(0.10, 0.23)	5.1	< 0.001
Number of children	-0.12	(-0.21, -0.04)	-2.8	0.005
Smoking	-0.37	(-0.52, 0.21)	-4.6	< 0.001

Adjusted R^2 = 33.5%, n = 592 †Per 5 year increase *Scaled to a standard deviation of 1

As a graphical indication of the individual effects of each of these predictors, women were assigned to four equal-sized groups, according to quarters of the prudent diet score. The proportion of women with prudent diet scores in the lowest quarter is used as a summary measure. For each predictor a graph in Figure 23 shows the proportion of women in the lowest quarter of the prudent diet score by levels of the predictor, adjusted for the other factors in the model. Thus Figure 23 illustrates how a woman in the lowest quarter of prudent diet scores is likely to be younger, have a lower Cambridge score, have fewer educational qualifications, live with more children and smoke.

All the predictors are highly statistically significant, but it is notable that the effects of age and education are particularly strong. Figure 23 and further investigation revealed that the positive effects of age and qualifications are linear. Thus equal increases in age (or qualifications) result in approximately equal increases in prudent diet score. The model in Table 61 indicates that a high Cambridge score and not smoking are slightly less important predictors of a prudent diet, but are nonetheless highly significant. The effect of Cambridge score is greatest amongst those with scores above the median. The number of children a woman lives with is the least important of the factors in influencing her food choices on a prudent dietary scale.

Figure 23 Proportion of women in the lowest quarter of PAH early pregnancy prudent diet score, by significant predictors of the score



5.4 Predictors of other dietary scores

5.4.1 Late pregnancy prudent diet score

The question of whether predictors of the prudent diet score are similar in late pregnancy to those in early pregnancy is relevant. Also, the pattern of predictors of the high-energy diet score in both early and late pregnancy is of interest. Therefore the same procedure as that in Section 5.3 was used to generate linear regression models summarising predictors of these dietary outcomes. The final models are presented below, alongside figures illustrating the proportion of women in the lowest quarter of the diet scores by the relevant predictors.

Table 62 All significant predictors of PAH late pregnancy prudent diet score

Predictor	β	95% CI for β	t	P
Age†	0.26	(0.19, 0.34)	6.6	< 0.001
Cambridge score*	0.17	(0.09, 0.26)	3.9	< 0.001
Qualifications	0.20	(0.13, 0.26)	6.1	< 0.001
Number of children	-0.20	(-0.28, -0.12)	-4.8	< 0.001
Townsend index*	-0.08	(-0.15, -0.01)	-2.2	0.03
Smoking	-0.43	(-0.60, -0.27)	-5.2	< 0.001

Adjusted $R^2 = 46.7\%$, n = 500 †Per 5 year increase *Scaled to a standard deviation of 1

Figure 24 Proportion of women in the lowest quarter of PAH late pregnancy prudent diet score, by significant predictors of the score

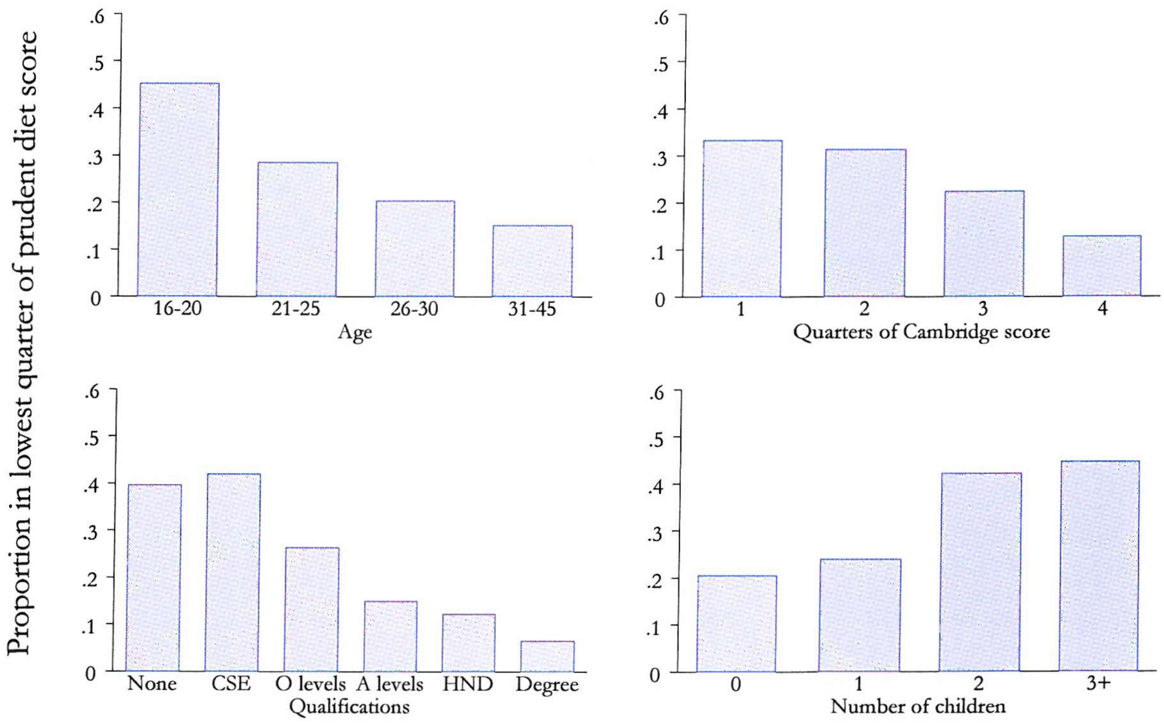
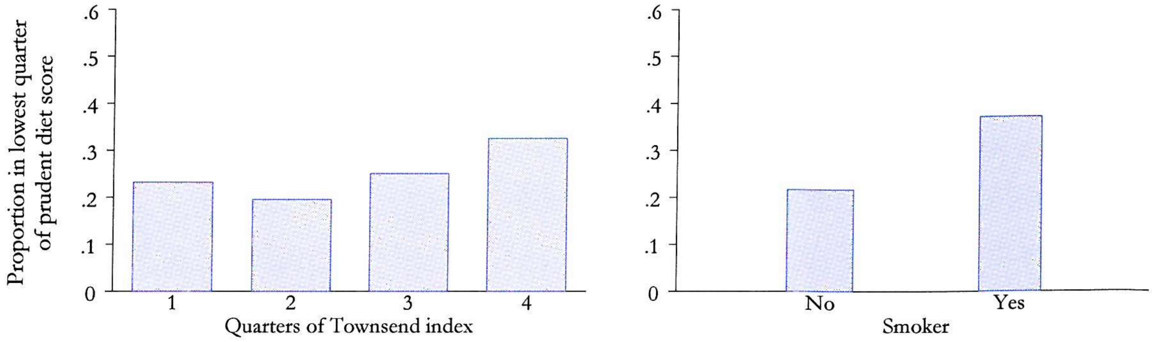


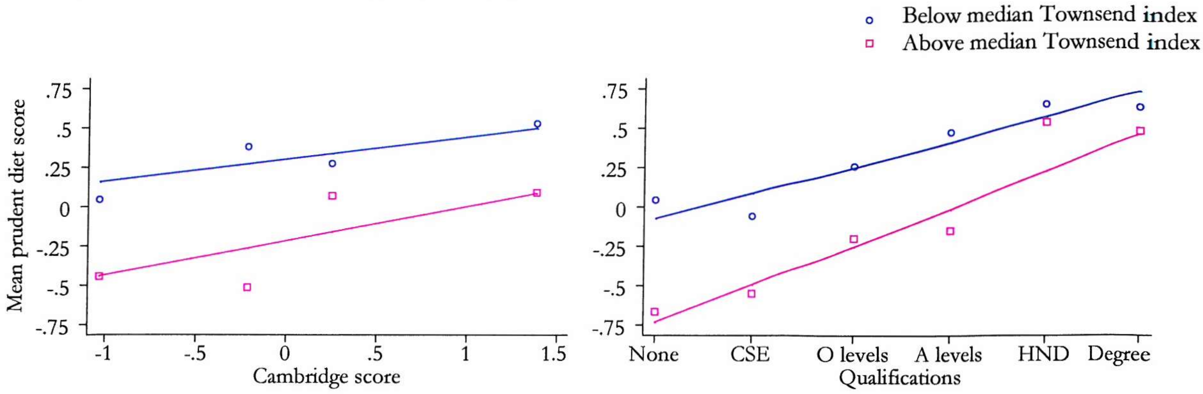
Figure 24 (continued) Proportion of women in the lowest quarter of PAH late pregnancy prudent diet score, by significant predictors of the score



The similarity between the significant effects on prudent diet scores in late and early pregnancy is striking (Table 61 and Table 62). The effects of age and qualifications are again the strongest with coefficients virtually identical to those in early pregnancy. The positive effect of Cambridge score and negative effect of smoking are also of very similar magnitude in late compared to early pregnancy, and the influence of Cambridge score is again greatest amongst those women with Cambridge scores above the median. The effect of number of children is somewhat stronger in late pregnancy. There is an additional effect of Townsend index although this is the least important of the factors influencing late pregnancy prudent diet score. While the coefficients are very similar between early and late pregnancy, all except that for Cambridge score are larger in late pregnancy. Therefore the predictors explain a larger proportion of variation in the prudent diet score in late pregnancy than in early pregnancy; in fact the R^2 statistic is strikingly high in late pregnancy (46.7%).

Two-way interactions between variables in Table 62 were tested, and the interactions between the Townsend index score and both measures of social position were statistically significant. As an illustration of this effect the women were divided into two groups, according to whether their Townsend score was above or below the median. The effects of Cambridge score and qualifications on average prudent diet score (adjusted for other factors in the model) amongst those with high and low Townsend index scores are illustrated in Figure 25.

Figure 25 Interactions between Townsend index and measures of social position as predictors of PAH late pregnancy prudent diet score



When the regression model in Table 62 (without Townsend index) is fitted separately for those above and below the median Townsend index score, a one standard deviation increase in Cambridge score is associated with an increase in late pregnancy prudent diet score of 0.14 amongst women who are less deprived, and with an increase of 0.22 amongst women who are more deprived. Similarly an increase in qualifications by one level is associated with an increase in late pregnancy prudent diet score of 0.16 amongst women who are less deprived and with an increase of 0.24 amongst women who are more deprived. Therefore the significant interactions show that the positive effect of a woman's social position on her late pregnancy prudent diet score is greater amongst those who live in areas with a higher Townsend index, indicating more deprivation. The converse is also true: the negative effect of deprivation is greater amongst those of a lower social position.

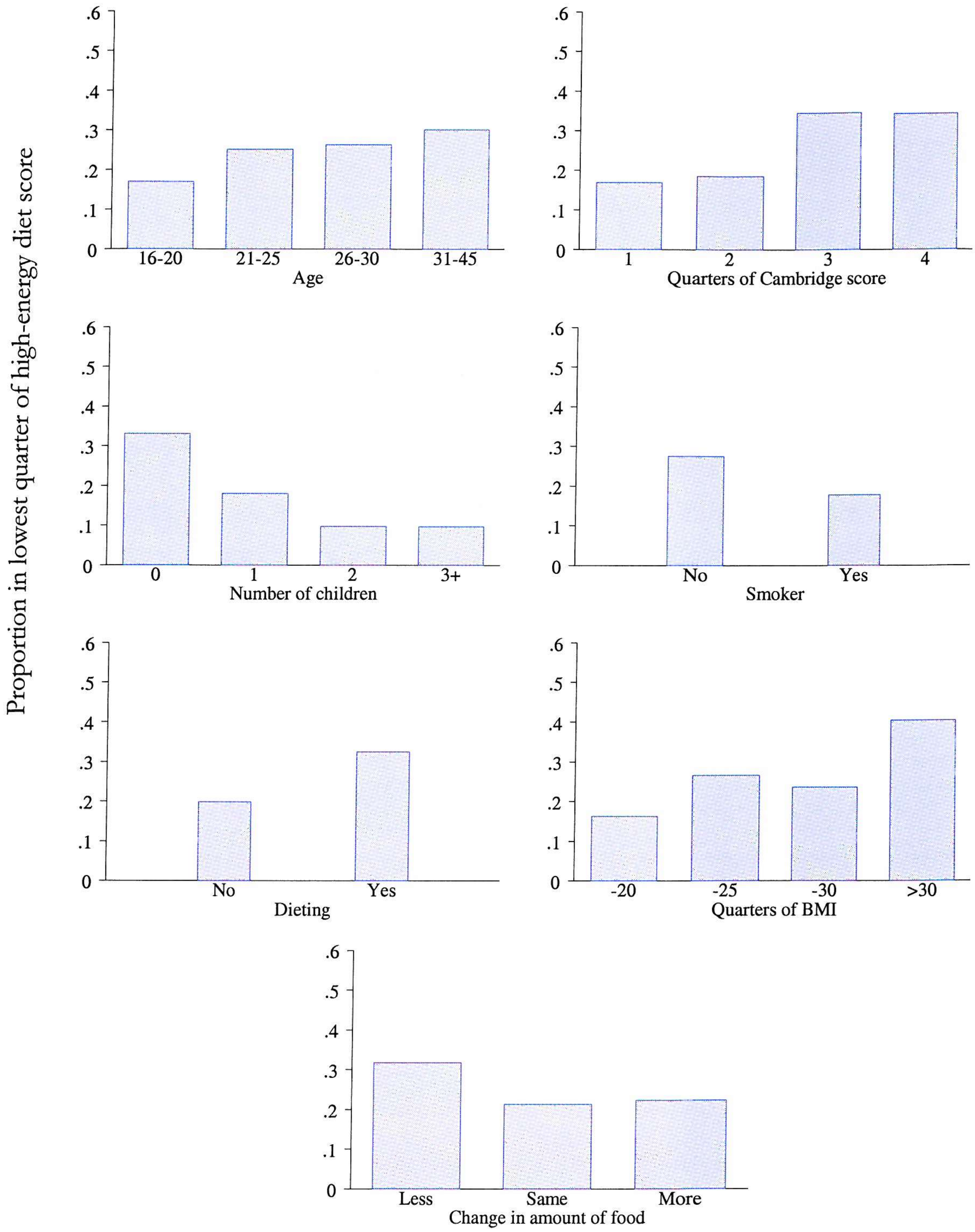
5.4.2 Early pregnancy high-energy diet score

Table 63 All significant predictors of PAH early pregnancy high-energy diet score

Predictor	**β**	**95% CI for β**	**t**	**P**
Age†	-0.13	(-0.22, -0.05)	-3.1	0.002
Cambridge score*	-0.19	(-0.28, -0.11)	-4.6	< 0.001
Number of children	0.25	(0.16, 0.34)	5.3	< 0.001
Smoking	0.21	(0.03, 0.39)	2.3	0.02
Dieting	-0.24	(-0.40, -0.08)	-3.0	0.003
BMI‡	-0.13	(-0.22, -0.05)	-3.1	0.002
Change in amount of food	0.15	(0.05, 0.24)	3.1	0.002

Adjusted R^2 = 20.0%, n = 579 †Per 5 year increase *Scaled to a standard deviation of 1
‡Logged and scaled to a standard deviation of 1

Figure 26 Proportion of women in the lowest quarter of PAH early pregnancy high-energy diet score, by significant predictors of the score

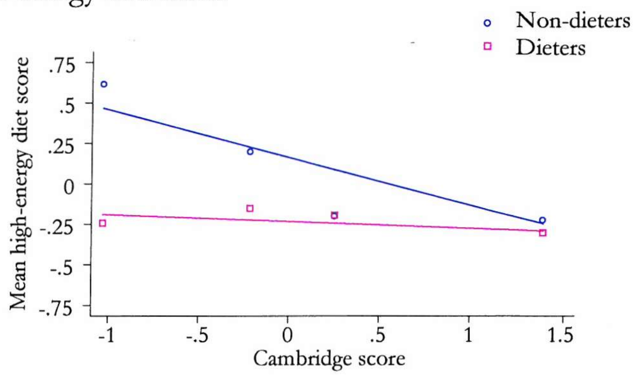


Number of children and Cambridge score have the strongest effect on early pregnancy high-energy score; women living with more children and having a lower Cambridge score report a higher-energy diet; the effects of Cambridge score are particularly marked amongst those with scores above the median. Age, dieting, BMI and change in amount of food are less important factors, and smoking has the least significant effect in predicting a high-energy diet in early pregnancy. The predictors of the early pregnancy high-energy diet score explain less variation

(20.0%) than the predictors of the prudent diet score in early (33.5%) and late (46.7%) pregnancy.

Two-way interactions between variables in Table 63 were tested, and the interaction between Cambridge score and dieting was statistically significant. The effect of Cambridge score on average high-energy diet score (adjusted for other factors in the model) amongst those dieting and not dieting is illustrated in Figure 27.

Figure 27 Interaction between dieting and Cambridge score as predictors of PAH early pregnancy high-energy diet score



When the regression model in Table 63 (without dieting) is fitted separately in dieters and non-dieters, a one standard deviation increase in Cambridge score is associated with a decrease in early pregnancy high-energy diet score of 0.29 amongst non-dieters, and with a decrease of 0.04 amongst dieters. Thus the negative effect of Cambridge score on early pregnancy high-energy diet score is greater amongst non-dieters. Similarly the negative effect of dieting is greater amongst those with lower Cambridge scores.

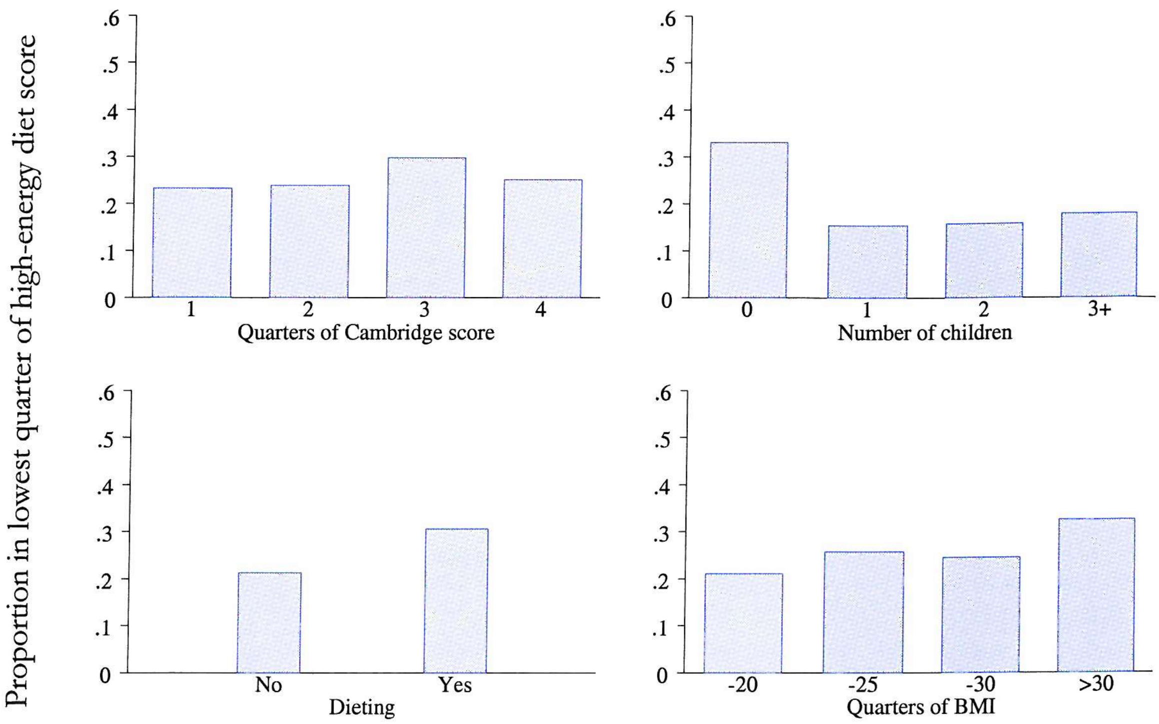
5.4.3 Late pregnancy high-energy diet score

Table 64 All significant predictors of PAH late pregnancy high-energy diet score

Predictor	β	95% CI for β	t	P
Cambridge score*	-0.11	(-0.19, -0.03)	-2.7	0.006
Number of children	0.16	(0.07, 0.26)	3.4	0.001
Dieting	-0.21	(-0.39, -0.04)	-2.4	0.02
BMI‡	-0.13	(-0.22, -0.05)	-3.0	0.003

Adjusted R^2 = 6.6%, n = 570 *Scaled to a standard deviation of 1
‡Logged and scaled to a standard deviation of 1

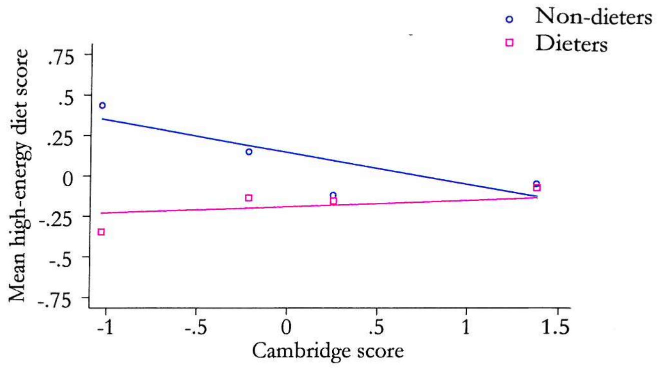
Figure 28 Proportion of women in the lowest quarter of PAH late pregnancy high-energy diet score, by significant predictors of the score



Four of the strongest predictors of early pregnancy high-energy diet score are significant predictors of late pregnancy high-energy diet score; women with a lower Cambridge score, living with more children, not dieting in the year prior to their pregnancy and having a lower BMI report a higher-energy diet. The coefficients for the effects of Cambridge score, number of children, dieting and BMI on high-energy diet score are in the same direction in early and late pregnancy (Table 63 and Table 64), although those for Cambridge score and number of children are of slightly lesser magnitude in late pregnancy. Similarly to early pregnancy, the negative effect of Cambridge score is notable principally amongst those with scores above the median. Age, smoking and change in amount of food, which were significant predictors of the early pregnancy high-energy score, are not important factors in influencing the late pregnancy score. The predictors of the late pregnancy score explain substantially less variation (6.6%) than the predictors of the early pregnancy score (20.0%).

Two-way interactions between variables in Table 64 were tested and, as in early pregnancy, the only significant interaction in late pregnancy was that between Cambridge score and dieting. The effect of Cambridge score on average high-energy diet score (adjusted for other factors in the model) amongst dieters and non-dieters is illustrated in Figure 29.

Figure 29 Interaction between dieting and Cambridge score as predictors of PAH late pregnancy high-energy diet score



When the regression model in Table 64 (without BMI) is fitted separately in dieters and non-dieters, a one standard deviation increase in Cambridge score is associated with a decrease in late pregnancy high-energy diet score of 0.20 amongst non-dieters, and with an increase of 0.04 amongst dieters. Thus the interpretation of the significant interaction is the same as for early pregnancy high-energy score: the negative effect of Cambridge score is greater amongst non-dieters, and similarly the negative effect of dieting is greater amongst those with lower Cambridge scores.

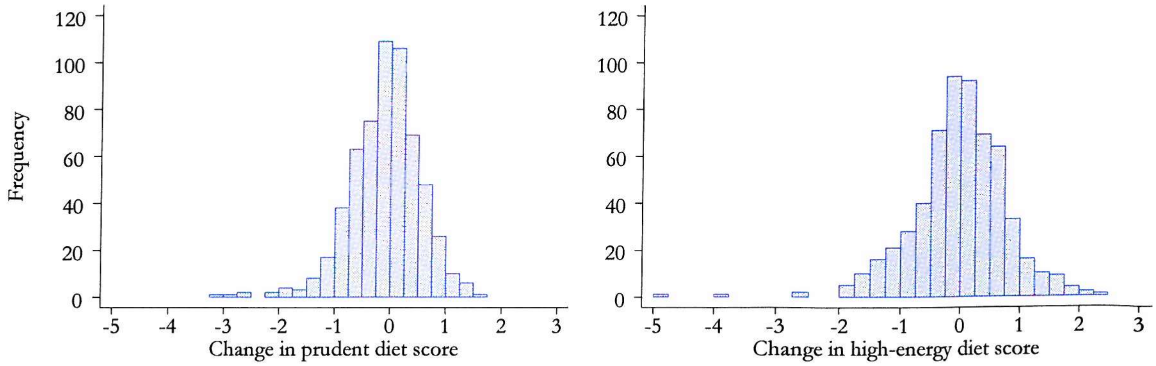
5.5 Predictors of changes in dietary scores between early and late pregnancy

5.5.1 Data preparation

A moderate degree of agreement was observed between early and late pregnancy diet scores based on early pregnancy coefficients (Section 4.1.4). Since some women's diets do change notably during this period, sociodemographic and lifestyle predictors of these changes are considered here. These may be of particular public health interest since many women's first contact with health professionals during their pregnancy is around the time of the PAH study early pregnancy interview. It is therefore helpful to identify those women whose diets change after this period, to understand how advice in early pregnancy might best be targeted.

Early and late pregnancy scores were derived based on early pregnancy coefficients, and the change in score calculated by subtracting the early from the late pregnancy score. Since both early and late pregnancy scores were divided by the standard deviation of the early pregnancy score (Section 4.1.4), the change is again interpreted in standard deviations of the early pregnancy score. The distributions of change in prudent and high-energy diet scores between early and late pregnancy are given in Figure 30.

Figure 30 Distributions of change in PAH prudent and high-energy diet score between early and late pregnancy



A regression to the mean effect was noted; change in score is correlated with early pregnancy score (r_S = -0.34 for prudent diet score and r_S = -0.46 for high-energy diet score). This effect is commonly observed in measurement of change since there is a natural tendency for particularly high or low measurements to be nearer the average when repeated[210]. The tendency is likely to reflect measurement error, as well as actual change.

In Prevost et al.'s[211] analysis of change in diet across the Health and Lifestyle Survey, change in score was adjusted for initial score using a regression technique to prevent associates of the initial score spuriously contributing to the estimation of the associates of score change. This regression technique involves subtracting from the change in score that which is linearly predictable from the initial score. A regression model is used to discover the average change (β) associated with a one unit increase of the initial value; the adjusted change in score is calculated by multiplying each woman's initial value by β and subtracting this from the woman's change in score. For example, the average change in prudent diet score (β) associated with an increase of one standard deviation in the initial score is –0.26 standard deviations. Therefore a woman with an initial score of 2 would have 0.52 (= 2 × 0.26) added to her change score to give an adjusted change score. This regression adjustment technique is used here, and thus the issue addressed is: did women with the same dietary score in early pregnancy change their scores differentially according to sociodemographic and lifestyle predictors?

These adjusted changes in prudent and high-energy diet scores were transformed to normality using Fisher-Yates normal scores[207], and the resulting scores scaled to have the same standard deviation as the regression-adjusted scores, so that changes can be interpreted in terms of standard deviations of the original data. Multivariate linear regression models with change in diet score as the outcome were utilised. The same procedure as that used in Section 5.3 was employed to investigate sociodemographic and lifestyle predictors of changes in scores.

5.5.2 Increase in prudent diet score

Older age, higher SC and living with fewer children were all associated with an increase in prudent diet score (Table 65). A linear effect across the six SC groups is assumed here. Thus a change in SC from I to II is associated with a decrease in prudent diet score between early and late pregnancy of 0.10 standard deviations. A change in SC from I to V is associated with a decrease in prudent diet score of 0.50 standard deviations. Analysis using a linear effect of social class was appropriate since it is in agreement with the concept of a graded hierarchy upon which the scheme is based[168].

Table 65 All significant predictors of increase in PAH prudent diet score

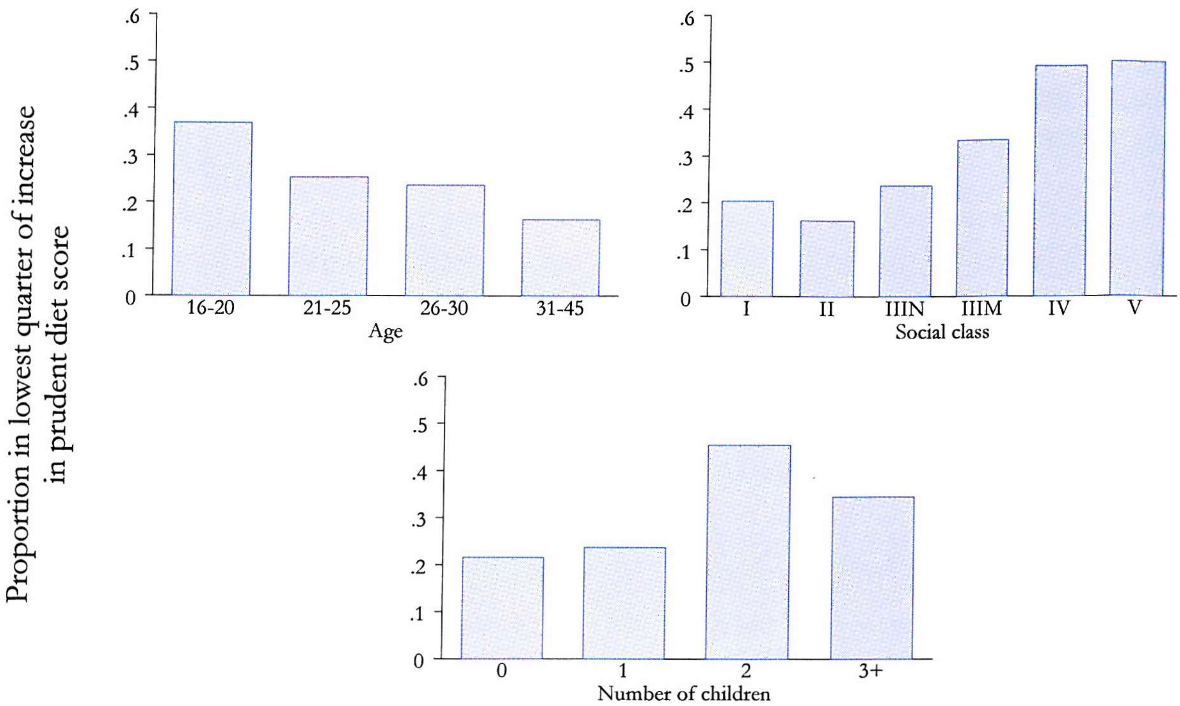
Predictor	**β**	**95% CI for β**	**t**	**P**
Age†	0.10	(0.05, 0.15)	3.7	< 0.001
Social Class	-0.10	(-0.14, -0.06)	-4.8	< 0.001
Number of children	-0.11	(-0.17, -0.06)	-4.1	< 0.001

Adjusted R^2 = 11.1%, n = 581 †Per 5 year increase

The percentage of variation explained is small (11.1%) in comparison with the variation explained in early and late pregnancy prudent diet scores (33.5% and 46.7% respectively). Figure 31 illustrates the proportion of women in the lowest quarter of the increases in prudent diet score, by levels of the significant predictors, adjusted for other terms in the model.

Older age, living with fewer children and high social position (as defined by Cambridge score and qualifications, rather than SC) were also significant predictors of early pregnancy prudent diet score. Therefore women with these attributes are not only expected to eat a prudent diet in early pregnancy, but are likely to improve this diet during the course of their pregnancy.

Figure 31 Proportion of women in the lowest quarter of increase in PAH prudent diet score, by significant predictors of the change in score



5.5.3 Increase in high-energy diet score

Significant predictors of change in high-energy score were Townsend index and employment in early pregnancy. Not being employed and living in an area of less deprivation were associated with an increasing score (Table 66). Figure 32 illustrates the proportion of women in the lowest quarter of the increases in high-energy score, by levels of the significant predictors, adjusted for other terms in the model. However, the percentage of variation in change in high-energy score explained is very small (1.6%), and the explanatory power of the model is therefore negligible.

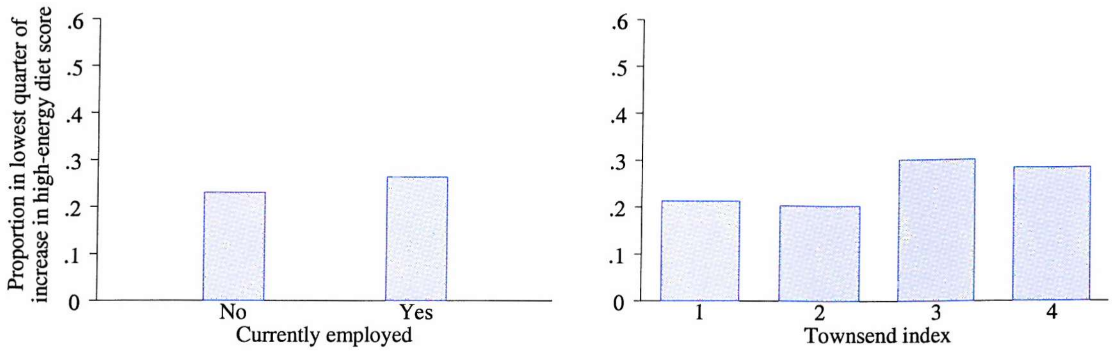
Table 66 All significant predictors of increase in PAH high-energy diet score

Predictor	β	95% CI for β	t	P
Townsend index*	-0.07	(-0.13, -0.02)	-2.7	0.01
Currently employed	-0.13	(-0.24, -0.02)	-2.4	0.02

Adjusted R^2 = 1.6%, n = 504 *Scaled to a standard deviation of 1

No two-way interactions between predictors of change in prudent or high-energy diet scores were statistically significant.

Figure 32 Proportion of women in the lowest quarter of increase in PAH high-energy diet score, by significant predictors of the change in score



5.6 Summary

- Women in the PAH study who were older, had a higher Cambridge score, were better qualified, lived with fewer children and didn't smoke were likely to report a more prudent diet in both early and late pregnancy. There is substantial agreement between the size and direction of these effects. In late pregnancy there is an additional marginally significant effect; those who lived in areas that were less deprived according to the Townsend index reported a more prudent diet. In early and late pregnancy the predictors identified explain substantial amounts of the variation in prudent diet score (adjusted R^2 = 33.5% and 46.7% respectively).

- The positive effect of a woman's social position on her late pregnancy prudent diet score is greater amongst those who live in areas with more deprivation. Similarly the negative effect of deprivation is greater amongst those of a lower social position.

- Women in the PAH study who had lower Cambridge scores, lived with more children, did not diet in the year prior to their pregnancy and had lower BMIs were more likely to report a higher-energy diet in both early and late pregnancy. There is again similarity in the size and direction of these effects. In early pregnancy there are additional effects; those who were younger, who smoked and who ate more food since becoming pregnant were also likely to report a higher-energy diet. The predictors identified explain less variation in the high-energy diet score than those for the prudent diet score.

- In both early and late pregnancy the negative effect of Cambridge score on high-energy diet score is greater amongst non-dieters, and similarly the negative effect of dieting is greater amongst those with lower Cambridge scores.

- Increases in prudent diet score between early and late pregnancy are associated with older age, being of a higher social class, and living with fewer children.

- Increases in high-energy diet score between early and late pregnancy are associated with not being employed in early pregnancy and living in an area of less deprivation. However, these effects only explain 1.6% of the variation in the change variable, and are therefore relatively inconsequential.

6 Dietary assessment in the SWS

The SWS provides an opportunity to study dietary patterns using principal component analysis on a contemporary sample of non-pregnant women; the resulting patterns are again interpreted as prudent and high-energy diets. The impact of including the 'extra foods' available in the SWS on these dietary patterns is investigated in Section 6.2.

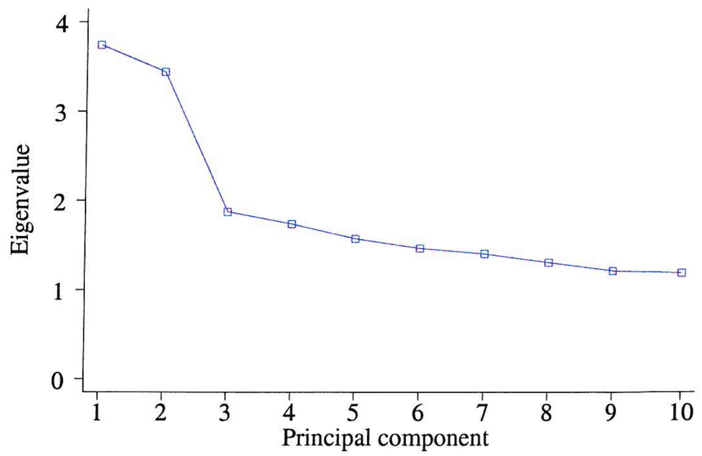
Fruit and vegetable intake is considered in Section 6.3, as reported using both the FFQ and the additional 'cross-check' questions. The effect of weighting FFQ data using responses to the cross-check questions on a subsequent PCA is investigated. The SWS stability study on 94 women provides an opportunity to assess the stability of dietary patterns over two years in a non-pregnant population; Section 6.4 describes the patterns found and compares women's scores at the two time points.

6.1 Principal component analysis

The initial 100 SWS foods and food groups differ slightly from the initial 100 PAH foods and food groups: in the SWS gravy granules and stock cubes are split into two separate categories, and wine is all included in one question (Section 2.3.3). However, when these 100 food groups are collapsed using the pragmatic food grouping (Table 32) they map to the same 49 food groups used in the PAH analysis, and thus results from the two studies are directly comparable.

Principal component analysis was therefore performed on the SWS FFQ data using the 49 pragmatic food groups. A scree plot demonstrates that the first two components explain notably more variation than the remaining components (Figure 33).

Figure 33 Scree plot from PCA using pragmatic food groups in the SWS data



Since these two components were also the most interpretable they were retained, and their coefficients are displayed alongside those from the PAH early pregnancy analysis in Table 67. Coefficients are coloured using the same scheme as that used in Chapter 3, to highlight particularly high and low values:

≥ 0.16
0.06 to 0.15
-0.05 to 0.05
-0.15 to -0.06
≤ -0.16

Table 67 Prudent and high-energy diet coefficients in the SWS and PAH early pregnancy

Food	Prudent diet		High-energy diet	
	SWS	PAH early pregnancy	SWS	PAH early pregnancy
White bread	-0.22	-0.20	0.10	0.14
Wholemeal bread	0.23	0.26	**0.04**	**0.04**
Breakfast cereals	0.13	0.06	0.06	0.09
Crackers	0.07	0.10	0.10	0.13
Rice and pasta	0.21	0.22	0.13	0.09
Quiche and pizza	**-0.04**	0.08	0.12	0.08
Yorkshire pudding and savoury pancakes	-0.18	-0.06	0.12	0.25
Cakes and biscuits	-0.13	**-0.01**	0.19	0.24
Puddings	**-0.05**	**0.05**	0.23	0.27
Full-fat milk (pints)	-0.20	-0.15	0.10	0.19
Reduced-fat milk (pints)	**0.05**	0.15	**0.02**	-0.09
Cream	**0.03**	0.09	0.13	0.14
Cheese and cottage cheese	0.08	0.21	0.14	**0.03**
Yoghurt	0.16	0.23	0.13	0.06
Eggs and egg dishes	**-0.04**	**0.00**	0.21	0.18
Full-fat spread	-0.17	**-0.03**	0.16	0.18
Reduced-fat spread	0.06	**0.03**	**0.01**	**-0.01**
Cooking fats and salad oils	0.11	**0.01**	0.16	0.13
Chicken and turkey	**-0.02**	0.13	0.12	0.10
Red meat	-0.17	**0.05**	0.23	0.23
Processed meat	-0.21	**-0.04**	0.21	0.28
Offal	-0.10	**-0.03**	0.14	0.15
Fish and shell fish	0.09	0.20	0.18	0.06
Salad vegetables	0.25	0.30	0.21	**0.05**
Green vegetables	0.17	0.21	0.25	0.09
Root vegetables	0.08	0.17	0.23	0.08
Other vegetables	0.23	0.27	0.23	0.10
Beans and pulses	**0.02**	**0.04**	0.15	0.12
Vegetable dishes and vegetarian foods	0.21	0.22	**0.03**	**-0.02**
Tinned vegetables	-0.14	-0.16	0.13	0.18
Boiled potatoes	**-0.05**	**0.01**	0.16	0.19
Roast potatoes and chips	-0.27	-0.11	0.10	0.22
Crisps	-0.19	-0.15	0.10	0.16
Citrus fruit and fruit juices	0.14	0.19	0.13	**0.04**
Other fruit	0.23	0.20	0.16	**0.02**
Other fruit juices	**0.05**	0.10	0.09	**0.00**
Nuts	**0.05**	0.11	0.09	0.09
Dried fruit	0.14	0.18	0.07	**0.00**
Cooked and tinned fruit	**0.03**	0.13	0.17	0.15
Confectionery	-0.16	-0.10	0.14	0.22
Sweet spreads	0.06	0.12	0.11	0.09
Added sugar (teaspoons)	-0.23	-0.20	0.07	0.22
High-energy soft drinks	-0.11	-0.09	0.10	0.15
Diet coke	**-0.02**	**-0.01**	**0.00**	**0.01**
Tea and coffee	-0.16	-0.13	0.07	0.17

Table 67 (continued) Prudent and high-energy diet coefficients in the SWS and PAH early pregnancy

Food	Prudent diet		High-energy diet	
	SWS	PAH early pregnancy	SWS	PAH early pregnancy
Decaffeinated tea and coffee	0.10	0.10	**0.02**	**-0.03**
Drinking chocolate	**0.01**	**0.01**	0.06	0.11
Wine	0.12	**0.04**	**0.01**	**-0.03**
Miscellaneous	**-0.04**	0.09	0.26	0.21
Proportion of variation explained	7.6%	8.7%	7.0%	7.1%

The greatest difference between the SWS first component and the PAH early pregnancy prudent dietary pattern is the lower coefficients for meat in the SWS, particularly red and processed meat. However, there is generally a remarkable similarity between these two patterns, and therefore the first SWS component is labelled a 'prudent' diet. The greatest differences between the SWS second component and the PAH early pregnancy high-energy dietary patterns are the higher coefficients for fruit and vegetables in the SWS, and the lower coefficient for added sugar. However, it is striking that all the coefficients for the SWS second component are positive, and again the 'high-energy' label is considered appropriate.

The lower coefficients for red and processed meats in the SWS may be due to health messages arising from concerns about associations between red and processed meat consumption and cancers, particularly colorectal cancer, between 1991 and 1998[212]. The higher coefficients for fruit and vegetables in the SWS, and the lower coefficient for added sugar may partly reflect national trends in diet between 1991 and 2000. Comparison of intakes in the DNSBA in 1986-7[66] and 2000-1[213] indicates that average fruit and vegetable intake increased by 478g per week, and average table sugar intake decreased by 35g per week.

To investigate the robustness of the prudent and high-energy patterns in the SWS, principal component analysis was also performed using the original 100 SWS food variables. The Spearman's correlation coefficients for individual's scores using the two methods were 0.96 and 0.97 for the prudent and high-energy diets respectively. Thus the robust nature of the two scores across grouping methods that was seen in the PAH study (Section 4.1.3.3) is replicated in the SWS.

Serum vitamin C measurements were reported in Section 4.3.3 as evidence of the validity of the prudent diet score in the PAH study. Although serum vitamin C measurements are not available in the SWS, red cell folate measurements were made on 3,823 women who were willing to provide a blood sample. Spearman's correlation coefficient between the prudent diet score and red cell folate levels is 0.30, demonstrating that the prudent diet score is positively associated with an alternative external measure of micronutrient intake.

6.2 Extra foods

The SWS questionnaire included an additional feature to the PAH questionnaire, asking women to specify any food or drinks not included within the FFQ that they had eaten once a week or more (Appendix 3). These foods were coded as closely as possible to the FFQ categories. For example, couscous was coded to category (20) (Brown and white rice) and avocados were coded to category (34) (Green salad). Foods used in very small amounts and for which there was no relevant FFQ category, such as lemon juice and garlic, were not coded. The impact of the inclusion of extra foods is investigated here.

Up to 5 extra foods were reported by 606 SWS participants. Of these, 94 women's extra foods could not be coded. For the remaining 512, the extra frequencies were added to the 100 frequencies reported in the FFQ. PCA on these updated frequencies was performed in the same way as in Section 6.1. The resulting coefficients are given in Table 68, alongside the coefficients for the PCA without extra foods.

Table 68 Prudent and high-energy diet coefficients in the SWS with and without extra foods

Food	Prudent diet		High-energy diet	
	With extra foods	Without extra foods	With extra foods	Without extra foods
White bread	-0.22	-0.22	0.11	0.10
Wholemeal bread	0.23	0.23	0.03	0.04
Breakfast cereals	0.13	0.13	0.05	0.06
Crackers	0.07	0.07	0.10	0.10
Rice and pasta	0.21	0.21	0.12	0.13
Quiche and pizza	-0.04	-0.04	0.12	0.12
Yorkshire pudding and savoury pancakes	-0.18	-0.18	0.13	0.12
Cakes and biscuits	-0.12	-0.13	0.20	0.19
Puddings	-0.04	-0.05	0.23	0.23
Full-fat milk (pints)	-0.20	-0.20	0.11	0.10
Reduced-fat milk (pints)	0.05	0.05	0.02	0.02
Cream	0.03	0.03	0.13	0.13
Cheese and cottage cheese	0.08	0.08	0.14	0.14
Yoghurt	0.16	0.16	0.12	0.13
Eggs and egg dishes	-0.03	-0.04	0.21	0.21
Full-fat spread	-0.16	-0.17	0.16	0.16
Reduced-fat spread	0.06	0.06	0.01	0.01
Cooking fats and salad oils	0.12	0.11	0.16	0.16
Chicken and turkey	-0.02	-0.02	0.12	0.12
Red meat	-0.17	-0.17	0.23	0.23
Processed meat	-0.20	-0.21	0.22	0.21
Offal	-0.09	-0.10	0.14	0.14
Fish and shell fish	0.10	0.09	0.18	0.18
Salad vegetables	0.26	0.25	0.20	0.21
Green vegetables	0.18	0.17	0.25	0.25

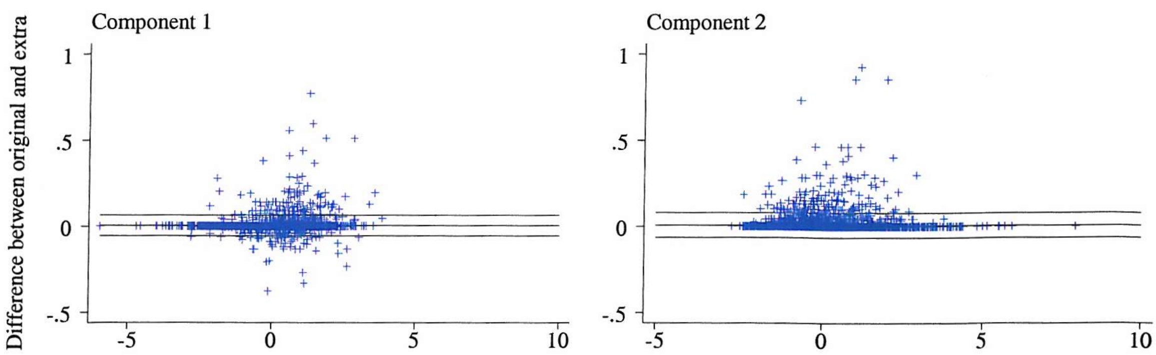
Table 68 (continued) Prudent and high-energy diet coefficients in the SWS with and without extra foods

Food	Prudent diet		High-energy diet	
	With extra foods	Without extra foods	With extra foods	Without extra foods
Root vegetables	0.09	0.08	0.22	0.23
Other vegetables	0.23	0.23	0.22	0.23
Beans and pulses	**0.03**	**0.02**	0.14	0.15
Vegetable dishes and vegetarian foods	0.22	0.21	**0.02**	**0.03**
Tinned vegetables	-0.13	-0.14	0.13	0.13
Boiled potatoes	**-0.05**	**-0.05**	0.16	0.16
Roast potatoes and chips	-0.26	-0.27	0.11	0.10
Crisps	-0.19	-0.19	0.11	0.10
Citrus fruit and fruit juices	0.14	0.14	0.13	0.13
Other fruit	0.24	0.23	0.15	0.16
Other fruit juices	0.06	**0.05**	0.09	0.09
Nuts	0.07	**0.05**	0.08	0.09
Dried fruit	0.14	0.14	0.07	0.07
Cooked and tinned fruit	**0.03**	**0.03**	0.17	0.17
Confectionery	-0.15	-0.16	0.15	0.14
Sweet spreads	0.07	0.06	0.11	0.11
Added sugar (teaspoons)	-0.23	-0.23	0.08	0.07
High-energy soft drinks	-0.10	-0.11	0.11	0.10
Diet coke	**-0.02**	**-0.02**	**0.00**	**0.00**
Tea and coffee	-0.16	-0.16	0.08	0.07
Decaffeinated tea and coffee	0.10	0.10	**0.02**	**0.02**
Drinking chocolate	**0.01**	**0.01**	0.06	0.06
Wine	0.12	0.12	**0.01**	**0.01**
Miscellaneous	**-0.02**	**-0.04**	0.26	0.26
Proportion of variation explained	7.6%	7.6%	7.0%	7.0%

The dietary patterns produced by inclusion of the extra foods are clearly extremely similar to those produced without; only the coefficients for nuts and miscellaneous in the prudent diet score differs by more than 0.01. This demonstrates that the principal component technique is robust to small additional dietary intakes.

Changes in individual women's scores were investigated using Bland and Altman plots (Figure 34), based on coefficients from the PCA without extra foods; the limits of agreement are –0.06 to 0.06 for the prudent diet score, and –0.07 to 0.08 for the high-energy diet score. However, it is notable that all women's high-energy diet scores increase, as would be expected since all the PCA coefficients are positive (Table 68). The woman with the highest increase in prudent diet score reported eating fresh dates twice a day in addition to her FFQ answers. Three of the four largest increases in high-energy diet score were due to women reporting consumption of seeds such as pumpkin, linseed and sesame. The fourth woman reported eating ice-lollies twice a day as an extra food.

Figure 34 Bland and Altman plots comparing SWS participants' scores with and without the extra foods, based on coefficients from the PCA without extra foods



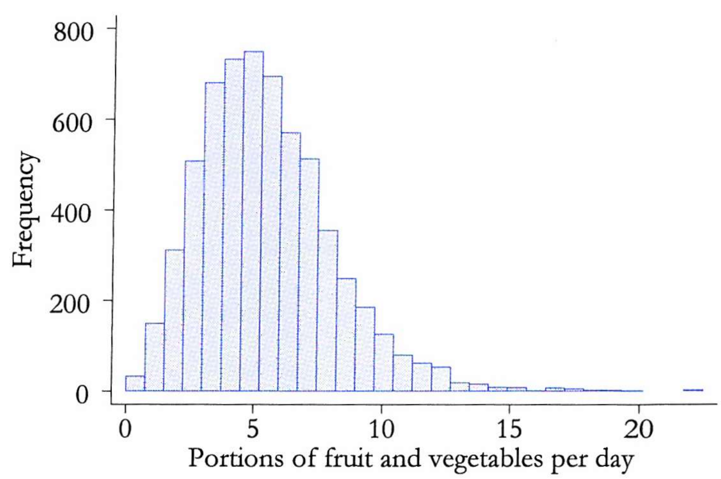
Generally individual women's scores are virtually unchanged by the inclusion of the extra foods; the use of the extra foods data within a PCA is likely to change a woman's dietary score by less than 0.08 standard deviations. Indeed 5,613 women's scores were completely unchanged. In the light of these very minor differences, the results from the PCA without extra foods will be used, because these are based on an FFQ that is most comparable to that used in the PAH study, and which has been validated by Robinson et al.[127].

6.3 Fruit and vegetables

6.3.1 Fruit and vegetable intake from FFQ

Fruit and vegetable intake was calculated from the FFQ, using the criteria described in Section 2.4.4. The result was divided by seven to give portions per day, and the distribution of responses is shown in Figure 35. Fifty-two per cent of the women ate at least five portions of fruit and vegetables per day.

Figure 35 Distribution of portions of fruit and vegetables per day in the SWS FFQ



The distribution is similar to that for the PAH early and late pregnancy FFQs (Figure 18), although summary statistics (Table 69) indicate that fruit and vegetable consumption is slightly higher amongst the non-pregnant SWS women than the PAH participants. This may be partly because national fruit and vegetable consumption increased over the period between the two studies (Section 6.1), although this increase was almost entirely due to an increase in fruit consumption, and the increase between the two studies is attributed to both fruit and vegetables. Also, women in the SWS have somewhat higher qualifications than those in the PAH study (Table 9), and fruit and vegetable intake is strongly positively correlated with educational qualifications in both studies.

Table 69 Fruit and vegetable portions per day reported using the FFQ by PAH and SWS women

	Lower quartile	Median	Upper quartile	n
PAH early pregnancy	3.2	4.5	6.0	620
PAH late pregnancy	2.8	4.1	5.5	593
SWS	3.6	5.1	6.9	6129

As was the case for the PAH data (Table 44), fruit and vegetable intake in the SWS is strongly positively correlated with the prudent diet score (Table 70). However, the correlation with the high-energy diet score is just as strong. This is because the SWS high-energy fruit and vegetable coefficients are higher than the PAH high-energy fruit and vegetable coefficients (Table 67).

Table 70 Spearman's correlations between SWS FFQ fruit and vegetable intake and prudent and high-energy diet scores

	Prudent diet	High-energy diet	n
FFQ fruit and vegetable intake	0.62	0.62	6125

6.3.2 Fruit and vegetable intake from cross-check questions

The SWS also included additional questions about consumption of vegetables (excluding potatoes), and fruit and fruit juices in the past week (Appendix 3). These 'cross-check' questions aimed to assess the accuracy of individual fruit and vegetable intake reports using the FFQ. Responses to the cross-check questions were summed to generate a total for fruit and vegetables, and divided by seven to give frequencies per day.

The Spearman's correlation coefficient between fruit and vegetable intake assessed using the FFQ and the cross-check questions is 0.64. However, the correlation coefficient does not assess agreement, and intake recorded using the FFQ is on average 2.5 portions greater than that using the cross-check question, with Bland and Altman limits of agreement of –1.5 to 6.5. Thus fruit and vegetable intake reported using the FFQ tends to be considerably greater than that using the cross-check question.

Studies have suggested that the FFQ may be more likely to over-report fruits and vegetables than other food groups[131,132]. This would be irrelevant within the context of a PCA if all women over-reported by the same amount, since variables are standardised before being entered into the analysis. However, there is variation between women. Calvert et al.[132] suggest an individual-based correction based on weighting the FFQ data by a factor reflecting responses to the cross-check questions. The factor is calculated as

$$\frac{\text{Number of servings per day from cross - check question}}{\text{Number of servings per day from FFQ}}$$

An individual's response to a fruit question within the FFQ is weighted using that individual's fruit weight, and similarly for vegetables. Thus Calvert et al.'s method assumes that the cross-check question is a more accurate measure of fruit and vegetable intake than the FFQ, and corrects the FFQ responses accordingly. In the few cases where the number of servings from the FFQ were zero, all the responses to the cross-check question were also zero, and thus the weighting was set to zero. A weighting factor could not be calculated for the nine subjects who did not provide a response to the fruit or vegetable cross-check question.

The impact of weighting the fruit and vegetable FFQ data entered into a PCA was explored (Table 71).

Table 71 Prudent and high-energy diet coefficients in the SWS with and without weightings from cross-check questions

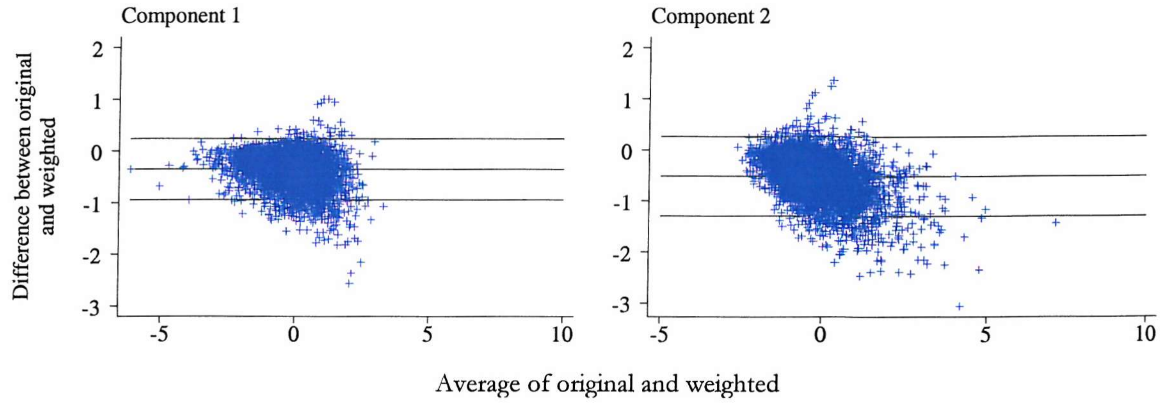
Food	Prudent diet: Using weightings	Prudent diet: Without using weightings	High-energy diet: Using weightings	High-energy diet: Without using weightings
White bread	-0.21	-0.22	0.11	0.10
Wholemeal bread	0.20	0.23	**0.04**	**0.04**
Breakfast cereals	0.11	0.13	0.06	0.06
Crackers	**0.04**	0.07	0.08	0.10
Rice and pasta	0.16	0.21	0.10	0.13
Quiche and pizza	-0.06	**-0.04**	0.11	0.12
Yorkshire pudding and savoury pancakes	-0.16	-0.18	0.13	0.12
Cakes and biscuits	-0.14	-0.13	0.20	0.19

Table 71 (continued) Prudent and high-energy diet coefficients in the SWS with and without weightings from cross-check questions

Food	Prudent diet: Using weightings	Prudent diet: Without using weightings	High-energy diet: Using weightings	High-energy diet: Without using weightings
Puddings	-0.07	-0.05	0.22	0.23
Full-fat milk (pints)	-0.18	-0.20	0.11	0.10
Reduced-fat milk (pints)	0.05	0.05	0.03	0.02
Cream	0.01	0.03	0.14	0.13
Cheese and cottage cheese	0.05	0.08	0.14	0.14
Yoghurt	0.13	0.16	0.12	0.13
Eggs and egg dishes	-0.07	-0.04	0.19	0.21
Full-fat spread	-0.16	-0.17	0.18	0.16
Reduced-fat spread	0.05	0.06	0.00	0.01
Cooking fats and salad oils	0.08	0.11	0.16	0.16
Chicken and turkey	-0.03	-0.02	0.12	0.12
Red meat	-0.18	-0.17	0.23	0.23
Processed meat	-0.21	-0.21	0.22	0.21
Offal	-0.10	-0.10	0.15	0.14
Fish and shell fish	0.06	0.09	0.18	0.18
Salad vegetables	0.29	0.25	0.21	0.21
Green vegetables	0.23	0.17	0.26	0.25
Root vegetables	0.17	0.08	0.25	0.23
Other vegetables	0.27	0.23	0.23	0.23
Beans and pulses	0.12	0.02	0.17	0.15
Vegetable dishes and vegetarian foods	0.20	0.21	0.00	0.03
Tinned vegetables	-0.05	-0.14	0.15	0.13
Boiled potatoes	-0.04	-0.05	0.18	0.16
Roast potatoes and chips	-0.24	-0.27	0.11	0.10
Crisps	-0.19	-0.19	0.10	0.10
Citrus fruit and fruit juices	0.14	0.14	0.08	0.13
Other fruit	0.25	0.23	0.10	0.16
Other fruit juices	0.07	0.05	0.06	0.09
Nuts	0.04	0.05	0.09	0.09
Dried fruit	0.16	0.14	0.07	0.07
Cooked and tinned fruit	0.06	0.03	0.15	0.17
Confectionery	-0.16	-0.16	0.15	0.14
Sweet spreads	0.04	0.06	0.11	0.11
Added sugar (teaspoons)	-0.21	-0.23	0.08	0.07
High-energy soft drinks	-0.12	-0.11	0.09	0.10
Diet coke	-0.03	-0.02	-0.01	0.00
Tea and coffee	-0.14	-0.16	0.09	0.07
Decaffeinated tea and coffee	0.09	0.10	0.03	0.02
Drinking chocolate	-0.01	0.01	0.06	0.06
Wine	0.11	0.12	0.02	0.01
Miscellaneous	-0.06	-0.04	0.26	0.26
Proportion of variation explained	8.3%	7.6%	6.1%	7.0%

The coefficients are little changed by the inclusion of Calvert et al.'s weighting term. Changes in individual women's scores were investigated using Bland and Altman plots (Figure 36) based on coefficients from the PCA without weighting. For the prudent diet score the average difference of the score with weighting minus the score without weighting is -0.35 with limits of agreement from –0.95 to 0.24. For the high-energy diet score the average difference of the score with weighting minus that without weighing is -0.51 with limits of agreement –1.29 to 0.26. These drops in score are expected because on average the weighting terms were substantially less than one; women tend to report more fruit and vegetable consumption using the FFQ than the cross-check question. The correlations between blood red cell folate and prudent diet scores calculated with and without weighting are both 0.30.

Figure 36 Bland and Altman plots comparing SWS participants' scores with and without weightings based on coefficients from the PCA without weightings



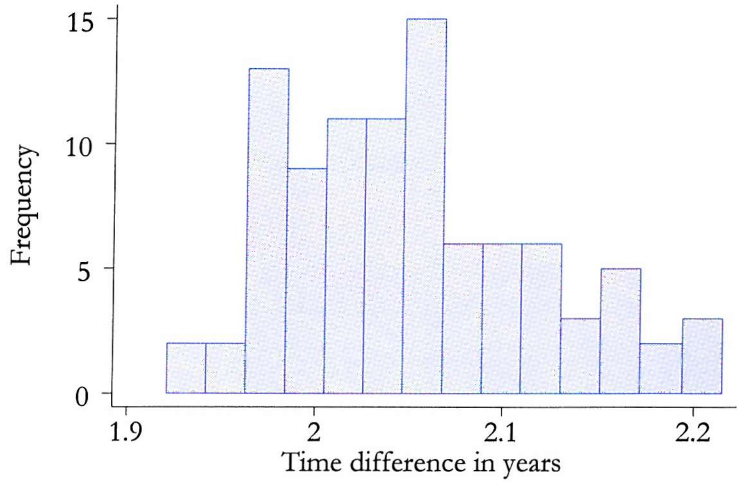
The use of cross-check weightings rests on the assumption that the cross-check questions are a better measure of habitual fruit and vegetable intake than the FFQ. This is not justified by Calvert et al., and the wording of the SWS questionnaire (Appendix 3) makes it less likely since the cross-check question refers to the past week, whereas the FFQ refers to the past three months. Furthermore, as Calvert et al. note, the weighting factor does not take account of differential mis-reporting of individual food items; subjects may be more likely to over-estimate strawberry consumption than apple consumption, for example. An additional minor consideration is that the weighted PCA cannot be performed on subjects who did not respond to one of the cross-check questions. Therefore analysis will proceed using the unweighted data.

6.4 Stability study

A further component of the SWS was a 'stability study', involving re-interviewing 94 non-pregnant women approximately two years after their initial interview. These data are useful in this thesis to assess the stability of dietary patterns over time. A similar analysis was performed on the PAH data (Section 4.1.4), but the advantage of these data is that the SWS women were not pregnant at any point during the two year interval. Therefore they were not subject to the metabolic changes experienced by the PAH women through the course of their pregnancy.

Figure 37 shows that the time interval from first to second interview varied between 1 year, 11 months, and 2 years, 2½ months.

Figure 37 Histogram of time interval between first and second interview in the SWS stability study



The subsample of 94 was found to be representative of the main study sample of 6,129 (Table 72).

Table 72 Characteristics of stability subsample, and main SWS sample

	Stability subsample (n = 94)	Main sample (n = 6129)
Mean (SD) age (years)	28.2 (4.16)	27.8 (4.24)
% Not living with children	62 %	56 %
% Currently smoking	26 %	33 %
Median (IQR) strenuous exercise (hours)	0.17 (0 to 1.5)	0.25 (0 to1.5)
% A-level qualifications or above	60 %	56 %
% Social Class I & II	44 %	44 %
Median (IQR) prudent diet score	0.1 (-0.7 to 0.6)	0.0 (-0.6 to 0.7)
Median (IQR) high-energy diet score	-0.2 (-0.7 to 0.5)	-0.1 (-0.7 to 0.6)
Median (IQR) portions of fruit and veg. per week	4.7 (3.7 to 6.3)	5.1 (3.6 to 6.9)

A principal component analysis was performed on FFQ data from the second interview of the 94 women in the stability subsample (Table 73). There are some differences between the prudent and high-energy diet patterns found. The coefficients for the prudent diet are very similar, although notably the contrast between full-fat and reduced-fat milk is not as great in the subsample as in the main sample. For the high-energy diet the coefficients in the subsample are less similar, but all are positive, except for chicken and turkey, and diet coke, and the pattern still appears to represent a high-energy diet. Although there are differences between the subsample and the main sample, a PCA on the original data from the 94 women also revealed differences of a similar magnitude. Since the coefficients from the PCA on the original data from the 94 women were generally no more similar to the main sample coefficients than the subsample coefficients in Table 73 are, it may be that some of the differences in Table 73 are due to the much smaller size of the subsample, rather than the subsample itself having markedly different eating patterns.

Table 73 Prudent and high-energy diet coefficients in the SWS from the stability subsample and main sample

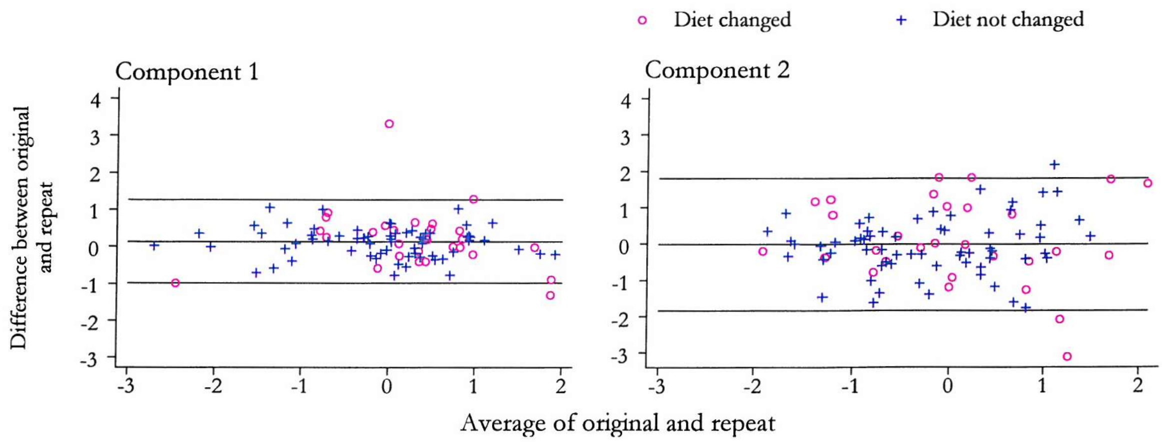
Food	Prudent diet		High-energy diet	
	Subsample	Main sample	Subsample	Main sample
White bread	-0.27	-0.22	0.06	0.10
Wholemeal bread	0.17	0.23	0.12	0.04
Breakfast cereals	0.11	0.13	0.18	0.06
Crackers	0.14	0.07	0.14	0.10
Rice and pasta	0.08	0.21	0.05	0.13
Quiche and pizza	0.03	-0.04	0.20	0.12
Yorkshire pudding and savoury pancakes	-0.19	-0.18	0.11	0.12
Cakes and biscuits	-0.15	-0.13	0.17	0.19
Puddings	-0.01	-0.05	0.17	0.23
Full-fat milk (pints)	-0.24	-0.20	0.19	0.10
Reduced-fat milk (pints)	-0.13	0.05	0.16	0.02
Cream	0.06	0.03	0.11	0.13
Cheese and cottage cheese	0.12	0.08	0.14	0.14
Yoghurt	0.01	0.16	0.07	0.13
Eggs and egg dishes	0.02	-0.04	0.20	0.21
Full-fat spread	-0.22	-0.17	0.09	0.16
Reduced-fat spread	0.03	0.06	0.00	0.01
Cooking fats and salad oils	0.01	0.11	0.07	0.16
Chicken and turkey	-0.02	-0.02	-0.08	0.12
Red meat	-0.20	-0.17	0.12	0.23
Processed meat	-0.26	-0.21	0.07	0.21
Offal	-0.05	-0.10	0.11	0.14
Fish and shell fish	0.15	0.09	0.12	0.18
Salad vegetables	0.20	0.25	0.08	0.21
Green vegetables	0.04	0.17	0.26	0.25
Root vegetables	0.08	0.08	0.23	0.23
Other vegetables	0.11	0.23	0.18	0.23
Beans and pulses	0.04	0.02	0.15	0.15
Vegetable dishes and vegetarian foods	0.17	0.21	0.15	0.03
Tinned vegetables	0.01	-0.14	0.10	0.13
Boiled potatoes	-0.04	-0.05	0.21	0.16
Roast potatoes and chips	-0.26	-0.27	0.08	0.10
Crisps	-0.26	-0.19	0.06	0.10
Citrus fruit and fruit juices	0.07	0.14	-0.01	0.13
Other fruit	0.12	0.23	0.08	0.16
Other fruit juices	0.13	0.05	0.15	0.09
Nuts	0.18	0.05	0.10	0.09
Dried fruit	0.16	0.14	0.18	0.07
Cooked and tinned fruit	0.10	0.03	0.16	0.17
Confectionery	-0.15	-0.16	0.04	0.14
Sweet spreads	0.06	0.06	0.13	0.11
Added sugar (teaspoons)	-0.28	-0.23	0.21	0.07
High-energy soft drinks	0.03	-0.11	0.23	0.10
Diet coke	0.01	-0.02	-0.09	0.00
Tea and coffee	-0.24	-0.16	0.13	0.07

Table 73 (continued) Prudent and high-energy diet coefficients in the SWS from the stability subsample and main sample

Food	Prudent diet		High-energy diet	
	Subsample	Main sample	Subsample	Main sample
Decaffeinated tea and coffee	0.16	0.10	**0.05**	**0.02**
Drinking chocolate	0.07	**0.01**	0.11	0.06
Wine	**0.01**	0.12	0.07	**0.01**
Miscellaneous	**0.02**	**-0.04**	0.30	0.26
Proportion of variation explained	9.2%	7.6%	8.2%	7.0%

Although Table 73 indicates that the first two components in the subsample of 94 women are similar to the prudent and high-energy diets found in the main sample, there remains the possibility that entirely different women are eating prudent and high-energy diets at the two time points. Therefore participants' scores were calculated from the original and repeat FFQ responses, based on the PCA coefficients from the main sample. In the same way as for the PAH study, scores are divided by the standard deviation of the relevant original score (Section 4.1.4), so that a difference of –1 is interpreted as a decrease of one standard deviation from her original score. Agreement of these participants' scores was assessed using Bland and Altman plots. Points on the graph are coloured according to whether the woman considered that her diet had undergone any major changes in the previous two years.

Figure 38 Bland and Altman plots comparing SWS participant's original and repeat scores by whether they changed their diet, based on original coefficients



The limits of agreement for the prudent diet score are similar to those from the comparison between early and late pregnancy in the PAH study (Table 74), despite the fact that the interval between the two time points is two years, rather than approximately 4 months. The limits of agreement for the high-energy diet score are slightly wider for the SWS stability data than the PAH data.

Also included in Table 74 are limits of agreement for the 65 women who said their diets hadn't undergone any major changes. These show a clear improvement on those for all 94 SWS women in the stability study, as would be expected since the more outlying points in Figure 38 are mainly those women who said their diets had changed. It is notable that the outlying woman with a high increase in prudent diet score (Figure 38) gave a description of her dietary change as "eating less fat". The woman with the greatest decrease in high-energy diet score (Figure 38) commented that she was "eating less". Therefore it appears that these women were reporting dietary changes in line with those seen from the changes in prudent and high-energy scores.

Table 74 Limits of agreement comparing PAH and SWS participants' original and repeat scores, based on original coefficients

	Prudent	High-energy	n
PAH late-early pregnancy	(-1.3, 1.1)	(-1.5, 1.5)	589
SWS repeat-original	(-1.0, 1.3)	(-1.8, 1.8)	94
SWS repeat-original no change	(-0.7, 0.9)	(-1.6, 1.5)	65

Table 74 thus shows that there is reasonable agreement between the prudent and high-energy diet scores for women when they repeat an FFQ two years later. This agreement is improved amongst those who say they have not made any major changes to their diet, indicating that women are to some extent able to report whether their diet has changed.

There is a tendency for women's prudent diet scores to increase by 0.13 standard deviations over the two years of the stability study. This is statistically significant ($P = 0.02$). The very slight drop in the high-energy diet score of 0.02 standard deviations is non-significant ($P = 0.82$).

6.5 Summary

- Principal component analysis on FFQ dietary data from the SWS yielded prudent and high-energy dietary patterns very similar to those found in the PAH study. Therefore these patterns appear to be robust across two datasets collected at different times, and on non-pregnant as well as pregnant women.

- The inclusion of extra foods makes very little change to the coefficients from a PCA, or individual women's scores resulting from it.

- Fruit and vegetable consumption is somewhat higher in the SWS than the PAH study, and correlates with both the prudent and high-energy diet scores in the SWS. Weighting the fruit and vegetable data using responses from the cross-check questions provides an alternative analysis, but differences are slight and not necessarily any more helpful.

- The prudent and high-energy dietary patterns are stable over a two year period in a subsample of 94 non-pregnant women. The women's scores show reasonable agreement over this period. The agreement is improved amongst participants who report that their diet hasn't undergone any major changes, indicating that women are to some extent able to report whether their diet has changed.

- In conclusion, PCA prudent and high-energy diet scores are useful summaries of diet in the SWS, and sociodemographic and lifestyle predictors of these scores will be investigated in Chapter 7. Predictors of change in diet will be considered in the 94 individuals who took part in the stability study, although this analysis is limited because numbers are small.

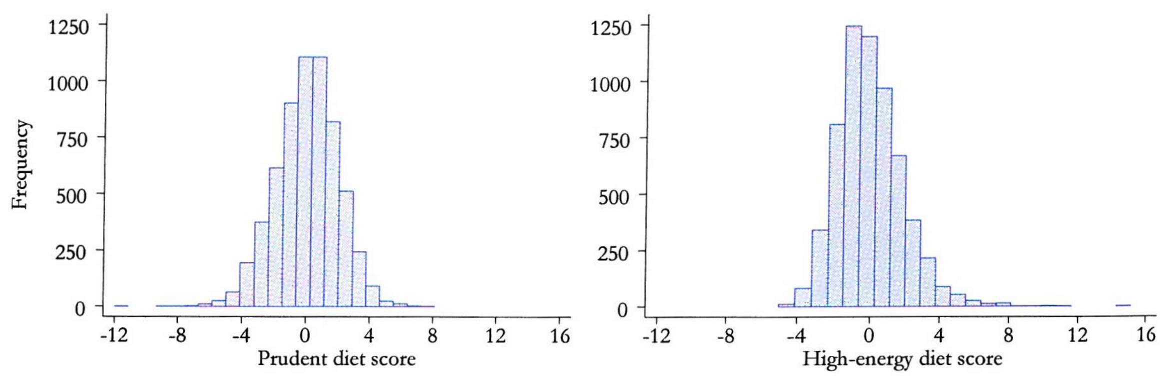
7 Sociodemographic and lifestyle predictors of diet in the SWS

The SWS provides an opportunity to study sociodemographic and lifestyle predictors of diet in a large, contemporary sample of non-pregnant young women from Southampton. Prudent and high-energy diet scores are used as measures of diet. The procedure for investigating the most important sociodemographic and lifestyle predictors of the prudent diet score is presented comprehensively in Section 7.3. Only the final model for predictors of the high-energy diet score is shown in Section 7.4. Predictors of changes in prudent and high-energy diet scores are also considered amongst the 94 women who participated in the SWS stability study in Section 7.5.

7.1 Data preparation

Prudent and high-energy scores in the SWS are illustrated in Figure 39.

Figure 39 Distributions of SWS prudent and high-energy diet scores

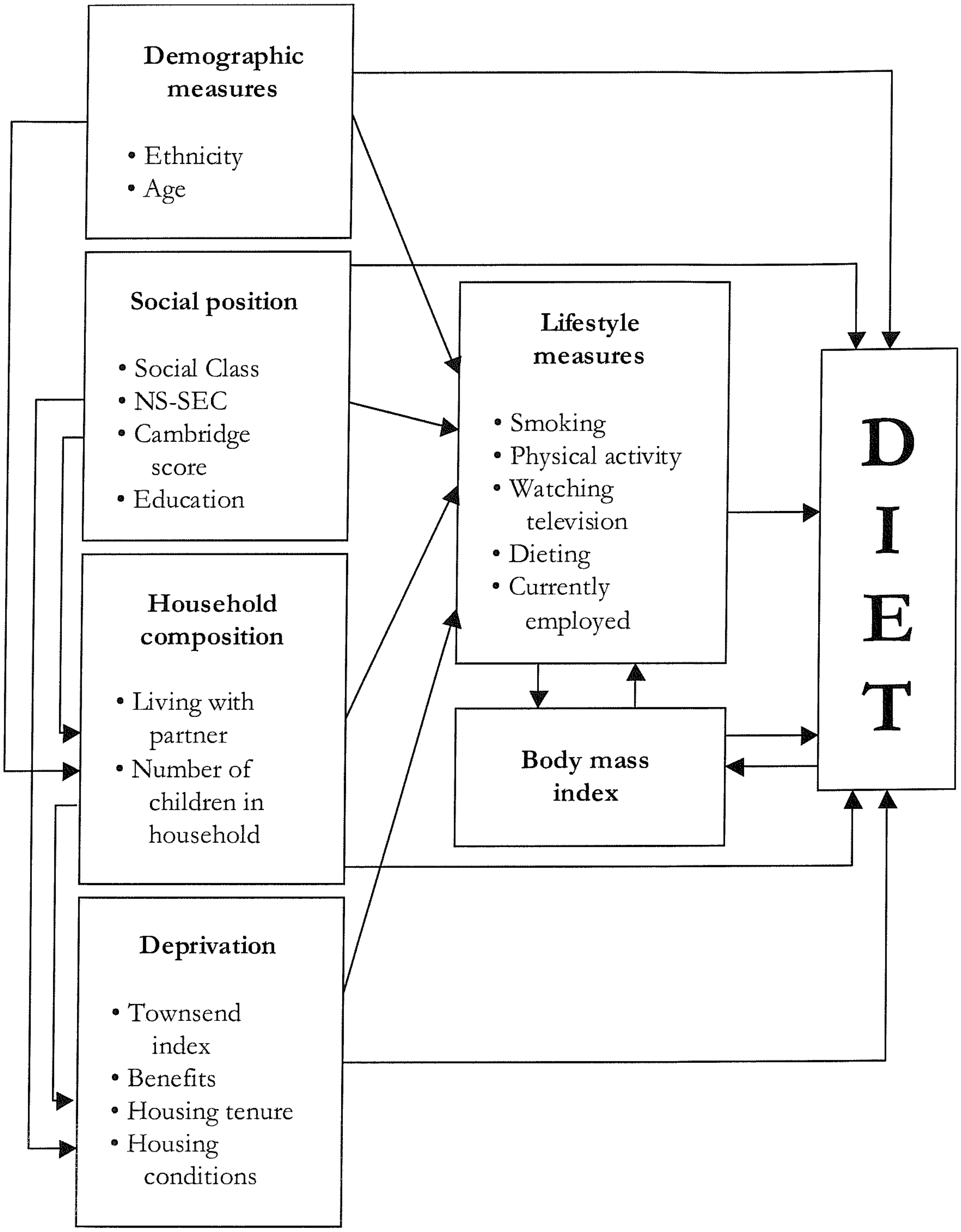


Since the high-energy diet score is not normally distributed, and both scores require standardisation, the prudent and high-energy diet scores were Fisher-Yates transformed[207] using the method described in Section 5.1.

7.2 Framework of pathways of influence on diet

Since the SWS includes the same groups of variables as the PAH study (except that pregnancy effects are not relevant here), Figure 40 is a similar framework of the major pathways of influence on diet to Figure 20. The SWS also provides the opportunity to include additional predictors: ethnicity as an additional demographic measure, NS-SEC as an assessment of social position, benefits, housing tenure and housing conditions to assess deprivation, and watching television as a lifestyle variable. The same reasoning as that described in Section 5.2 was used to generate the framework in Figure 40. It is the basis upon which the SWS models of predictors of diet are built in the same way as Figure 20 in the PAH study.

Figure 40 Framework of proposed major pathways of influence of dietary predictors in the SWS



Considering a P-value of less than 0.05 as statistically significant (as used in analysis of the PAH study) is not appropriate in analysis of the SWS because the number of subjects is so large that a P-value of less than 0.05 does not imply material significance. Instead, therefore, it was calculated that a t-statistic of 1.96 (equivalent to a P-value of 0.05) indicates the same strength of relationship on a sample of 600 as a t-statistic of 6.21 on a sample of 6,000. In the same way, a t-statistic of 1.04 (equivalent to a P-value of 0.3) indicates the same strength of relationship on a sample of 600 as a t-statistic of 3.28 on a sample of 6,000. Therefore predictors are judged worthy of further consideration in the initial SWS analyses if the t-statistic is greater than 3.28. When the final model is being developed, a t-statistic of 6.21 is used to identify important predictors. Nominal categorical measures are considered statistically significant if at least one of the categories is significantly different from the baseline category. All tables of regression models presented in this chapter are mutually adjusted.

7.3 Predictors of prudent diet score

7.3.1 Demographic measures

There is no association between ethnicity and age (Table 75), $P = 0.12$ by analysis of variance.

Table 75 Ethnicity and age in the SWS

Ethnic group	Mean age	n
White	27.8	5691
Black	27.9	84
Indian	28.6	150
Pakistani/Bangladeshi	27.2	69
Other	28.0	118
Total	27.8	6112

There is a strong tendency for women in Indian or Pakistani/Bangladeshi ethnic groups to live with more children and to live in areas of more deprivation as defined by the Townsend index (Table 76).

Table 76 Ethnicity, living with children and Townsend index in the SWS

Ethnic group	Percentage living with children	Mean Townsend index*
White	44%	-0.02
Black	37%	0.14
Indian	59%	0.27
Pakistani/Bangladeshi	71%	0.79
Other	45%	-0.13
Total	45%	0.00

*Scaled to a standard deviation of 1

There are also strikingly strong associations between ethnicity and lifestyle measures (Table 77). Women from ethnic minority groups, particularly Indian or Pakistani/Bangladeshi, were substantially less likely to smoke, to take strenuous exercise and to be currently employed.

Table 77 Ethnicity, living with children and Townsend index in the SWS

Ethnic group	Percentage current smokers	Percentage taking strenuous exercise	Percentage currently employed
White	35%	62%	74%
Black	24%	61%	65%
Indian	7%	45%	63%
Pakistani/Bangladeshi	6%	30%	41%
Other	20%	58%	61%
Total	33%	61%	73%

Of the demographic measures, ethnicity and age were both significant predictors of the prudent diet score (Table 78). Ethnicity was analysed as a nominal categorical variable. Since most women are white this was used as the baseline category, and hence the coefficients in Table 78 are interpreted as increases in the prudent diet score for each ethnic minority group, compared to white women. Older women and those in any ethnic minority group ate more prudent diets.

Table 78 Ethnicity and age as predictors of the SWS prudent diet score

Predictor	β	95% CI for β	t	P
Black^W	0.43	(0.22, 0.65)	4.0	< 0.001
Indian^W	0.42	(0.27, 0.59)	5.2	< 0.001
Pakistani/Bangladeshi^W	0.64	(0.41, 0.88)	5.4	< 0.001
Other^W	0.58	(0.40, 0.76)	6.3	< 0.001
Age†	0.06	(0.03, 0.09)	3.9	< 0.001

Adjusted R^2 = 1.9%, n = 6112 †Per 5 year increase ^{W}Compared to White

The Southampton FFQ was initially developed for use amongst white women[127], and hence the appropriateness of this questionnaire in capturing the diets of women of other ethnic groups in the SWS was uncertain. Women were therefore given the opportunity to report extra foods at the end of the FFQ; those from ethnic minorities indeed reported more extra foods than white women (Table 79). Nevertheless, inclusion of the extra foods did not affect the prudent and high-energy dietary patterns found in this thesis (Section 6.2).

Table 79 Ethnicity and extra foods in the SWS

Ethnic group	Percentage reporting extra foods
White	9%
Black	14%
Indian	41%
Pakistani/Bangladeshi	26%
Other	17%
Total	10%

However, it is still possible that the FFQ does not assess the diet of a particular ethnic minority woman as accurately as that of a white woman. Seven percent of SWS women are from ethnic minorities (Table 13), and therefore errors in their dietary assessment are unlikely to have large influences on the statistical relationships found. Even so the final regression models below are repeated on white women only to ensure the associations are similar.

7.3.2 Social position

Cambridge score, SC, qualifications and NS-SEC are all available in the SWS as measures of social position. As they are measuring similar concepts, they are all strongly associated. Correlations between Cambridge score, SC and qualifications are similar to those seen in the PAH study (Section 5.3.2); that between Cambridge score and SC is –0.81, between Cambridge score and qualifications is 0.50, and between SC and qualifications is –0.46. Relationships with NS-SEC cannot be summarised using correlation coefficients because it is a nominal categorical measure (Section 3.2.3). Instead Figure 41, Figure 42 and Figure 43 indicate that all other measures of social position are strongly associated with NS-SEC. There is a tendency for the qualification levels to be more evenly spread across the NS-SEC groups; all qualification levels are represented in each NS-SEC category, and an analysis of variance indicated that there is less difference between the average qualification level in each NS-SEC group than between the average Cambridge score or SC in each NS-SEC group. Therefore qualification level is less strongly related to other measures of social position, as might be expected since the other measures are all employment-based.

Figure 41 Social class and NS-SEC in the SWS

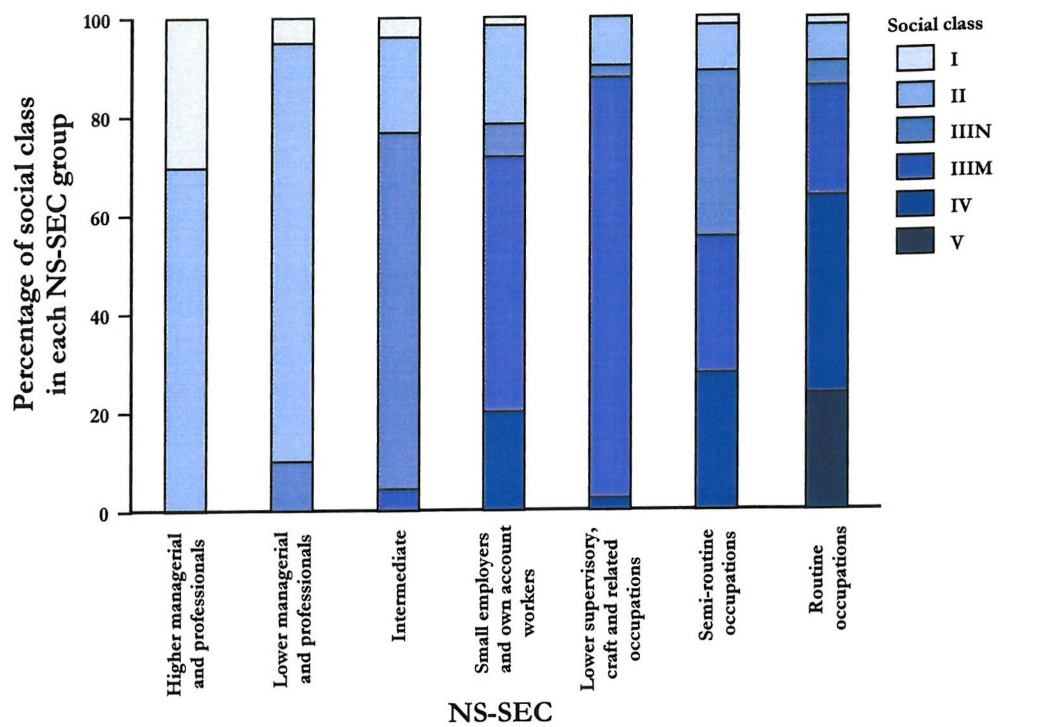


Figure 42 Cambridge score and NS-SEC in the SWS

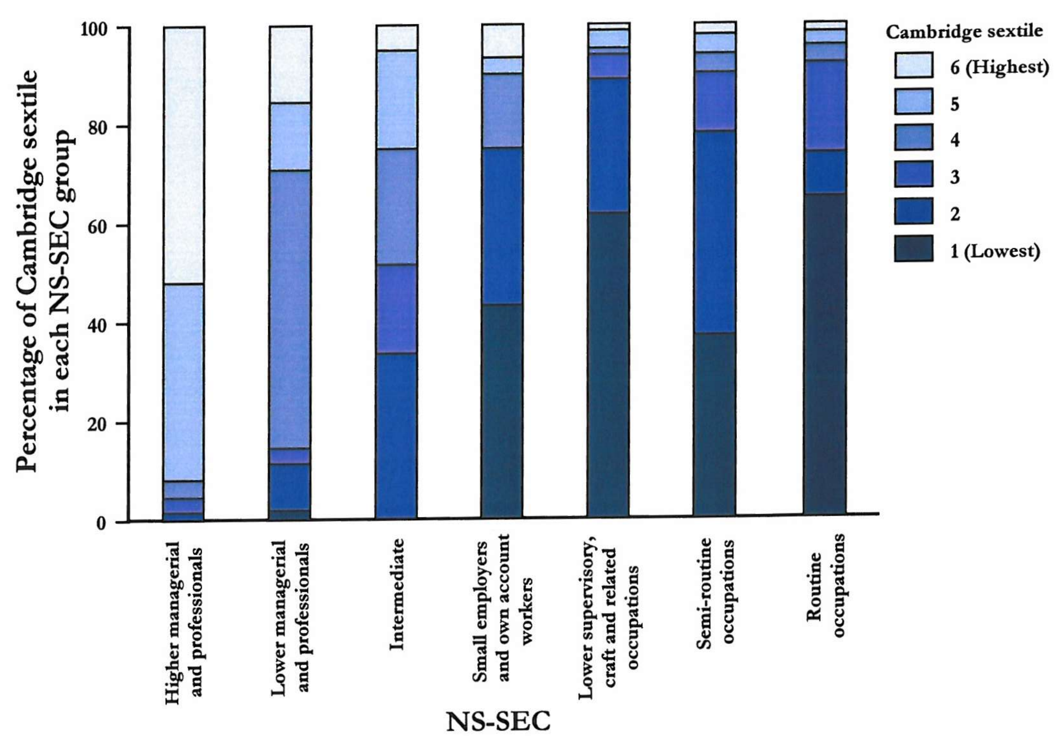
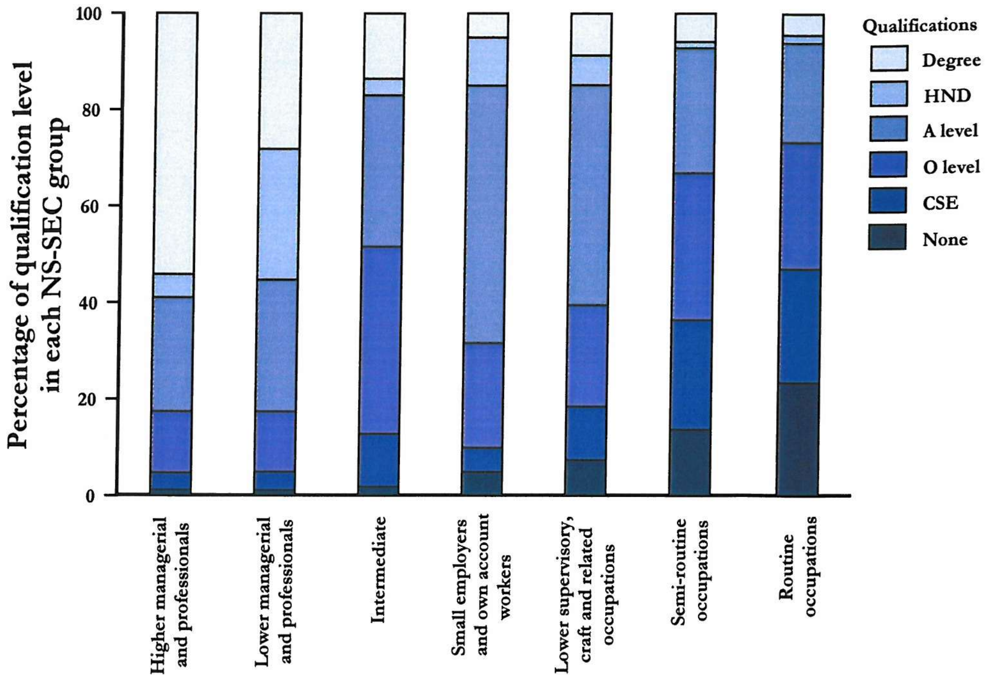


Figure 43 Qualifications and NS-SEC in the SWS



All the measures of social position were univariately highly significant predictors of the prudent diet score. However, when they were all entered into a model together, SC was not significant, and nor were any of the NS-SEC categories once SC had been removed; the resulting model without SC and NS-SEC is shown in Table 80. Women of a higher social position as defined by the Cambridge score and qualification level ate more prudently. These effects were more important than SC and NS-SEC indicating that they have the strongest relationship with prudent diet score. Educational attainment has a particularly strong association (t = 28.4).

Table 80 Social position predictors of the SWS prudent diet score

Predictor	β	95% CI for β	t	P
Cambridge score*	0.13	(0.10, 0.16)	9.7	< 0.001
Qualifications	0.26	(0.24, 0.27)	28.4	< 0.001

Adjusted R^2 = 21.1%, n = 5871 *Scaled to a standard deviation of 1

7.3.3 Household composition

There is a notable trend for women in the SWS living with a partner to have more children (Table 81).

Table 81 Living with a partner and number of children in the SWS

Live with partner	Number of children 0	1	2	3+	Total
No	1893 (**76%**)	284 (**11%**)	196 (**8%**)	123 (**5%**)	2496 (**100%**)
Yes	1509 (**42%**)	735 (**20%**)	885 (**24%**)	504 (**14%**)	3633 (**100%**)
Total	3402 (**56%**)	1019 (**17%**)	1081 (**18%**)	627 (**10%**)	6129 (**100%**)

Women not living with children and living with a partner had a higher prudent diet score when both were included in the model (Table 82). Investigation of the effect of living with children revealed a linear trend across the four categories, rather than a difference between simply living with children and not living with children.

Table 82 Household composition predictors of the SWS prudent diet score

Predictor	**β**	**95% CI for β**	**t**	**P**
Number of children	-0.27	(-0.30, -0.25)	-22.2	< 0.001
Living with a partner	0.20	(0.15, 0.25)	7.7	<0.001

Adjusted $R^2 = 7.4\%$, n = 6125

7.3.4 Deprivation

There were highly significant associations between all the measures of deprivation. Those between housing tenure and Townsend index, and housing tenure and receipt of benefits were particularly strong (Table 83). Women who rented from the council or housing association lived in areas with a notably higher Townsend index score and were substantially more likely to receive benefits than women with other housing tenures.

Table 83 Housing tenure, Townsend index and receipt of benefits in the SWS

Housing tenure	**Mean (SD) Townsend index**	**Percentage receiving benefits**
Owns outright or buying with mortgage	-0.26 (0.95)	5%
Rent from private landlord	0.05 (0.86)	15%
Rent from council or housing association	0.69 (0.83)	60%
Lives with parents, partner or friend	-0.31 (1.06)	10%
Total	0.00 (1.00)	19%

All the measures of deprivation have an independent effect on prudent diet score (Table 84). Owning a home outright or buying with a mortgage (the largest housing tenure group) was used as the baseline housing tenure category. Women with higher Townsend index scores, in receipt of benefits, renting from the council or housing association, or living with parents, partner or friends, and having at least one big housing problem tended to eat a less prudent diet.

Table 84 Deprivation predictors of the SWS prudent diet score

Predictor	β	95% CI for β	t	P
Townsend index*	-0.06	(-0.08, -0.03)	-4.4	< 0.001
Receipt of benefits	-0.35	(-0.42, -0.28)	-9.6	< 0.001
Rent from private landlord[O]	0.19	(0.13, 0.25)	5.9	< 0.001
Rent from council or housing assoc.[O]	-0.41	(-0.49, -0.34)	-10.4	< 0.001
Lives with parents, partner or friend[O]	-0.29	(-0.37, -0.22)	-7.7	< 0.001
At least one big housing problem	-0.19	(-0.25, -0.14)	-7.0	< 0.001

Adjusted R^2 = 12.0%, n = 6014
[O]Compared to Own outright or buying with mortgage
*Scaled to a standard deviation of 1

7.3.5 Lifestyle measures

Associations between lifestyle predictors showed strong similarity with those in the PAH study. Taking strenuous exercise and being currently employed were again more common amongst non-smokers (Table 85). Dieting was again not associated with any of the other lifestyle variables (including hours spent watching television which was not available in the PAH study).

Table 85 Smoking, strenuous exercise and current employment in the SWS

Smoker	Percentage taking strenuous exercise	Percentage currently employed
No	65%	76%
Yes	53%	65%
Total	61%	73%

However, SWS participants who were currently employed were significantly more likely to engage in strenuous activity (Table 86).

Table 86 Current employment and strenuous exercise in the SWS

Currently employed	Percentage taking strenuous exercise
No	50%
Yes	65%
Total	61%

Women in the SWS who smoked, did no strenuous exercise or were not currently employed watched more hours of television (Table 87).

Table 87 Hours watching television, smoking, strenuous exercise and current employment in the SWS

	Smoker		Strenuous exercise		Currently employed	
	No	Yes	No	Yes	No	Yes
Average hours of television per day	2.5	2.9	2.9	2.4	3.1	2.4

All the lifestyle predictors except current employment were independent predictors of the prudent diet score (Table 88). Women who did not smoke, took strenuous exercise, spent less time watching television and were currently dieting had a more prudent diet. Women who worked were more likely to eat a prudent diet in a univariate analysis, but the effect was not significant in the multivariate analysis; this effect was not reduced by the inclusion of just one of the other variables in the model, but rather all (except dieting) somewhat reduced the effect.

Table 88 Lifestyle predictors of the SWS prudent diet score

Predictor	β	95% CI for β	t	P
Smoking	-0.51	(-0.56, -0.46)	-20.8	< 0.001
Strenuous exercise	0.41	(0.36, 0.45)	17.0	< 0.001
Hours watching television	-0.17	(-0.18, -0.15)	19.9	< 0.001
Currently dieting	0.26	(0.20, 0.31)	9.4	< 0.001

Adjusted $R^2 = 20.8\%$, n = 6062

7.3.6 Body mass index

BMI is positively skewed (Figure 12), and was therefore logged before analyses. There was no linear association between BMI and prudent diet score (Table 89). However, further investigation indicated that a quadratic relationship was present (Table 90).

Table 89 Body mass index as a predictor of the SWS prudent diet score

Predictor	β	95% CI for β	t	P
BMI‡	-0.01	(-0.03, 0.02)	-0.6	0.54

Adjusted $R^2 = 0.0\%$, n = 6059 ‡Logged and scaled to a standard deviation of 1

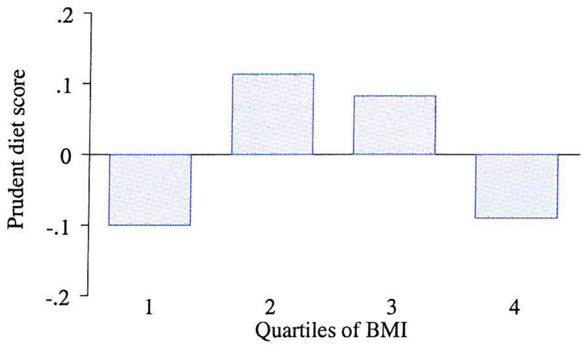
Table 90 Body mass index as a quadratic predictor of the SWS prudent diet score

Predictor	β	95% CI for β	t	P
BMI‡	0.04	(0.02, 0.07)	3.0	0.003
BMI‡ squared	-0.06	(-0.08, -0.05)	-7.0	< 0.001

Adjusted $R^2 = 0.8\%$, n = 6059 ‡Logged and scaled to a standard deviation of 1

Since BMI is squared the effect is more difficult to interpret and therefore Figure 44 shows the relationship between BMI and the prudent diet score. Those with particularly low or high BMIs have lower prudent diet scores than those with average BMIs.

Figure 44 Body mass index and SWS prudent diet score



7.3.7 Summary model

A model of the joint effects of demographic measures and social position is fitted below. Household composition terms are added to this model to assess whether the effects of demographic measures and social position are mediated through household composition (Figure 40). It may be that the effects of demographic measures, social position and household composition act partially through deprivation, and therefore these measures are subsequently added to the model. All measures then significant may have direct effects or may be mediated via lifestyle measures, and so the significant lifestyle variables are fitted. Finally the quadratic effect of BMI is assessed. Decisions about variables to retain are based solely on statistical significance, now indicated by a t-statistic of 6.21 (Section 7.2).

There are independent effects of ethnicity, age, and Cambridge score and qualifications as measures of social position (Table 91).

Table 91 Demographic and social position predictors of the SWS prudent diet score

Predictor	β	95% CI for β	t	P
Black^{W}	0.40	(0.21, 0.61)	4.0	< 0.001
Indian^{W}	0.38	(0.23, 0.53)	5.1	< 0.001
Pakistani/Bangladesh^{W}	0.85	(0.62, 1.08)	7.2	< 0.001
Other^{W}	0.43	(0.27, 0.60)	5.1	< 0.001
Age†	0.10	(0.08, 0.13)	7.4	< 0.001
Cambridge score*	0.11	(0.09, 0.14)	8.4	< 0.001
Qualifications	0.27	(0.25, 0.28)	29.5	< 0.001

Adjusted R^2 = 23.3%, n = 5863 ^{W}Compared to White †Per 5 year increase
*Scaled to a standard deviation of 1

There is an independent effect of living with children (Table 92), but similarly to the PAH study (Section 5.3.8) the effect of living with a partner is non-significant when age is included in the model because living with a partner is strongly associated with age; the average age of those living with a partner in the SWS is 28.9, whereas those not living with a partner are 26.2 years on average.

Table 92 Demographic, social position and household composition predictors of the SWS prudent diet score

Predictor	β	95% CI for β	t	P
Black^{W}	0.37	(0.17, 0.57)	3.6	< 0.001
Indian^{W}	0.44	(0.29, 0.59)	5.9	< 0.001
Pakistani/Bangladeshi^{W}	0.94	(0.71, 1.17)	8.0	< 0.001
Other^{W}	0.45	(0.29, 0.62)	5.4	< 0.001
Age†	0.18	(0.15, 0.21)	11.6	< 0.001
Cambridge score*	0.10	(0.07, 0.13)	7.5	< 0.001
Qualifications	0.23	(0.21, 0.25)	23.7	< 0.001
Number of children	-0.16	(-0.19, -0.13)	-11.7	< 0.001
Living with a partner	0.03	(-0.02, 0.08)	1.1	0.27

Adjusted R^2 = 25.0%, n = 5863 ^{W}Compared to White †Per 5 year increase
*Scaled to a standard deviation of 1

When measures of deprivation are added to the model without living with a partner, the only one that is significant is housing tenure (Table 93), and here it is only the 'renting from council or housing association' category that is significantly different from the baseline category of 'own outright or buying with a mortgage'. The effect of the other measures of deprivation are not significant when measures of social position are included in the model, indicating that lower prudent diet scores for those with a higher Townsend index, in receipt of benefits and who have at least one big housing problem are in fact reflecting the lower prudent diet scores of women of a lower social position.

Table 93 Demographic, social position, household composition and deprivation predictors of the SWS prudent diet score

Predictor	β	95% CI for β	t	P
Black[W]	0.38	(0.18, 0.57)	3.7	< 0.001
Indian[W]	0.46	(0.31, 0.61)	6.2	< 0.001
Pakistani/Bangladeshi[W]	0.94	(0.71, 1.17)	8.1	< 0.001
Other[W]	0.39	(0.22, 0.55)	4.6	< 0.001
Age†	0.16	(0.13, 0.19)	9.9	< 0.001
Cambridge score*	0.08	(0.06, 0.11)	6.3	< 0.001
Qualifications	0.21	(0.19, 0.23)	21.4	< 0.001
Number of children	-0.14	(-0.17, -0.11)	-9.8	< 0.001
Rent from private landlord[O]	0.03	(-0.03, 0.10)	1.1	0.28
Rent from council or housing assoc.[O]	-0.27	(-0.30, -0.17)	-7.2	< 0.001
Lives with parents, partner or friend[O]	-0.16	(-0.24, -0.08)	-3.9	< 0.001

Adjusted R^2 = 25.8%, n = 5824
*Scaled to a standard deviation of 1
[W]Compared to White
†Per 5 year increase
[O]Compared to Own outright or buying with mortgage

When lifestyle measures are added to this model, Cambridge score and housing tenure are no longer significant. Thus the effects of Cambridge score and deprivation as measured by housing tenure are mediated via lifestyle measures. Further investigation indicated that smoking, strenuous exercise and hours watching television, rather than dieting, removed the effects of Cambridge score and housing tenure. Once all these terms were in the model quadratic BMI terms were also no longer significant; this is due to the presence of both qualifications and strenuous exercise in the model. As seen in the PAH study, there is a quadratic relationship between social position and BMI. SWS women of a high social position (as measured by qualification level) are unlikely to have a very high BMI, but there is less evidence than in the PAH study of SWS women of a high social position being unlikely to have a low BMI.

The final model of all independent predictors of SWS prudent diet score is given in Table 94; virtually identical relationships were seen when the same model (without ethnicity) was fitted in the subgroup of 5,691 white women.

Table 94 All significant predictors of the SWS prudent diet score

Predictor	β	95% CI for β	t	P
Black[W]	0.37	(0.19, 0.55)	4.1	< 0.001
Indian[W]	0.37	(0.23, 0.50)	5.4	< 0.001
Pakistani/Bangladeshi[W]	0.86	(0.66, 1.06)	8.5	< 0.001
Other[W]	0.44	(0.29, 0.59)	5.7	< 0.001
Age[†]	0.18	(0.15, 0.21)	12.8	< 0.001
Qualifications	0.19	(0.18, 0.21)	23.4	< 0.001
Number of children	-0.13	(-0.15, -0.11)	-10.4	< 0.001
Smoking	-0.32	(-0.36, -0.27)	-13.6	< 0.001
Strenuous exercise	0.29	(0.25, 0.34)	12.7	< 0.001
Hours watching television	-0.11	(-0.13, -0.09)	-13.8	< 0.001
Currently dieting	0.31	(0.26, 0.36)	12.4	< 0.001

Adjusted R^2 = 33.2%, n = 6033 [W]Compared to White †Per 5 year increase

As a graphical illustration of the individual effects of each predictor, women were assigned to four equal-sized groups, according to quarters of the prudent diet score. As for the PAH study, the proportion of women with prudent diet scores in the lowest quarter is used as a summary measure. Figure 45 shows the proportion of women in the lowest quarter of the prudent diet score by levels of each predictor, adjusted for other factors in the model.

Figure 45 Proportion of women in the lowest quarter of SWS prudent diet score, by significant predictors of the score

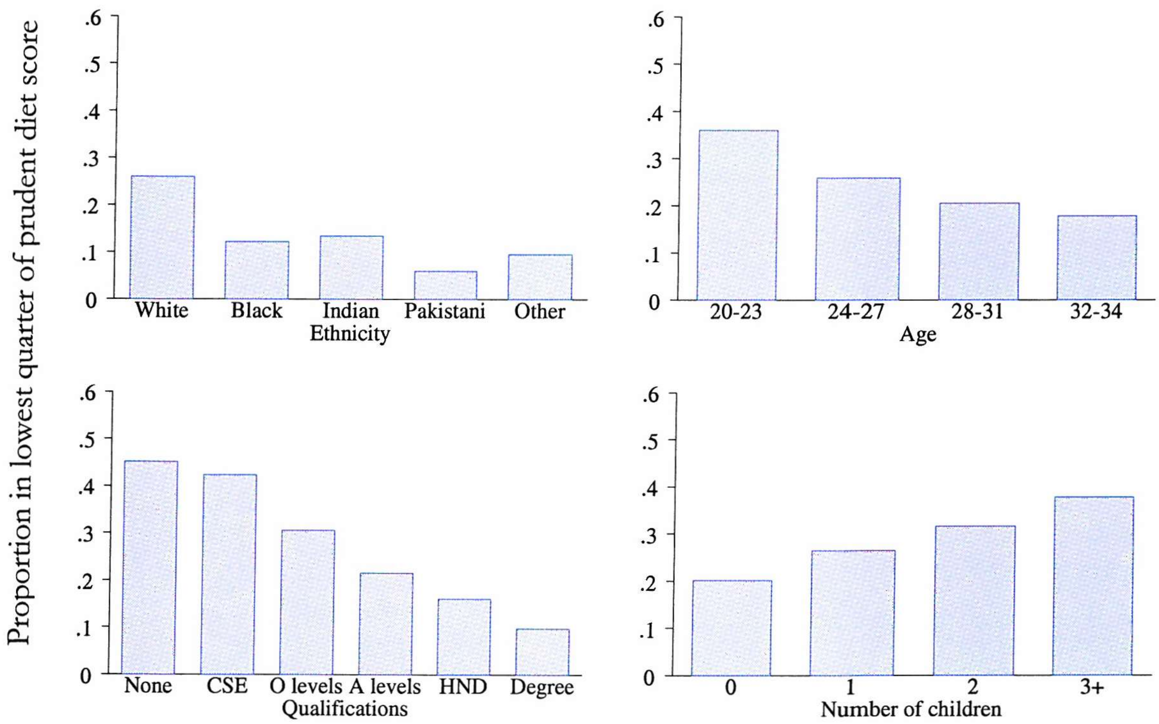
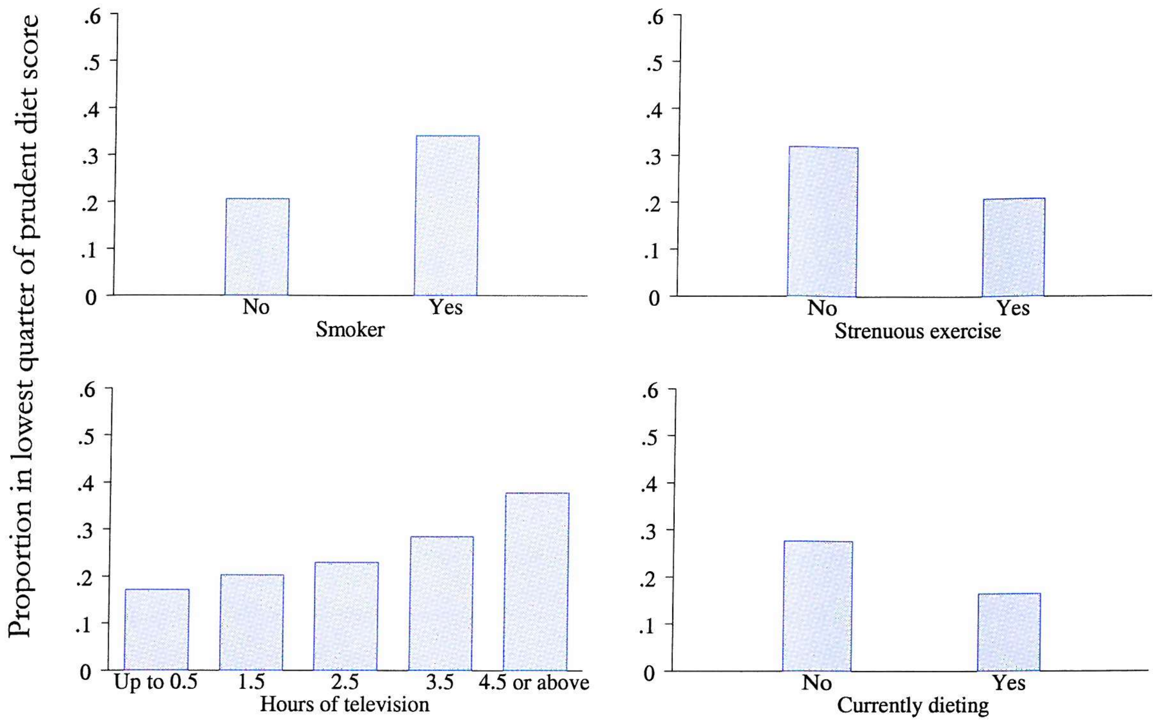


Figure 45 (continued) Proportion of women in the lowest quarter of SWS prudent diet score, by significant predictors of the score



Qualifications is the strongest predictor of prudent diet score, and has a coefficient very similar to those seen in the PAH study early and late pregnancy data (Table 61 and Table 62). The effects of age, number of children and smoking are also very similar to those seen in the PAH study. Cambridge score is not a significant predictor in the SWS, although it was in the PAH study. However, strenuous exercise and dieting are significant in the SWS but were not in the PAH study. Also significant in the SWS are ethnic group and hours watching television, which are not available in the PAH study. The effects of age, qualifications, number of children and hours watching television in the SWS were found to be linear when investigated further. None of the two-way interactions between the predictor variables were statistically significant.

7.4 Predictors of high-energy diet score

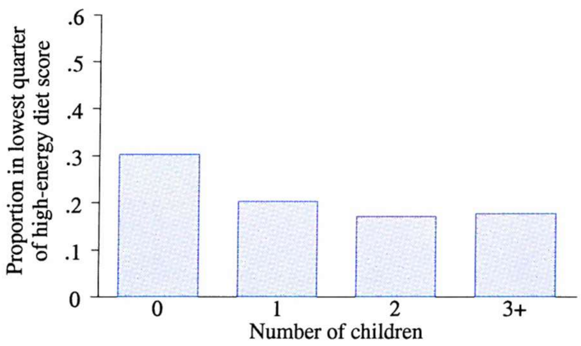
The pattern of predictors of the high-energy score in the SWS is also of interest. Therefore the same procedure as that used in Section 7.3 was used to initially investigate associations within each group of predictors (Figure 40), and then to explore the multivariate associations. Only number of children was a significant multivariate predictor of high-energy diet score (Table 95). Women who lived with more children ate a higher-energy diet (Figure 46). This same relationship was observed when the model was fitted to the subgroup of 5,691 white women.

Table 95 All significant predictors of the SWS high-energy diet score

Predictor	β	95% CI for β	t	P
Number of children	0.17	(0.15, 0.20)	14.5	< 0.001

Adjusted R^2 = 3.3%, n = 6125

Figure 46 Proportion of women in the lowest quarter of SWS high-energy diet score, by significant predictors of the score



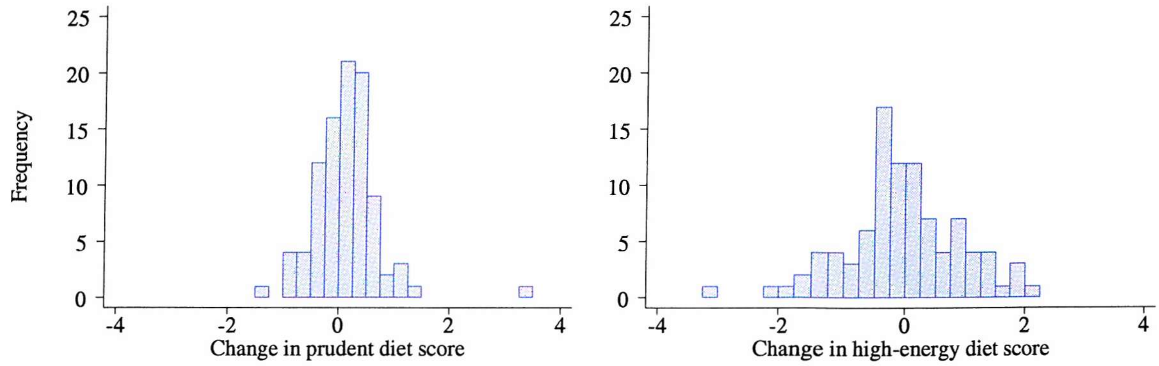
The effect of number of children on high-energy diet score is similar in the SWS to the PAH study (Table 63 and Table 64); number of children was the most significant predictor of high-energy diet score in both early and late pregnancy PAH data. However, Cambridge score, dieting and BMI, which were seen to be significant predictors of the high-energy diet score in both early and late pregnancy PAH data were not significant in the SWS. Cambridge score had a negative relationship with high-energy diet score (t = -3.6, P < 0.001) in a univariate linear regression model, but this relationship was completely non-significant (P = 0.44) when number of children was included in the model. Dieting also had a negative relationship, but was non-significant both with and without number of children included in the model. BMI is an important linear predictor of high-energy diet score in the PAH study. However, in the SWS there was a suggestion of a quadratic relationship, but this was not significant enough to be included in the final model (t = 4.0).

7.5 Predictors of changes in dietary scores in the stability study

7.5.1 Data preparation

Sociodemographic and lifestyle predictors of changes in prudent and high-energy scores, based on the original coefficients, are considered here amongst the 94 women who were reinterviewed approximately two years after their initial interview for the SWS 'stability study'. Since both the original and repeat scores were divided by the standard deviation of the original score (Section 6.4), the change is again interpreted in standard deviations of the original score. The distributions of change in prudent and high-energy diet scores between original and repeat questionnaires are shown in Figure 47.

Figure 47 Distributions of change in SWS prudent and high-energy diet scores between original and repeat questionnaires



A regression to the mean effect similar to that in the PAH study (Section 5.5.1) was observed (r_S = -0.38 for prudent diet score and r_S = -0.37 for high-energy diet score). Therefore Prevost et al's[211] regression technique was again used to derive a measure of a woman's change in score assuming she started at a standard original score. These adjusted changes in scores were transformed to normality using Fisher-Yates normal scores[207], and the resulting scores scaled to have the same standard deviation as the regression-adjusted scores, so that changes can be interpreted in terms of standard deviations of the original data. The same procedure as that used in Section 7.3 was employed to investigate sociodemographic and lifestyle predictors of changes in prudent and high-energy diet scores. Since there were only 94 subjects it was more appropriate to use the conventional cut-off of 0.05 to denote statistical significance. Graphical illustrations of the effects described in the final regression models are not included because the small numbers lead to inaccurate summaries using the lowest quarter of increase in scores.

7.5.2 Increase in prudent diet score

Strenuous exercise at initial interview was the only significant predictor of increase in prudent diet score (Table 96), such that women taking strenuous exercise were likely to change to a more prudent diet in the subsequent two years. Women taking strenuous exercise were also likely to eat a more prudent diet at the initial SWS interview (Table 94).

Table 96 All significant predictors of increase in SWS prudent diet score

Predictor	β	95% CI for β	t	P
Strenuous exercise	0.36	(0.16, 0.56)	3.6	0.001

Adjusted $R^2 = 11.2\%$, n = 94

Only 3 of the 94 women in the stability study were from an ethnic minority group; the same relationship between change in score and strenuous exercise was observed in the 91 white women.

7.5.3 Increase in high-energy diet score

NS-SEC and BMI squared were significant predictors of high-energy diet score, as described in Table 97. 'Higher managerial and professionals' was chosen as the baseline category of NS-SEC. Note that the 'small employees and own account workers', and 'lower supervisory, craft and related occupations' categories are omitted since the first of these contained no participants, and the second contained only one participant. Women with 'lower managerial and professional', 'intermediate', 'semi-routine' and 'routine' occupations were more likely than 'higher managerial and professionals' to change to a higher-energy diet. The linear BMI term should not be interpreted alone, but rather the linear and quadratic terms must be considered together. Again, since BMI is squared the effect is more difficult to interpret and therefore Figure 48 shows the relationship between change in high-energy diet score and BMI adjusted for NS-SEC. Women with average BMIs had a greater reduction in high-energy diet score than those with high and low BMIs. The same relationships were seen when the analysis was restricted to white women.

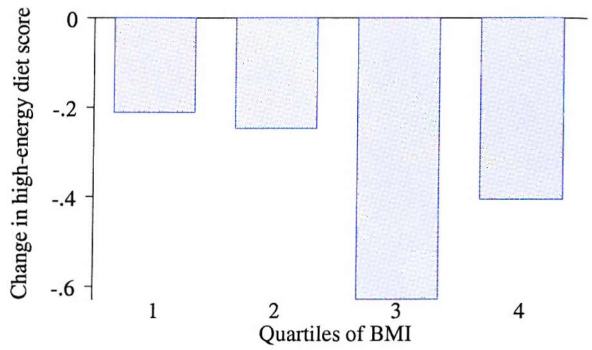
Table 97 All significant predictors of increase in SWS high-energy diet score

Predictor	β	95% CI for β	t	P
Lower managerial and professionals^{H}	0.28	(-0.25, 0.81)	1.1	0.29
Intermediate^{H}	0.49	(-0.03, 1.01)	1.9	0.07
Semi-routine occupations^{H}	0.31	(-0.25, 0.86)	1.1	0.28
Routine occupations^{H}	0.89	(0.08, 1.70)	2.2	0.03
BMI‡	-0.21	(-0.41, -0.01)	-2.1	0.04
BMI‡ squared	0.20	(0.05, 0.36)	2.7	0.008

Adjusted $R^2 = 11.1\%$, n = 83 ^{H}Compared to Higher managerial and professionals

‡Logged and scaled to a standard deviation of 1

Figure 48 Body mass index and change in SWS high-energy diet score



7.6 Summary

- Women in the SWS who belonged to an ethnic minority group, were older, better qualified, lived with fewer children, didn't smoke, took strenuous exercise, watched fewer hours of television, and were currently dieting were likely to report a more prudent diet. These effects explain 33.2% of the variation in prudent diet score. The effects of age, qualifications, number of children and smoking were very similar to those seen in the PAH study.

- Women in the SWS who lived with more children reported eating a higher-energy diet. This effect was similar to that seen in the PAH study, although it only explained 3.3% of the variation in the diet score. None of the other predictors in the PAH study were significant in the SWS.

- Amongst the 94 women participating in the stability study, increases in prudent diet score over the two year period of the stability study were associated with taking strenuous exercise.

- Increases in high-energy diet score were associated with NS-SEC and BMI. Women with 'lower managerial and professional', 'intermediate', 'semi-routine' and 'routine' occupations were more likely than 'higher managerial and professional' women to change to a higher-energy diet. Those with average BMIs had a greater reduction in high-energy diet score than those with high and low BMIs.

8 Discussion

Women's diets are important in terms of their own health, and that of their children. These motivations, and the ensuing objectives of this thesis are reviewed in Section 8.1. Robust dietary patterns were found in the PAH study and the SWS, and these are described in Section 8.2 with reference to other patterns in the literature; sociodemographic and lifestyle predictors of these patterns are summarised.

The strengths and limitations of the data on which this thesis is based, as well as the methods employed, are considered in Section 8.3. Finally, the implications of the work and suggestions for future research are discussed in Section 8.4, with particular reference to barriers to healthy eating that may exist amongst young women today.

8.1 Motivation

The diet of a woman both before and during pregnancy is an important determinant of her own health and of the growth and development of her child (Section 1.1.3). Impaired fetal growth is known to be linked to higher rates of adult disease in later life (Section 1.1.4). Socioeconomic differentials in morbidity and mortality have been widely reported, and it is thought that diet throughout the life-course may be a mediator of such inequalities in health. Indeed, the 1998 Independent Inquiry into Inequalities in Health[15] noted the importance of a mother's nutrition, both before and during pregnancy, in terms of effects on the health of the next generation. Eating patterns vary with sociodemographic and lifestyle information, and therefore it is valuable to identify characteristics of women most likely to have inadequate diets so that public health messages can be most appropriately targeted.

Traditionally diet has been assessed using nutrients, but this approach has many limitations, such as the large number of nutrients that are present in foods. A major aim of this thesis was therefore to perform dietary pattern analyses using principal component analysis to characterise overall diet in young women.

Two datasets were used: the Princess Anne Hospital study of 620 pregnant women, and the Southampton Women's Survey of 6,129 non-pregnant women aged 20 to 34. Occupational-based characterisation of women's social position is often problematic because relatively few women work full-time. These datasets therefore provided an opportunity to explore the accurate measurement of women's social position, as well as other sociodemographic and lifestyle characteristics.

Diet was assessed using a food frequency questionnaire, but the parallel use of four-day food diaries in the PAH study enabled comparison of dietary patterns resulting from diaries with those resulting from the FFQ. Additional questions in the SWS provided an opportunity to investigate weighting for cross-check fruit and vegetable questions, and to examine the impact of asking respondents to report on extra foods at the end of an FFQ. Both the PAH and SWS datasets had longitudinal components, and a further aim of the thesis was therefore to investigate changes in women's diets, between early and late pregnancy in the PAH study, and over a two-year period in the SWS. The final thesis objective was to examine sociodemographic and lifestyle predictors of the dietary patterns using a framework of pathways of influence on diet.

8.2 Main findings

8.2.1 Dietary patterns

8.2.1.1 Prudent and high-energy patterns

The 'pragmatic food grouping' was developed in this thesis as a method of grouping the original 100 dietary variables from the Southampton MRC FFQ such that the resulting PC scores had a high correlation with PC scores derived from the original data (Section 4.1.3.3 and Section 6.1). Principal component analysis on the PAH and SWS data using these 49 food groups produced dietary patterns that were called 'prudent' and 'high-energy' diets. These coefficients are combined below (Table 98) from Table 34 and Table 67, for ease of reference.

Table 98 Prudent and high-energy diet coefficients in the PAH study and SWS

Food	Prudent diet			High-energy diet		
	PAH early	**PAH late**	**SWS**	**PAH early**	**PAH late**	**SWS**
White bread	-0.20	-0.21	-0.22	0.14	0.19	0.10
Wholemeal bread	0.26	0.27	0.23	**0.04**	**0.01**	**0.04**
Breakfast cereals	0.06	0.06	0.13	0.09	0.08	0.06
Crackers	0.10	0.11	0.07	0.13	**0.05**	0.10
Rice and pasta	0.22	0.20	0.21	0.09	0.12	0.13
Quiche and pizza	0.08	**0.05**	**-0.04**	0.08	0.10	0.12
Yorkshire pudding and savoury pancakes	**-0.06**	**-0.11**	-0.18	0.25	0.17	0.12
Cakes and biscuits	**-0.01**	**0.01**	**-0.13**	0.24	0.28	0.19
Puddings	**0.05**	0.07	**-0.05**	0.27	0.19	0.23
Full-fat milk (pints)	**-0.15**	-0.22	-0.20	0.19	0.13	0.10

Table 98 (continued) Prudent and high-energy diet coefficients in the PAH study and SWS

Food	Prudent diet			High-energy diet		
	PAH early	PAH late	SWS	PAH early	PAH late	SWS
Reduced-fat milk (pints)	0.15	0.19	0.05	-0.09	-0.08	0.02
Cream	0.09	0.13	0.03	0.14	0.14	0.13
Cheese and cottage cheese	0.21	0.22	0.08	0.03	0.12	0.14
Yoghurt	0.23	0.24	0.16	0.06	0.10	0.13
Eggs and egg dishes	0.00	0.00	-0.04	0.18	0.15	0.21
Full-fat spread	-0.03	-0.05	-0.17	0.18	0.22	0.16
Reduced-fat spread	0.03	0.06	0.06	-0.01	-0.06	0.01
Cooking fats and salad oils	0.01	0.01	0.11	0.13	0.18	0.16
Chicken and turkey	0.13	0.09	-0.02	0.10	0.14	0.12
Red meat	0.05	0.01	-0.17	0.23	0.25	0.23
Processed meat	-0.04	-0.07	-0.21	0.28	0.28	0.21
Offal	-0.03	-0.07	-0.10	0.15	0.13	0.14
Fish and shell fish	0.20	0.14	0.09	0.06	0.14	0.18
Salad vegetables	0.30	0.29	0.25	0.05	0.09	0.21
Green vegetables	0.21	0.17	0.17	0.09	0.15	0.25
Root vegetables	0.17	0.15	0.08	0.08	0.10	0.23
Other vegetables	0.27	0.23	0.23	0.10	0.11	0.23
Beans and pulses	0.04	0.02	0.02	0.12	0.09	0.15
Vegetable dishes and vegetarian foods	0.22	0.17	0.21	-0.02	0.00	0.03
Tinned vegetables	-0.16	-0.18	-0.14	0.18	0.14	0.13
Boiled potatoes	0.01	-0.06	-0.05	0.19	0.15	0.16
Roast potatoes and chips	-0.11	-0.19	-0.27	0.22	0.19	0.10
Crisps	-0.15	-0.12	-0.19	0.16	0.17	0.10
Citrus fruit and fruit juices	0.19	0.15	0.14	0.04	0.08	0.13
Other fruit	0.20	0.23	0.23	0.02	0.11	0.16
Other fruit juices	0.10	0.08	0.05	0.00	-0.01	0.09
Nuts	0.11	0.07	0.05	0.09	0.02	0.09
Dried fruit	0.18	0.18	0.14	0.00	0.01	0.07
Cooked and tinned fruit	0.13	0.06	0.03	0.15	0.15	0.17
Confectionery	-0.10	-0.07	-0.16	0.22	0.22	0.14
Sweet spreads	0.12	0.10	0.06	0.09	0.10	0.11
Added sugar (teaspoons)	-0.20	-0.23	-0.23	0.22	0.17	0.07
High-energy soft drinks	-0.09	-0.07	-0.11	0.15	0.15	0.10
Diet coke	-0.01	-0.01	-0.02	0.01	-0.03	0.00
Tea and coffee	-0.13	-0.13	-0.16	0.17	0.18	0.07
Decaffeinated tea and coffee	0.10	0.12	0.10	-0.03	-0.07	0.02
Drinking chocolate	0.01	0.03	0.01	0.11	0.06	0.06
Wine	0.04	0.15	0.12	-0.03	-0.03	0.01
Miscellaneous	0.09	0.06	-0.04	0.21	0.25	0.26
Proportion of variation explained	8.7%	9.0%	7.6%	7.1%	6.1%	7.0%

The prudent and high-energy dietary patterns were very similar in early and late pregnancy in the PAH study. The SWS prudent dietary pattern was comparable to that found in the PAH study, although the coefficients for meat were lower in the SWS. This may be due to health messages arising from concerns about associations between red and processed meat consumption and cancers, particularly colorectal cancer, between 1991 and 1998[212]; if those women who eat more prudent diets are choosing to eat less meat then the axis along which women's diets vary most will have the additional feature of contrasting the low meat-eating of those who eat prudent diets with the higher meat-eating of the non-prudent consumers. The high-energy dietary pattern was also comparable in the two studies, although the coefficients for fruit and vegetables in the SWS were notably higher, and those for added sugar markedly lower. National increases in fruit and vegetable consumption and decreases in sugar consumption were seen in the DNSBA between 1986-7 and 2000-2001 (Section 6.1). Consistent trends were observed when the original consumption frequencies in the PAH study and the SWS were tabulated (data not shown), although the differences in coefficients resulting from the principal component analyses indicate that the trends are greatest amongst those consuming a high-energy diet.

The prudent and high-energy patterns were consistent across different schemes of food grouping (Section 4.1.3), substantiating the results of McCann et al.[61]. The pragmatic food grouping (Section 4.1.3.3) was chosen because it was designed as an appropriate grouping of the Southampton FFQ. Furthermore, it produced very high correlations with the PCA scores from all 100 foods and thus little information was lost by grouping the foods; all correlations were 0.96 or greater in the PAH study and the SWS.

The prudent diet is a major international dietary pattern (Section 1.3.1), and the pattern found in this thesis agrees well with the descriptions and coefficients in the literature. In the PAH study it was strongly positively correlated with all micronutrients and protein, showed small negative correlations with fat and starch, and small positive correlations with sugars (Table 35). The prudent dietary pattern was also positively correlated with serum vitamin C measurements in the PAH study and with red cell folate levels in the SWS. These external measures of micronutrient intake corroborate Fung et al.'s[71] finding that the prudent dietary pattern was associated with a favourable biomarker profile.

In addition to the prudent dietary pattern, research in the United States has consistently revealed a Western dietary pattern, indicating high intakes of meat, especially red meat, refined grains, chips, and high-fat products, and low levels of fruit and vegetables (Section 1.3.1). The high-energy pattern in Table 98 certainly has similarities with the Western pattern, although comparisons can be difficult because complete tables of coefficients are not presented in the literature. The same high intakes of red and processed meat, refined grains and sweet foods are seen in Table 98 as in the Western pattern. However, the high-energy pattern, particularly in the SWS, does not show the low levels of fruits and vegetables mentioned by Slattery et al.[56] who initially identified the Western pattern. Furthermore, it is notable that virtually all coefficients for the high-energy dietary pattern are positive; it is common in principal component analysis to find one of the first components indicating overall 'amount', and therefore the high-energy interpretation would fit this precedent.

In the PAH study the high-energy pattern had a correlation of 0.87 with energy intake in both early and late pregnancy. Amongst the literature describing Western dietary patterns only Slattery et al.[56] and McCann et al.[61] provide correlations with energy intake. Slattery et al.'s Western diet had a correlation with energy intake of 0.76 amongst men, and 0.67 amongst women. McCann et al.'s 'high-fat' diet looks very similar to the Western diet, and correlations between this factor and energy intake range between 0.32 and 0.63, depending on the grouping used. The high-energy diet pattern in this thesis thus demonstrated a higher correlation with energy intake than the Western pattern. Therefore, although the high-energy diet has similarities with the Western dietary pattern, there are noticeable differences; it may be that the Western dietary pattern was not found in the PAH and SWS data because this pattern is particularly prominent in the United States.

The prudent and high energy diet scores together account for between 14.6% and 15.8% of the variation in the 49 foods and food groups (Table 98). Direct comparisons cannot be made across the literature since the proportion of variation explained by a component is highly dependent on the number of variables entered into a PCA, such that less variation is likely to be explained by the first two components from a PCA with more variables. However, if the variation is spread equally across all components, the proportion expected to be explained by the first two components is calculated as two divided by the number of variables. Therefore multiples of the percentage of variation explained compared to the percentage expected can be compared, as described by Whichelow and Prevost[65]. Table 99 displays these values for those papers in the literature finding prudent and Western patterns and providing information about percentage of variation explained, alongside the PAH and SWS study results. The percentage of variation explained by the PAH and SWS patterns is comparable with that in the literature, and is in fact greater than all other studies except the 56 and 168 FFQ item patterns of McCann et al.[61]. Therefore, although the dietary patterns found in the PAH study and the SWS may ostensibly appear to explain a relatively small percentage of variation in diet, they are in fact highly successful compared to other similar analyses.

Much of the literature presenting dietary patterns uses factor analysis, rather than PCA, and therefore various methods of factor analysis were considered as an additional level of sophistication to PCA (Section 4.2). Only miniscule differences in individual scores were seen between methods of extraction used. Rotation also had a relatively small impact on individual scores and did not appear to improve interpretability. Moreover, a desirable property of unrotated PCs is in defining the axes along which participants vary the most. Therefore factor analysis was not considered further.

Table 99 Percentage of variation explained by first two components in PCA and factor analysis in the literature, the PAH study and the SWS.

Study	Percentage explained	No. of FFQ items	Percentage expected	Multiple
Slattery et al.[56] (males)	18.1	35	5.7	3.2
Slattery et al. (females)[56]	15.2	35	5.7	2.7
Gex-Fabry et al.[58]	13.8	33	6.1	2.3
Osler et al.[60]	19.5	26	7.7	2.5
McCann et al.[61]: 168 FFQ items	7.7	168	1.2	6.5
McCann et al.[61]: 56 FFQ items	13.4	56	3.6	3.8
McCann et al.[61]: 36 FFQ items	16.9	36	5.6	3.0
Terry et al.[63]	17.6	24	8.3	2.1
PAH early pregnancy	15.8	49	4.1	3.9
PAH late pregnancy	15.1	49	4.1	3.7
SWS	14.6	49	4.1	3.6

8.2.1.2 Diaries

Four-day food diaries were kept by 588 women in early pregnancy in the PAH study. Principal component analysis of the resulting data revealed patterns again described as prudent and high-energy diets (Table 38). The question of whether individual women's diets differ between FFQ and diary records according to the prudent and high-energy patterns was addressed by calculating each woman's score from the two dietary assessment methods, both based on the FFQ coefficients. The prudent diet scores using diaries tended to be 0.18 standard deviations lower than the FFQ with limits of agreement –1.69 to 1.32. The high-energy diet scores using diaries tended to be 0.33 standard deviations lower than the FFQ with limits of agreement –2.43 to 1.78. This could be described as moderate agreement for the prudent diet score, and moderate-to-weak agreement for the high-energy diet score. FFQs have been seen to over-report fruit and vegetable intake[131,132], a phenomenon which may have contributed to the tendency for individual's prudent diet scores to be lower using diary data than FFQ data. The drop in high-energy diet score between FFQs and diaries is likely to be related to known overestimation of energy using FFQs[126].

However, since the main aim of an FFQ is to rank participants according to nutrient intake[127], scores on the diary and FFQ components were also compared using Spearman's correlation coefficients. These were 0.68 for the prudent diet and 0.24 for the high-energy diet, which compare well with the correlations found between 20 nutrient intakes that ranged from 0.26

to 0.44[127]. Hu et al.[57] contrasted results from PCA analyses on FFQ and diary data in a cohort of 127 male health professionals living in the Boston area. Two FFQs were administered one year apart and correlations between the diary prudent pattern and the two FFQ prudent patterns were 0.34 and 0.41, whilst those between the diary Western pattern and the two FFQ Western patterns were 0.51 and 0.64. Thus the prudent pattern in this thesis showed a better correlation than that of Hu et al., whereas the high-energy pattern correlation was worse than Hu's Western pattern. It may be that the PAH correlations would have been improved if two food diaries had been collected and within-person variation thus reduced by averaging the intake from the two diaries, as implemented by Hu et al.

Correlations between diary and FFQ scores were not influenced by the woman's nausea in early pregnancy (Table 39) or whether she considered she had changed the amount she was eating since becoming pregnant (Table 40). However, the nurses' subjective recording of diary quality showed a strong association, such that women whose diaries were coded as 'very poor, probably incomplete' had much lower correlations between both their prudent and high-energy FFQ and diary scores (Table 41). Thus discrepancies seen between diaries and FFQs are likely to be due in part to poor diary quality.

There was little effect of over-reporting on the FFQ or under-reporting in the diaries on correlations between diary and FFQ scores (Table 42 and Table 43). This indicates that discrepancies between diaries and FFQs are not principally due to the systematic under or over-reporting issues inherent in these methods of dietary assessment.

Instead, variations between patterns resulting from diaries and FFQs would be expected since these methods of collection have many differences. The PAH diaries were only collected over four days and are thus inevitably less representative of general food intake than FFQs which refer to a three month period. The short diary time frame also means that the lowest possible frequency response was once every four days. The period of diary recording occurred after the three months referred to by the FFQ, and in pregnancy, when women may change their diets relatively rapidly, this difference in time period may be significant. Also, FFQ reports cover weekends in the correct proportion of two weekend days to every five week days, whereas the diary method was less accurate by asking women to include one weekend day amongst the four days on which they kept diaries. Finally, the coding of food diaries was initially performed to generate nutrient intake, and for this reason a few foods were entered as ingredients, making the FFQ and diary categories less comparable.

It is therefore remarkable that there was so much agreement for the prudent dietary pattern between the diaries and the FFQ, both in the pattern of coefficients produced and the association of individual scores. Agreement for the high-energy pattern was also evident, but not as strong.

No method of dietary assessment is error-free, and thus diary data should not be considered a gold standard. As Tseng[214] notes, even if diaries were to assess food intake more accurately they do not necessarily identify food patterns any better. Feskanich et al.[131] suggest that comparisons between FFQ and diary data are called 'calibration' rather than 'validation' studies. Low correlations between FFQ and diary scores amongst diaries coded as 'very poor, probably incomplete' indicate that a proportion of the disparity is likely to be due to errors in diary as well as FFQ data. These analyses confirm that an FFQ yields useful dietary patterns that are associated with patterns from diary data, particularly for the prudent diet score. Therefore FFQs are likely to be a valuable tool in further research where diet is analysed using principal component analysis.

8.2.1.3 Extra foods

The SWS questionnaire allowed women to report additional foods not included in the FFQ that they had eaten once a week or more. Matching these extra foods to the FFQ categories enabled assessment of their impact on dietary patterns. Principal component analysis on data including information about extra foods made negligible differences to the dietary patterns (Table 68). Individual women's dietary scores were also virtually unchanged by the use of data about intake of extra foods; Bland and Altman limits of agreement were –0.06 to 0.06 for the prudent diet score and –0.07 to 0.08 for the high-energy diet score. Therefore allowing women to include extra foods in their responses to the FFQ is not considered essential to future questionnaires in terms of principal component analyses.

8.2.1.4 Fruit and vegetables

Fruit and vegetable intake is frequently used in epidemiological research as a simple measure of diet because the consumption of five fruit and vegetables per day is an important government health message (Section 1.2.1) due to strong evidence for a protective effect in studies of important diseases such as cancer, and coronary heart disease[144]. High fruit and vegetable intake is recognised by European citizens to be associated with a healthy diet[145].

In the PAH study 41% of women reported achieving the recommended five portions of fruit and vegetable intake per day in their early pregnancy FFQ. Fruit and vegetables intake was found to decrease slightly between early and late pregnancy by 0.4 portions per day on average (Bland and Altman limits of agreement –4.1 to 3.2).

The known tendency for fruit and vegetables to be over-reported using FFQs[131,132] was seen in comparisons with diary data in the PAH study; on average each participant reported 1.2 more portions per day using the FFQ than the food diary. This may also be attributed to under-reporting using food diaries, by reducing intake during the four day period, or not recording all foods consumed.

Fruit and vegetable intake in the SWS was slightly higher than in the PAH study (Table 69); 52% of SWS women reported consuming at least five portions of fruit and vegetables per day. As described in Section 6.3.1 this may be due to an increase in national consumption between the two study periods, or the fact that women in the SWS have somewhat higher qualifications.

Fruit and vegetable intake is highly correlated with the prudent diet score in both the PAH study and the SWS (Table 44 and Table 70), as expected from the high coefficients for fruit and vegetable intakes in the prudent diet pattern. The high-energy dietary pattern is also positively correlated with fruit and vegetable intake. This association is strongest in the SWS and weakest in the PAH diaries, reflecting the coefficients for fruit and vegetables in each of the principal component analyses.

Fruit and vegetable cross-check questions were included in the SWS to provide information about over-reporting of this aspect of diet in the FFQ. Calvert et al.[132] suggest using cross-check questions to weight FFQ responses as described in Section 6.3.2. The impact of this weighting on a PCA was investigated and coefficients were seen to change very little by the inclusion of Calvert et al.'s terms (Table 71), again demonstrating the robust nature of the technique. Bland and Altman limits of agreement for change in individuals' scores from with to without weighting were –0.95 to 0.24 for the prudent diet and –1.29 to 0.26 for the high-energy diet, and thus agreement was generally good. The use of the weighting terms was not pursued because the validity of the cross-check questions was not justified by Calvert et al., and the SWS cross-check question only refers to consumption over the past week, rather than the past three months. Furthermore, the weighting factor does not take account of differential mis-reporting within the fruit and vegetable categories. Therefore cross-check questions are not recommended as a method of addressing over-reporting on FFQs in future questionnaires in terms of impact on principal component analysis.

8.2.1.5 Changes in dietary scores

Diet in the PAH study was assessed in both early and late pregnancy. In the SWS a second interview was conducted approximately two years after the initial interview in the 94 non-pregnant women comprising the 'stability study'. These repeated measurements enabled assessment of changes in prudent and high-energy diets over time.

Bland and Altman limits of agreement for changes in prudent and high-energy diet scores in the PAH study and the SWS are shown in Table 74. Agreement between prudent diet scores was generally good, and that between high-energy diets could be described as moderate. The limits were narrower for those in the SWS reporting that they hadn't made any major changes to their diet in the previous two years, indicating that women are to some extent able to report whether their diet has changed.

Whilst high-energy diet scores did not consistently increase or decrease over time, there were changes in average prudent diet scores. PAH women's diets showed a significant decrease in prudent diet score by an average of 0.10 standard deviations between early and late pregnancy (Section 4.1.4), whilst the prudent diet scores of SWS women tended to increase by 0.13 standard deviations over the two years of the stability study (Section 6.4). In analyses of predictors of the prudent diet score, older age was significantly associated with higher SWS scores (Table 94), although this could be a cohort rather than an ageing effect. Furthermore this effect would predict an increase of only 0.07 standard deviations over a two year period. However, the increase in age over the two years of the SWS stability study may explain some of the increase seen in prudent diet scores. The remaining increase may be due to national changes in diet; for example, further investigation revealed the increase in prudent diet score amongst the 94 women in the repeatability study to be due in part to increased intake of salad vegetables, which has been a feature of changes in diet over recent years as recorded in the DNSBA[66,213]. The decrease in prudent diet score amongst PAH study women is interesting and less expected, although the effect size is small. It is corroborated by the average decrease in fruit and vegetable consumption of 0.4 portions per day between early and late pregnancy described in Section 4.4.1. Further investigation also revealed increases in consumption of red and processed meat, cakes, biscuits, puddings, confectionery, tea and coffee.

That the Bland and Altman limits of agreement for changes in score accurately represent true agreement rests on the assumption that the Southampton FFQ is perfectly reproducible. However, FFQs are not perfectly reproducible (Section 2.3.2) and therefore some of the lack of agreement between each interview must be attributed to measurement error on each occasion. Nevertheless, lack of reproducibility is unlikely to cause the consistent increases and decreases in prudent diet score. Therefore the consistent decreases in prudent diet score amongst PAH participants, and the consistent increases amongst SWS participants, are important outcomes of the objective of examining changes in women's diets, and merit replication in other studies.

8.2.1.6 Summary

Throughout these analyses of dietary patterns the prudent diet score is seen to be more robust than the high-energy diet score; in particular the coefficients are more comparable between early PAH, late PAH and SWS analyses, and between diary and FFQ analyses. The prudent dietary pattern is also robust internationally, being found in all analyses described in Section 1.3.1 except Barker et al.[69] and van Dam et al.[70]. Results regarding the prudent diet scores are therefore likely to be most relevant to other populations, although the high-energy diet score is nonetheless of interest.

Therefore, the objective of employing suitable statistical methods for characterising the diets of young women has been accomplished. The first two components from a PCA both have useful interpretations, and could be valuable when considering predictors of dietary behaviour, as utilised in this thesis, or when studying diet as a predictor of health outcomes.

8.2.2 Predictors of prudent diet score

Predictors of the prudent diet score were investigated using a framework of major pathways of influence (Figure 20 and Figure 40). Many similarities were seen amongst the significant predictors in the PAH study and the SWS, although differences were also apparent, as summarised in Table 100. In general effect sizes tended to be slightly larger in the late pregnancy PAH analyses than the early pregnancy or SWS analyses, particularly for number of children and smoking. This results in the large adjusted R^2 of 46.7% in late pregnancy. In the case of smoking this indicates that women who continue to smoke into late pregnancy are those with worse prudent diet scores than those who give up smoking.

Table 100 β coefficients of significant predictors of prudent diet score

Predictor	PAH early pregnancy	PAH late pregnancy	SWS
Black[w]	–	–	0.37
Indian[w]	–	–	0.37
Pakistani/Bangladeshi[w]	–	–	0.86
Other[w]	–	–	0.44
Age[†]	0.25	0.26	0.18
Cambridge score	0.18	0.17	
Qualifications	0.17	0.20	0.19
Number of children	-0.12	-0.20	-0.13
Townsend index*		-0.08	
Smoking	-0.37	-0.43	-0.32
Strenuous exercise			0.29
Hours watching television	–	–	-0.11
Currently dieting			0.31
Adjusted R^2	33.5%	46.7%	33.2%

– Not available in PAH study
[w]Compared to White [†]Per 5 year increase *Scaled to a standard deviation of 1

8.2.2.1 Demographic measures

All women in the PAH study were white, but ethnicity but was seen to have an important effect on the SWS prudent diet score such that women from ethnic minorities, particularly those who considered they belonged to Pakistani/Bangladeshi ethnic groups, reported more prudent diets. These effects were mediated very little by other variables in the model. This result is notable, particularly in light of the mixed conclusions of the effects of ethnicity on diet summarised in Section 1.4.1.4. Further investigation revealed the effect to be mainly due to markedly greater consumption of salad vegetables, vegetables other than green or root vegetables, fruit and fruit juices, vegetable dishes and vegetarian foods amongst the ethnic minorities, alongside less frequent intakes of Yorkshire pudding and savoury pancakes, processed meat and offal, and tea and coffee. However, these positive influences on the prudent diet score were offset somewhat by lower intakes of breakfast cereals, cheese and wine, and higher intakes of added sugar and high-energy soft drinks by women from ethnic minorities. These differences are likely to be due to cultural influences on food choice.

Older age is associated with a more prudent diet in both the PAH study and the SWS. It is the most significant predictor of a prudent diet in the PAH study and has a slightly stronger effect here than in the SWS. Older age has been seen to be associated with greater intake of fruit and vegetables, fibre, and vitamins and minerals, as well as women's scores on a 'health-conscious' component (Section 1.4.1.5), and thus this result is in agreement with the consistent pattern in the literature.

Macintyre and Anderson[195] suggest that age differences may be due to biological states and processes, historical processes, or psychological and social processes (Section 3.4.1.2). In this context, historical processes would indicate that a cultural change caused women in younger age groups to choose different eating patterns from their parents. On the other hand biological, psychological or social processes may act to change women's food choices as they age. Unfortunately it is not possible to differentiate between these effects with cross-sectional data. Considering the longitudinal data available, there is a slight tendency for women's prudent diet scores to drop over the three months of the PAH study and to rise over the two years of the SWS (Section 8.2.1.5). Since the 3 months of the PAH is a relatively short period this is likely to be related to pregnancy effects. The increase over the two years of the SWS stability sample is larger than that expected from the regression model. Although the size of the stability sample is small this gives an indication that the association between age and higher prudent diet score may be largely due to biological, psychological or social processes occurring as the women age, rather than historical processes.

8.2.2.2 Social position

Measures of social position were strongly associated with each other (Section 5.3.2 and Section 7.3.2) but qualifications was the least strongly related to the others. Cambridge score and qualifications were significant predictors of the prudent diet scores in the PAH study, with very similar positive effect sizes in early and late pregnancy. Qualifications was the strongest predictor of the prudent diet score in the SWS, but Cambridge score was not as important here. The positive relationships between prudent diet scores and social position are in agreement with the widely recognised associations between other measures of diet and social position in the literature (Section 1.4.1.3).

NS-SEC was available in the SWS as a measure of social position, but was not a significant predictor of the prudent diet score. This indicates that the employment aspects of a woman's social position, such as her autonomy at work, are not as important as other measures of social position in influencing her diet along a prudent axis.

Social Class is a measure of household social standing and is available in both the PAH study and the SWS. It was not significant in predicting the prudent diet score in either study, indicating that SC is not as important as Cambridge score and qualifications in terms of explaining the prudent dietary pattern. This may be due at least in part to the fact that the SC scheme is based on a male occupational structure, and is therefore a less accurate means of classifying women's employment (Section 3.2.2).

Cambridge score is significant as a predictor of the prudent diet score in the PAH dataset indicating that this measure of household social standing is strongly associated with the prudent diet score. Since Cambridge scale scores are available separately for men and women (Section 3.2.4) it may more accurately reflect women's social position than SC. The Cambridge Scale is also based on friendship patterns and it might be that this dimension of social standing has more influence on eating habits than the occupational skill referred to by SC.

Although the Cambridge score is not a significant predictor of the SWS prudent diet score, it is only marginally insignificant when added to the final model ($t = 5.27$, compared to the cutoff used for inclusion in this large sample of $t = 6.21$). Further investigation indicated that the exclusion of hours of television from this model increased the t-statistic to $t = 6.26$. Therefore if information about hours of television watching had been available in the PAH study it may have been that the effect of Cambridge score would be partially mediated through this lifestyle variable. However, this would be unlikely to account for all of the difference in Cambridge score effect and the disparity could also be due to lack of comparability of the datasets or chance findings. Furthermore, the Cambridge score was developed on data collected before 1990 and thus changes in employment characteristics over time may mean that the Cambridge score provides a more accurate representation of women's social position in the PAH study than the SWS.

The effect of education on prudent diet score is prominent; it was the second most significant predictor of a high score in the PAH study, and by far the strongest in the SWS. This result concurs with Thompson et al.[215] who observed that education was more predictive of fruit and vegetable intake than manual/non-manual occupation, and Groth et al.[91] who found that education was a stronger predictor of increased fruit and vegetable intake and reduced percentage energy from fat than Danish occupational status. It is unlikely that the effect of education is through teaching about food and nutrition since prudent diet scores are seen to increase across all levels of education (Figure 23, Figure 24 and Figure 45), whereas little is taught about food technology after the age of 16. It is also unlikely that the association between prudent diet score and education is due to the higher incomes of those with more education since there was little effect of adding measures of deprivation, which most closely reflect income, to a model containing education.

Instead there is some evidence that people with more education live under less stress[216,217] and thus it may be that more educated women have a greater capacity to apply health messages to their lives. They may also have greater problem-solving abilities to enable them to meet the challenge of improving their diet. Moreover, education could alter individuals' values, such that preventative health measures are considered of greater importance as education increases[218,219]. This may be due to a greater understanding of the meaning and implications of health messages. Fuchs[220] suggests that both education and health-related activities, such as eating a more prudent diet, involve incurring current costs for the sake of future benefits. Individuals who are more willing or able to invest in education may similarly be more willing to invest in a healthier diet.

8.2.2.3 Household composition

Living with a partner was univariately associated with a more prudent diet in both the PAH study and the SWS. However, this became non-significant when age was added to the model, reflecting the fact that living with a partner is strongly related to age.

Living with children was a significant predictor of the prudent diet score, such that every child a woman lived with decreased her score by between 0.12 and 0.20 standard deviations, after controlling for other factors in the models (Table 100). It is of concern that women living with children, and presumably often catering for them, are themselves eating a poorer diet. The direction of this effect is in agreement with observations that participants living with children had lower intakes of fruit and vegetables in Denmark[91] and a deprived area of Leeds[93], but is at variance with studies in Norway[103] and Finland[88] where adults living with children were seen to eat more fruit and have healthier diets in general.

8.2.2.4 Deprivation

The Townsend index was the only measure of deprivation available in the PAH study, and was marginally statistically significant as a predictor of the prudent diet score in the late pregnancy final regression model ($P = 0.03$). In the SWS all measures of deprivation available were initially predictive of the prudent diet score. However, further exploration revealed that the lower prudent diet scores for those with a higher Townsend index, in receipt of benefits, and who had at least one big housing problem, were in fact reflecting the lower prudent diet scores of women of a lower social position. The lower scores of those renting from the council or housing association, or living with parents, partners or friends, were mediated through several lifestyle variables: smoking, strenuous exercise and hours watching television.

Thus since measures of deprivation are relatively unimportant in predicting the prudent diet score, it could be more effective to target public health messages at those of a lower social position, particularly as defined by education, and those with less healthy lifestyles.

8.2.2.5 Lifestyle measures

Smoking was an important predictor of a lower prudent diet score in both the PAH study and the SWS. Smokers had a prudent diet score between 0.32 and 0.43 standard deviations less than non-smokers (Table 100). This is in agreement with Dallongeville et al.'s meta-analysis[105] of 51 studies which concluded that smokers have unhealthier patterns of nutrient intake than non-smokers. The effect of smoking might be causal through regulation of appetite, mood, or taste[105], or alternatively smoking and food choice may both be linked to attitudes towards a healthy lifestyle. Whatever the cause, the poorer diets of smokers are of particular concern since they are likely to exacerbate the detrimental effects of smoking itself[105,198].

SWS participants who took strenuous exercise in the previous three months had prudent diet scores 0.29 standard deviations higher than those who took no strenuous exercise, concurring with other findings in the literature (Section 1.4.1.7). The effect of physical activity in the PAH study was not significant once measures of social position were included in the model. This may be because the PAH question enquires about physical activity in the previous week, which could be a less sensitive discriminator of women who choose to exercise, or because women who choose to exercise in pregnancy have different characteristics to those who choose exercise when they are not pregnant. Like smoking, the effect of exercise may be causal through effects on appetite, or both exercise and diet may be a result of individual attitudes about healthy lifestyles.

Hours spent watching television was only measured in the SWS, and was a strongly negative predictor of the prudent diet score. Watching television is likely to indicate a more sedentary lifestyle, which may be associated with diet. Furthermore, food choices are prone to influence by the content of messages received through the television media, that are often at variance with healthy eating guidelines[221]. Thompson et al.[215] found that participants eating most meals in the living room in front of the television were significantly more likely to be low fruit and vegetable consumers.

Women in the SWS who report that they are currently dieting to lose weight have a prudent diet score on average 0.31 standard deviations higher than women not currently dieting. This effect is not surprising, since weight-loss diets are generally more prudent. The effect of dieting was not significant in the PAH study, but this is likely to be because the PAH question about dieting relates to the previous year, and is therefore not very specific to the FFQ data which refers to the previous three months.

If a typical pattern of dieting involved exchanging less prudent foods for more prudent ones, then it could be considered tautological to include 'currently dieting' as a predictor of a prudent diet outcome. However, it is also possible that patterns of dieting to lose weight amongst young women involve curtailing consumption of all foods to reduce energy intake, and since the prudent diet score is not associated with energy intake (Table 35), this would not cause a less prudent diet directly. In this case the association between dieting to lose weight and a higher prudent diet score may be due to an interest in pursuing healthy dietary habits. Nevertheless, the final model of all significant predictors of the SWS prudent diet score without currently dieting is virtually identical to that in Table 100.

Current employment did not make any independent contribution to the final models of predictors of prudent diet scores; in the PAH study it was strongly related to education, and in the SWS there were clear associations with other lifestyle variables.

8.2.2.6 Body mass index

In both the PAH study and the SWS there was a univariate quadratic relationship between BMI and the prudent diet score, such that women in the higher or lowest quartiles of BMI were likely to have a less prudent diet than those in the central two quartiles (Figure 22 and Figure 44). However, a quadratic relationship between social position and BMI indicated that women of a high social position were less likely to have very high or very low BMIs. Thus the quadratic effect of BMI was not significant in the final models of independent predictors of the prudent diet score that included social position (Table 100).

8.2.2.7 Pregnancy effects

Since women in the PAH study were pregnant it was important to investigate how effects specific to pregnancy may have influenced diet. Cravings and aversions are very common in pregnancy[222] and are mainly attributed to factors such as nausea or changes in taste[222]. Thus questions in the PAH study about nausea and change in amount of food consumed since becoming pregnant were analysed as influences on diet in the PAH study. Although there was a significant effect of nausea when only pregnancy effects were considered (Table 55), the result did not remain in the final model since nausea is negatively associated with age[203].

8.2.2.8 Interactions

The only two-way interactions that were statistically significant for predictors of the prudent diet score were the interactions between Townsend index and the two measures of social position (qualifications and Cambridge score) in late pregnancy in the PAH study (Section 5.4.1). This interaction demonstrates that the positive effect of social position is greater amongst women living in more deprived areas. In late pregnancy public health messages would therefore best be targeted at women who are both living in a deprived area and of a lower social position. However, since the effect of Townsend index and this interaction are only relevant in late pregnancy, a more pragmatic approach relevant to most young women would be to direct public health messages primarily towards women of a lower social position.

8.2.3 Predictors of high-energy diet score

Predictors of the high-energy diet score were investigated using the same framework of major pathways of influence as the prudent diet score (Figure 20 and Figure 40). Similarities were seen between the predictors in early and late pregnancy in the PAH study, but the only significant predictor in the SWS was the number of children the woman lived with (Table 101). Effect sizes were generally largest in the PAH early pregnancy analyses, and seven variables were significant predictors here, resulting in the largest adjusted R^2 of 20.0%.

Table 101 β coefficients of significant predictors of high-energy diet score

Predictor	PAH early pregnancy	PAH late pregnancy	SWS
Age†	-0.13		
Cambridge score*	-0.19	-0.11	
Number of children	0.25	0.16	0.17
Smoking	0.21		
Dieting	-0.24	-0.21	
BMI‡	-0.13	-0.13	
Change in amount of food	0.15		–
Adjusted R^2	20.0%	6.6%	3.3%

– Not relevant to the SWS
*Scaled to a standard deviation of 1
†Per 5 year increase
‡Logged and scaled to a standard deviation of 1

8.2.3.1 *Demographic measures*

Ethnicity was measured in the SWS, but did not have a significant effect on the high-energy diet score. Younger age was associated with a higher-energy diet score in early pregnancy in the PAH study. However, this effect was not seen in late pregnancy or in the SWS. In late pregnancy there was a negative association univariately, but this was not strong enough to be significant once other variables were included in the model. In the SWS the effect of age was positive, but there was also a strong positive association between age and the number of children a woman lived with, such that number of children was more important than age in explaining the high-energy diet score. The negative association between age and high-energy diet score in the PAH study is supported by a similar association with Beaudry et al.'s[68] high-energy density factor.

8.2.3.2 *Social position*

The Cambridge score was not significant in the SWS as a predictor of the high-energy score, but was important in the PAH study. The Cambridge score was similarly seen to be a significant predictor of the prudent diet score in the PAH study, but not in the SWS (Section 8.2.2.2); the disparity for high-energy diet score could likewise in part be attributed to the greater relevance of the Cambridge score to data collected nearer 1990.

In the PAH study women with lower Cambridge scores tended to eat higher-energy diets. A similar negative association between social position and energy intake was reported by Smith and Baghurst[100]. Once Cambridge score was included, no other measures of social position were significant predictors of the high-energy diet score in the final models (Table 101).

The Cambridge score was a significant predictor of the prudent and high-energy diet score in both early and late pregnancy in the PAH study, and it is notable in every case that graphical summaries appear to indicate a dichotomous effect around the median. Models with this dichotomous variable explain very similar amounts of variation to those with the continuous variable. Therefore, rather than there being a continuous relationship between Cambridge score and diet, the influence of having a 'high' score may be particularly different from having a 'low' score. However, the use of such a dichotomy in future studies would not reduce the employment information required from each woman. As all information would be available to calculate the continuous variable this would be recommended since the Cambridge scale is based on the concept of a finely graded hierarchy of social order (Section 3.2.4).

8.2.3.3 Household composition

Living with a partner was not a significant predictor of the high-energy diet score in any of the final models. However, the number of children a woman lives with was the most significant predictor of her high-energy diet score in all models, and the only significant predictor in the SWS (Table 101). This result is of interest, and may indicate a lifestyle effect resulting from living with children.

8.2.3.4 Deprivation

No measures of deprivation were significant predictors of the high-energy diet score. In the PAH study this was due to the presence of the Cambridge score as an assessment of social position in the model. In the SWS being in receipt of benefits, renting from the council or housing association, or having a big housing problem were all associated with eating a higher-energy diet. However, they were also all associated with living with more children, and therefore once living with children was included in the model these measures of deprivation were no longer significant. Thus the effect of deprivation acts on the high-energy diet score mainly through the effects of living with children. It may be that since the number of children a woman lives with is easier to quantify than her level of deprivation, it has a stronger effect for this reason.

8.2.3.5 Lifestyle measures

Smoking is positively associated with a high-energy diet in the PAH early pregnancy data. Perkins[223] reviewed seven large studies and found that caloric intake was greater amongst smokers than non-smokers in five of them, in agreement with the PAH results. However, Haste et al.'s[111] study of pregnant women, included in Perkins's review, found higher caloric intake amongst non-smokers, although this difference was not statistically significant. A positive relationship between high-energy diet score and smoking in the SWS was not strong enough to be included in the final model.

The negative association between dieting and high-energy diet score in the PAH study is largely unsurprising since most diets to lose weight involve a reduction in caloric intake, although the fact that the question about dieting relates to the year prior to interview and the FFQ refers to the three months before may have been less predictable. A negative relationship between these two variables in the SWS was not large enough to be included in the final model.

Physical activity was unrelated to high-energy diet score. A positive association between energy intake and a measure of energy expenditure might be expected if the participants are not gaining or losing weight; more sedentary individuals have lower energy needs, as evidenced by the Physical Activity Ratio calculations in the Dietary Reference Values for Food Energy and Nutrients for the United Kingdom[47]. Watching television is another measure that can indicate reduced physical activity, but was similarly found to be unrelated to the high-energy diet score in the SWS.

In the PAH early pregnancy data women who did not work were initially found to have a higher-energy diet, but this effect was not significant in the final model. In late pregnancy current employment was not even significant in the initial model, probably because the variable is not a good discriminator at this stage in pregnancy since 80% of women are not working (Table 20). Current employment in the SWS is strongly negatively related to living with children and therefore the initial association in the same direction as the PAH study was not significant once number of children was added to the model.

8.2.3.6 Body mass index

Pre-pregnant BMI is negatively related to high-energy diet score in the PAH data. Women with a higher weight have a greater demand for energy, assuming that they have the same activity levels as women with lower weight, and are gaining weight through pregnancy at the same rate. The high-energy diet score was not related to activity levels. However, lower weight gains in pregnancy were seen amongst more obese women ($r_S = -0.08$), and this indicates that high BMI is to some extent genuinely associated with lower high-energy diet scores. Nevertheless, obese individuals are often observed to underreport intake[126,133], and such reporting biases may also strengthen the negative associations between high-energy diet score and pre-pregnant BMI in the PAH study, although if such biases were present in the SWS they were not strong enough to result in a significant association.

8.2.3.7 Pregnancy effects

As might be expected, women who claim to be eating more since becoming pregnant (at the early pregnancy interview) have higher-energy diets in early pregnancy. At the late pregnancy interview this effect is no longer relevant, almost certainly because the question asking about changes in amounts eaten relates to the period from conception to the early pregnancy interview.

8.2.3.8 Interactions

The only two-way interaction that was statistically significant for predictors of the high-energy diet score was the interaction between dieting and Cambridge score in both early and late pregnancy in the PAH study (Section 5.4.2 and Section 5.4.3). Figure 27 and Figure 29 demonstrate that the effect of Cambridge score on high-energy diet score was negligible amongst dieters, but was strongly negative amongst non-dieters; the effects were very similar in early and late pregnancy. This implies that the decision to diet is having a greater impact on the diets of those of a lower Cambridge score, whereas dieting is relatively ineffective in reducing the high-energy diet score in women with a higher Cambridge score, whose diets show a lower energy intake anyway.

8.2.4 Predictors of change in prudent diet score

There are no similarities between predictors of change in the prudent diet score in the PAH study and the SWS (Table 102), which is perhaps unsurprising since the three month longitudinal period of the PAH study is relatively short and is more likely to be influenced by factors associated with pregnancy.

Table 102 β coefficients of significant predictors of increase in prudent diet score

Predictor	PAH	SWS
Age†	0.10	
Social Class	-0.10	
Number of children	-0.11	
Strenuous exercise		0.36
Adjusted R^2	11.1%	11.2%

†Per 5 year increase

Older women, those in a higher Social Class, and women living with fewer children tended to increase their prudent diet score between early and late pregnancy. Older women and those with a higher social position (as defined by Cambridge score and qualifications) and women living with fewer children tended to eat more prudent diets in early pregnancy (Table 100). However, in this analysis change in score was adjusted for initial score to prevent associates of the initial score spuriously correlating with the change in score (Section 5.5.1). Thus these results indicate that some characteristics of those women who eat prudently at early pregnancy are also associated with improvements in diets over the course of their pregnancy. Similarly the increase in prudent diet score over the two year longitudinal period of the SWS amongst those who took strenuous exercise at the initial interview reflects the positive association between initial prudent diet score and strenuous exercise.

A comparable pattern was noted by Prevost et al.[211], where non-manual groups and those with desirable lifestyle habits (non-smoking and modest alcohol consumption) responded positively to dietary health guidelines over the seven year period of the Health and Lifestyle Survey. The concerns the authors raise about the implications of these findings for the widening of the health divide is thus enhanced by the data in this thesis. These results imply that dietary advice targeted at women who are younger, of a lower social class, living with more children and not taking strenuous exercise should therefore include assistance in maintaining a healthy diet in the longer term, since these women are the ones whose dietary prudence is more likely to deteriorate.

8.2.5 Predictors of change in high-energy diet score

There are also no similarities between predictors of change in the high-energy diet score in the PAH study and the SWS (Table 103), again probably due to the relatively shorter 3 month PAH time period, which is more likely to be influenced by factors associated with pregnancy.

Table 103 β coefficients of significant predictors of increase in high-energy diet score

Predictor	PAH	SWS
Townsend index*	-0.07	
Currently employed	-0.13	
Lower managerial and professionals^H		0.28
Intermediate^H		0.49
Semi-routine occupations^H		0.31
Routine occupations^H		0.89
BMI‡		-0.21
BMI‡ squared		0.20
Adjusted R^2	1.6%	11.1%

*Scaled to a standard deviation of 1 ^H Compared to Higher managerial and professionals
‡Logged and scaled to a standard deviation of 1

In the PAH study women living in areas of less deprivation and who were not currently employed had greater increases in high-energy diet scores. However, these effects only explain 1.6% of the variation in the change variable and are therefore relatively inconsequential.

In the SWS an effect of NS-SEC was seen on increase in the high-energy diet score; participants in 'lower managerial and professional', 'intermediate', 'semi-routine', and 'routine' occupations increased their high-energy diet scores notably more than those in 'higher managerial and professional' occupations. This effect is actually only statistically significant for those in 'routine' occupations, in this small sample of 83 (Table 97). Despite the small sample size a notable statistically significant quadratic effect of BMI was seen. The effect is hard to interpret since BMI is squared, so it is illustrated in Figure 48; women with average BMIs at initial interview had a greater reduction in high-energy diet scores than those in the highest and lowest quartiles of BMIs. However, these results require replication on a larger sample.

8.3 Strengths and limitations

8.3.1 The PAH and SWS datasets

The PAH study was conducted amongst 620 white women aged 16 or over with singleton pregnancies who attended an antenatal booking clinic at less than 17 weeks gestation and agreed to take part in the study. These criteria mean that the women will not be completely representative of the general population. The 6,129 SWS participants were all aged 20 to 34 and resident in the city of Southampton. An important selection criterion in this study was that the women had to be willing to give an hour and a half interview at a time when they did not have other relevant contact with health professionals. They were perhaps less willing to agree to this over the telephone than the PAH participants who were approached face-to-face, in the context of prenatal healthcare provision; this is reflected by the SWS response rate of approximately 75%, as compared to the PAH response rate of 95%.

Analyses of the representativeness of the PAH and SWS cohorts are summarised in Section 3.5. Both the PAH and SWS women were representative of the deprivation of enumeration districts of Southampton, as measured by the Townsend index, although Southampton is generally slightly more deprived than all of England and Wales. Also, women with no qualifications were under-represented in the PAH study and the SWS. However, these biases would not affect the internal comparisons between Townsend index or education and diet in this thesis, and it is notable that all but 0.9% of enumeration districts in England and Wales have scores represented in Southampton, as well as all educational levels being represented in the studies. These selection biases could affect the principal component scores if women living in less deprived areas or with no qualifications have diets with most variation along different axes compared to other women. However, the principal component analyses have generally been seen to be very robust to small changes. Both the PAH and SWS participants were representative of the UK in terms of smoking, and the SWS women had a very similar distribution of ethnicity to the GHS.

Despite the potential biases outlined above, the data analysed in this thesis consist of two large cohorts of pregnant and non-pregnant young women containing interviewer-collected information. The size and thorough data-collection methods are significant strengths of the datasets used. The studies also benefited from high response rates and inter-observer variability studies improved the accuracy of the anthropometric measurements obtained. A further strength lies in the longitudinal elements of the PAH and SWS studies which enable changes in diet to be examined.

8.3.2 Food frequency questionnaires

The use of FFQs to assess dietary intake has limitations. As described in Section 2.3.2, the reproducibility and validity of FFQs is acceptable, but not perfect. The prudent and high-energy dietary patterns may have accounted for more of the total variation in food intake if the food assessment method had been more precise. Analysis of diary data produced comparable patterns accounting for a similar amount of variation in food intake; nevertheless, the diaries may not be a more accurate representation of food intake than the FFQ since the validation of the Southampton MRC FFQ found energy intakes determined by the food diaries to be low compared to BMR[127].

Research indicates that FFQs tend to over-report fruit and vegetables (Section 2.3.2). However, an attempt to correct for this using cross-check questions in the SWS made little difference to the dietary patterns (Section 6.3.2), again demonstrating their robust nature. The over-reporting of fruit and vegetable consumption is therefore not a major concern when using FFQs to derive principal component scores.

A further issue involving self-reported methods of dietary assessment is that women may falsely report their intake. In particular, relationships such as that between education and the prudent diet score could be a result of more educated women reporting the foods they know they should eat, rather than actual consumption. However, other research has found a positive relationship in the SWS between women's education and their red cell folate concentrations[224], which is an objective measurement. Also, women with higher educational qualifications are seen in the literature to report greater alcohol consumption[91,225], a behaviour that is not in line with guidelines for a healthy diet. Moreover, increased fruit and vegetable consumption was mentioned more commonly as a characteristic of healthy eating by those with primary education than those with tertiary education in a study by Margetts et al.[145]. Therefore, although there may be some effect of more educated women reporting a diet they know is more healthy, this is unlikely to account for a large proportion of the association seen with the prudent diet score.

Despite the limitations of the FFQ data, this technique was chosen for its strengths. In particular, virtually complete data reflecting consumption over a three month time period was collected from a large sample of the population. The interviewer-administered technique improved participation amongst those from ethnic minorities or with few educational qualifications whose written English language skills were limited.

8.3.3 Social position

A key strength of this thesis has been the careful measurement of social position. Social Class, NS-SEC, the Cambridge Scale and education were all considered as alternative measures of social position, so that the different facets of the phenomenon each purports to measure could be investigated. The employment-based measurement of social position of women can be problematic (Section 3.3.1) and therefore the most suitable approach was carefully investigated (Section 3.3). The dominant approach was found to be the most appropriate method of characterising a women's Cambridge score or her SC, whereas the individual approach was chosen to denote a woman's working conditions as encapsulated by the NS-SEC.

Education was seen to be the most important method of assessing a woman's social position in terms of impact on the prudence of her diet. This variable has the added advantages of being easy to collect, relatively constant across the adult life-course, and available for all adults regardless of whether they are employed. The importance of education as the most significant social position predictor of the prudent diet score, particularly in the SWS, was a notable finding from the analyses in this thesis. Thus education is recommended as the most useful and easy to obtain measure of social position to collect when studying dietary prudence.

The Cambridge score provided useful additional information about a woman's generalised social advantage, and was an important predictor of a woman's prudent and high-energy diet score. The Cambridge score was more predictive than the SC, and thus it seems that Prandy's[173] assertion that the Cambridge scale is preferred to SC on empirical grounds is verified in the field of dietary pattern analysis. This may be due to the availability of Cambridge scores for women, separately from men. SC and the NS-SEC were the most predictive social position measures of change in prudent diet score and high-energy diet score respectively, although these analyses require replication on a larger sample.

8.3.4 Principal component analysis

An important aim of this thesis has been the characterisation of diet using principal component analysis. The advantages of multivariate techniques are summarised in Section 2.4.1, and include relevance to real-life food choice, and the focus on combinations of foods, which may have more influence on health than individual nutrients.

PCA and factor analysis were chosen in preference to cluster analysis because they provide greater discrimination of individuals in terms of a continuous score, rather than into a few relatively crude groups. Furthermore, greater consistency has been found across studies using PCA and factor analysis, than amongst those using cluster analysis (Sections 1.3.1 and 1.3.2). PCA and factor analysis were also preferred to diet quality indices due to their more objective approach. When analyses were performed using both PCA and factor analysis the additional complexity of factor analysis made little difference to the scores, and produced no improvement in the interpretability of the patterns (Section 4.3.1); thus PCA was chosen to characterise diet.

However, several disadvantages are associated with PCA. In particular it is often observed that, along with other multivariate techniques, PCA provides little information about specific nutrient-disease relationships. It is details of these relationships that inform investigations into mechanisms between diet and health, although such analyses were not an objective of this thesis. A further disadvantage of PCA is the subjective nature of the decisions involved, as emphasised by Martínez et al.[226]. However, as Slattery and Boucher[227] note, most statistical analyses in epidemiological research involve arbitrary decisions. An advantage of PCA over factor analysis is that fewer subjective decisions are necessary, but nevertheless choices have to be made about the selection and grouping of foods, and the labelling of components. In this thesis care has therefore been taken to ensure that the groupings used were robust

(Section 4.1.3). Full descriptions of the prudent and high-energy dietary patterns have also been given, as recommended by Slattery and Boucher[227], using coloured coefficients to highlight important features, thus enabling the reader to understand the entirety of each pattern rather than relying completely on the label given.

A further disadvantage of PCA is that the patterns resulting from the technique are based purely on behaviour, and are not necessarily optimal. Therefore the prudent dietary pattern may not score participants along a truly 'most prudent' axis, and the high-energy score may not accurately represent the highest-energy intake. However, the 'most prudent' axis would certainly be difficult to define because it depends on the health outcome considered and is limited by the extent of research knowledge to date. The most prudent diet also varies between individuals, such that a diet high in wholemeal bread would not be suitable for someone with gluten intolerance, for example.

McCann et al.[147] questioned the advantages of PCA over simply measuring fruit and vegetable intake. They found that fruit and vegetables provided slightly improved discrimination between endometrial cancer cases and controls over the more sophisticated technique of PCA-derived dietary patterns. Indeed fruit and vegetables may be particularly important in the aetiology of endometrial cancer. However, in the context of this thesis sociodemographic and lifestyle predictors of principal component scores were investigated since these scores give a description of overall diet, rather than of one aspect. In fact high fruit and vegetable intake may reflect high intake in general, as demonstrated by the positive correlations with the high-energy diet score, particularly in the SWS, and would therefore be inadequate as a measure of dietary prudence.

The advantage of using the high-energy score over a measure of calorie intake could also be questioned. The correlation between both the early and late pregnancy high-energy dietary patterns in the PAH study is indeed very high ($r_S = 0.87$). However, the advantage of the high-energy diet score is its ready availability directly from FFQ data without the requirement of nutrient calculations.

8.3.5 Predictors of dietary scores

This thesis is an exploratory study into predictors of prudent and high-energy diet scores. Models of major pathways of influence were used to determine the order in which variables were entered into the models rather than simply using statistical significance alone; the advantage of this method is that it reveals the mediating effect of one variable on another. However, the summary models may have been different if an alternative decision-making process had been followed, such as one based on the public-health utility of variables. Similarly, the availability of other variables providing further information about causal pathways would have altered the final models.

The term 'predictor' has been used throughout this thesis in a statistical sense. However, it is plausible that the direction of influence is from diet to the characteristic in some instances. For example, an effect of eating a more prudent diet amongst women in the SWS could be to incline them to take more strenuous exercise. Alternatively there may be a common external cause leading to an association between a sociodemographic characteristic and diet; in particular a positive attitude towards a healthy lifestyle is likely to motivate both adherence to dietary guidelines and other health-related behaviours, as discussed in Section 8.2.2.

As has been noted by other authors[98,100], differences in dietary patterns across sociodemographic and lifestyle groups are not large in comparison with overall variations in intake. For example, the most significant predictor of the prudent diet score in the PAH study in early pregnancy was age; the average prudent diet score of women in the lowest quartile of age was –0.60 standard deviations, whereas that in the highest quartile of age was 0.53 standard deviations. The most significant predictor of the prudent diet score in the SWS was education; the average prudent diet score for women with no qualifications was –0.75 standard deviations, whereas that for those with degrees was 0.65 standard deviations. As well as being a result of measurement error, these relatively small effects are likely to be due to other influences on diet, data about which are unavailable in the SWS and PAH studies. For example, women's food choice is also affected by access to appropriate shops[119,228], resources[228], knowledge[121], cooking confidence[119,229], taste[122,230], attitudes[215] and cost[231,232]. Such influences on healthy eating are discussed in Section 8.4.2.

8.4 Implications and future research

8.4.1 Implications

Work in this thesis has addressed the objectives stated in Section 1.5. Objective 1 was examined by considering the information available in the PAH study and the SWS suitable for characterising the women's sociodemographic and lifestyle situation (Chapter 3). In particular the complex issue of defining women's social position was carefully investigated as described in Section 8.3.3.

Analyses in this thesis have added to the evidence that the use of principal component analysis is a helpful method of assessing diet. Although nutrient-based analyses will still be important, particularly to assess mechanisms, PCA is a complimentary means of studying dietary intake. In particular the technique has been seen to be stable across different groupings considered, and in the two studies. This is especially true for the prudent dietary pattern. Thus objective 2 has been fulfilled.

It is important to note that PCA is simple to implement in current statistical packages. For example in Stata version 7.0[140] only 4 lines are required to perform an analysis to generate and label the first two principal components:

```
factor food1 food2 food3 food4 food5 food6 food7 food8 food9 /*
      */ food10 food11 food12 food13 food14 food15 food16 food17 /*
      */ food18 food19 food20 food21 food22 food23 food24 food25 /*
      */ food26 food27 food28 food29 food30 food31 food32 food33 /*
      */ food34 food35 food36 food37 food38 food39 food40 food41 /*
      */ food42 food43 food44 food45 food46 food47 food48 food49, /*
      */ pc factors(2)
score pc1 pc2
label variable pc1 "Principal component 1"
label variable pc2 "Principal component 2"
```

The speed of modern computers provides almost instantaneous results, even when many variables and subjects are included in the analysis.

The additional complexity of factor analysis provided little benefit beyond PCA, and thus further work would best utilise the simpler technique. In future research it may also be unnecessary to enquire about extra foods, or use fruit and vegetable cross-check questions; the additional data these questions provided little altered the dietary patterns found, although the implications for these questions in assessing nutrient intake is beyond the scope of this thesis. It may also be that questions about extra foods are of greater importance in populations with a larger proportion of participants from ethnic minorities since it is their diets that could be captured less adequately by the FFQ.

Research into the results of principal component analysis on diary data as described in objective 3 found that similar patterns were seen to those derived from the FFQ data. The correlation between women's dietary scores as calculated using FFQ and diary data was also good, particularly for the prudent diet score. This was despite the differences between the

FFQ and diary methods of data collection as described in Section 8.2.1.2. Thus this calibration study gives the results from the FFQ data additional potency.

Objective 4 was addressed by examining the changes in prudent and high-energy diet score in the PAH and SWS women. The reduction in prudent diet score amongst pregnant women is of interest, and suggests that dietary advice during pregnancy may be useful to improve diet. This might also reduce the negative association seen between the number of children a woman lives with and her prudent diet score. Anderson[233] notes that pregnancy might be a good time for dietary change because it generally occurs in relatively early adulthood when women may still be forming adult dietary patterns; the average age of a mother at childbirth in 2003 was 29.3 years[234]. Women may also be more willing to change in order to enhance their babies' health, and antenatal care provides the opportunity to reach large numbers of women with the potential to influence the health of the next generation. However, Anderson describes how many women still associate dietary change with slimming, but that this should be discouraged during pregnancy. Nausea could be a further barrier to adopting dietary messages. It may be that optimum nutritional status before pregnancy begins is more important than that during pregnancy, and therefore the most constructive interventions would be directed towards non-pregnant women of peak child-bearing age, who are therefore most likely to become pregnant. Anderson concludes that pregnancy may best be viewed as an opportunity for maintaining good dietary selections.

Objective 5 was addressed by fitting multiple regression models with prudent and high-energy diet scores as the outcomes. Although it is not always possible to define the direction of causality using cross-sectional data (Section 8.3.5), the associations between sociodemographic and lifestyle measures with the prudent diet score provide information about where health interventions should be targeted. Nutritional health messages should thus be aimed primarily at younger women with fewer qualifications and lower Cambridge scores. White women and those living with more children are also more likely to be at risk of a poor diet. Clustering of unhealthy behaviours occurred, such that women who smoked, took no strenuous exercise and watched several hours of television per day also tended to have less prudent diets.

It may be naïve to assume that increasing a woman's educational qualifications would automatically improve her prudent diet score. Instead, further research into influences on healthy eating (Section 8.4.2) that mediate this association is necessary. Similarly the clustering of unhealthy behaviours observed does not imply that changing one behaviour, such as smoking, will necessarily have a positive impact on diet. Research into the reasons for the clustering would be useful, and health-related interventions might be most effective if they address diet, smoking and exercise simultaneously. Other sociodemographic variables cannot be changed, such as age and ethnicity, but instead provide information on appropriate target groups for dissemination of healthy eating messages.

Current government dietary initiatives for tackling inequalities in health involve the 5 A DAY programme to improve fruit and vegetable intake[235]. Such schemes are valuable, particularly since the message is easy to understand, although aspects of diet other than fruit and vegetable consumption need to be addressed. Implementation of the 5 A DAY programme is especially encouraged amongst those in the lowest quintile of household income distribution. However, data from this thesis would suggest that, amongst young women, those with no qualifications would most benefit from this intervention.

The high-energy diet score does not have such a qualitative interpretation as the prudent diet score, particularly as it is independent of this score. Restriction of energy intake in pregnancy should be discouraged[233]. However amongst obese non-pregnant women, a reduced high-energy score could be beneficial.

8.4.2 Influences on healthy eating

In reaching women with dietary advice it is important to understand why nutritional guidance may not be followed. Limited access to appropriate shops is thought to make healthy eating more difficult; food in the UK is now primarily sold in large superstores, designed mainly for car access[228]. Bus access may be possible, but Dowler[228] notes that people with disabilities or those with young children find it particularly difficult to use buses, and are therefore more restricted to small local shops where food is more expensive, and the price differentials are particularly marked for foods recommended as part of a healthy diet. Worries about wastage when offering foods less palatable to children, and insufficient food storage space, were also mentioned by participants in studies of those on a low income[228]. Fruit and vegetables are particularly prone to go off, leading to more wastage[119].

Lack of knowledge may also be a barrier to healthy eating. Wardle et al.[121] found that nutrition knowledge was positively associated with fruit and vegetable intake, and negatively associated with fat intake; interestingly it was seen partially to mediate the effects of education and SC. There is also evidence that greater confidence about cooking skills is related to a more healthy diet[229]. Leather[119] notes that vegetables are generally associated with a traditional 'sit-down' main meal, and thus increases in solitary snacking are a barrier to vegetable consumption. This may be even more pertinent in pregnancy when women are encouraged to snack to avoid nausea. In a large Swedish cohort Lindström et al.[236] observed that high social participation as measured by involvement in study circles, unions, churches, sports events etc. was associated with higher fruit and vegetable consumption. The authors conclude that 'a high level of social participation seems to constitute a resource that makes it easier to choose a healthy lifestyle'.

The question of whether a limited income puts women at risk of a poorer diet has been widely disputed[237]. Cade et al.[232] found that the cost of the diets eaten by women in the very highest group of a healthy diet indicator were on average £1.48 higher than those in the very lowest group. However, this analysis did not assess whether a similarly healthy diet could be purchased more economically. Using a novel modelling technique Darmon et al.[231] inferred that the nutritional quality of the diet of a rational French individual would decrease as their food budget reduced. Financial constraints are therefore likely to be a determinant of food choice, particularly amongst those with low incomes, and the authors conclude that 'nutrition education alone may prove ineffective unless it is combined with economic measures aimed at improving the affordability of a healthy diet'. In a qualitative study using focus groups in a deprived area of Leeds mothers with younger children were seen to be primarily preoccupied with the cost of their food[238]. However, it is clear that cost is not the only motivator of behaviour since smoking evidently costs more than not smoking, but financial difficulties are strongly positively associated with smoking[116].

Lennernäs et al.[120] found that price was an important consideration in the food choices of Europeans, but not as important as quality and freshness. In a further analysis of this survey, Kearney et al.[122] observed that lack of time, and compromising taste were perceived as the greatest barriers to healthy eating in this cohort, rather than cost. Similarly the cost of fruit and vegetables was perceived as a relatively unimportant barrier to fruit and vegetable consumption in the Health Education Authority's Health and Lifestyle Survey[215]. Interestingly, Kearney et al.[122] also observed that 71% of European respondents (62% in the UK) did not believe they needed to make changes to their diets since it was already healthy enough. Thus an inability to assess personal intake, or an indifferent attitude towards diet, are likely to be significant barriers to healthy eating. Indeed Thompson et al.[215] found that respondents' attitudes towards food were more predictive of fruit and vegetable intake than their nutritional knowledge. Steptoe and Wardle[239] observed that motives and attitudes partially mediated the association between diet and education.

The literature cited above indicates that improving knowledge and changing attitudes and motivations may improve diets. Steptoe et al.[240] found that behavioural counselling (giving personalised specific advice along with the setting of short and long term goals) to a group of patients from a primary health centre in a deprived inner city area was particularly successful in increasing fruit and vegetable consumption. Similarly, the use of the brief negotiation method has been seen to increase self-reported fruit and vegetable consumption and reduce blood pressure[241].

However, Leather[119] notes that not all barriers to increasing fruit and vegetable consumption are individual; some are economic and political. The UK currently has inadequate production, marketing, advertising and retailing structures to support a national diet in accordance with the Balance of Good Health (Figure 2).

8.4.3 Future research

To further the work in this thesis it would be informative to perform PCA on other datasets, to see whether the prudent and high-energy dietary patterns persist. For example, research is planned to investigate patterns resulting from FFQs administered in early and late pregnancy in the SWS. Work amongst pregnant and non-pregnant young women will broaden knowledge relevant to the health of the next generation, and studies on wider cohorts will add to the literature on dietary patterns resulting from PCA. Larger studies of change in diet amongst non-pregnant women are particularly necessary since it was only possible to assess this amongst a small group of SWS participants in this thesis. Longer follow-ups would enable more accurate assessment of biological, psychological and social processes occurring over the life course, and to distinguish these from historical processes.

Further research into the robustness of dietary patterns resulting from PCA is also planned. In particular the prospect of shortening the FFQ for use in further studies where the main dietary outcome is defined using PCA is being considered; if the length of the FFQ could be significantly reduced by collecting information about a few, important variables, and the resulting prudent diet score was nevertheless highly correlated with that from the full FFQ, the resources required to collect the dietary information could be greatly reduced. Schulze et al.[242] were able to use this technique to create a correlation greater than 0.95 between the original dietary score and the simplified pattern variable.

The PAH and SWS datasets were limited in the data available to assess the mechanisms by which associations found may operate. In particular, the literature described in Section 8.4.2 indicates that psychosocial variables including knowledge, attitudes, motivations, and financial and other resources may be important influences on dietary choices and may mediate some of the associations found in this thesis. Thus future work is planned as part of the SWS to re-interview a sample of women and ask about psychosocial factors alongside the types of information in the main study; the results of this research will aid understanding of the causal pathways involved in dietary choice. The prudent diet score will be used in other analyses of the SWS; studies of pre-menstrual syndrome, depression, fetal growth and neonatal bone density will utilise the score as a dietary predictor of these important outcomes.

Strong gradients between measures of social position and the prudent diet have been seen in this thesis. This provides more evidence that diet is a candidate mediator of inequalities in health (Section 1.1.2). That this evidence comes from young women may implicate their diets in affecting the health of their children (Section 1.1.3). Further studies in datasets containing both dietary information and birth outcomes would be useful to examine the influence of the prudent and high-energy diet scores on neonatal measurements. The SWS will eventually provide important data in this area, and information about pre-pregnant as well as pregnant diets will be key in assessing the importance of the timing of nutritional effects. Ultimately it is thought that early growth and development influences adult disease (Section 1.1.4) and therefore long-term studies linking maternal dietary intake with later disease in the offspring would be most informative in terms of improving the health of the population.

PRINCESS ANNE HOSPITAL NUTRITION STUDY
EARLY PREGNANCY QUESTIONNAIRE

Name: ____________________ Hosp. no.______________

Address: ____

____

Interviewer: ☐ MRC serial no. ☐☐☐☐

Date of interview: d d ☐☐ m m ☐☐ y y ☐☐

1) In which county/country were you born ? ☐☐

____

2) In which county/country was the baby's father born ? ☐☐

____

3) In which county/country was your mother born ? ☐☐

____

4) How many babies did your mother have (including any still born children) ? ☐☐

5) How many were born before you ? ☐☐

6) In what year was your mother born ? 1 9 ☐☐

7) What was your father's full-time job when you were born ? (or if unemployed, last job)
(probe industry & self-employed/manager/foreman) ☐☐☐☐

____

8) What is the baby's father's current full time job ?
(or if unemployed, last full time job)
(probe industry & self-employed/manager/foreman) ☐☐☐☐

____

9) Are you currently employed in paid work ? ☐
0. No GO TO Q9a & Q11
1. Yes GO TO Q9b & Q10

a) What was your last full-time job ?
b) What is your current fulltime job ?
(probe industry & self-employed/manager/foreman) ☐☐☐☐

____

10) a) Over the last week how many hours did you spend in paid work ? ☐☐.☐

b) On a typical working day how many hours did you spend in paid work ? ☐☐.☐

and roughly how many of these hours in a typical day were spent
i) sitting/driving ☐☐.☐

ii) standing/walking ☐☐.☐

other: specify ____________________________

11) Since becoming pregnant, over a 24 hour period how many hours have you generally spent sleeping or lying down with your feet up ?
(include any in paid work !) ☐☐.☐

12) Including any paid work, how many hours a day have you generally spent driving or sitting down since becoming pregnant ?
e.g. sewing, reading or watching television ☐☐.☐

13) Over the last week have you engaged in any sporting activity, aerobics or dancing long enough to work up a sweat or get out of breath ?
0. No
1. Yes How many times ? ☐☐

Activity(s) ______________ ☐☐

______________ ☐☐

______________ ☐☐

Appendix 1

14) As regards your general level of activity since you became pregnant, has the last week been fairly typical ?
0. No
1. Yes

I would like now to ask you a few questions about your drinking and smoking habits.

15a) Have you changed the amount of alcohol you drink during this pregnancy ?
0. Decreased
1. No change
2. Increased

b) **At present** how often do you drink alcohol ?
0. never GO TO Q16
1. one day a month or less
2. 2-3 days a month
3. 1-2 days a week
4. 3-4 days a week
5. 5-7 days a week

c) When you drink, how much would you normally have in one day ?

1. beer / stout / cider / lager	pints ______	
&/or		
2. low alcohol wine	glasses______	
&/or		
3. wine / sherry / martini / cinzano	glasses______	.
&/or		
4. spirits / liqueurs	measures______	

16a) Have you ever smoked regularly ? (i.e. at least once a day for a year or more)
1. Yes ASK Q16b
0. No GO TO Q17

b) How old were you when you first smoked regularly ?

c) Which is/was your usual brand ? ________________

d) Did you smoke at the time of your last menstrual period ? 1. Yes / 0. No
If Yes, how many per day ______

e) Are you currently smoking ?
1. Yes / 0. No
If Yes, how many per day ______

17) Does the baby's father smoke regularly ?
1. Yes / 0. No

I would like to ask you briefly about your school and higher education.

18) How old were you when you left school ? yrs

19) Do you have any formal qualifications/exams ?
1. None
2. CSEs / School certificate
3. O levels / Matric
4. A levels
5. HND
6. Degree
7. Other ____________________

I would like to ask you now about any medicines you may have taken.

20) What, if any, medicines/inhalers/pills/tablets/indigestion remedies have you taken since your last menstrual period ?
PLEASE USE BLOCK CAPITALS & COPY NAMES DIRECTLY OFF BOTTLES IF POSSIBLE

1 ____________________
2 ____________________
3 ____________________
4 ____________________
5 ____________________
6 ____________________
7 ____________________
8 ____________________

21) Did you ever use the Pill (oral contraceptive) ?
1. Yes / 0. No

If yes, at what age did you first start ? . yrs

at what age did you last stop? . yrs

have you ever had to stop the Pill because of high blood pressure ?
1. Yes / 0. No

was your last period a Pill withdrawal bleed ? 1. Yes / 0. No

22) How old were you when you had your first period ?
. yrs

23) Is this your first pregnancy ? 1. Yes GO TO Q25
0. No GO TO Q24

I would now like to ask you a few questions about any <u>ILLNESSES</u> you may have suffered from.

24) During your previous pregnancies were you ever treated by a doctor for
a) high blood pressure 1. Yes / 0. No
b) diabetes 1. Yes / 0. No
c) anaemia 1. Yes / 0. No

25) When not pregnant, have you ever been treated by a doctor for
a) high blood pressure 1. Yes / 0. No
b) diabetes 1. Yes / 0. No
c) anaemia 1. Yes / 0. No

26) Have you been a blood donor at any time in the last year ? 1. Yes / 0. No
If yes, on how many occasions in the last year ______

27) Either as a child or an adult, have you ever suffered from asthma ?
1. Yes / 0. No

28) Have you had wheezing or whistling in your chest at any time in the last year ? 1. Yes / 0. No

29) Have you ever been told by a doctor that have hay fever ? 1. Yes / 0. No

30) When you DON'T have a cold or the 'flu', do you experience
a) Frequent sneezing 1. Yes / 0. No
b) Runny nose 1. Yes / 0. No

31) Did you suffer from eczema in childhood ?
1. Yes / 0. No

32) Have you had eczema affecting your elbows or knees in the last year ? 1. Yes / 0. No

33) To your knowledge has the baby's father ever had
a) Asthma 1. Yes / 0. No / 9. Don't know
b) Hay fever 1. Yes / 0. No / 9. Don't know
c) Childhood eczema 1. Yes / 0. No / 9. Don't know

<u>ADDITIONAL DIETARY QUESTIONS</u>

34) How much did you weigh 4-6 months ago ? (i.e. before you became pregnant)
__________ st __________ lb
__________ kg
1. Imperial 2. Metric 9. Don't know

35) Have you experienced any nausea or sickness since becoming pregnant ?
1. Yes / 0. No (if no, go to Q37)

36) If yes, as this been
1. Mild (nausea only)
2. Moderate (sometimes sick)
3. Severe (regularly sick, unable to retain meals)

37) Since you became pregnant, are you eating:
1. more / 2. the same / 3. less <u>in amount</u>
1. more / 2. the same / 3. less <u>milk, eggs and dairy products</u>
1. more / 2. the same / 3. less <u>green vegetables</u>
1. more / 2. the same / 3. less <u>fresh fruit</u>
1. more / 2. the same / 3. less <u>red meat</u>
1. more / 2. the same / 3. less <u>liver</u>
1. more / 2. the same / 3. less <u>cereal, cakes & biscuits</u>
1. more / 2. the same / 3. less <u>tea or coffee</u>

38) In the past year, have you tried to lose weight by dieting ?
0. No 1. Yes

39) Since becoming pregnant, have you been taking any pills, tonics or tablets to supplement your diet ? (e.g. vitamins, iron tablets, cod liver oil capsules etc.)

0. No 1. Yes

(if no, go to frequency questionnaire)

40) Please state which:-

Supplement name	Dose	How long have you been taking the supplement ?	Is it prescribed ? Yes / No	Code	Amount

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN: Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
1	White bread/french stick/rolls/hamburger buns	1	2	3	4	5	6	7	☐
2	Wholemeal bread/rolls /granary/rye/brown bread/pumpernickel	1	2	3	4	5	6	7	☐
3	Crackers eg cream crackers	1	2	3	4	5	6	7	☐
4	Wholemeal and rye crackers eg ryvita	1	2	3	4	5	6	7	☐
5	'Bran' breakfast cereals	1	2	3	4	5	6	7	☐
6	Other breakfast cereals eg cornflakes,museli	1	2	3	4	5	6	7	☐
7	Added bran to foods	1	2	3	4	5	6	7	☐
8	Cakes eg battenburg,sponge fruit,chocolate,scones	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN: Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
9	Buns eg crumpets,currant doughnut,muffins	1	2	3	4	5	6	7	☐
10	Pastries eg danish,eccles,pastry, chocolate eclairs	1	2	3	4	5	6	7	☐
11	Biscuits eg digestives,wrapped biscuits,ginger	1	2	3	4	5	6	7	☐
12	Other biscuits	1	2	3	4	5	6	7	☐
13	Fruit based puddings eg apple crumble,lemon meringue pie	1	2	3	4	5	6	7	☐
14	Milk based puddings eg rice pud,instant whip white sauces	1	2	3	4	5	6	7	☐
15	Other puddings eg treacle tart,sponge pud,cheesecake	1	2	3	4	5	6	7	☐
16	Yogurts/fruit fools /excluding fromage frais	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN							
		Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
17	Potatoes- boiled/mashed/baked/ tinned new potatoes	1	2	3	4	5	6	7	☐
18	Potatoes- roast/chips/waffles	1	2	3	4	5	6	7	☐
19	Yorkshire puddings /savoury pancakes	1	2	3	4	5	6	7	☐
20	Rice- brown/white,risotto	1	2	3	4	5	6	7	☐
21	Pasta/dumplings /macaroni/tinned spaghetti/noodles	1	2	3	4	5	6	7	☐
22	Processed & tinned peas,tinned carrots, tinned mixed veg	1	2	3	4	5	6	7	☐
23	Fresh and frozen peas/french beans/broad beans	1	2	3	4	5	6	7	☐
24	Fresh and frozen carrots	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN							
		Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
25	Fresh and frozen parsnips/swedes /turnips	1	2	3	4	5	6	7	☐
26	Fresh and frozen sweetcorn/mixed veg /beetroot	1	2	3	4	5	6	7	☐
27	Baked beans/red kidney beans/chick peas/other pulses/lentils	1	2	3	4	5	6	7	☐
28	Tomatoes/tomato juice/ tinned tomatoes	1	2	3	4	5	6	7	☐
29	Spinach	1	2	3	4	5	6	7	☐
30	Brussel sprouts /broccoli /spring greens	1	2	3	4	5	6	7	☐
31	Cabbage/cauliflower	1	2	3	4	5	6	7	☐
32	Green pepper /watercress	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN							
		Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
33	Onion-raw/boiled /spring onion	1	2	3	4	5	6	7	☐
34	'Salad' ie. lettuce, cucumber,celery,radish	1	2	3	4	5	6	7	☐
35	Coleslaw/salad in dressing eg Waldorf,celery & nut	1	2	3	4	5	6	7	☐
36	Courgettes/marrow /leeks	1	2	3	4	5	6	7	☐
37	Mushrooms/fried	1	2	3	4	5	6	7	☐
38	Vegetable dishes-ratatouille/curry /pasties	1	2	3	4	5	6	7	☐
39	Vegetarian foods-tofu/TVP/soya substitutes	1	2	3	4	5	6	7	☐
40	Tinned fruit salad etc	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN							
		Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
41	Cooked fruit (stewed) excluding blackcurrants	1	2	3	4	5	6	7	☐
42	Dried fruit	1	2	3	4	5	6	7	☐
43	Apple/pears	1	2	3	4	5	6	7	☐
44	Orange and orange juice	1	2	3	4	5	6	7	☐
45	Grapefruit and grapefruit juice	1	2	3	4	5	6	7	☐
46	Blackcurrants/ribena	1	2	3	4	5	6	7	☐
47	Other fruit juices (not squashes)	1	2	3	4	5	6	7	☐
48	Diet Coke/Pepsi excluding caffeine free	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN							
		Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
49	Coke/Pepsi	1	2	3	4	5	6	7	☐
50	Fizzy drinks/squashes not diet drinks ice lollies/sorbets	1	2	3	4	5	6	7	☐
51	Banana	1	2	3	4	5	6	7	☐
52	Peaches/nectarine /plums/apricots /cherries/grapes	1	2	3	4	5	6	7	☐
53	Strawberry/raspberry	1	2	3	4	5	6	7	☐
54	Melon/pineapple /fresh mango/kiwi	1	2	3	4	5	6	7	☐
55	Nuts-peanuts/cashews/peanut butter/nut dishes	1	2	3	4	5	6	7	☐
56	Bacon/gammon	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN							
		Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
57	Pork-roast/chops/casseroles /curries/kebabs	1	2	3	4	5	6	7	☐
58	Chicken/turkey-roast/burgers/nuggets/ casseroles/curries	1	2	3	4	5	6	7	☐
59	Lamb-roast/chops/cutlets /casseroles/hot pot	1	2	3	4	5	6	7	☐
60	Beef steaks/roast beef/corned beef /veal/stews	1	2	3	4	5	6	7	☐
61	Minced meat dishes-bolognaise/lasagne /meat balls	1	2	3	4	5	6	7	☐
62	Meat pies/steak and kidney pie/chicken pie	1	2	3	4	5	6	7	☐
63	Liver/kidney	1	2	3	4	5	6	7	☐
64	Pate/liver sausage	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN							
		Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
65	Black pud/faggots	1	2	3	4	5	6	7	☐
66	Sausage/sausage roll/salami/hot dogs	1	2	3	4	5	6	7	☐
67	Ham/canned meats/luncheon meat/Spam	1	2	3	4	5	6	7	☐
68	White fish- grilled poached/steamed /in crumbs or batter	1	2	3	4	5	6	7	☐
69	Fish pie/kedgeree/fish fingers/Duets/fish in sauces	1	2	3	4	5	6	7	☐
70	Oily fish- tuna,sardines,trout, salmon,mackerel	1	2	3	4	5	6	7	☐
71	Shell fish- crab,prawns,mussels	1	2	3	4	5	6	7	☐
72	Eggs/boiled/poached	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN							
		Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
73	Omelette/scambled /fried/Scotch egg	1	2	3	4	5	6	7	☐
74	Cottage Cheese	1	2	3	4	5	6	7	☐
75	Cheese-all types eg hard cheddar types, soft cheese processed, fromage frais, cheese sauce	1	2	3	4	5	6	7	☐
76	Quiche/pizza	1	2	3	4	5	6	7	☐
77	Soup	1	2	3	4	5	6	7	☐
78	Mayonnaise/salad cream	1	2	3	4	5	6	7	☐
79	Pickles/chutney /ketchup	1	2	3	4	5	6	7	☐
80	Chocolate/chocolate bars- mars,bounty,milky way etc	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN							
		Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
81	Other confectionery- liquorice allsorts, boiled sweets etc	1	2	3	4	5	6	7	☐
82	Ice cream/cornetto/ choc ice	1	2	3	4	5	6	7	☐
83	Creams- all types, Elmlea, UHT	1	2	3	4	5	6	7	☐
84	Crisps/wotsits etc	1	2	3	4	5	6	7	☐
85	Sweet spreads- jams, marmalade, honey, lemon curd	1	2	3	4	5	6	7	☐
86	Meat Stock cubes Oxo/Knorr/Bovril	1	2	3	4	5	6	7	☐
87	Drinking chocolate/ovaltine /cocoa	1	2	3	4	5	6	7	☐
88	Decaffeinated coffee/ tea	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN							
		Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
89	Tea	1	2	3	4	5	6	7	☐
90	Coffee	1	2	3	4	5	6	7	☐
91	Red wine	1	2	3	4	5	6	7	☐
92	White wine/cider	1	2	3	4	5	6	7	☐
93	Spreading fat (1) ________ ☐☐	1	2	3	4	5	6	7	☐
94	Spreading fat (2) ________ ☐☐	1	2	3	4	5	6	7	☐
95	Spreading fat (3) ________ ☐☐	1	2	3	4	5	6	7	☐
96	Frying fat/oil (1) ________ ☐☐	1	2	3	4	5	6	7	☐

FOOD CODE	FOOD DESCRIPTION	FREQUENCY EATEN							
		Never	Once every 2-3 Mths	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
97	Frying fat/oil (2) ________	1	2	3	4	5	6	7	
98	Frying fat/oil (3) ________	1	2	3	4	5	6	7	
99	Other vegetable oil(1) eg salad dressing ________ marinades	1	2	3	4	5	6	7	
100	Other vegetable oil(2)______	1	2	3	4	5	6	7	

1. Which types of milk have you used regularly in tea, coffee and added to breakfast cereals over the last 3 months ?

   Milk 1____________________ ☐☐☐☐☐

   Milk 2____________________ ☐☐☐☐☐

   Milk 3____________________ ☐☐☐☐☐

2. On average over the last 3 months how much of each milk have you consumed per day ?

   Milk 1__________ pts ☐.☐☐

   Milk 2__________ pts ☐.☐☐

   Milk 3__________ pts ☐.☐☐

3. When you eat bread/toast/sandwiches, how many slices/rolls do you eat at a typical meal ?

   White bread/rolls (counts as 2 slices)/ white toast (FFQ category 1) __________ ☐☐

   Brown bread/rolls (counts as 2 slices)/ brown toast/granary/wholemeal bread/ rye breads (FFQ category 2) __________ ☐☐

4. Do you add sugar to breakfast cereals, tea & coffee, puddings etc ? 0. No 1. Yes ☐
   (If NO go to Q6)

5. Approximately how many teaspoons of sugar do you add each day ? ______ ☐☐

6. When you eat meat, how much of the fat do you usually cut off (including chicken skin) ? ☐

   1. all 100%
   2. most 60%
   3. some 30%
   4. none 0%

7. When you eat potatoes, how many do you eat at a typical meal ?

   Old/baking (counts as 3)/ new (counts as 0.5) (FFQ category 17) ☐☐

   Roast (FFQ category 18) ☐☐

Head circumference ________ . ___ cm

________ . ___ cm

________ . ___ cm

PRINCESS ANNE HOSPITAL NUTRITION STUDY
LATE PREGNANCY QUESTIONNAIRE

Name: ____________________ Hosp. no. ______________

Address: ____

____

Interviewer: [ ] MRC serial no. [ ][ ][ ][ ]

Date of interview: d d [ ][ ] m m [ ][ ] y y [ ][ ]

I'd just like to ask you briefly again about your level of activity to see if this has changed since we saw you earlier in the pregnancy.

1) Are you currently employed in paid work ? [ ]
0. No GO TO Q1a & Q3
1. Yes GO TO Q1b & Q2

a) If **no**, what date did you give up work ? d d [ ][ ] m m [ ][ ] y y [ ][ ]

b) If **yes**, what is your current job ? [ ][ ][ ][ ]
(probe industry & self-employed/manager/foreman)

____

2) a) Over the last week how many hours did you spend in paid work ? [ ][ ][ ].[ ]

b) On a typical working day how many hours did you spend in paid work ? [ ][ ].[ ]

and roughly how many of these hours in a typical day were spent
i) sitting/driving [ ][ ].[ ]

ii) standing/walking [ ][ ].[ ]

other: specify ___________________________

3) Over the 3 months since we last saw you, over a 24 hour period how many hours have you generally spent sleeping or lying with your feet up ? (include any in paid work !) [ ][ ].[ ]

4) Including any paid work, how many hours a day have you generally spent driving or sitting down over the last 3 months ? [ ][ ].[ ]
e.g. sewing, reading or watching television

5) Over the last week have you engaged in any sporting activity, aerobics or dancing long enough to work up a sweat or get out of breath ?
0. No
1. Yes How many times ? [ ][ ]

Activity(s) ______________ [ ][ ]

______________ [ ][ ]

______________ [ ][ ]

6) As regards your general level of activity over the last 3 months, has the last week been fairly typical ? [ ]
0. No
1. Yes

I would like now to ask you a few questions about your current drinking and smoking habits.

7a) **At present** how often do you drink alcohol ? [ ]
0. never GO TO Q8
1. one day a month or less
2. 2-3 days a month
3. 1-2 days a week
4. 3-4 days a week
5. 5-7 days a week

b) When you drink, how much would you normally have in one day ? [ ][ ].[ ]
1. beer / stout / cider / lager pints ______
&/or
2. low alcohol wine glasses______
&/or
3. wine / sherry / martini / cinzano glasses______
&/or
4. spirits / liqueurs measures_____

8a) Are you currently smoking ? [ ] [ ][ ]
1. Yes / 0. No
If Yes, how many per day ______

b) Which is your current brand ? ______________ [ ][ ][ ]

9) Is the baby's father currently smoking ? [ ]
1. Yes / 0. No

10) What was your own birth weight ? [ ][ ].[ ][ ] lbs.oz

[ ][ ][ ][ ] grams

1. Imperial / 2. Metric / 9. Don't know birth weight - Go to Q12 [ ]

11) Were you born at home or in hospital ? [ ]
1. Home
2. Hospital - specify ____________________

12) What is your shoe size ? [ ][ ].[ ][ ] imperial

[ ][ ] metric

1. Imperial / 2. Metric [ ]

13) What was the baby's father's birth weight ? [ ][ ].[ ][ ] lbs.oz

[ ][ ][ ][ ] grams

1. Imperial / 2. Metric / 9. Don't know birth weight [ ]

14) Do you know how tall the baby's father is ? [ ].[ ][ ] ft.ins

[ ][ ][ ] cm

1. Imperial / 2. Metric / 9. Don't know father's height [ ]

15) Do you know the baby's father's shoe size ? [ ][ ].[ ][ ] imperial

[ ][ ] metric

1. Imperial / 2. Metric / 9. Don't know [ ]

16) Part of this research may shed some light on the effects of being a single mother. Can I ask you if you are sharing a home with your partner or living as a single person ?
1. Sharing / 2. Single [ ]

17) Have you experienced any nausea or sickness **over the last 3 months since we last saw you** ? [ ]
0. No 1. Yes (if no, go to Q 19)

18) If yes, has this been [ ]
1. Mild (nausea only)
2. Moderate (sometimes sick)
3. Severe (regularly sick, unable to retain meals)

I would like to ask you now about any <u>medicines</u> you may have taken.

19) What, if any, medicines/inhalers/pills/tablets/indigestion remedies have you taken **over the 3 months since we last saw you** ? PLEASE USE BLOCK CAPITALS & COPY NAMES DIRECTLY OFF BOTTLES IF POSSIBLE

1 ____________________ [ ]
2 ____________________ [ ]
3 ____________________ [ ]
4 ____________________ [ ]
5 ____________________ [ ]
6 ____________________ [ ]
7 ____________________ [ ]
8 ____________________ [ ]

20) Over the last 3 months, have you been taking any pills, tonics or tablets to supplement your diet ? (e.g. vitamins, iron tablets, cod liver oil capsules etc.) [ ]

0. No 1. Yes

(if no, go to frequency questionnaire)

21) Please state which:-

Supplement name	Dose	How long have you been taking the supplement ?	Is it prescribed ? Yes / No	Code	Amount
				[ ][ ][ ]	[ ][ ][ ]
				[ ][ ][ ]	[ ][ ][ ]
				[ ][ ][ ]	[ ][ ][ ]
				[ ][ ][ ]	[ ][ ][ ]
				[ ][ ][ ]	[ ][ ][ ]

FOOD FREQUENCY QUESTIONNAIRE IDENTICAL TO THAT USED IN EARLY PREGNANCY

1. Which types of milk have you used regularly in tea, coffee and added to breakfast cereals over the last 3 months ?

   Milk 1____________________ ☐☐☐☐☐

   Milk 2____________________ ☐☐☐☐☐

   Milk 3____________________ ☐☐☐☐☐

2. On average over the last 3 months how much of each milk have you consumed per day ?

   Milk 1__________ pts ☐.☐☐

   Milk 2__________ pts ☐.☐☐

   Milk 3__________ pts ☐.☐☐

3. When you eat bread/toast/sandwiches, how many slices/rolls do you eat at a typical meal ?

   White bread/rolls (counts as 2 slices)/ white toast (FFQ category 1) ________ ☐☐

   Brown bread/rolls (counts as 2 slices)/ brown toast/granary/wholemeal bread/ rye breads (FFQ category 2) ________ ☐☐

4. Do you add sugar to breakfast cereals, tea & coffee, puddings etc ? 0. No 1. Yes (If NO go to Q6) ☐

5. Approximately how many teaspoons of sugar do you add each day ? _____ ☐☐

6. When you eat meat, how much of the fat do you usually cut off (including chicken skin) ? ☐

   1. all 100%
   2. most 60%
   3. some 30%
   4. none 0%

7. When you eat potatoes, how many do you eat at a typical meal ?

   Old/baking (counts as 3)/ new (counts as 0.5) (FFQ category 17) ☐☐

   Roast (FFQ category 18) ☐☐

8. What is your usual brand of tea ? ☐☐☐

   ____

9. What is your usual brand of coffee ? ☐☐☐

   ____

URINE SAMPLE Date ☐☐ (d d) ☐☐ (m m) ☐☐ (y y) Time ☐☐☐☐

0. Fasting / 1. Non-fasting ☐

Mid upper arm circumference ☐☐.☐ cm

☐☐.☐ cm

☐☐.☐ cm

PRE-PREGNANCY QUESTIONNAIRE

SOUTHAMPTON
omen's
URVEY

Name: ____________________

Address: ____________________

____

Postcode: [ ][ ][ ][ ] [ ][ ][ ]

Phone No: [ ][ ][ ][ ][ ] [ ][ ][ ][ ][ ][ ]

Interviewer: [ ][ ] Date of interview: d d [ ][ ] m m [ ][ ] y y [ ][ ]

If the woman wants to have a cup of tea/coffee with you and has not eaten or drunk anything in the past hour, do the mouthwash sample first but remember to obtain the woman's ***consent****. If not, go to section 1.*

Mouthwash sample provided (0 = No, 1 = Yes) [ ]

Time of mouthwash sample
(24 hr clock) [ ][ ][ ][ ]

223

1: OCCUPATION

I would like to start by talking about any paid work that you do.

1.1 Were you in paid employment or self-employed in the week ending last Sunday? [ ]
- 0. No, *go to 1.3*
- 1. Yes, *go to 1.2*

1.2 Were you working full time or part time? [ ]
- 0. Full time (more than 30 hours) *go to 1.6b*
- 1. Part time (30 hours or fewer) *go to 1.3*

1.3 Are you going to college full time? [ ]
- 0.No *if working part-time go to 1.6a*
  if not working go to 1.5
- 1.Yes

1.4 *If yes*, what are you studying? ____________________
If working part time go to 1.7
If not working go to section 2

1.5 *If not working or studying* were you [ ]
- Unemployed ? (1)
- Permanently unable to work because of long term sickness or disability ? (2)
- looking after home or family? (3)
- other ? (specify)____________________ (4)

1.6a *If* ***not*** *working or working part-time*, what was your last **full-time** job ?
If only ever part-time ask for last part time job.
Then if currrently working part time go to 1.7, otherwise go to section 2.

Job Position ____________________
Self-employed/manager/foreman/employee
Industry ____________________

1.6b *If working full-time*, what is your job ? *(Then go to section 2)*
Probe industry & self-employed/manager/foreman/employee

Job Position
Self-employed/manager/foreman/employee
Industry____________________

1.7 *If working part-time now*, what is your current job?

Job Position ____________________
Self-employed/manager/foreman/employee
Industry ____________________

1.8 *If working part time,* how many hours per week do you work?

[ ][ ] hrs [ ][ ] mins

2: ACTIVITY AND EXERCISE

Now I'm going to ask you about your activity and exercise patterns over the last three months. We would like you to divide up a "typical" day into three types of activity. These are:

(1) sleeping or lying, (2) sitting, (3) standing or walking.

2.1 Over a typical 24 hour day how many hours do you generally spend sleeping or lying with your feet up? [ ][ ] hrs [ ][ ] mins

(*ask time usually go to bed & wake up, including any at work!*)

This would indicate xx hours sitting or on your feet.

Appendix 3

2.2 Of those hours how many on a typical day do you spend sitting down? (*e.g. includes sitting at work, mealtimes, driving, reading, watching TV*)

☐☐ hrs ☐☐ mins

2.3 This would mean that you spend about xx hours a day on your feet. Does this sound about right?

☐☐ hrs ☐☐ mins

2.4 Out of these xx hours spent on your feet, about how much of the time are you **actively on the move** (rather than standing fairly still)?

*

1.	Very little	10%
2.	Some	30%
3.	About half	50%
4.	Most	70%
5.	Almost all	90%

☐

2.5 **During the past three months,** how often have you done the following kinds of exercise or activities?
*

a) **strenuous exercise which** normally makes your heart beat rapidly **AND** leaves you breathless e.g. jogging, vigorous swimming or cycling, aerobics.

FFQ categories 1-7 ☐ >x1 ☐

and **on average** about how long does each period of activity last? ☐ hrs ☐☐ mins

b) **moderate exercise** which normally leaves you exhausted but not breathless, e.g. brisk walking, dancing, easy swimming or cycling, badminton, sailing.

FFQ categories 1-7 ☐ >x1 ☐☐

and **on average** about how long does each period of activity last? ☐ hrs ☐☐ mins

c) **gentle exercise** which normally leaves you tired but not exhausted, e.g. walking, heavy housework (including washing windows and polishing), gardening, DIY, golf.

FFQ categories 1-7 ☐ >x1 ☐☐

and **on average** about how long does each period of activity last? ☐ hrs ☐☐ mins

2.6 On a typical day, how many hours do you generally spend watching television?
*

1. More than 5 hours
2. 4-5 hours
3. 3-4 hours
4. 2-3 hours
5. 1-2 hours
6. Less than one hour
7. None

☐

2.7 Which of the following best describes your walking speed?
*

1. Very slow
2. Stroll at an easy pace
3. Normal speed
4. Fairly brisk
5. Fast

☐

3: DIETARY QUESTIONS

3.1 Now I am going to ask you about the foods you eat. To do this I have a list of foods and I would like you to tell me how often you have eaten each food during the past 3 months. The list may include foods you never eat or you may find foods which you eat a lot are missing. These can be added on at the end. *(Define the 3 month period)*

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
1	White bread	1	2	3	4	5	6	7	☐
	When you eat bread/toast/sandwiches, how many slices/rolls do you eat at a typical meal? *Rolls (count as 2 slices)* ☐☐.☐ *French bread (2″ counts as 1 slice)*								
2	Brown and wholemeal bread/rolls	1	2	3	4	5	6	7	☐
	How many slices/rolls do you eat at a typical meal? *Rolls (count as 2 slices)* ☐☐.☐								
3	Crackers and cheese biscuits	1	2	3	4	5	6	7	☐
4	Wholemeal and rye crackers	1	2	3	4	5	6	7	☐

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
5	'Bran' breakfast cereals	1	2	3	4	5	6	7	☐
6	Other breakfast cereals	1	2	3	4	5	6	7	☐
7	Added bran to foods	1	2	3	4	5	6	7	☐
8	Cakes and gateaux	1	2	3	4	5	6	7	☐
9	Buns	1	2	3	4	5	6	7	☐
10	Pastries	1	2	3	4	5	6	7	☐
11	Biscuits – chocolate, digestive and ginger	1	2	3	4	5	6	7	☐☐
12	Other biscuits	1	2	3	4	5	6	7	☐☐
13	Fruit puddings	1	2	3	4	5	6	7	☐

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
14	Milk based puddings and sauces	1	2	3	4	5	6	7	☐
15	Other puddings	1	2	3	4	5	6	7	☐
16	Yogurt and fruit fools	1	2	3	4	5	6	7	☐
17	Potatoes - boiled and jacket	1	2	3	4	5	6	7	☐
	When you eat these how many potatoes do you eat at a typical meal? ☐☐.☐ *Large baking (count as 3)/new (count as 0.5)*								
18	Roast potatoes and chips	1	2	3	4	5	6	7	☐
	When you eat these how many potatoes do you eat at a typical meal? ☐☐.☐								
19	Yorkshire puddings and savoury pancakes	1	2	3	4	5	6	7	☐

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
20	Brown and white rice	1	2	3	4	5	6	7	☐
21	Pasta and dumplings	1	2	3	4	5	6	7	☐
22	Tinned vegetables	1	2	3	4	5	6	7	☐
23	Peas and green beans	1	2	3	4	5	6	7	☐
24	Carrots	1	2	3	4	5	6	7	☐
25	Parsnips, swede and turnip	1	2	3	4	5	6	7	☐
26	Sweetcorn and mixed veg	1	2	3	4	5	6	7	☐
27	Beans and pulses	1	2	3	4	5	6	7	☐
28	Tomatoes	1	2	3	4	5	6	7	☐

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
29	Spinach	1	2	3	4	5	6	7	☐
30	Broccoli, Brussels sprouts and spring greens	1	2	3	4	5	6	7	☐
31	Cabbage and cauliflower	1	2	3	4	5	6	7	☐
32	Peppers and watercress	1	2	3	4	5	6	7	☐
33	Onion	1	2	3	4	5	6	7	☐
34	Green salad	1	2	3	4	5	6	7	☐
35	Side salads in dressing	1	2	3	4	5	6	7	☐
36	Courgettes, marrow and leeks	1	2	3	4	5	6	7	☐
37	Mushrooms	1	2	3	4	5	6	7	☐

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
38	Vegetable dishes	1	2	3	4	5	6	7	☐
39	Vegetarian foods	1	2	3	4	5	6	7	☐
40	Tinned fruit not including grapefruit, prunes, figs or blackcurrants	1	2	3	4	5	6	7	☐
41	Cooked fruit not including blackcurrants	1	2	3	4	5	6	7	☐
42	Dried fruit	1	2	3	4	5	6	7	☐
43	Fresh apples and pears	1	2	3	4	5	6	7	☐
44	Fresh oranges and orange juice	1	2	3	4	5	6	7	☐☐
45	Grapefruit and grapefruit juice	1	2	3	4	5	6	7	☐☐
46	Blackcurrants, ribena and hi-juice blackcurrant drinks	1	2	3	4	5	6	7	☐☐

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
47	*Other fruit juices (not squashes)*	1	2	3	4	5	6	7	
48	Diet Coke and Pepsi not including caffeine free	1	2	3	4	5	6	7	
49	Coke and Pepsi	1	2	3	4	5	6	7	
50	Soft drinks not including diet drinks (low calorie or low sugar)	1	2	3	4	5	6	7	
51	Bananas	1	2	3	4	5	6	7	
52	Fresh peaches, plums, cherries and grapes	1	2	3	4	5	6	7	
53	Strawberries and raspberries	1	2	3	4	5	6	7	
54	Fresh pineapple, melon, kiwi fruit and other tropical fruits	1	2	3	4	5	6	7	
55	Nuts	1	2	3	4	5	6	7	

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
56	Bacon and gammon	1	2	3	4	5	6	7	
57	Pork	1	2	3	4	5	6	7	
58	Chicken and turkey	1	2	3	4	5	6	7	
59	Lamb	1	2	3	4	5	6	7	
60	Beef	1	2	3	4	5	6	7	
61	Minced meat dishes	1	2	3	4	5	6	7	
62	Meat pies	1	2	3	4	5	6	7	
63	Liver and kidney	1	2	3	4	5	6	7	
64	Paté and liver sausage	1	2	3	4	5	6	7	

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
65	Faggots and black pudding	1	2	3	4	5	6	7	☐
66	Sausages	1	2	3	4	5	6	7	☐
67	Ham and luncheon meat	1	2	3	4	5	6	7	☐
68	White fish	1	2	3	4	5	6	7	☐
69	Fish fingers and fish dishes	1	2	3	4	5	6	7	☐
70	Oily fish	1	2	3	4	5	6	7	☐
71	Shellfish	1	2	3	4	5	6	7	☐
72	Boiled and poached eggs	1	2	3	4	5	6	7	☐
73	Omelette and fried eggs	1	2	3	4	5	6	7	☐

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
74	Cottage Cheese	1	2	3	4	5	6	7	☐
75	Cheese	1	2	3	4	5	6	7	☐
76	Pizza, quiches and cheese flans	1	2	3	4	5	6	7	☐
77	Soup	1	2	3	4	5	6	7	☐
78	Mayonnaise and salad cream	1	2	3	4	5	6	7	☐
79	Pickles, chutney, tomato ketchup and brown sauce	1	2	3	4	5	6	7	☐
80	Chocolate	1	2	3	4	5	6	7	☐☐
81	Other sweets	1	2	3	4	5	6	7	☐☐
82	Ice cream and chocolate desserts	1	2	3	4	5	6	7	☐

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
83	Cream	1	2	3	4	5	6	7	☐
84	Crisps and savoury snacks	1	2	3	4	5	6	7	☐☐
85	Sweet spreads	1	2	3	4	5	6	7	☐
86A	Gravy granules and powders	1	2	3	4	5	6	7	☐
86B	Stock cubes and Marmite	1	2	3	4	5	6	7	☐
87	Drinking chocolate and milk shakes not including McDonald style milkshakes	1	2	3	4	5	6	7	☐
88	Decaffeinated coffee and tea	1	2	3	4	5	6	7	☐☐
89	Tea	1	2	3	4	5	6	7	☐☐
90	Coffee	1	2	3	4	5	6	7	☐☐

	FOOD DESCRIPTION	FREQUENCY EATEN							
FOOD CODE		Never	Once every 2-3 Months	Once a Month	Once a Fort-night	1-2 Times per Week	3-6 Times per Week	Once a day	More than once a day
93	Spreading fat (1) ________ F☐☐☐☐	1	2	3	4	5	6	7	☐
94	Spreading fat (2) ________ F☐☐☐☐	1	2	3	4	5	6	7	☐
95	Spreading fat (3) ________ F☐☐☐☐	1	2	3	4	5	6	7	☐
96	Frying fat or oil (1) ________ F☐☐☐☐	1	2	3	4	5	6	7	☐
97	Frying fat or oil (2) ________ F☐☐☐☐	1	2	3	4	5	6	7	☐
98	Frying fat or oil (3) ________ F☐☐☐☐	1	2	3	4	5	6	7	☐
99	Other vegetable oil (1) e.g. salad dressings, marinades ________ F☐☐☐☐	1	2	3	4	5	6	7	☐
100	Other vegetable oil (1) e.g. salad dressings, marinades ________ F☐☐☐☐	1	2	3	4	5	6	7	☐

3.2 Are there food or drinks which you have eaten or drunk **once a week or more** which are not on the list? (Include breakfast bars such as Nutrigrain and Kellogs).

0. No/1. Yes ☐

If Yes

Name of food/drink	1-2 times per week	3-6 times per week	Once a day	More than once a day
☐☐☐☐☐				☐
☐☐☐☐☐				☐
☐☐☐☐☐				☐
☐☐☐☐☐				☐
☐☐☐☐☐				☐

Now I would like to ask in more detail about some specific foods

3.3 Which types of milk have you used regularly in drinks and added to breakfast cereals over the last 3 months?

1. Whole pasteurised
2. Semi-skimmed pasteurised
3. Skimmed pasteurised
4. Whole UHT
5. Semi-skimmed UHT
6. Skimmed UHT
7. Other

Milk 1 ☐ Other (specify)________________

Milk 2 ☐ Other (specify)________________

Milk 3 ☐ Other (specify)________________

3.4 * On average over the last 3 months how much of each milk have you consumed per day?

Milk 1 ☐ . ☐☐ pints

Milk 2 ☐ . ☐☐ pints

Milk 3 ☐ . ☐☐ pints

3.5 Do you add sugar to breakfast cereals, tea & coffee, puddings etc.? ☐

0. No *go to 3.7*
1. Yes

3.6 Approximately how many teaspoons of sugar do you add each day? ☐☐

3.7 When you eat meat, how much of the fat do you usually cut off (including chicken skin)? ☐

1. all 100%
2. most 60%
3. some 30%
4. none 0%
9. not applicable

4: FOOD SUPPLEMENTS

4.1 During the past three months have you taken any pills, tonics or tablets to supplement your diet? (e.g. vitamins, minerals, iron tablets, folic acid, fish oils etc.) ☐

0. No 1. Yes

If yes, please state which:
(for number per day record number of tablets/capsules/teaspoons per day, as appropriate)

Supplement	Number per day	How many days in the last 90?
☐☐☐☐		
☐☐☐☐		
☐☐☐☐		
☐☐☐☐		
☐☐☐☐		

5: GENERAL DIET QUESTIONS

5.1 Are the **past three months** typical of the way you generally eat? ☐

0. No
1. Yes
2. Reasonably

5.2 Still thinking about your normal pattern of eating - in a typical **week** how often do you:

*	Never	< once/ week	1-2 times	3-6 times	everyday
eat breakfast					
eat lunch					
eat an evening meal					
go out in the evening not necessarily to eat but also to socialise					

5.3 Just thinking about the **past week** how many servings did you eat of:

vegetables and vegetable-containing dishes (excluding potatoes)?	
fruit and pure fruit juices?	
meat and fish and their dishes?	

6: DIETING

6.1 * Which of the following describes you best?
1. I have **NEVER** been on a diet to lose weight
2. I have **ONLY ONCE** been on a diet to lose weight
3. I **USED TO** diet **REGULARLY** to lose weight but **DON'T ANYMORE**
4. I go on a diet to lose weight **EVERY NOW AND AGAIN**
5. I am **USUALLY** on a diet to lose weight

If 2, 4 or 5 ask 6.2 otherwise go to section 7

6.2 Are you currently trying to lose weight by dieting?
0. No
1. Yes

7: ALCOHOL CONSUMPTION

I'd like to ask you a few questions about your drinking and smoking habits.

7.1 Do you ever drink alcohol? 0. No *go to section 8* 1. Yes

During the past three months:

7.2 a) How often have you drunk **Shandy or Low Alcohol Beer/Lager/Cider**? (don't include alcohol **free** lager etc) FFQ 1-7 >x1

b) When you drank these how many pints did you normally have? *(if range given code mid-point)*

7.3 a) How often have you drunk **Beer/Stout/Lager/Cider/Alcopops?** FFQ 1-7 >x1

b) When you drank these how many pints did you normally have? *(if range given code mid-point)*

7.4 a) How often have you drunk **Low alcohol wine?** FFQ 1-7 >x1

b) When you drank this how many glasses did you normally have? *(if range given code mid-point)*

233

7.5 a) How often have you drunk **Wine/Sherry/Martini/Cinzano?** FFQ 1-7 > x1

b) When you drank these how many glasses did you normally have? *(if range given code mid-point)*

7.6 a) How often have you drunk **Spirits/Liqueurs?** FFQ 1-7 > x1

b) When you drank these how many measures did you normally have? *(if range given code mid-point)*

8: SMOKING

8.1 Have you <u>ever</u> smoked regularly (at least once a day for a year or more) ?
0. No *go to section 9*
1. Yes

8.2 How old were you when you first smoked regularly ?

8.3 Are you currently smoking ?
0. No *go to section 9*
1. Yes *go to 8.4*

8.4 How many per day? *Record maximum stated*

9: FAMILY BACKGROUND

Now I'd like to ask some questions about your family.

Tell the woman that she may find some of these questions difficult or impossible to answer. Explain that you would like to leave a form for her to complete where possible by asking her parents for the details. Answers that she can give us now (even approximately) are useful but if she can supplement them later that would be extremely helpful.

Starting with your **FATHER**:

9.1 Is your father still alive?
0.No, 1.Yes, 7. Adopted, 8. Don't talk about him, 9. Don't know

9.2 What was his full-time job when you were born?
or if unemployed or part time, last full time job before that time.
Probe industry & self-employed/manager/foreman/employee.

If full time student give subject.

Job Position

Self-employed/manager/foreman/employee

Industry

9.3 Approximately what is/was his height?

In feet and inches? ☐ ft ☐☐ . ☐ ins

OR In centimetres ☐☐☐ . ☐ cms

9.4 Approximately what is/was his current/latest weight?

In stones and pounds? ☐☐ st ☐☐ lbs

OR In kilograms? ☐☐☐ . ☐ kg

9.5 What was his birth weight?

In pounds and ounces? ☐☐ lbs ☐☐ . ☐ oz

OR In grams? ☐☐☐☐ grams

Now your **MOTHER**:

9.6 Is your mother still alive? ☐
0. No, 1.Yes, 7. Adopted, 8. Don't talk about her, 9. Don't know

9.7 and what was her **full** name when you were born? ____________________

____

9.8 What is/was her date of birth? d d ☐☐ m m ☐☐ y y ☐☐

9.9 Where was she born?

If in UK: Town/Village ____________________

County ____________________

If abroad: Country ____________________

9.10 What is/was her height?

In feet and inches ☐ ft ☐☐ . ☐ ins

OR In centimetres? ☐☐☐ . ☐ cm

9.11 What did she weigh before you were conceived?

In stones and pounds? ☐☐ st ☐☐ lbs

OR In kilograms? ☐☐☐ . ☐ kg

9.12 What was her birth weight?

In pounds and ounces? ☐☐ lbs ☐☐ . ☐ oz

OR In grams? ☐☐☐☐ grams

Returning to **YOURSELF**:

9.13 What is your date of birth? d d ☐☐ m m ☐☐ y y ☐☐

9.14 What was your birth weight?

In pounds and ounces? ☐☐ lbs ☐☐ . ☐ oz

Or In grams? ☐☐☐☐ grams

9.15 Where were you born?

If in UK: Town/Village____________________

County ____________________

If abroad: Country ____________________

9.16 Were you born at home or in hospital ?
1. Home
2. Hospital - specify ______________________

9.17 Were you part of a multiple birth (twin, triplet etc.)?
0. No
1. Yes

9.18 Were you born early, late or when you were expected?
1. Early
2. When expected *go to 9.20*
3. Late
9. Don't know

9.19 How early/ late were you? ☐☐ weeks ☐ days
99. Don't know

1. Certain
2. Not certain or mid point of a range

9.20 How many children did your mother have before you were born (including stillbirths)?

9.21 Do you have any sisters aged 20 or over?
0.No, 1.Yes

10: EDUCATION

I would like to ask you briefly about your education.

10.1 How old were you when you left full-time education ? ☐☐ yrs
(don't round up; enter current age if still studying)
(count a year or less out as continuous education)

10.2 Have you passed any exams or do you have any formal qualifications ?
1. None
2. CSE/ School cert/ GCSE grade D or lower/ NVQ1/ Foundation GNVQ
3. O levels/ Matric/ GCSE grade A,B,C/ RSA secretarial/ NVQ2/ Intermediate GNVQ
4. A levels/ City & Guilds/ EN(G)/ ONC/ NNEB/ BTech (day release)/ NVQ3/ Advanced GNVQ/ OND / HNC
5. HND/ RGN/ Teaching Cert/ NVQ4
6. Degree/ NVQ5
7. Other (specify)

11: ETHNIC GROUP

11.1 To which of the ethnic groups listed on this card do you consider you belong?
*
1. White
2. Black Caribbean
3. Black African
4. Black Other
5. Indian
6. Pakistani
7. Bangladeshi
8. Chinese
9. Other Asian group
10. Other (specify)______________________

12: MARITAL STATUS

12.1 What is your marital status?
1. Single (never married)
2. Married (living with husband)
3. Separated
4. Divorced
5. Widowed

13: HOUSING

13.1 What type of accommodation do you live in?
1. Detached house/bungalow
2. Semi-detached house/bungalow
3. End terraced house
4. Terraced house
5. Purpose built flat/maisonette
6. Converted flat/maisonette
7. Dwelling with business premises
8. Bedsitter in multiple occupation
9. Bedsitter other
10. Hostel
11. Hall of residence
12. Other student accommodation
13. Other (specify)______________________

13.2 On what floor is the main part of living accommodation?
(If more than one code the lowest)

1. Basement
2. Ground floor/street level
3. 1st floor
4. 2nd floor
5. 3rd floor
6. 4th to 9th floor
7. 10th to 19th floor
8. 20th floor or higher

☐

13.3 Do you own your own home, or are you buying it on a mortgage, or do you rent it in some way?
*
1. Owns outright or buying with mortgage
2. Rent from private landlord
3. Rent from council or housing association
4. Other rented accommodation (hostel, hall of residence, B& B)
5. Lives with parents
6. Other (specify)________________________

☐

13.4 Here is a list of some problems that people often have with their homes. Please tell me if you think that each one is a big problem, a small problem or not a problem for you and your family? *(Tick appropriate boxes)*

*	Big problem	Small problem	Not a problem
Condensation			
Rising or penetrating damp			
Difficulty in keeping home warm			
Leaking roof			
Rot in window frames, timbers or floorboards			
Not enough space			

14: HOUSEHOLD COMPOSITION AND CHILDREN

14.1. Does anyone else live in the house with you?
0 = No *go to 14.2* 1 = Yes

☐

For each person living in the household (apart from the woman herself) complete one line.
A household is defined as a group of people who share a living room or eat together for at least one meal a day. People living in hostels or halls of residence are classed as living alone.
*For **all children** (see younger generation list) record **date of birth** (or age if d.o.b. is not available).*
*For the **woman's own children** give the child's **birthweight**.*
For all adults, record whether they currently smoke at least once a day. 0=No, 1=Yes
***Days per week is for anyone who is only in the household part-time.** Record the average number of days per week that person lives in the household.*

KEY: ***Own Generation***

H	= *Husband*
C	= *Cohabitee*
S	= *Sibling (brother/sister)*
AS	= *Adopted sibling*
SIL	= *Sibling-in-law (sister/brother-in-law)*
SS	= *Stepsibling*
FS	= *Foster sibling*

Younger Generation

OC	= *Own child (son/daughter)*
SC	= *Step child*
AC	= *Adopted child*
FC	= *Foster child*
CIL	= *Child-in-law (son/daughter–in-law)*
CC	= *Cohabitee's child*
GC	= *Grandchild*
SB	= *Still born child*

Older Generation

P	= *Parent*
FP	= *Foster parent*
SP	= *Step parent*
PIL	= *Parent-in-law*
GP	= *Grandparent*

Other

OR	= *Other relative*
ON	= *Other non-relative*

Person number	Relationship to woman	sex		Date of birth			Age (yrs)	Birthweight			Smoker	Days per week
		M	F	Day	Mth	Yr		lb	oz	grams		
1												
2												
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												

14.2. How many children have you had, including any stillbirths?
(Any not included above add to the table with 0 days/week)

14.4 *If the woman has a child under the age of two years:* Are you breastfeeding your (youngest) child? *(Any amount of breastfeeding counts as yes)*
0. No 1. Yes

15: PARTNER'S OCCUPATION

If there is a husband or partner living in the house (if not go to 16):

15.1 Was your husband/partner in paid employment or self-employed in the week ending last Sunday?
0. No *go to 15.3*
1. Yes *go to to 15.2*

15.2 Was he working full time or part time?
0. Full time (more than 30 hours) *go to 15.6b*
1. Part time (30 hours or fewer) *go to 15.3*

15.3 Was he going to college full time?
0.No *if working part-time go to 15.6a*
if not working go to 15.5
1.Yes

15.4 *If yes,* what is he studying? ____________________
If working part time go to 15.7
If not working go to section 16.

15.5 *If not working or studying* was he
Unemployed ? (1)
Permanently unable to work because of long term sickness or disability ? (2)
looking after home or family? (3)
other ? (specify)____________ (4)

15.6a *If **not** working or working part-time,* what was his last **full-time** job?
If only ever part-time give last part time job.
Then if currently working part time go to 15.7, otherwise go to section 16

Job Position ______________
Self-employed/manager/foreman/employee
Industry ______________

15.6b *If working full-time,* what is his job ? *(Then go to section 16)*
Probe industry & self-employed/manager/foreman/employee

Job Position
Self-employed/manager/foreman/employee
Industry ______________

15.7 *If working part-time now,* what is his current job?

Job Position ______________
Self-employed/manager/foreman/employee
Industry ______________

15.8 *If working part time,* how many hours per week does he work?
hrs mins

16: CHILDCARE ARRANGEMENTS

16.1 *If the woman works (part-time or full-time) and has children at home under the age of twelve years: (if not go to section 17)*

Which of the following best describes the way you arrange for your children aged 12 or under to be looked after while you are at work?
Tick up to three boxes.

*	1st mention	2nd mention	3rd mention
1. I work only while they are at school			
2. They look after themselves until I get home			
3. I work from home			
4. My husband/partner looks after them			
5. A nanny or mother's help looks after them at home			
6. They go to a work-place nursery			
7. They go to a day nursery			
8. They go to a child minder			
9. A relative looks after them			
10. A friend or neighbour looks after them			
11. Other (specify)			

17: BENEFITS

17.1 * Are you (or your husband/partner) receiving any of the following benefits? (Income support/job seekers allowance/family credit/housing benefit) ☐

0 = No *go to section 18*
1 = Yes

17.2 How long have you been receiving them? (0=No, 1=<1 year, 2=1-2 years, 3=2+years, 9=Don't know)

(a) Income support ☐

(b) Job seekers allowance ☐

(c) Family credit ☐

(d) Housing benefit ☐

*If not done before, get **consent** here*

18: BODY MEASUREMENTS

18.1 Pulse (30sec) ☐☐☐
(Double the value to give pulse for 1 minute)

18.2 Which hand do you write with? ☐

1. Right
2. Left
3. Completely ambidextrous

18.3 Weight ☐☐☐ . ☐ kg

18.4 Height ☐☐☐ . ☐ cm

Mark and measure up the non-dominant arm and side of the body (measure the left if completely ambidextrous)

18.5 Leg length ☐☐☐ . ☐ cm

18.6 Waist circumference ☐☐☐ . ☐ cm

18.7 Hip circumference ☐☐☐ . ☐ cm

18.8 Mid-thigh circumference ☐☐ . ☐ cm

18.9 Calf circumference ☐☐ . ☐ cm

18.10 Mid-upper arm circumference (non-dominant side) ☐☐ . ☐ cm

18.11 Triceps skinfold (non-dominant side)

☐☐ . ☐ mm ☐☐ . ☐ mm
☐☐ . ☐ mm ☐☐ . ☐ mm
☐☐ . ☐ mm

18.12 Biceps skinfold (non-dominant side)

☐☐ . ☐ mm ☐☐ . ☐ mm
☐☐ . ☐ mm ☐☐ . ☐ mm
☐☐ . ☐ mm

18.13 Subscapular skinfold (non-dominant side)

☐☐ . ☐ mm ☐☐ . ☐ mm
☐☐ . ☐ mm ☐☐ . ☐ mm
☐☐ . ☐ mm

18.14 Upper suprailiac skinfold (non-dominant side)

☐☐ . ☐ mm ☐☐ . ☐ mm
☐☐ . ☐ mm ☐☐ . ☐ mm
☐☐ . ☐ mm

18.15 Skinfold calipers used

18.16 Time (24 hr clock)

19: MOUTHWASH SAMPLE

If the mouthwash sample was obtained at the beginning, go to section 20

19.1 Mouthwash sample provided

(0=No, 1=Yes)

19.2 Time of mouthwash sample (24 hr clock)

20: GENERAL HEALTH

20.1 How is your health in general? Would you say it was:
*
1. Very good
2. Good
3. Fair
4. Bad
5. Very bad

20.2 Do you have any long-standing illness, disability or infirmity? By long standing, I mean anything that has troubled you over a period of time or that is likely to affect you over a period of time.

0. No *go to 20.4*
1. Yes

20.3 What is the illness/disability/infirmity?________________________________

(Do not record headaches, indigestion, aches and pains. We are interested in major problems such as diabetes , multiple sclerosis, rheumatoid arthritis, muscular dystrophy – anything which might affect growth or body composition)

20.4 To what extent do you feel that the stress or pressure you have experienced in your life has affected your health?
*
1. None
2. Slightly
3. Moderately
4. Quite a lot
5. Extremely

20.5 In general, how much stress or pressure have you experienced in your daily living in the last 4 weeks?
*
1. None
2. Just a little
3. A good bit
4. Quite a lot
5. A great deal

21: MENSTRUAL CYCLE AND PREGNANCIES

21.1 What was the date of the first day of your last menstrual period? d d m m y y

21.2 How long is your usual cycle between the start of one period and the start of the next period? days
(Don't know 99)

21.3 Is your usual cycle regular, or has it varied by more than 5 days between periods in the last 6 months?
1: Regular
2: Varied by more than 5 days

21.4 How old were you when you had your first period ? yrs
(Don't know 99.9)

21.5 Within the last 3 months have you taken the oral contraceptive pill or had the Depot injection or other hormonal treatment?

0. No *go to 21.8*
1. Yes

21.6 Which? Specify (most recent if several) ________________________

21.7 Are you currently taking this?

0. No
1. Yes

21.8 Do you anticipate trying for a baby within the next 12 months?

0. No
1. Yes

That is the end of the questionnaire but we would be grateful for your help with some extra items.

Use the explanations in fieldworker notes for the following items but please mark the results below:

Have you left a birth details form?
0. No
1. Yes

Have you left a food diary?
0. No
1. Yes

Is there agreement to a blood sample?
(Remember to mark the woman's record card as well)
0. No
1. Yes

*Has **consent** been obtained for the GP to notify us if the woman becomes pregnant?*
0. No
1. Yes

Is the woman willing to be approached for other studies related to the SWS?
0. No
1. Yes

Don't forget to leave a fridge magnet, pregnancy reply card, two prepaid envelopes (one large and one small), and, if the woman is interested, an information leaflet.

THANK YOU VERY MUCH FOR ALL YOUR HELP. THE INFORMATION YOU HAVE GIVEN US IS VERY IMPORTANT FOR IMPROVING THE HEALTH OF WOMEN. THE MORE WOMEN WHO TAKE PART, THE MORE VALUABLE ALL THE DATA BECOME SO WE WOULD BE VERY GRATEFUL IF YOU WOULD ENCOURAGE YOUR FRIENDS TO TAKE PART.

MANY THANKS AGAIN

Local Research Ethics
Committee No 276/97

References

1. Thurnham DI, Bender DA, Scott J, Halsted CH. Water-soluble vitamins. In: Garrow JS, James WPT, Ralph A, eds. Human Nutrition and Dietetics. Edinburgh: Churchill Livingstone,2000: 249-287.

2. Messina M, Lampe JW, Birt DF, et al. Reductionism and the narrowing nutrition perspective: time for reevaluation and emphasis on food synergy. *Journal of the American Dietetic Association* 2001; **101**: 1416-1419.

3. Hegsted DM. A look back at lessons learned and not learned. *Journal of Nutrition* 1994; **124**: 1867S-1870S.

4. Sandhu MS, White IR, McPherson K. Systematic review of the prospective cohort studies on meat consumption and colorectal cancer risk: a meta-analytical approach. *Cancer Epidemiology, Biomarkers & Prevention* 2001; **10**: 439-446.

5. Ness A, Powles JW. Fruit and vegetables, and cardiovascular risk: a review. *International Journal of Epidemiology* 1997; **26**: 1-13.

6. Hooper L, Summerbell CD, Higgins JPT, et al. Dietary fat intake and prevention of cardiovascular disease: systematic review. *British Medical Journal* 2001; **322**: 757-763.

7. Beal VA. Nutritional studies during pregnancy: changes in intakes of calories, carbohydrate, fat, protein and calcium. *Journal of the American Dietetic Association* 1971; **58**: 312-320.

8. Anderson AS, Campbell D, Shepherd R. Nutrition knowledge, attitude to healthier eating and dietary intake in pregnant compared to non-pregnant women. *Journal of Human Nutrition and Dietetics* 1993; **6**: 335-353.

9. Davey Smith G, Brunner E. Socio-economic differentials in health: the role of nutrition. *Proceedings of the Nutrition Society* 1997; **56**: 75-90.

10. Marmot MG, Shipley MJ, Rose G. Inequalities in death - specific explanations of a general pattern. *Lancet* 1984; **i**: 1003-1006.

11. Mackenbach JP, Kunst AE, Cavelaars AEJM, Groenhof F, Geurts JJM, EU Working Group on Socioeconomic Inequalities in Health. Socioeconomic inequalities in morbidity and mortality in Western Europe. *Lancet* 1997; **349**: 1655-1659.

12. James WPT, Nelson M, Ralph A, Leather S. The contribution of nutrition to inequalities in health. *British Medical Journal* 1997; **314**: 1545-1549.

13. Dubois L, Girard M. Social position and nutrition: a gradient relationship in Canada and the USA. *European Journal of Clinical Nutrition* 2001; **55**: 366-373.

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

15. Acheson D. Independent inquiry into inequalities in health report. London: The Stationery Office 1998.

16. de Onis M, Villar J, Gülmezoglu M. Nutritional interventions to prevent intrauterine growth retardation: evidence from randomized controlled trials. *European Journal of Clinical Nutrition* 1998; **52**: Suppl 1, S83-S93.

17. Godfrey K, Robinson S, Barker DJP, Osmond C, Cox V. Maternal nutrition in early and late pregnancy in relation to placental and fetal growth. *British Medical Journal* 1996; **312**: 410-414.

18. Godfrey KM, Barker DJP, Robinson S, Osmond C. Maternal birthweight and diet in pregnancy in relation to the infant's thinness at birth. *British Journal of Obstetrics and Gynaecology* 1997; **104**: 663-667.

19. Ramakrishnan U, Manjrekar R, Rivera J, Gonzáles-Cossío T, Martorell R. Micronutrients and pregnancy outcome: a review of the literature. *Nutrition Research* 1999; **19**: 103-159.

20. Belizán JM, Villar J, Bergel E, et al. Long term effect of calcium supplementation during pregnancy on the blood pressure of offspring: follow up of a randomised controlled trial. *British Medical Journal* 1997; **315**: 281-285.

21. Campbell DM, Hall MH, Barker DJP, Cross J, Shiell AW, Godfrey KM. Diet in pregnancy and the offspring's blood pressure 40 years later. *British Journal of Obstetrics and Gynaecology* 1996; **103**: 273-280.

22. Shiell AW, Campbell-Brown M, Haselden S, Robinson S, Godfrey KM, Barker DJP. High-meat, low-carbohydrate diet in pregnancy: relation to adult blood pressure in the offspring. *Hypertension* 2001; **38**: 1282-1288.

23. Antonov AN. Children born during the siege of Leningrad. *Journal of Pediatrics* 1947; **30**: 250-259.

24. Smith CA. The effect of wartime starvation in Holland upon pregnancy and its product. *American Journal of Obstetrics and Gynecology* 1947; **53**: 599-608.

25. Prentice AM, Whitehead RG, Watkinson M, Lamb WH, Cole TJ. Prenatal dietary supplementation of African women and birth-weight. *Lancet* 1983; **ii**: 489-492.

26. Viegas OAC, Scott PH, Cole TJ, Eaton P, Needham PG, Wharton BA. Dietary protein energy supplementation of pregnant Asian mothers at Sorrento, Birmingham. II: Selective during third trimester only. *British Medical Journal* 1982; **285**: 592-595.

27. Rao S, Yajnik CS, Kanade A, et al. Intake of micronutrient-rich foods in rural Indian mothers is associated with the size of their babies at birth: Pune maternal nutrition study. *Journal of Nutrition* 2001; **131**: 1217-1224.

28. Brown JE. Preconceptional nutrition and reproductive outcomes. *Annals of the New York Academy of Sciences* 1993; **678**: 286-292.

29. MRC Vitamin Study Research Group. Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. *Lancet* 1991; **338**: 131-137.

30. Jackson AA, Robinson SM. Dietary guidelines for pregnancy: a review of current evidence. *Public Health Nutrition* 2001; **4**: 625-630.

31. Barker DJP. Mother's, babies and health in later life. Edinburgh: Churchill Livingstone 1998.

32. Rose G. Familial patterns in ischaemic heart disease. *British Journal of Preventative and Social Medicine* 1964; **18**: 75-80.

33. Forsdahl A. Are poor living conditions in childhood and adolescence an important risk factor for arteriosclerotic heart disease? *British Journal of Preventive and Social Medicine* 1977; **31**: 91-95.

34. Barker DJP, Osmond C. Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. *Lancet* 1986; **i**: 1077-1081.

35. Barker DJP, Winter PD, Osmond C, Margetts B, Simmonds SJ. Weight in infancy and death from ischaemic heart disease. *Lancet* 1989; **ii**: 577-580.

36. Osmond C, Barker DJP, Winter PD, Fall CHD, Simmonds SJ. Early growth and death from cardiovascular disease in women. *British Medical Journal* 1993; **307**: 1519-1524.

37. Forsén T, Eriksson JG, Tuomilehto J, Teramo K, Osmond C, Barker DJP. Mother's weight in pregnancy and coronary heart disease in a cohort of Finnish men: follow up study. *British Medical Journal* 1997; **315**: 837-840.

38. Hales CN, Barker DJP, Clark PMS, et al. Fetal and infant growth and impaired glucose tolerance at age 64. *British Medical Journal* 1991; **303**: 1019-1022.

39. Huxley R, Shiell AW, Law CM. The role of size at birth and postnatal catch-up growth in determining systolic blood pressure: a systematic review of the literature. *Journal of Hypertension* 2000; **18**: 815-831.

40. Barker DJP, Godfrey KM, Fall C, Osmond C, Winter PD, Shaheen SO. Relation of birth weight and childhood respiratory infection to adult lung function and death from chronic obstructive airways disease. *British Medical Journal* 1991; **303**: 671-675.

41. Barker DJP, Forsén T, Uutela A, Osmond C, Eriksson JG. Size at birth and resilience to effects of poor living conditions in adult life: longitudinal study. *British Medical Journal* 2001; **323**: 1-5.

42. Brooks AA, Johnson MR, Steer PJ, Pawson ME, Abdalla HI. Birth weight: nature or nurture? *Early Human Development* 1995; **42**: 29-35.

43. Carr-Hill R, Campbell DM, Hall MH, Meredith A. Is birth weight determined genetically? *British Medical Journal* 1987; **295**: 687-689.

44. Nance WE, Kramer AA, Corey LA, Winter PM, Eaves LJ. A causal analysis of birth weight in the offspring of monozygotic twins. *American Journal of Human Genetics* 1983; **35**: 1211-1223.

45. Margetts BM, Nelson M. Design concepts in nutritional epidemiology. Oxford: Oxford University Press 1997.

46. Department of Health and Social Security. Committee on Medical Aspects of Food Policy. Recommended daily amounts of food energy and nutrients for groups of people in the United Kingdom. London: HMSO 1979.

47. Department of Health. Committee on Medical Aspects of Food Policy. Dietary reference values for food energy and nutrients for the United Kingdom. London: The Stationery Office 1991.

48. Department of Health. The health of the nation. London: HMSO 1992.

49. Department of Health. Eat well: an action plan from the Nutrition Task Force to achieve the health of the nation targets on diet and nutrition. London: Department of Health 1994.

50. Department of Health. Eat well II: a progress report from the Nutrition Task Force on the action plan to achieve the health of the nation targets on diet and nutrition. London: Department of Health 1996.

51. Department of Health. Committee on Medical Aspects of Food Policy. Nutritional aspects of cardiovascular disease. London: HMSO 1994.

52. Williams C. Healthy eating: clarifying advice about fruit and vegetables. *British Medical Journal* 1995; **310**: 1453-1455.

53. Health Education Authority. Eight guidelines for a healthy diet. 1997;

54. Health Education Authority. The pregnancy book. London: Health Education Authority 1997.

55. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Current Opinion in Lipidology* 2002; **13**: 3-9.

56. Slattery ML, Boucher KM, Caan BJ, Potter JD, Ma KN. Eating patterns and risk of colon cancer. *American Journal of Epidemiology* 1998; **148**: 4-15.

57. Hu FB, Rimm E, Smith-Warner SA, et al. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *American Journal of Clinical Nutrition* 1999; **69**: 243-249.

58. Gex-Fabry M, Raymond L, Jeanneret O. Multivariate analysis of dietary patterns in 939 Swiss adults: sociodemographic parameters and alcohol consumption profiles. *International Journal of Epidemiology* 1988; **17**: 548-555.

59. Fung TT, Willett WC, Stampfer MJ, Manson JE, Hu FB. Dietary patterns and the risk of coronary heart disease in women. *Archives of Internal Medicine* 2001; **161**: 1857-1862.

60. Osler M, Heitmann BL, Gerdes LU, Jørgensen LM, Schroll M. Dietary patterns and mortality in Danish men and women: a prospective observational study. *British Journal of Nutrition* 2001; **85**: 219-225.

61. McCann SE, Marshall JR, Brasure JR, Graham S, Freudenheim JL. Analysis of patterns of food intake in nutritional epidemiology: food classification in principal component analysis and the subsequent impact on estimates for endometrial cancer. *Public Health Nutrition* 2001; **4**: 989-997.

62. Tseng M, DeVellis RF. Fundamental dietary patterns and their correlates among US whites. *Journal of the American Dietetic Association* 2001; **101**: 929-932.

63. Terry P, Hu FB, Hansen H, Wolk A. Prospective study of major dietary patterns and colorectal cancer risk in women. *American Journal of Epidemiology* 2001; **154**: 1143-1149.

64. Sánchez-Villegas A, Delgado-Rodríguez M, Martínez-González MÁ, De Irala-Estévez J, for the SUN group. Gender, age, socio-demographic and lifestyle factors associated with major dietary patterns in the Spanish Project SUN (Seguimiento Universidad de Navarra). *European Journal of Clinical Nutrition* 2003; **57**: 285-292.

65. Whichelow MJ, Prevost AT. Dietary patterns and their association with demographic, lifestyle and health variables in a random sample of British adults. *British Journal of Nutrition* 1996; **76**: 17-30.

66. Gregory J, Foster K, Tyler H, Wiseman M. The Dietary and Nutritional Survey of British Adults. London: HMSO 1990.

67. Williams DEM, Prevost AT, Whichelow MJ, Cox BD, Day NE, Wareham NJ. A cross-sectional study of dietary patterns with glucose intolerance and other features of the metabolic syndrome. *British Journal of Nutrition* 2000; **83**: 257-266.

68. Beaudry M, Galibois I, Chaumette P. Dietary patterns of adults in Québec and their nutritional adequacy. *Canadian Journal of Public Health* 1998; **89**: 347-351.

69. Barker ME, McClean SI, Thompson KA, Reid NG. Dietary behaviours and sociocultural demographics in Northern Ireland. *British Journal of Nutrition* 1990; **64**: 319-329.

70. van Dam RM, Grievink L, Ocké MC, Feskens JM. Patterns of food consumption and risk factors for cardiovascular disease in the general Dutch population. *American Journal of Clinical Nutrition* 2003; **77**: 1156-1163.

71. Fung TT, Rimm EB, Spiegelman D, et al. Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease. *American Journal of Clinical Nutrition* 2001; **73**: 61-67.

72. Hu FB, Rimm E, Stampfer MJ, Ascherio A, Spiegelman D, Willett WC. Prospective study of major dietary patterns and risk of coronary heart disease in men. *American Journal of Clinical Nutrition* 2000; **72**: 912-921.

73. Osler M, Helms Andreasen A, Heitmann B, et al. Food intake patterns and risk of coronary heat disease: a prospective cohort study examining the use of traditional scoring techniques. *European Journal of Clinical Nutrition* 2002; **56**: 568-574.

74. van Dam RM, Rimm EB, Willett WC, Stampfer MJ, Hu FB. Dietary patterns and risk for type 2 diabetes mellitus in U.S. men. *Annals of Internal Medicine* 2002; **136**: 201-209.

75. Fung T, Hu FB, Fuchs C, et al. Major dietary patterns and the risk of colorectal cancer in women. *Archives of Internal Medicine* 2003; **163**: 309-314.

76. Hulshof KFAM, Wedel M, Löwik MRH, et al. Clustering of dietary variables and other lifestyle factors (Dutch Nutrition Surveillance System). *Journal of Epidemiology and Community Health* 1992; **46**: 417-424.

77. Wirfält AKE, Jeffery RW. Using cluster analysis to examine dietary patterns: nutrient intakes, gender and weight status differ across food pattern clusters. *Journal of the American Dietetic Association* 1997; **97**: 272-279.

78. Greenwood DC, Cade JE, Draper A, Barrett JH, Calvert C, Greenhalgh A. Seven unique food consumption patterns identified among women in the UK Women's Cohort Study. *European Journal of Clinical Nutrition* 2000; **54**: 314-320.

79. Pryer JA, Nichols R, Elliott P, Thakrar B, Brunner E, Marmot M. Dietary patterns among a national random sample of British adults. *Journal of Epidemiology and Community Health* 2001; **55**: 29-37.

80. Chen H, Ward MH, Graubard BI, et al. Dietary patterns and adenocarcinoma of the esophagus and distal stomach. *American Journal of Clinical Nutrition* 2002; **75**: 137-144.

81. Farchi G, Mariotti S, Menotti A, Seccareccia F, Torsello S, Fidanza F. Diet and 20-y mortality in two rural populations groups. *American Journal of Clinical Nutrition* 1989; **50**: 1095-1103.

82. Wirfält E, Hedblad B, Gullberg B, et al. Food patterns and components of the metabolic syndrome in men and women: a cross-sectional study within the Malmö diet and cancer cohort. *American Journal of Epidemiology* 2001; **154**: 1150-1159.

83. Tucker KL, Chen H, Hannan MT, et al. Bone mineral density and dietary patterns in older adults: the Framingham Osteoporosis Study. *American Journal of Clinical Nutrition* 2002; **76**: 245-252.

84. Kant AK. Indexes of overall diet quality: a review. *Journal of the American Dietetic Association* 1996; **96**: 785-791.

85. Kennedy ET, Ohls J, Carlson S, Fleming K. The Healthy Eating Index: design and applications. *Journal of the American Dietetic Association* 1995; **95**: 1103-1108.

86. Hann CS, Rock CL, King I, Drewnowski A. Validation of the Healthy Eating Index with use of plasma biomarkers in a clinical sample of women. *American Journal of Clinical Nutrition* 2001; **74**: 479-486.

87. Prättälä R, Berg MA, Puska P. Diminishing or increasing contrasts? Social class variations in Finnish food consumption patterns, 1979-1990. *European Journal of Clinical Nutrition* 1991; **46**: 279-287.

88. Roos E, Lahelma E, Virtanen M, Prättälä R, Pietinen P. Gender, socioeconomic status and family status as determinants of food behaviour. *Social Science and Medicine* 1998; **46**: 1519-1529.

89. Hupkens CLH, Knibbe RA, Drop MJ. Social class differences in women's fat and fibre consumption: a cross-national study. *Appetite* 1997; **28**: 131-149.

90. Billson H, Pryer JA, Nichols R. Variation in fruit and vegetable consumption among adults in Britain. An analysis from the dietary and nutritional survey of British adults. *European Journal of Clinical Nutrition* 1999; **53**: 946-952.

91. Groth MV, Fagt S, Brøndsted L. Social determinants of dietary habits in Denmark. *European Journal of Clinical Nutrition* 2001; **55**: 959-966.

92. Johansson L, Thelle DS, Solvoll K, Bjørneboe GEA, Drevon CA. Healthy dietary habits in relation to social determinants and lifestyle factors. *British Journal of Nutrition* 1999; **81**: 211-220.

93. Wrigley N, Warm D, Margetts B, Whelan A. Assessing the impact of improved retail access on diet in a 'food desert': a preliminary report. *Urban Studies* 2002; **39**: 2061-2082.

94. Roos E, Prättälä R, Lahelma E, Kleemola P, Pietinen P. Modern and healthy?: socioeconomic differences in the quality of diet. *European Journal of Clinical Nutrition* 1996; **50**: 753-760.

95. Roos G, Johansson L, Kasmel A, Prättälä R. Disparities in vegetable and fruit consumption: European cases from the north to the south. *Public Health Nutrition* 2000; **4**: 35-43.

96. Pryer J, Brunner E, Elliott P, Nichols R, Dimond H, Marmot M. Who complied with COMA 1984 dietary fat recommendations among a nationally representative sample of British adults in 1986-7 and what did they eat? *European Journal of Clinical Nutrition* 1995; **49**: 718-728.

97. De Irala-Estévez J, Groth M, Johansson L, Oltersdorf U, Prättälä R, Martínez-González MA. A systematic review of socio-economic differences in food habits in Europe: consumption of fruit and vegetables. *European Journal of Clinical Nutrition* 2000; **54**: 706-714.

98. Baghurst KI, Record SJ, Baghurst PA, Syrette JA, Crawford D, Worsley A. Sociodemographic determinants in Australia of the intake of food and nutrients implicated in cancer aetiology. *Medical Journal of Australia* 1990; **153**: 444-452.

99. Hulshof KFAM, Löwik MRH, Kok FJ, et al. Diet and other lifestyle factors in high and low socio-economic groups (Dutch Nutrition Surveillance System). *European Journal of Clinical Nutrition* 1991; **45**: 441-450.

100. Smith AM, Baghurst KI. Public health implications of dietary differences between social status and occupational category groups. *Journal of Epidemiology and Community Health* 1992; **46**: 409-416.

101. Murphy SP, Rose D, Hudes M, Viteri FE. Demographic and economic factors associated with dietary quality for adults in the 1987-88 Nationwide Food Consumption Survey. *Journal of the American Dietetic Association* 1992; **92**: 1352-1357.

102. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies and guidelines. *Annual Review of Public Health* 1997; **18**: 341-378.

103. Wandel M. Dietary intake of fruits and vegetables in Norway: influence of life phase and socio-economic factors. *International Journal of Food Sciences and Nutrition* 1995; **46**: 291-301.

104. Gillman MW, Pinto BM, Tennstedt S, Glanz K, Marcus B, Friedman RH. Relationships of physical activity with dietary behaviors among adults. *Preventive Medicine* 2001; **32**: 295-301.

105. Dallongeville J, Marécaux N, Fruchart J-C, Amouye P. Cigarette smoking is associated with unhealthy patterns of nutrient intake: a meta-analysis. *Journal of Nutrition* 1998; **128**: 1450-1457.

106. Palaniappan U, Jacobs Starkey L, O'Loughlin J, Gray-Donald K. Fruit and vegetable consumption is lower and saturated fat intake is higher among Canadians reporting smoking. *Journal of Nutrition* 2001; **131**: 1952-1958.

107. Lissner L, Heitmann BL. Dietary fat and obesity: evidence from epidemiology. *European Journal of Clinical Nutrition* 1995; **49**: 79-90.

108. Togo P, Osler M, Heitmann BL. Food intake patterns and body mass index in observational studies. *International Journal of Obesity* 2001; **25**: 1741-1751.

109. Smithells RW, Ankers C, Carver M, Lennon D, Schorah CJ, Sheppard S. Maternal nutrition in early pregnancy. *British Journal of Nutrition* 1977; **38**: 497-506.

110. Haste FM, Brooke OG, Anderson HR, Bland JM, Peacock JL. Social determinants of nutrient intake in smokers and non-smokers during pregnancy. *Journal of Epidemiology and Community Health* 1990; **44**: 205-209.

111. Haste FM, Brooke OG, Anderson HR, et al. Nutrient intakes during pregnancy: observations on the influence of smoking and social class. *American Journal of Clinical Nutrition* 1990; **51**: 29-36.

112. Schofield C, Stewart J, Wheeler E. The diets of pregnant and post-pregnant women in different social groups in London and Edinburgh: calcium, iron, retinol, ascorbic acid and folic acid. *British Journal of Nutrition* 1989; **62**: 363-377.

113. Schofield C, Wheeler E, Stewart J. The diets of pregnant and post-pregnant women in different social groups in London and Edinburgh: energy, protein, fat and fibre. *British Journal of Nutrition* 1987; **58**: 369-381.

114. Mathews F, Yudkin P, Smith RF, Neil A. Nutrient intakes during pregnancy: the influence of smoking status and age. *Journal of Epidemiology and Community Health* 2000; **54**: 17-23.

115. Trygg K, Lund-Larsen K, Sandstad B, Hoffmann HJ, Jacobsen G, Bakketeig LS. Do pregnant smokers eat differently from pregnant non-smokers? *Paediatric and Perinatal Epidemiology* 1995; **9**: 307-319.

116. Rogers I, Emmett P, Baker D, Golding J. Financial difficulties, smoking habits, composition of diet and birthweight in a population of pregnant women in the South West of England. *European Journal of Clinical Nutrition* 1998; **52**: 251-260.

117. Pollard J, Kirk SFL, Cade JE. Factors affecting food choice in relation to fruit and vegetable intake: a review. *Nutrition Research Reviews* 2002; **15**: 373-387.

118. Brug J, Debie S, van Assema P, Weijts W. Psychosocial determinants of fruit and vegetable consumption among adults: results of focus group interviews. *Food Quality and Preference* 1995; **6**: 99-107.

119. Leather S. Reaching the recommendations: what are the barriers? In: Sharp I, ed. At least five a day: strategies to increase vegetable and fruit consumption. London: The Stationery Office,1997: 35-44.

120. Lennernäs M, Fjellström C, Becker W, Schmitt A, Remaut de Winter AM, Kearney M. Influences on food choice perceived to be important by nationally-representative samples of adults in the European Union. *European Journal of Clinical Nutrition* 1997; **51**: Suppl 2, S8-S15.

121. Wardle J, Parmenter K, Waller J. Nutrition knowledge and food intake. *Appetite* 2000; **34**: 269-275.

122. Kearney JM, McElhone S. Perceived barriers in trying to eat healthier - results of a pan-EU consumer attitudinal survey. *British Journal of Nutrition* 1999; **81**: Suppl 2, S133-S137.

123. Bartley M, Sacker A, Firth D, Fitzpatrick R. Understanding social variation in cardiovascular risk factors in women and men: the advantage of theoretically based measures. *Social Science and Medicine* 1999; **49**: 831-845.

124. Godfrey KM, Barker DJP. Fetal programming and adult health. *Public Health Nutrition* 2001; **4**: 611-624.

125. Inskip H. The Southampton Women's Survey. *MIDIRS Midwifery Digest* 1999; **9**: 445-447.

126. Nelson M, Bingham SA. Assessment of food composition and nutrient intake. In: Margetts BM, Nelson M, eds. Design concepts in nutritional epidemiology. Oxford: Oxford University Press,1997: 123-169.

127. Robinson S, Godfrey K, Osmond C, Cox V, Barker D. Evaluation of a food frequency questionnaire used to assess nutrient intakes in pregnant women. *European Journal of Clinical Nutrition* 1996; **50**: 302-308.

128. Willett W. Nutritional epidemiology. New York: Oxford University Press 1998.

129. Margetts BM, Thompson RL. Validation of dietary intake estimation. *European Journal of Clinical Nutrition* 1995; **49**: 934

130. Margetts BM, Cade JE, Osmond C. Comparison of a food frequency questionnaire with a diet record. *International Journal of Epidemiology* 1989; **18**: 868-873.

131. Feskanich D, Rimm E, Giovannucci EL, et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *Journal of the American Dietetic Association* 1993; **93**: 790-796.

132. Calvert C, Cade J, Barrett JH, Woodhouse A. Using cross-check questions to address the problem of mis-reporting of specific food groups on Food Frequency Questionnaires. *European Journal of Clinical Nutrition* 1997; **51**: 708-712.

133. Heitmann BL, Lissner L. Dietary reporting by obese individuals - is it specific or non-specific. *British Medical Journal* 1995; **311**: 986-989.

134. Holland B, Welch AA, Unwin ID, Buss DH, Paul AA, Southgate DAT. McCance and Widdowson's The composition of foods. UK: Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food 1991.

135. Crawley H. Food portion sizes. London: HMSO 1988.

136. Holland B, Unwin ID, Buss DH. Cereals and cereal products. Third supplement to McCance and Widdowson's The composition of foods. UK: Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food 1988.

137. Holland B, Unwin ID, Buss DH. Milk products and eggs. Fourth supplement to McCance and Widdowson's The composition of foods. UK: Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food 1989.

138. Holland B, Unwin ID, Buss DH. Vegetables, herbs and spices. Fifth supplement to McCance and Widdowson's The composition of foods. UK: Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food 1991.

139. Jacques PF, Tucker KL. Are dietary patterns useful for understanding the role of diet in chronic disease? *American Journal of Clinical Nutrition* 2001; **73**: 1-2.

140. StataCorp. Stata statistical software: release 7.0. College Station, TX: Stata Corporation 2001.

141. Joliffe IT, Morgan BJT. Principal component analysis and exploratory factor analysis. *Statistical Methods in Medical Research* 1992; **1**: 69-95.

142. Kim J-O, Mueller CW. Introduction to factor analysis: what it is and how to do it. Beverly Hills: Sage Publications 1978.

143. Kim J-O, Mueller CW. Factor analysis: statistical methods and practical issues. Beverly Hills: Sage Publications 1978.

144. Willett WC. Diet and health: what should we eat? *Science* 1994; **264**: 532-537.

145. Margetts BM, Martinez JA, Saba A, Holm L, Kearney M. Definitions of 'healthy' eating: a pan-EU survey of consumer attitudes to food, nutrition and health. *European Journal of Clinical Nutrition* 1997; **51**: Suppl 2, S23-S29.

146. Hunt CJ, Nichols RN, Pryer JA. Who complied with national fruit and vegetable population goals? Findings from the dietary and nutritional survey of British adults. *European Journal of Public Health* 2000; **10**: 178-184.

147. McCann SE, Weiner J, Graham S, Freudenheim JL. Is principal component analysis necessary to characterise dietary behaviour in studies of diet and disease. *Public Health Nutrition* 2001; **4**: 903-908.

148. Rowntree BS. Poverty: a study of town life. London: Macmillan 1902.

149. Mack J, Lansley S. Poor Britain. London: George Allen & Unwin 1985.

150. Townsend P. Deprivation. *Journal of Social Policy* 1987; **16**: 125-146.

151. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the North. Beckenham: Croom Helm 1988.

152. Jarman B. Identification of underprivileged areas. *British Medical Journal* 1983; **286**: 1705-1709.

153. Jarman B. Underprivileged areas: validation and distribution of scores. *British Medical Journal* 1984; **289**: 1587-1592.

154. Carstairs V, Morris R. Deprivation and mortality: an alternative to social class? *Community Medicine* 1989; **11**: 210-219.

155. Carstairs V. Multiple deprivation and health state. *Community Medicine* 1981; **3**: 4-13.

156. Carstairs V, Morris R. Deprivation: explaining differences in mortality between Scotland and England and Wales. *British Medical Journal* 1989; **299**: 886-889.

157. Department of Environment Transport and the Regions. Indices of deprivation. 2000

158. Morris R, Carstairs V. Which deprivation? A comparison of selected deprivation indexes. *Journal of Public Health Medicine* 1991; **13**: 318-326.

159. Jarman B, Townsend P, Carstairs V. Deprivation indices. *British Medical Journal* 1991; **303**: 523

160. Martin D, Brigham P, Roderick P, Barnett S, Diamond I. The (mis) representation of rural deprivation. *Environment and Planning A* 2000; **32**: 735-751.

161. Arber S. Gender and class inequalities in health: understanding the differentials. In: Fox J, ed. Health inequalities in European countries. Aldershot: Gower,1989: 250-279.

162. Macintyre S, Hiscock R, Kearns A, Ellaway A. Housing tenure and health inequalities: a three dimensional perspective on people, homes and neighbourhoods. In: Graham H, ed. Understanding health inequalities. Buckingham: Open University Press,2000: 129-142.

163. Filakti H, Fox J. Differences in mortality by housing tenure and by car access from the OPCS Longitudinal Study. *Population Trends* 1995; **81**: 27-30.

164. OPCS. Standard occupational classification. London: HMSO 1990.

165. Rose D, O'Reilly K. The ESRC review of government social classifications. Swindon: Economic and Social Research Council and Office for National Statistics 1998.

166. Prandy K. The revised Cambridge scale of occupations. *Sociology* 1990; **24**: 629-655.

167. Bartley M, Sacker A, Firth D, Fitzpatrick R. Social position, social roles and women's health in England: changing relationships 1984-1993. *Social Science and Medicine* 1999; **48**: 99-115.

168. Rose D, O'Reilly K. Constructing classes. In: Rose D, O'Reilly K, eds. Constructing classes: towards a new social classification for the UK. Swindon: Economic and Social Research Council and Office for National Statistics,1997: 1-19.

169. Jones IG, Cameron D. Social class analysis - an embarrassment to epidemiology. *Community Medicine* 1984; **6**: 37-46.

170. Macintyre S, Hunt K. Socio-economic position, gender and health: how do they interact? *Journal of Health Psychology* 1997; **2**: 315-334.

171. Stewart A, Prandy K, Blackburn RM. Social stratification and occupations. London: Macmillan Press 1980.

172. Prandy K. Cambridge scale scores for CASOC groupings. *Working Paper No.11, Sociological Research Group, University of Cambridge* 1992;

173. Prandy K. Class, stratification and inequalities in health: a comparison of the registrar-general's social classes and the Cambridge scale. *Sociology of Health and Illness* 1999; **21**: 466-484.

174. Arber S. Revealing women's health: re-analysing the General Household Survey. In: Roberts H, ed. Women's Health Counts. London: Routledge,1990: 63-92.

175. Arber S. Comparing inequalities in women's and men's health: Britain in the 1990s. *Social Science and Medicine* 1997; **44**: 773-787.

176. Goldthorpe JH. Women and class analysis: in defence of the conventional view. *Sociology* 1983; **17**: 465-488.

177. Joshi H, Davies H, Land H. The tale of Mrs Typical. London: Family Policy Studies Centre 1996.

178. Dex S. Women's work histories: an analysis of the women and employment survey. Department of Employment 1984.

179. Stanworth M. Women and class analysis: a reply to John Goldthorpe. *Sociology* 1984; **18**: 159-170.

180. Heath A, Britten N. Women's jobs do make a difference: a reply to Goldthorpe. *Sociology* 1984; **18**: 475-490.

181. Pugh H, Moser K. Measuring women's mortality differences. In: Roberts H, ed. Women's Health Counts. London: Routledge,1990: 93-112.

182. Erikson R. Social class of men, women and families. *Sociology* 1984; **18**: 500-514.

183. Krieger N, Chen JT, Selby JV. Comparing individual-based and household-based measures of social class to assess class inequalities in women's health: a methodological study of 684 US women. *Journal of Epidemiology and Community Health* 1999; **53**: 612-623.

184. Moser KA, Pugh HS, Goldblatt PO. Inequalities in women's health: looking at mortality differentials using an alternative approach. *British Medical Journal* 1988; **296**: 1221-1224.

185. Martikainen P. Socioeconomic mortality differentials in men and women according to own and spouse's characteristics in Finland. *Sociology of Health and Illness* 1995; **17**: 353-375.

186. Vågerö D, Lundberg O. Socio-economic differentials among adults in Sweden. In: Lopez AD, Caselli G, Valkonen T, eds. Adult mortality in developed countries. Oxford: Clarendon Press,1995: 223-242.

187. Dahl E. Inequalities in health and the class position of women - the Norwegian experience. *Sociology of Health and Illness* 1991; **13**: 492-505.

188. Arber S, Ginn J. Gender and inequalities in health in later life. *Social Science and Medicine* 1993; **36**: 33-46.

189. Arber S. Opening the 'Black box': inequalities in women's health. In: Abbott P, Payne G, eds. New directions in the sociology of health. London: The Falmer Press,1990: 37-56.

190. Krieger N. Women and social class: a methodological study comparing individual, household, and census measures as predictors of black/white differences in reproductive history. *Journal of Epidemiology and Community Health* 1991; **45**: 35-42.

191. Chandola T. Social inequality in coronary heart disease: a comparison of occupational classifications. *Social Science and Medicine* 1998; **47**: 525-533.

192. Sacker A, Firth D, Fitzpatrick R, Lynch K, Bartley M. Comparing health inequality in men and women: prospective study of mortality 1986-96. *British Medical Journal* 2000; **320**: 1303-1307.

193. Bartley M, Sacker A, Firth D, Fitzpatrick R. Dimensions of inequality and the health of women. In: Graham H, ed. Understanding health inequalities. Buckingham: Open University Press,2000: 58-74.

194. Jary D, Jary J. Collins dictionary of sociology. London: HarperCollins 1991.

195. Macintyre S, Anderson A. Socio-demographic and psycho-social variables. In: Margetts BM, Nelson M, eds. Design concepts in nutritional epidemiology. Oxford: Oxford University Press,1997: 273-288.

196. ONS. Living in Britain: results from the 1998 General Household Survey. London: The Stationery Office 2000.

197. ONS. Social trends 29. London: The Stationery Office 1999.

198. Margetts BM, Jackson AA. Interactions between people's diet and their smoking habits: the Dietary and Nutritional Survey of British Adults. *British Medical Journal* 1993; **307**: 1381-1384.

199. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. *Canadian Journal of Applied Sport Sciences* 1985; **10**: 141-146.

200. Garn SM, Leonard WR, Hawthorne VM. Three limitations of the body mass index. *American Journal of Clinical Nutrition* 1986; **44**: 996-997.

201. OPCS. 1992 General Household Survey. London: HMSO 1994.

202. ONS. Living in Britain: results from the 1996 General Household Survey. London: The Stationery Office 1998.

203. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1**: 307-310.

204. Hu FB, Willett WC. Validation of dietary patterns assessed with a food-frequency questionnaire - reply to M Tseng. *American Journal of Clinical Nutrition* 1999; **70** : 422

205. Weigel MM, Weigel RM. Nausea and vomiting of early pregnancy and pregnancy outcome. An epidemiological study. *British Journal of Obstetrics and Gynaecology* 1989; **96**: 1304-1311.

206. Prentice AM, Spaaij CJK, Goldberg GR, et al. Energy requirements of pregnant and lactating women. *European Journal of Clinical Nutrition* 2002; **50**: Suppl1, S82-S111.

207. Armitage P, Berry G. Statistical methods in medical research. Oxford: Blackwell Science Ltd 2002.

208. Berrington A. Transition to adulthood in Britain. In: Corijn M, Klijzing E, eds. Transition to adulthood in Europe. Dordrecht: Kluwer Academic Publishers,2001: 67-102.

209. O'Brien B, Zhou Q. Variables related to nausea and vomiting during pregnancy. *Birth* 1995; **22**: 93-100.

210. Petrie A, Sabin C. Medical statistics at a glance. Oxford: Blackwell 2000.

211. Prevost AT, Whichelow MJ, Cox BD. Longitudinal dietary changes between 1984-5 and 1991-2 in British adults: associations with socio-demographic, lifestyle and health factors. *British Journal of Nutrition* 1997; **78**: 873-888.

212. Department of Health. Committee on Medical Aspects of Food Policy. Nutritional aspects of the development of cancer. London: The Stationery Office 1998.

213. Henderson L, Gregory J, Swan G. The National Diet and Nutrition Survey: adults aged 19 to 64 years. London: HMSO 2002.

214. Tseng M. Validation of dietary patterns assessed with a food-frequency questionnaire. *American Journal of Clinical Nutrition* 1999; **70**: 422

215. Thompson RL, Margetts BM, Speller VM, McVey D. The Health Education Authority's health and lifestyle survey 1993: who are the low fruit and vegetable consumers? *Journal of Epidemiology and Community Health* 1999; **53**: 294-299.

216. Leclerc A, Pietri F, Boitel L, Chastang J-F, Carval P, Blondet M. Level of education, lifestyle, and morbidity in two groups of white collar workers. *Journal of Epidemiology and Community Health* 2003; **46**: 403-408.

217. Marmot MG, Davey Smith G, Stansfeld S, et al. Health inequalities among British civil servants: the Whitehall II study. *Lancet* 1991; **337**: 1387-1393.

218. Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. *Epidemiologic Reviews* 1988; **10**: 87-121.

219. Galobardes B, Morabia A, Bernstein MS. Diet and socioeconomic position: does the use of different indicators matter? *International Journal of Epidemiology* 2001; **30**: 334-340.

220. Fuchs VR. Economics, health and post-industrial society. *Milbank Memorial Fund Quarterly / Health and Society* 1979; **57**: 153-182.

221. Østbye J, Pomerleau J, White M, Coolich M, McWhinney J. Food and nutrition in Canadian 'prime time' television commercials. *Canadian Journal of Public Health* 1993; **84**: 370-374.

222. Tierson FD, Olsen CL, Hook EB. Influence of cravings and aversions on diet in pregnancy. *Ecology of Food and Nutrition* 1985; **17**: 117-129.

223. Perkins KA. Effects of tobacco smoking on caloric intake. *British Journal of Addiction* 1992; **87**: 193-205.

224. Inskip HM, Borland SE, Bayly RA, Hammond JF, Robinson SM, Godfrey KM. Influences on red cell folate concentrations in non-pregnant women aged 20-34 years in Southampton, UK. *Pediatric Research* **53**: 43A, P713.2003

225. Braddon FEM, Wadsworth MEJ, Davies JMC, Cripps HA. Social and regional differences in food and alcohol consumption and their measurement in a national birth cohort. *Journal of Epidemiology and Community Health* 1988; **42**: 341-349.

226. Martínez ME, Marshall JR, Sechrest L. Invited commentary: factor analysis and the search for objectivity. *American Journal of Epidemiology* 1998; **148**: 17-19.

227. Slattery ML, Boucher KM. The senior author's response: factor analysis as a tool for evaluating eating patterns. *American Journal of Epidemiology* 1998; **148**: 20-21.

228. Dowler E. Budgeting for food on a low income in the UK: the case of lone-parent families. *Food Policy* 1997; **22**: 405-417.

229. Lawrence JM, Thompson RL, Margetts BM. Young women's confidence about cooking skills in relation to food choice. *Proceedings of the Nutrition Society* 2000; **59**: Suppl, 34A.

230. Drenowski A. Taste preferences and food intake. *Annual Review of Nutrition* 1997; **17**: 237-253.

231. Darmon N, Ferguson E, Briend A. A cost constraint alone has adverse effects on food selection and nutrient density: an analysis of human diets by linear programming. *Journal of Nutrition* 2002; **132**: 3764-3771.

232. Cade J, Upmeier H, Calvert C, Greenwood D. Costs of a healthy diet: analysis from the UK Women's Cohort Study. *Public Health Nutrition* 1999; **2**: 505-512.

233. Anderson A. Pregnancy as a time for dietary change? *Proceedings of the Nutrition Society* 2001; **60**: 497-504.

234. ONS. Population Trends 113. London: HMSO 2003.

235. Department of Health. Tackling health inequalities: a programme for action. London: Department of Health Publications 2003.

236. Lindström M, Hanson BS, Wirfält E, Östergren P-O. Socioeconomic differences in the consumption of vegetables, fruit and fruit juices. *European Journal of Public Health* 2001; **11**: 51-59.

237. Murcott A. Nutrition and inequalities: a note on sociological approaches. *European Journal of Public Health* 2002; **12**: 203-207.

238. Whelan A, Wrigley N, Warm D, Cannings E. Life in a 'food desert'. *Urban Studies* 2002; **39**: 2083-2100.

239. Steptoe A, Wardle J. Motivational factors as mediators of socioeconomic variations in dietary intake patterns. *Psychology and Health* 1999; **14**: 391-402.

240. Steptoe A, Perkins-Porras L, McKay C, Rink E, Hilton S, Cappuccio FP. Behavioural counselling to increase consumption of fruit and vegetables in low income adults: randomised trial. *British Medical Journal* 2003; **326**: 855-860.

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

242. Schulze MB, Hoffman K, Kroke A, Boeing H. An approach to construct simplified measures of dietary patterns from exploratory factor analysis. *British Journal of Nutrition* 2003; **89**: 409-418.