General Enquiries on the form should be made to:
Defra, Procurements and Commercial Function (Evidence Procurement Team)
E-mail: research.competitions@defra.gsi.gov.uk

defra
Department for Environment Food and Rural Affairs

Evidence Project Final Report

- **Note**
  In line with the Freedom of Information Act 2000, Defra aims to place the results of its completed research projects in the public domain wherever possible.
  The Evidence Project Final Report is designed to capture the information on the results and outputs of Defra-funded research in a format that is easily publishable through the Defra website
  An Evidence Project Final Report must be completed for all projects.
- This form is in Word format and the boxes may be expanded, as appropriate.
- **ACCESS TO INFORMATION**
  The information collected on this form will be stored electronically and may be sent to any part of Defra, or to individual researchers or organisations outside Defra for the purposes of reviewing the project. Defra may also disclose the information to any outside organisation acting as an agent authorised by Defra to process final research reports on its behalf. Defra intends to publish this form on its website, unless there are strong reasons not to, which fully comply with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.
  Defra may be required to release information, including personal data and commercial information, on request under the Environmental Information Regulations or the Freedom of Information Act 2000. However, Defra will not permit any unwarranted breach of confidentiality or act in contravention of its obligations under the Data Protection Act 1998. Defra or its appointed agents may use the name, address or other details on your form to contact you in connection with occasional customer research aimed at improving the processes through which Defra works with its contractors.

Project identification

1. Defra Project code: FO0414 (Part 2)
2. Project title: Evaluation of the potential of using Family Food Survey data to assess dietary patterns
3. Contractor organisation(s): Warwick HRI
   Wellesbourne
   Warwick CV35 9EF
4. Total Defra project costs (agreed fixed price): £ 186,894.00
5. Project: start date 01/07/2009
   end date 31/07/2010

6. It is Defra's intention to publish this form.
Please confirm your agreement to do so.. **YES** x☐ **NO** ☐

(a) When preparing Evidence Project Final Reports contractors should bear in mind that Defra intends that they be made public. They should be written in a clear and concise manner and represent a full account of the research project which someone not closely associated with the project can follow.
Defra recognises that in a small minority of cases there may be information, such as intellectual property or commercially confidential data, used in or generated by the research project, which should not be disclosed. In these cases, such information should be detailed in a separate annex (not to be published) so that the Evidence Project Final Report can be placed in the public domain. Where it is impossible to complete the Final Report without including references to any sensitive or confidential data, the information should be included and section (b) completed. NB: only in exceptional circumstances will Defra expect contractors to give a "No" answer.
In all cases, reasons for withholding information must be fully in line with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.

(b) If you have answered NO, please explain why the Final report should not be released into public domain

Executive Summary

7. The executive summary must not exceed 2 sides in total of A4 and should be understandable to the intelligent non-scientist. It should cover the main objectives, methods and findings of the research, together with any other significant events and options for new work.

<u>Consumer insight into food prices and food security - Part 2</u>

The context to this research was the increased interest in food security and the rise in food prices in 2008, following decades of falling food costs and generally rising incomes, and the onset of the current economic recession. The project was separated into two distinct parts (Objectives 1-4 and Objectives 5-7). The aim of the first part of the project (Objectives 1-4) was to assess consumer understanding of, and reactions to, changing food prices and food security, and their expectations of Government. This first part of the project has been published as a separate report and has also been considered subsequently in two peer-reviewed publications (Dowler *et al*., 2011, *Critical Public Health*, 21(4), 403-416; Kneafsey *et al*., 2012, *Journal of Rural Studies*, 29, 101-112). The aim of this second part of the project (a feasibility study) was to undertake statistical analysis of the Defra Family Food Survey (FFS) data to explore whether it is possible to use these data to assess the cost of a healthy diet and associations between dietary patterns and socio-economic factors. The specific objectives were to:

5. Explore possibilities of using Defra Family Food Survey data (2002 to 2007) to assess the cost of a healthy diet.
6. Explore possibilities of using Defra Family Food Survey data (2002 to 2007) to examine patterns in the diets of those at risk of social exclusion.
7. Make comparisons between Defra Family Food Survey data (2002 to 2007) and data for 2008 to explore whether (*there is any evidence that the start of*) the economic downturn has affected food purchases and thereby the diet of different social groups.

These objectives were concerned with the investigation of whether (how) the available data could be used to address such issues, rather than with answering particular research or policy questions.

Objective 5

To address this objective it was first necessary to characterise healthier (and therefore less healthy) diets. The 'eatwell plate' (DH) characterises a healthy, balanced diet, but recognises that there are many combinations of food choices that fit within the overall model of a healthy diet that it presents. Advice from nutritionists cautioned against being prescriptive about what constitutes healthy and unhealthy diets, therefore statistical clustering approaches were applied to identify groups of households based on similar and commonly occurring dietary patterns. These dietary groups are labelled in terms of the dominant dietary components, but cannot easily be labelled as either healthy or unhealthy. One possible approach to measuring the 'healthiness' of a household diet was developed by calculating how much the distribution of dietary components differed from the 'eatwell plate' (DH) ideal, referred to as the "eatwell plate distance". This "eatwell plate distance" essentially combines the difference between the household percentage and the ideal percentage in each of the five 'eatwell plate' categories as a Euclidean distance. A larger distance indicated a greater difference from the ideal diet, and hence a potentially less healthy diet. Households were also classified into five dietary groups based on the values of this distance.

Households were classified into different cost categories, after first adjusting expenditure for household composition, and the associations between different dietary groups and different cost categories were assessed using a log-linear model analysis. This approach has the advantage of allowing consideration of the variation in, or distribution of, expenditure for different diet groups, rather than just considering differences in the means, as would be provided by a conventional analysis of variance.

The analysis approach was extended to include socio-economic classifications of the households to assess whether expenditure for different diets varied between different socio-economic groups. A combination of the size of the available dataset and statistical software memory constraints meant that a number of separate statistical models needed to be constructed to address a range of example questions about the impacts of different (groups of) socio-economic factors.

Analysis of the FFS data therefore allowed the identification of households with different dietary patterns (*assuming that purchase data are a good surrogate for consumption data*), **and hence an assessment of the relative costs of these different dietary patterns. However, it is important to note that these groups were defined based on statistical rather than nutritional arguments, illustrating the potential of the statistical approach, so that further nutritional expertise is needed to interpret the health benefits of different diets and therefore to compare the costs of healthy and less healthy diets. It is also important to note that the project was focussed on the development of statistical approaches to address these questions using the FFS data, and not on drawing any firm conclusions based on any presented analyses (which are only presented to illustrate the potential of the developed approaches).**

Objective 6

To compare patterns in the diets of different socio-economic groups, and therefore to identify dietary patterns potentially associated with groups of households at risk of social exclusion, appropriate adjustment of nutrient and food component intakes was first needed, based on household composition. Household intakes were compared with the household requirements, as specified by published population dietary reference values (DRVs) for different age/gender combinations, so that each household could be classified relative to the household DRVs for each nutrient and food component. Application of the same log-linear model analysis approach allowed assessment of the associations between different socio-economic groups and different household adjusted intake categories for each nutrient and food component. This approach again allowed consideration of the variation in intake distributions for different socio-economic groups, rather than just considering differences in means, as would be provided by a conventional analysis of variance.

Using the same diet groups defined for Objective 5, the analysis approach was further extended to assess for associations between different diet groups and different socio-economic groups, potentially identifying how different dietary patterns might be more common among particular socio-economic groups.

A potential flaw with this approach relates to the use of population-based DRVs to adjust for household composition differences. The DRVs define the distribution of intakes required by individuals within the population, not the distribution of intakes that might be appropriate for each individual. However, these differences should cancel out across the large number of households included in the study, so that changes in the distribution of intakes between socio-economic groups should reflect real differences. Other limitations of the study include the validity of using the collected socio-economic data to identify households at potential risk of social exclusion, and the reliability of using the FFS data to represent a regular household food component and nutrient intake (there are potential issues with over- and under-estimates of bulk-bought products).

The study identified the potential for analysis of the FFS data to be used to examine variation in dietary patterns associated with different socio-economic factors. However, the size of the data set again imposed constraints on the complexity of the statistical models of socio-economic factors that could be considered, limiting the strength of any conclusions that can be drawn from the data. Careful identification of appropriate statistical models and socio-economic factors to address particular questions is therefore required, together with assessment of the partial confounding between the different socio-economic factors included in the data set, and careful interpretation of the trends indicated for statistical models with partially confounded socio-economic factors. The methods were developed and illustrated through application to eight case studies using hypothetical statistical models.

Objective 7

Almost all of the models considered in the log-linear model analyses of associations between socio-economic factors and nutrients/food components/diets in Objective 6 included financial year as an explanatory factor. Assessment of any trends in purchase (intake) patterns over time was therefore possible within these existing analyses. For most nutrient/food component/dietary group classifications analysed, there was evidence for variation in distributions between financial years. For some variables these differences appeared to depend on other socio-economic factors. Changes in the distributions of households across categories between years were therefore explored in more detail for the classifications with the most variation. The distribution for 2008/9 was often different from that for the previous financial year. However, this was either part of a gradual continuing trend across the whole of the time period, or the difference between these years was no larger than other differences between successive financial years. As the dataset was only for the period from 2001 to 2008/9, it was not possible to assess the impact of the economic downturn starting in 2008/9. Data for later time periods would be needed in addressing this question.

The same limitations identified for Objectives 5 and 6 are, of course, relevant for these analyses, with concerns about the impact of bulk buying potentially of more importance when considering the impact of the economic downturn. Certainly statistical analysis of further data on the impact of the economic downturn on shopping patterns is needed in determining whether the 2-week data collection period for each household provides a reliable representation of food purchase patterns. In this study, it is also possible that some important interactions between financial year and other socio-economic factors might have been missed because of the limited socio-economic data collected in the FFS. Finally, an improved identification of the health benefits of different diets, based on more detailed nutritional guidance, would potentially provide stronger evidence for the relative impact of the economic downturn on dietary patterns for different social groups.

Overall

This feasibility study has demonstrated the potential for analysis of the FFS data to be used to assess differences in dietary patterns and the costs of different diets for different socio-economic groups. However, this study has also highlighted a number of possible limitations of the FFS data:

- *whether the nutrient/food component data based on purchases reflects consumption,*
- *whether the data for a two-week period is sufficiently representative of longer-term consumption,* and
- *whether the socio-economic data collected in the FFS can provide appropriate classifications to address questions of interest about the diets of different socio-economic groups*.

Future application of the developed analysis methods is also highly dependent on accepting the assumption that household data can be related to the combined household requirements based on the sum of individual DRVs, and may also be limited by statistical software constraints. Effective application of the analysis approach depends on the clear identification of specific questions to be addressed using the FFS data, with the possible requirement for additional socio-economic data to be collected within the FFS. While development of the analysis approaches was guided by nutritional expertise, future effective application of the approaches to answer specific questions will require further such guidance, particularly with regards to the identification of different diets and the associated health benefits. Nutritional expertise will also be essential in interpreting the results of future analyses using the developed approaches.

The developed analysis approaches can be used to indicate statistically significant associations between household characteristics, food and nutrient intakes, diets and expenditure, but, because of the way in which the FFS data were collected, the identification of causality will almost certainly require the collection of more detailed data focussed on particular social groupings. Results are presented to illustrate the potential and versatility of the statistical approaches, constructing statistical models to consider hypothetical examples of different socio-economic groupings, and a better-informed construction of socio-economic groupings is needed before any firm conclusions can be drawn.

Key findings (objectives 5-7):

- This project has demonstrated the feasibility of analysing the Family Food Survey data to identify possible associations between diet and intake (purchase) levels of a range of nutrients and food components, socio-economic factors, and expenditure, based on household aggregated values and considering changes in the distributions of numbers of households between different categories.
- The ‘distance’ from the ‘eatwell plate’ “ideal” provides a potentially valuable approach as an indicator of a healthy diet, but needs further refining to better incorporate the nutritional importance of different dietary components, and possibly extension to understand and incorporate additional aspects of how household diets differ from the ‘eatwell plate’ “ideal”.
- However, there are a number of limitations associated with the data used in the study. These include that the Family Food Survey data record the purchases for a household over a two-week period, so that analyses are based on assumptions about these data being representative of

longer periods, and also reflecting the food intake of the whole household.

- A further limitation of the approach is the assumption that household food component and nutrient intake (purchase) data can be related to the aggregated requirements of the individuals within the household, based on the population-based thresholds of requirements (which relate to the requirements of particular proportions of the population).
- Applying the developed log-linear analysis approaches to the Family Food Survey data will be most effective if focussed on addressing specific questions (for well-defined socio-economic or dietary groups of households, related to particular dietary components). The hypothetical case studies considered in this project were only used to illustrate the potential of the analysis approach, with identified socio-economic groupings based on statistical features of the data. The potential value of the approach is therefore in the application to the FFS data to address clearly defined research questions.
- Practical quantitative characterisations of healthier and less healthy diets need to be developed that allow the identification of parameters that can be used to categorise households based on a quantitative assessment of household intakes (purchases) to allow interpretation of associations between different diets and socio-economic factors that might be identified using this analysis method.
- Hierarchical cluster analysis of the households based on dietary intake patterns, to identify common dietary groups, has the potential for further development, and hence improvement of the groups formed, by incorporating further nutritional knowledge and information.
- Annex 2 to this report provides a more detailed exploration and illustration of how the statistical analysis methods developed in this project could be used to analyse the FFS data to address questions about associations between socio-economic classifications and aspects of diet. The project has demonstrated the potential value of the developed statistical analysis approaches, but further research is needed to identify appropriate socio-economic classifications of households and nutritionally meaningful classifications of dietary components before these statistical analysis approaches can be used to draw any firm conclusions.

Project Report to Defra

8. As a guide this report should be no longer than 20 sides of A4. This report is to provide Defra with details of the outputs of the research project for internal purposes; to meet the terms of the contract; and to allow Defra to publish details of the outputs to meet Environmental Information Regulation or Freedom of Information obligations. This short report to Defra does not preclude contractors from also seeking to publish a full, formal scientific report/paper in an appropriate scientific or other journal/publication. Indeed, Defra actively encourages such publications as part of the contract terms. The report to Defra should include:
   - the objectives as set out in the contract;
   - the extent to which the objectives set out in the contract have been met;
   - details of methods used and the results obtained, including statistical analysis (if appropriate);
   - a discussion of the results and their reliability;
   - the main implications of the findings;
   - possible future work; and
   - any action resulting from the research (e.g. IP, Knowledge Exchange).

Background to the project (this report considers Part 2 of a two-part project)

Food security can be defined as all consumers having access at all times to sufficient, safe and nutritious food for an active and healthy life at affordable prices; to enable this, food supply must be reliable and resilient to shocks and crises, and food has to be produced in sustainable ways. The current concerns about UK and international food security stem from security of key inputs such as energy and water, potential impact of global climate change and volatility in food and fuel price rises. One key mechanism by which potential insecurity is indicated is through price: UK food price inflation peaked at 14.5% in August 2008 (DFID/Defra 2010), with variation for specific commodities, despite efforts by the retail sector to absorb increases in inputs and other costs. In the UK this followed several decades of falling food costs, generally rising incomes and thus a decreasing proportion of income spent on food even by the lowest income quintile. Such increases in food prices, coinciding with increases in oil and transport costs, financial and job insecurities, are likely to have differential effects on households in different circumstances. In addition, there is some evidence of project participant concern over the reliability of the food system because of environmental threats and emergent instabilities. The research (commissioned in 2009) , focused on project participant level in terms of perceptions and behaviours in relation to key components of food security at household level: *access* (where and how people obtain food), *affordability* (experience of, and practices in relation to, changes in food prices) and *availability* (what is bought), as well project participants' sense of food security (confidence in the food system; intended and actual changes in practice in the relatively short term.) The overall aim of Part 1 of the project (Objectives 1-4) was to assess consumer understanding of, and reactions to, changing food prices and food security, and consumer expectations of Government, and to gather evidence to inform the development of potential household food security indicators. Part 1 of the project has been published as a separate report and has also been considered subsequently in two peer-reviewed publications. Statistical analysis of the Defra Family Food Survey (FFS) data was undertaken in Part 2 of the project as a feasibility study (Objectives 5-7 reported here) to assess whether analysis of these data could be used in the future to assess the cost of a healthy diet, and associations between these costs, socio-economic characteristics of households, and household dietary patterns.

Objectives

The project objectives of this feasibility study, covered in this report, are as follows:

5) Explore possibilities of using Defra Family Food survey data (2002 to 2007) to assess the cost of a healthy diet.
6) Explore possibilities of using Defra Family Food survey data (2002 to 2007) to examine patterns in the diets of those at risk of social exclusion.
7) Make comparisons between Defra Family Food survey data (2002 to 2007) and data for 2008 to explore whether (*there is any evidence that the start of*) the economic downturn has affected food purchases and thereby the diet of different social groups.

The project was focussed on the development of statistical methods to address these objectives, using a series of case studies to illustrate how the statistical methods could be applied to the Defra Family Food Survey data. The research is described in much greater detail in Annex 2.

Objectives 5-7 were focussed on assessing whether the Defra Family Food Survey (FFS) data could be used to assess the cost of a healthy diet, to examine patterns in the diets of those at risk of social exclusion (by comparing different socio-economic groups), and explore whether the economic downturn in 2008 has affected food purchases and thereby the diets of different social groups. These, or similar data, had not been analysed to address such questions previously, so that this part of the project formed the basis of exploratory research to assess the potential to apply different statistical approaches to such data to address these questions. The output from this part of the project is therefore an assessment of the potential to apply different statistical analysis methods to these data to address these questions, and to illustrate (using a series of case studies) how the results of these analyses could be interpreted with regards to nutritional aspects and socio-economic characteristics. This report also highlights the limitations and constraints associated with both the FFS data and the applied statistical methods, identifying areas where further research, method development and analysis would be beneficial.

To put the research into context, it is useful to identify some of the key limitations of the study before describing the approaches in any detail. The FFS data are household level purchase data, so that intakes of different nutrient and food components for individuals within the household, and household dietary patterns, have been inferred from these data. Comparisons between different dietary (purchase) patterns within the study potentially allow the description of one pattern as being healthier than another, but it is important to realise that these are relative statements based on the comparisons and analytical approaches developed within this study, and do not provide information about the absolute healthiness of a diet (purchase pattern).

Methodology

The Defra Family Food Survey data considered in this project consist of records of household food purchases and expenditure over a two-week period, and associated household characteristics (demographics), from nearly 52,000 household cases collected in annual, cross-sectional surveys over a period from the start of the 2001/2 financial year to the end of the 2008 calendar year. The food purchase and expenditure data contained information on quantities and of purchased foods in 260 defined product categories and total expenditure for the household, with estimated food component and nutrient intake values calculated for each purchase for each of 47 food components and nutrients. In total, across the nearly 52,000 household cases, this comprised over 42,000,000 individual data values. These data were summarised to provide a total intake for each of the 47 food components and nutrients for each of the household cases.

Key to the approach is the need to find a mechanism to establish indicators to use as potential markers of healthier and less healthy diets. Traditionally, the proportions of individuals, within any group, meeting nutrient Dietary Reference Standards are used; in particular, those with nutrient intakes below the Lower Reference Nutrient Intake (LRNI) are more likely, if this continues over time, to be consuming insufficient of that nutrient. However, apart from the LRNI for certain nutrients, Dietary Reference Values (DRVs) do not easily lend themselves to categorising whole diets, since UK diets typically contain many nutrients from a very wide range of sources, and few in the population are likely to be below the LRNI for many nutrients. Therefore, an alternative approach was explored, drawing on the estimated distance from the 'eatwell plate' (DH) "ideal" which indicates the proportions of the diet which should come from each of five broad categorisations of foods (Fruit and Vegetables; Bread, Cereals and Potatoes; Meat, Fish and Alternatives; Fats and Sugars; Milk and Dairy) as shown in Figure 1. (The 'eatwell plate' is based on research which links dietary patterns to health outcomes, as well as research with ordinary people on how such visual representations are perceived.)

Figure 1 Illustration of the five 'eatwell plate' categories. The "ideal" diet is suggested to comprise 33% Fruit and vegetables, 33% Bread, rice, potatoes and pasta (Bread, Cereals and Potatoes), 15% Milk and dairy foods (Milk and Dairy), 12% Meat, fish, eggs and beans (Meat, Fish and Alternatives) and 8% Foods and drinks high in fat and/or sugar (Fats and Sugars). Note that the total of the percentages for the five categories is actually 101%. Note also that the image used is current at the time of publication; at the time of analysis (2009-10), the image came from the Food Standards Agency.)

Thus, for the purposes of exploring the construction and use of such an indicator, 230 of the 260 product categories (most of the 30 excluded product categories related to drinks) were allocated to one of the five 'eatwell plate' categories, enabling the calculation of the percentage of food purchases (by weight) and total financial expenditure for each of these five categories, and the subsequent calculation of a potential indicator of a healthy diet consumption in terms of the distance (difference) of this set of percentages for each household from the eatwell plate "ideal". One advantage of this approach is that differences in the absolute quantity per head of each food or drink item purchased by different households are removed (hence removing issues associated with the size and composition of different households – because the Family Food Survey data are recorded at household level, inferences about individual intakes can only be made based on division of the total intake by the number of household members), with comparisons made purely on the basis of the distribution of food purchases across the five categories – for this approach, it is the distribution that matters, and the level of departure from the "ideal", rather than the absolute amounts. There are two potential disadvantages however: First, that the five 'eatwell plate' category percentages are not independent variables (the sum of the five percentages must always be 100%, so that an increase in the percentage for one category must result in a decrease in the percentage for one or more of the other categories). While this constraint can be incorporated into statistical analyses of the combined data, or into a summary measure such as the distance from the ideal, care must be taken in interpreting the results of analyses of the percentages for single categories, as an observed increase may really be caused by a reduction in another category. The second disadvantage is that allocation of foods to one of each five categories is not always clear-cut, particularly for composite dishes. This issue demands further exploration, such as using a sensitivity analysis to consider the impact of different allocation rules, but there was insufficient time available in this project to address this.

A second approach to interpreting the potential of the diet to contribute to health is to compare individual intakes of nutrients and other dietary components with Dietary Reference Values (DRVs). DRVs are dietary guidelines for the UK population, summarising the distribution of daily requirements for the population for levels of energy, nutrients and other dietary components that are needed for good health, allowing for differences in requirements for different age and gender groups within the population. They are not intended as recommendations for individual intakes. The main challenge in using the Family Food Survey data for such an assessment is therefore in relating intake data, at a household scale, to requirement data, at an individual scale, with the added complication that DRVs are associated with the requirements of particular proportions of the population, rather than the distribution of requirements for any individual within that population. In our analysis we have calculated a series of 'household level DRVs', as set out below, for comparison with the household level intake of each nutrient and food component. In general, three DRVs are defined:

- the Estimated Average Requirement (EAR), defined as the median requirement within the population, the value such that (approximately) 50% of the population will require less and (approximately) 50% of the population will require more;
- the Reference Nutrient Intake (RNI), an upper threshold, defined to be such that only 2.5% of the population require more than this (in statistical terms, the 97.5th percentile), often calculated as the mean plus two standard deviations (based on an assumption that the requirement follows a Normal distribution across the population);
- the Lower Reference Nutrient Intake (LRNI), a lower threshold, defined to be such that only 2.5% of the population require less than this (in statistical term, the 2.5th percentile), often calculated as the mean minus two standard deviations (based on an assumption that the requirement follows a Normal distribution across the population).

The COMA (Committee on Medical Aspects of Food Policy) (1991) report generally provides values for each of these three DRVs, usually with different values by gender, often with different values for pregnant women, and usually with different values for different age groups (commonly, the required intakes change substantially during childhood years, and sometimes different values are given for adults who are older and younger than 50).

From discussions with other project team members, and DH and Defra officers, and from examination of the tabulated DRVs, it is clear that the assumption of Normality behind the DRV definitions above does not always hold, but we were unable to find strong evidence for other explicit distributional assumptions. Therefore, where a DRV was not given in the COMA report tables, the definitions above were used to infer the values of the missing DRV based on the assumption of Normality.

In order to simplify the analysis to meet the objectives (to demonstrate the feasibility of the approach), several key food components and nutrients were identified as being of initial interest, and the current UK Dietary Reference Values (DH, 1991) were used to construct the household level values – the values for EAR, RNI and LRNI (where they were defined) for the combined actual household compositions. Comparison of the observed household intakes with this series of calculated household DRVs then allows the classification of each household into one of a number of groups (e.g.

below the LRNI, between the LRNI and the EAR, between the EAR and the RNI, above the RNI). Alternatively the observed household intake can be expressed as a percentage of a specified household DRV (e.g. as a percentage of the household LRNI). Both the distribution of households across these classes and the percentage values can then be analyses to examine variability in the nutritional status of different socio-economic and demographic groups.

The requirements for a number of components (Total Fat, Saturated Fat, Mono-unsaturated Fat, Poly-unsaturated Fat, Starch, Non-milk extrinsic sugars [NMES]) are only expressed in terms of a percentage of the total dietary energy. As with the 'eatwell plate' categories, this expression as a percentage of some other intake provides a standardisation across different household compositions.

The COMA (1991) report does not provide a full set of DRVs for Fibre, but the report does indicate that requirements for children should be proportionately smaller based on body weight – an estimated value per day per person equivalent was therefore calculated to standardise across household compositions. More recent guidance (http://www.food.gov.uk/multimedia/pdfs/nutguideuk.pdf) suggests a value of 18g per day for adults and children over 5 years of age, and the analysis could be repeated incorporating this alternative standardisation.

The recommended intake of fruits and vegetables is for at least 400g per adult per day, which is translated in advice as 'at least 5 portions a day', where a portion is therefore assumed to be 80g. There is a suggestion that portion size for children should be smaller than this, but no clear guidance on the amount of reduction. An arbitrary portion size of 40g for children was used in this study, but, again, analyses could be repeated using alternative figures. Thus the intake of fruits and vegetables was converted based on these average portion weights to give a mean number of portions per person per day for the household.

The effect of using DRVs in this way is thus to introduce a notional standardisation in terms of household size and composition to dietary data collected at a household level. Such calculations are based on implicit assumptions about intra-household distributions (for instance, that all members have proportionate access to nutrients via household supply; that access is determined by physiological needs, etc). In reality, of course, the actual distribution of food, and thus nutrient intakes, within any given household is determined by a number of factors, in which physiological need plays only a part. These issues are extensively discussed elsewhere (e.g. Nelson & Bingham, 1997; Chesher, 1997; Vasdekis et al, 2001; Trichopoulou et al, 2003). Much of the literature discussing data to assess the healthiness or otherwise of diets is concerned with the status of individuals, either for epidemiological purposes, or for public health surveillance, and obtaining accurate records of food and nutrient intakes can be critically important for predicting the relationship between diet and health (or rather, ill-health) outcomes and/or assessing individual risk. In this instance, the purpose is to explore the possibilities of using regular household level survey data (FFS) to estimate the cost of a healthy diet, to examine dietary patterns in those at risk of social exclusion, and to monitor the impact of economic trends, particularly an increase in food prices. Thus, individual intakes and status are of less importance than the possibility of deriving indicators that can be used to characterise household food intakes. Thus, using DRVs as a tool to derive nutrient intakes 'equivalised' by physiological/health need seemed a reasonable procedure.

To develop the statistical approaches, and illustrate the potential of these approaches when applied to the FFS data, a set of eight separate statistical models were constructed to incorporate the different socio-economic and demographic characteristics recorded in the FFS data. These different statistical models generated a series of hypothetical groupings of the households, thereby allowing the assessment of the effects of the different household characteristics on dietary patterns through the comparison of patterns between different groups. Each of these statistical models were defined to have sufficient frequency of observations in each combined category, even when including both (financial) year and season in the model, to provide reliable assessment of differences in dietary patterns between categories. The hope was that these different statistical models might allow the identification of particular groups of "potentially at risk" households, therefore providing a potential approach to address Objective 6 by comparing the dietary responses for different defined groups. Of course, the different explanatory(socio-economic and demographic) variables may be partially correlated or confounded, and using a number of models, each with different household characteristics, may increase the likelihood of confusing or misleading results, since the explanatory variables included in different models may be associated with the same variability in a particular response. The extent of any such confounding can, to some extent, be assessed by consideration of two-way tables of the frequencies associated with particular pairs of explanatory variables from different models. Even with a dataset of this size (52,000 household records) it was not possible to consider the full complexity of a model involving all the household characteristics of interest (both because the number of possible combinations would exceed the number of observations and because of the computing power needed to fit such complex models).

Equivalised income was included in most of the models as it was considered likely both to have strong associations with many of the food component, nutrient and eatwell plate variables, and also to show some interaction effects with other household characteristics (such as geographical region and household composition). Season and financial year were included in all models which took more than one year into account, to explore the possibility of between-season differences in intakes, and to try to address Objective 7, of whether any changes in dietary patterns were detectable in relation to the economic downturn (i.e. comparing 2008 with earlier years). The fitted models allowed the broader assessment of differences between years from 2001 to 2008, with further model modification possible to assess the variation between 2008 and earlier years, compared with the variation among the seven earlier years. Each of these statistical models was combined with the food component and nutrient data, eatwell plate data, constructed diet data and expenditure data, using a range of statistical methodologies, to provide assessment of the impact of the different household characteristics (and combinations included within these models) on dietary patterns and expenditure. These analyses would then allow the comparison of dietary patterns and expenditure for different sub-populations, potentially identifying sub-populations with unusual dietary patterns.

The household socio-economic and demographic characteristics included in each of the eight case studies, and combined in the statistical models, are set out below, together with an illustrative, exploratory aim for applying each statistical model to the dietary response data:

- Case study (model) 1 – to explore associations between dietary patterns and both government region and equivalised household income (i.e. *do the data suggest there are geographical/income level groups with particular dietary intake patterns?*), allowing for variation over time (financial year, season)
    - 12 government regions, 5 equivalised income quintiles, 8 financial years, 2 seasons (Winter, Summer)
    - Government regions are the 12 administrative regions of the UK, identified in Table 2, used in the collection of the Family Food Survey data
- Case study (model) 2 – to explore associations between dietary patterns and both the minimum adult age and equivalised household income (i.e. *do the data suggest that old-age pensioners have different dietary patterns to other age groups, possibly also influenced by income?*), allowing for variation over time (financial year, season)
    - 4 ranges of minimum adult age (18-30, 31-50, 51-65, 66+), 5 equivalised income quintiles, 8 financial years, 2 seasons - initially considered 86+ age group separately, but frequencies too low to be reliable
- Case study (model) 3 – to explore associations between dietary patterns and both equivalised household income and employment status (i.e. *do the data suggest there are employment status/income groups with particular dietary intake patterns?*), allowing for variation over time (financial year, season)
    - 3 levels of economic status (Full-time employed, part-time employed, other), 5 equivalised income quintiles, 8 financial years, 2 seasons. Grouping of economic activity levels necessary to provide sufficiently large frequencies when combined with income and time classifiers
- Case study (model) 4 – to explore associations between dietary patterns, ethnic origin and equivalised household income (i.e. *do the data suggest that different ethnic groups have different diets, possibly also influenced by income?*), allowing for variation over time (financial year, season)
    - 2 levels of ethnicity (White British, other), 5 equivalised income quintiles, 8 financial years, 2 seasons
    - Low frequencies in all ethnic categories other than White British, so unable to include any further discrimination relating to ethnicity.

- Case study (model) 5 – to explore associations between dietary patterns and both the number of adults and presence of children in a household (i.e.*do the data suggest that lone parent households have different dietary patterns than either two-parent households or households without children?*), considering any influences of income, and allowing for variation over time (financial year, season)
    - Presence/absence of children (2 levels), Number of adults (1, 2, 3+), 5 equivalised income quintiles, 8 financial years, 2 seasons
- Case study (model) 6 – to explore associations between dietary patterns and household composition, including groups deemed potentially at risk of social exclusion (i.e. *do the data suggest that particular household composition groups have different dietary patterns to the rest of the population?*), considering any influences of income, and allowing for variation over time (financial year, season)
    - 15 household composition groups, 5 equivalised income quintiles, 8 financial years, 2 seasons
    - Household composition groups defined based on cross-tabulation of frequencies for combinations of the minimum and maximum adult age (5 categories for each), the presence/absence of children and the gender of household (all male, all female, mixed), with cohesive groups of potential interest identified with sufficient overall frequencies to provide reliable comparisons
- Case study (model) 7 – to explore associations between dietary patterns and both welfare benefit (in receipt of Income Support or not) and dwelling status (i.e. *do the data suggest that these socio-economic factors influence dietary patterns?*), allowing for variation over time (financial year, season)
    - 2 Income Support levels (yes/no), 3 dwelling status levels (Owned/Mortgaged, Rented, Rent-free/Squatting), 8 financial years, 2 seasons
    - The original 6 levels of dwelling status were appropriately combined to provide sufficient frequencies for meaningful categories. The 6 combinations of Income Support and Dwelling Status were eventually collapsed into 5 categories for analyses (combining the two Income Support categories within "Rent-free or Squatting")
- Case study (model) 8 – to explore differences between dietary patterns for urban and rural communities, allowing for variation over time (season) (2008 data only)
    - 3 urban/rural levels (Urban, Rural, Northern Ireland), 2 seasons
    - Data on urban/rural classification were collected in 2008 only, and were not available for Northern Ireland. Data combined across different urban and rural definitions to provide sufficient frequencies for reliable comparisons.

The statistical analysis approaches applied to address each of the objectives were as follows:

Objective 5: Explore possibilities of using Defra Family Food Survey (FFS) data (2002 to 2007) to assess the cost of a healthy diet.

This required the characterisation of a 'healthy diet' and therefore of a range of diets and indicators of the potential healthiness of different diets. In the absence of agreed diet definitions based on nutritional expertise, dietary groups were defined based on the observed household data using a number of statistical methods. The associations between these diet groups, expenditure and equivalised income were then assessed using two statistical approaches.

- Statistical definition of diets by three distinct methods
    - Classification of households into groups using the distance from the 'eatwell plate' "ideal", with groups approximately based on the quintiles of the observed data
    - Tabular cross-classifications of observed data, using household DRV threshold-based classifications for two (or potentially more) nutrients and food

components or ‘eatwell plate’ categories, to identify groups of households with similar dietary characteristics

      - Using two classifications allowed the identification of groups of households with extreme levels of both, as well as extreme levels of only one for comparison.

   - Hierarchical cluster analysis of observed data for a selected set of nutrients and food components to define groups of household with similar dietary characteristics. Similar approaches have been used previously to identify dietary groups amongst populations of individual based on measured food intakes (for example Pryer & Rogers, 2009; Pryer *et al*, 2001; Akin *et al*, 1986; Tucker *et al*, 1992, Hulshof *et al*, 1992)

- Log-linear analysis, assessment of contingency table contributions, and analysis of variance to assess associations between different diets and expenditure
- Log-linear analysis, assessment of contingency table contributions, and analysis of variance of expenditure levels for different diets in different income quintiles

Objective 6: Explore possibilities of using Defra Family Food survey data (2002 to 2007) to examine patterns in the diets of those at risk of social exclusion.

- Log-linear analysis and assessment of contingency table contributions to assess associations between household characteristics and threshold-based classifications of individual food components and nutrients/’eatwell plate’ components/distance from the ‘eatwell plate’ “ideal”
- Analysis of variance of household responses for individual food components and nutrients/’eatwell plate’ components/distance from the ‘eatwell plate’ “ideal”, for different household characteristics and combinations of household characteristics

Objective 7: Make comparisons between Defra Family Food survey data (2002 to 2007) and data for 2008 to explore whether (*there is any evidence that the start of*) the economic downturn has affected food purchases and thereby the diet of different social groups.

- Log-linear analysis to identify interactions in the associations between household characteristics and calendar/financial years for the threshold-based classifications of individual food components and nutrients/‘eatwell plate’ components/distance from the ‘eatwell plate’ “ideal”
- Analysis of variance of household responses for individual food components and nutrients/’eatwell plate’ components/distance from the ‘eatwell plate’ “ideal”, to assess for interactions between different household characteristics and calendar/financial year

A full description of the analytical methods is given in Annex 2. The following sections briefly indicate the analysis approaches for each of the main elements within these objectives, providing a few results to illustrate the application of the methods developed. All tables referred to are in Annex 2, together with some additional figures and tables.

Exploring Associations between Nutritional Indicators and Household Socio-economic Characteristics

It is important to emphasise that the classifications of households used in the eight case studies (models), and the classifications of dietary components, were developed solely to assess the potential application of the statistical analysis methods developed in this project, so that any associations identified in the analyses should only be considered as illustrations of the potential of the statistical approach, and not as firm inferences.

The first set of analyses started with the construction of frequency tables (contingency tables) summarising the numbers of households for each combination of the classes defined for each nutritional variable (eight key nutritional variables – Saturates; Calcium; Iron; Vitamin C; Sodium; Non-Milk Extrinsic Sugars; Fibre; Fruit & Vegetable Portions – and five 'eatwell plate' broad categories – Fruit & Vegetables; Bread, Cereals & Potatoes; Meat, Fish & Alternatives; Fats & Sugars; Milk & Dairy) and the combinations defined by each of the eight case study statistical models for household socio-economic characteristics. These contingency tables were then subjected to a log-linear model analysis, first fitting all model terms that only involved household socio-economic characteristic factors and the main effect of the nutritional variable. The analysis then adds model terms concerned with the interactions between the nutritional variable and each household characteristic factor, or combination of household characteristic factors, usually including interaction terms involving up to three or four household characteristic factors. Summary analysis of deviance tables for example analyses (for Calcium intake) are included in Annex 2, together with summaries of the statistical significance of each model term, for each of the variables analysed.

From these summaries of statistical significance levels, it is clear that each of the household characteristic main effects and a number of the combinations (interactions) appear to be strongly associated with most of the nutritional variables and 'eatwell plate' categories. For all the nutritional variables there is a strong association between nutrient intakes and household composition, particularly whether or not the household contains dependent children. There is also a strong association between nutrient intakes and minimum adult age . The equivalised income factor appears to have strong associations with some variables – notably Vitamin C, Non-Milk Extrinsic Sugars and Fruit & Vegetable portions – and statistically significant associations with all variables. Season and financial year generally appear to be less important, with the possible exception of a reasonably strong association of financial year with Sodium intake levels. Although statistically significant for all variables, Government Region only appears to be strongly associated with intake levels of Sodium. For all variables there are a number of strong associations with household characteristic interaction effects, usually involving interactions with the equivalised income quintile, suggesting that income changes the intake levels within the different socio-economic groups that the other household characteristic factors have attempted to define.

Household composition, minimum adult age, the presence or absence of children and the number of adults all appear to be strongly associated with each of the 'eatwell plate' category percentages, with the equivalised income factor strongly associated with all except the Meat, Fish & Alternatives category. Government Region has the strongest association with the percentage of Fruit & Vegetables, and is fairly strongly associated with the other four categories. Season and financial year appear to be less important for these measures of diet.

While these summaries of the statistical significance of the deviances associated with each term provide an indication of the important household characteristics associated with each nutritional variable, an assessment of the changes in the distribution of values for each variable across the levels of the household characteristic factors is needed to interpret the importance of these statistically significant associations. Illustrative and broad interpretations are given below for the various household characteristic factors and interactions for the variables for which the effects of these factors are most statistically significant – of course there may still be particularly interesting (and highly statistically significant) effects for variables that are not described below.

For Sodium intake it is clear that there has been a shift in levels over the course of the study, with a general decrease in the percentage of households in the highest Sodium intake category (> 2*RNI) from nearly 45% in 2001/2 to about 33% in 2008/9, and a corresponding general increase in the percentage of households in the lower Sodium intake categories (< RNI) from just below 13% in 2001/2 to nearly 20% in 2008/9 (Table 35a). While these data show an encouraging trend in the reduction of Sodium intake level over time, it is worth noting that the vast majority of households still appear to have an intake in excess of the RNI and therefore in excess of probable household requirements. Further discrimination of changes in the patterns of household intakes of Sodium might have been possible if the higher intake categories had been further split, maybe having additional thresholds at 1.5*RNI and 2.5*RNI.

For the 'eatwell plate' Fruit & Vegetable intake proportion (Table 35b) there are simple trends with financial year, in that the percentage of households in the higher intake categories (> 26%) generally increased up to 2006/7 (a slight decrease is shown in the last year of the study – which may reflect increased costs or reduced expenditure on food in this last year) and a corresponding general decrease in the percentage of households in the lower intake categories (< 13%). The trends are less clear for the ‘eatwell plate’ Bread, Cereals & Potatoes intake proportion (Table 35c) and also for Vitamin C intake (Table 35d) (some fruits, vegetables and potatoes are important sources of vitamin C in the diet). The general trend is for the ‘eatwell plate’ Bread, Cereals and Potatoes intake proportion to decrease in more recent survey years (higher than average percentages of households in the higher categories in the earlier years, and lower than average percentages of households in the higher categories in the later years, with the opposite patterns for the percentages of households in the lower categories), but the changes in the distributions of households year on year are relatively small. There are some obvious differences between financial years for the distribution of household Vitamin C levels – higher than average percentages of households with high levels (> 2*RNI) in 2006/7 and 2007/8, lower than average percentages of households with high levels in 2001/2 and 2004/5, with corresponding changes in the percentages for one or more of the low levels, but there is no evidence of any strong trends over time. (This difference is harder to interpret if absolute, rather than relative, levels of fruit and vegetable consumption did in fact decrease during 2007/8.)

While there are marked statistically significant differences between seasons (winter v. summer) for most of the variables analysed, the changes in the distributions of households between seasons are fairly simple (Table 36). The percentages of households with higher proportional intakes of Meat & Fish (‘eatwell plate’ category) are higher in winter than in summer, with the percentage with lower proportional intakes being lower in winter than in summer. The opposite pattern is seen for proportional intakes of Fats & Sugars (‘eatwell plate’ category) – note that the percentages of households with the lowest proportional intakes (< 8%) are fairly consistent across seasons. For Calcium intake, the percentage of households with the highest intakes (> 4*RNI) is higher in winter than in summer, with the percentage with lower intakes (notably between RNI and 3*RNI) being lower in winter than in summer – again note that the percentages of households with the lowest intake levels (< RNI) are fairly consistent across seasons, and that most households substantially exceed the RNI.

For the ‘eatwell plate’ Fruit & Vegetables intake proportion (Table 37a) there is clearly variation in the percentages of households in the different categories between regions, suggesting that households in some regions have relatively higher intakes of fruit and vegetables that those in other regions. Differences between regions in the distributions of households between categories are less distinct for the ‘eatwell plate’ Bread, Cereals & Potatoes variable (Table 37b), while for Sodium intake (Table 37c) there are a number of distinctly different distribution patterns for households across categories

For the period covered by the analysis, the patterns of intakes with income are again very clear (Table 38). For the ’eatwell plate’ Fruit & Vegetables intake proportion, the trend is for higher percentages of lower income households to have lower proportional intakes of Fruit & Vegetables (‘eatwell plate’ category), with a corresponding reduction in the percentage of households with higher proportional intakes compared with the overall average. The trend for households in the highest income quintile is the opposite: they are more likely to have higher proportional intakes of Fruit & Vegetables (‘eatwell plate’ category) compared with the overall average. A similar pattern is seen for the number of Fruit & Vegetable portions and for Vitamin C intake levels.

A number of the case studies explored the effect of household composition (considering different age-gender combinations) on intake levels. There were clear changes in patterns for Calcium and Iron intakes with adult age (Table 39), while for Calcium, Iron and NSP there were also differences between households with and without children. Of course, these differences ignore the effects of other household characteristics, such as the number of adults in the household, and it is reasonable to expect an interaction between the presence/absence of children and the number of adults for some variables. The overall effect of the numbers of adults in the household was also considered, and the most statistically significant associations were for the ‘eatwell plate’ Bread, Cereals & Potatoes intake percentage and the ‘eatwell plate’ Fruit & Vegetables intake percentage (Table 43), where ‘single adult households’ have different distribution patterns to ‘multi adult households’. Of course, these patterns ignore the effects of other household characteristics, and it is therefore important to also consider the interactions amongst these factors. The analyses for Case study (model) 5 do provide evidence of a statistically significant interaction in the combined effects of the presence or absence of children and the number of adults on the ‘eatwell plate’ Bread, Cereals and Potatoes intake proportions (Table 49), with the contrast between ‘single adult households’ and ‘multi adult households’ stronger in the absence of children than in the presence. Again, it is important to note that these analyses are ignoring other household characteristics that might be associated with these different household compositions, and therefore that there may be other household characteristics that can explain these different behaviours.

An alternative approach to exploring the impact of household composition on dietary intake patterns was included (Case study (model) 6), with 15 different compositional groups identified based on finding coherent groupings (each containing a sufficient number of households) with an obvious common characteristic. The identified household composition groupings are identified in Table 8. For both Calcium (Table 44a) and Iron (Table 44b) intakes there are some clear trends, but more detailed interpretation of these differences requires careful consideration of how the household composition groups were constructed – some groups will be quite diverse, while others are very uniform – though this more detailed (and therefore complex) classification of households does support the trends noted above, while illustrating the potential power of the analysis approach. Such more detailed classifications of households would potentially identify groups with unusual distributions across the categories for different food and nutrient variables, stimulating a more detailed exploration of the broader household characteristics associated with the households in these groups.

Although a number of statistically significant multi-way interactions were also identified by the analyses, these interactions are quite difficult to interpret because of the sizes of the summary contingency tables for these terms – in addition, there may be some doubt as to the reliability of these given that there may be relatively small numbers of households for some combinations of factors (by considering different factors in different models this problem should be minimised – but equally, different factors may be explaining some of the same variation in different models). With so many levels of each of the classifying explanatory variables, it is quite challenging to identify where patterns differ from the main effects (included at the top of each part of the tables – further details are given in Annex 2). One example from these analyses concerns possible associations between the presence/absence of children, equivalised income levels and Iron intake (Table 47). To aid interpretation it is useful to organise the contingency tables and calculate summary percentages relative to different margins – Table 47a allows a comparison for the distributions for different equivalised income levels relative to the overall distributions for households with or without children respectively, while Table 47b allows a comparison for households with or without children relative to the overall distributions for each equivalised income level. Table 48 shows a similar interaction for equivalised income and adult age, more clearly shown graphically in Annex 2 Figure 10.

A second set of analyses used a non-orthogonal, regression-based sequential analysis of variance approach to assess the impacts of the different household characteristic models on the mean responses for each nutritional variable. Where nutritional variables were initially expressed on a percentage scale (as percentages of total dietary energy, or, as for the 'eatwell plate' categories, as percentage of total food intake) these values could be directly analysed. The other variables needed to be adjusted for the size and composition (adults/children) of the household, and so were either expressed in an amount per person per day, or as percentages of the household RNI, prior to analysis. Examples of the analysis of variance tables produced as a summary of the analysis for each nutritional variable using each of the 8 household characteristic models (for the Calcium intake variable) are provided in Annex 2, together with summaries of the significance levels associated with each model term for each variable analysed, and tables of means for the levels of key household characteristics.

Levels of Sodium intake and percentages in the 'eatwell plate' Fruit & Vegetable and Bread, Cereals & Potatoes categories appear to be strongly affected by the variable 'Financial Year' (Table 60). Average Sodium intake appears to decrease over time, from 2001/2 to 2007/8, though with a slight "blip" in the downward trend in 2005/6. This is similar to findings from the FSA[1] and NatCen (2008) for reduction in salt intakes in recent years. A similar reduction over time is seen for the mean percentage in the 'eatwell plate' Bread, Cereals & Potatoes category, and an increase is seen for mean percentage in the 'eatwell plate' Fruit & Vegetables category over time, although in 2007 and 2008 the data show a slight reversal of these trends. Again, this would be in accordance with findings from the FSA[2]. Average levels of Vitamin C and Sodium intakes and percentages in the 'eatwell plate' Meat, Fish & Alternatives, and Fats & Sugars categories are the variables most affected by Season (Table 61). Average levels of Vitamin C and the mean percentage of Fats & Sugars appear to be lower in Winter than in Summer, with the opposite pattern seen for the mean percentage of Meat, Fish & Alternatives, and for average Sodium intake levels.

Average values for both Fruit and Vegetable variables ('eatwell plate' percentage and number of portions) and the Vitamin C intake are strongly influenced by income (Equivalised Income Quintile) in

[1] http://www.food.gov.uk/news/pressreleases/2008/jul/sodiumrep08
[2] http://www.food.gov.uk/news/newsarchive/2007/feb/cas2006

all models, with all three variables showing strong increases with increasing income levels (Table 62). Average Fruit and Vegetable intake (measured using both the ‘eatwell plate’ percentage and the number of portions), and the mean ‘eatwell plate’ Bread, Cereals & Potatoes percentage are strongly influenced by Government Region (Table 63).

The nutritional variables strongly influenced by each of the other main effects of household characteristics can be similarly identified. Although direct comparisons are not always possible, the findings of these analyses present a picture broadly consistent with findings from national individual level surveys of food patterns and dietary intakes (e.g. the National Diet and Nutrition Surveys) as well as smaller scale individual level surveys. The implication is that the statistical approaches and data used in this study give results that are reasonably reliable.

Using Statistical Approaches to Explore Ways of Grouping Different Dietary Patterns

It is important to emphasise that the groupings of households based on dietary components, described below, has been developed solely to illustrate the potential of the statistical analysis approach to use the FFS data to explore associations between diet and cost, so that any results described are for illustrative purposes only, with further research on the appropriateness of the dietary groupings needed before any firm conclusions can be drawn.

One of the objectives of this project (Objective 5) was to assess whether the Family Food Survey data could be used to assess the cost of a healthy diet. This requires some means of defining or characterising a ‘healthy diet’, and also the need to identify ‘unhealthy’ diets so that the costs of different diets can be compared. As discussed earlier, there are difficulties in being too prescriptive about which dietary patterns conform to a ‘healthy diet’ and which do not, as well as the statistical challenges of handling so large a dataset. Following consultation with experts, it was considered that it would be inappropriate to identify particular combinations of nutritional intakes or food patterns too rigidly as a means of characterising a healthy (or unhealthy) diet for the purposes of ascertaining costs from the FFS. Instead, guided by nutritional/public health considerations, a number of statistical approaches were applied to the FFS data (suitably adjusted for household composition as described above) to identify common dietary patterns across the range of nutritional variables included in the analyses described above. These common dietary patterns might then be considered as more or less ‘healthy diets’ – they can be used to identify groups of households with different dietary patterns, though there is often considerable variability in the dietary patterns of households included in the same group. Further exploration of these patterns and their utility for addressing Objective 5 could be undertaken.

For all methods used to develop groups with different dietary patterns, the household records from all years of the FFS data used in this study were included. This allowed assessment of how the distributions of households between groups varies with household characteristics (including ‘Financial Year’), as well as an assessment of how the costs associated with the different dietary patterns varied with the different household characteristics. An alternative approach would have been to develop models for dietary patterns for a subset of the data (such as for a particular year), but none of the statistical approaches considered would easily provide reliable models in this way, primarily because they are based on statistical rather than nutritional approaches, and hence are driven by the empirical data rather than by any mechanistic requirements. It would therefore be difficult to develop “rules” for the groups generated for a particular subset (year) of the data that could be applied to the rest of the data.

One approach, developed during discussions with Defra and DH is based directly on the “ideal” percentages for each of the ‘eatwell plate’ categories, i.e. a ‘distance from eatwell plate’, as already introduced above. This is an attempt to assess how far the dietary intake of a household differs from the ‘eatwell plate’ “ideal” (see Figure 2: 33% Fruit and Vegetables, 33%Bread, Cereals and Potatoes, 15% Milk and Dairy, 12%Meat, Fish and Alternatives, and 8% Fats and Sugars). Our statistical approach calculates the root mean squared distance between each observed distribution and this “ideal” across the five categories described. However, a number of contributions to this distance were set to zero to avoid over-penalising low levels of Fats & Sugars (zero contribution if percentage is less than the 8% ideal), high levels of Fruit & Vegetables (zero contribution if percentage is greater than the 33% ideal) and high levels of Bread, Cereals & Potatoes (zero contribution if percentage is greater than the 33% ideal). It was not clear whether any categories should be regarded as more important than others in constructing these distances, though it might be argued from epidemiological evidence that, given the lack of a category relating specifically to ‘salt’ the proportions of Fruit & Vegetables and of Fats & Sugars in the diet have the best established relationship with health outcomes, and should therefore be regarded as most significant in terms of characterising a 'healthy diet'. But in the absence of definitive information, deviations for all categories contributed equally to the 'eatwell plate distance'. These 'eatwell plate distances' can be analysed in two forms, either as a continuous numerical variable, or as a categorical variable, which is obtained by allocating households into different ranges of 'eatwell plate distances'. There was no prior experience of how different 'eatwell plate distances' could translate into verifiable indicators of 'healthy' and 'unhealthy' diets (i.e. what sizes of 'distance' would constitute 'healthy' and 'unhealthy' diets – though it might be reasonable to assume that a smaller distance should indicate a more healthy diet). Therefore, the households were divided into five roughly equal sized groups based on the 'eatwell plate distance', using thresholds obtained by rounding the empirical quintiles; this resulted in groups of households with 'eatwell plate distances' in the following ranges: <9, 9 – 12, 12 – 15, 15 – 18 and >18. It is interesting to note that even those households with the smallest 'eatwell plate distances' (14.5% of the households had values less than 9) were not particularly close to the 'eatwell plate' "ideal", overall deviation from the 'eatwell plate' ideal being up to 9 percentage points (combined across the five categories – note this is not the average deviation across the categories but a Euclidean distance based on the deviation contributions from each of the categories).

It is interesting to consider how this novel measure of diet relates to the individual 'eatwell plate' category percentage contributions (remembering that the contributions from the five categories must sum to 100%). Figure 2 shows the mean value for each of the household groups defined using the 'eatwell plate distance' for each 'eatwell plate' category component, suggesting that the diets differ mainly in the percentage contributions for the Fats & Sugars, Bread & Cereals, and Fruit & Vegetables components. On average, a healthier diet (with a smaller 'eatwell plate distance') contains low levels of the first of these components and higher levels of the latter two, with a less healthy diet (with a larger 'eatwell plate distance') having the opposite pattern).

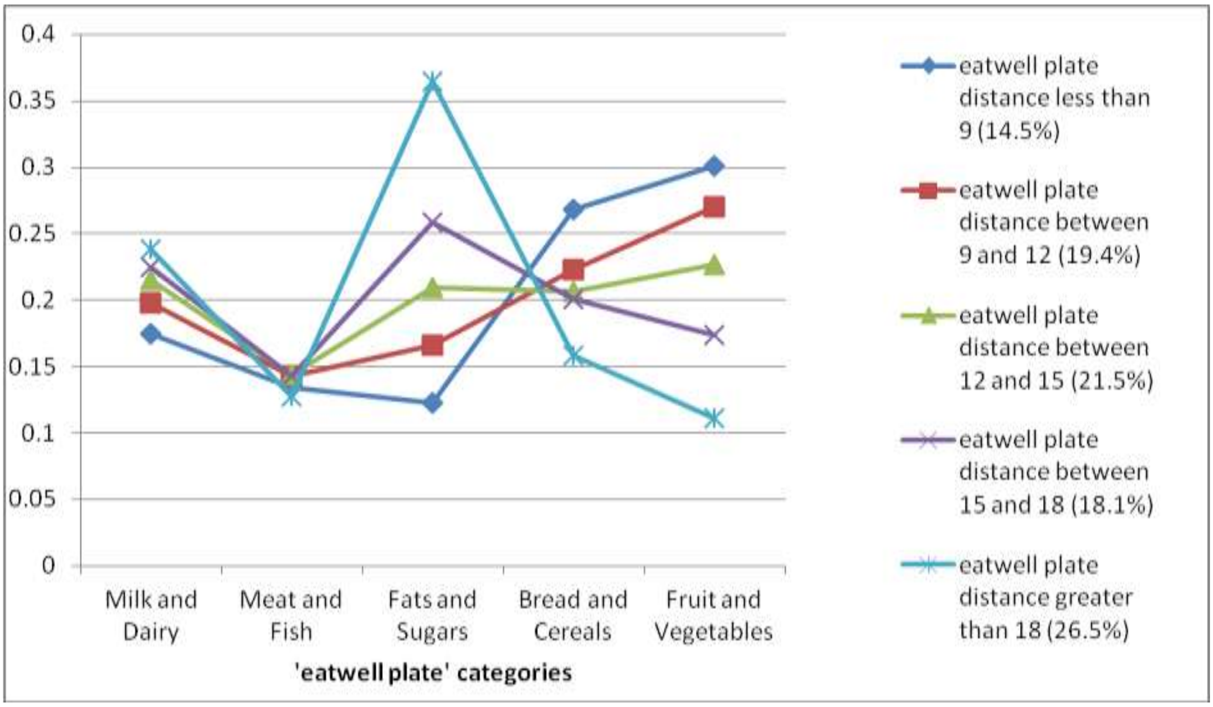
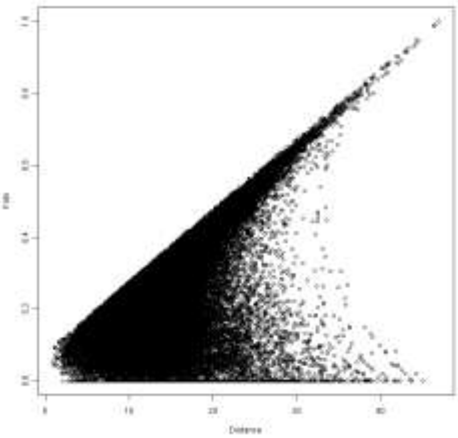


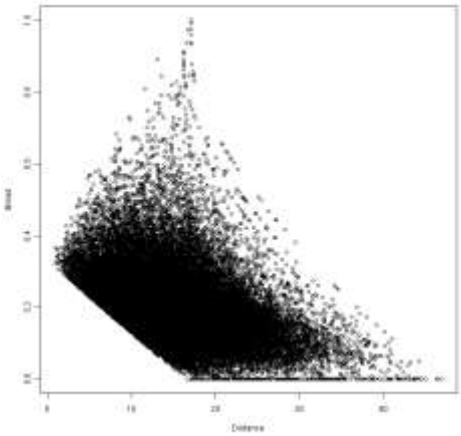
Figure 2 Mean proportions of each 'eatwell plate' component for the households allocated to each of the five diet groups defined based on the approximate quintiles of the observed distribution of 'eatwell distances' across the complete dataset. Legend also indicates the percentage of households in the dataset falling into each diet group.

Figure 3 then shows the relationship between the 'eatwell plate distance' and each of the five component percentages for the individual household observations – because of the large number of observations, individual points can only be seen towards the extreme values of each variable. These show that, as expected, the minimum 'eatwell plate distance' is achieved when the percentages in each category are close to the 'eatwell plate' ideal. Of particular interest is that larger 'eatwell plate distances' (i.e. household dietary patterns which are fairly unlike the 'eatwell plate' ideal proportions and thus interpreted as relatively unhealthy using this index) can be associated with percentages of each component both higher and lower than the 'eatwell plate' ideal, most notably for the Fats and Sugars category and the Milk and Dairy category (for both categories an 'eatwell plate distance' above 40 can be associated both with percentages less than 20% and greater than 80%).

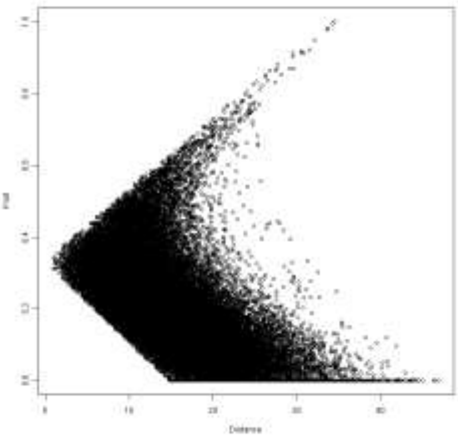
a: Fats and Sugars



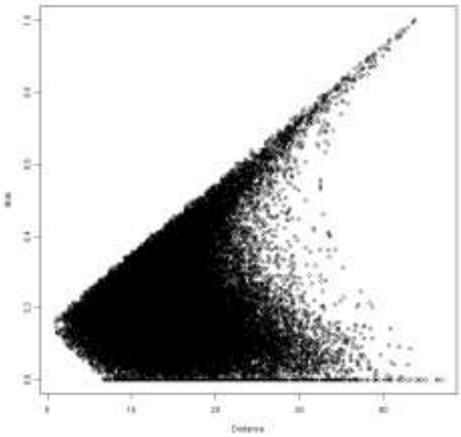
d: Bread, Cereals and Potatoes



b: Fruit and Vegetables



e: Milk and Dairy



c: Meat, Fish and Alternatives

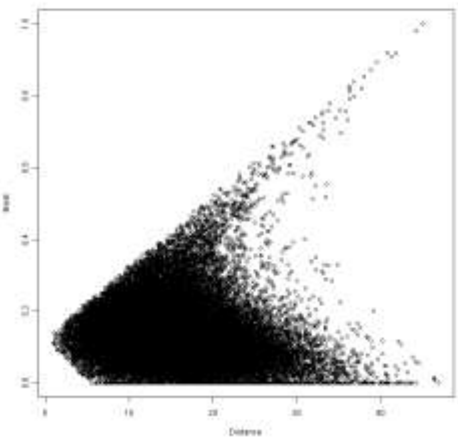


Figure 3 Relationship between 'eatwell plate distance' (horizontal axes) and the proportion intakes of each of the five 'eatwell plate' category components.

These associations are further illustrated by considering the numbers of households classified into different combinations of 'eatwell plate distance' diet categories and the classifications used for each of the 'eatwell plate' components (Table 71). Most notable are the strong associations between the 'eatwell plate distance' diet categories and the 'eatwell plate' Fats and Sugars percentage (a strong positive association – a diet further from the 'eatwell plate' ideal tends to include higher percentages of Fats and Sugars) and 'eatwell plate' Fruit and Vegetables percentage (a strong negative association – a diet further from the 'eatwell plate' ideal tends to have lower percentages of Fruit and Vegetables). Given the overall connections between the percentages in the five 'eatwell plate' categories, it is likely that these two patterns are related – diets containing high levels of Fruit and vegetables tend to have low levels of Fats and Sugars, and vice versa. But also note that 11% of households in the ">18" 'eatwell plate distance' diet group have less than 8% Fats and Sugars, and that 9% of those in the "15-18" 'eatwell plate distance' diet group have greater than 33% Fruit and Vegetables. There is a weaker negative association between the 'eatwell plate distance' diets and both the 'eatwell plate' Bread, Cereals and Potatoes percentage and 'eatwell plate' Meat Fish and Alternatives percentage, with each of the diet groups including households with a broad spread of category percentages. For the 'eatwell plate' Milk and Dairy categories, diets with smaller 'eatwell plate distances' tend to be associated with intermediate percentages (close to, or slightly higher than, the 15% ideal for the category), and diets with larger 'eatwell plate distances' tend to be associated with higher percentages (>30%), though there are also a relatively high proportion of households with lower percentages for this category for the ">18" diet group.

These results illustrate how each of the 'eatwell plate distance' diet groups can be associated with a range of percentages of each component, so that while on average a smaller 'eatwell plate distance' appears to be associated with a more healthy diet (more Fruit and Vegetables, less Fats and Sugars), and a larger 'eatwell plate distance' with a less healthy diet (more Fats and Sugars, less Fruit and Vegetables), this simple index is not able to capture all the multi-facetted complexity of diet. The implication of using this approach in deriving an indicator of the healthiness of a dietary pattern is that it should capture major differences in dietary pattern (e.g. concerned with levels of Fruit and Vegetables and of Fats and Sugars), but that some detail of the variation in dietary pattern will be hidden by this simple index. Further research into the characteristics of this index is therefore needed, possibly including the development of a second companion index to summarise a second form of deviation from the 'eatwell plate' ideal associated with the healthiness of a dietary pattern.

A second approach to the identification of common dietary patterns across households within the data set was based on the cross-tabulation of numbers of households for different combinations of the 'eatwell plate' category levels or nutritional component levels. A first set of diet definitions was obtained by considering the cross-tabulations of the 'eatwell plate' Fruit and Vegetable categories with the 'eatwell plate' Milk and Dairy categories, and of the 'eatwell plate' Fruit and Vegetable with the 'eatwell plate' Fats and Sugars categories. Three groups with extreme dietary patterns were identified – those with levels 1 or 2 for Fruit and Vegetables and level 6 for Milk and Dairy (high Milk & Dairy/low Fruit & Vegetables – 8.8% of the households), those with levels 1 or 2 for Fruit and Vegetables and level 6 for Fats and Sugars (high Fats & Sugars/low Fruit & Vegetables – 9.1% of the households), and those with level 6 for Fruit and Vegetables and levels 1 or 2 for Fats and Sugars (low Fats & Sugars/high Fruit & Vegetables – 11.2% of the households). The remaining households were collected together into a "non-specified" group (70.9% of households) – ideally this group of households would have been further divided based on the 'eatwell plate' category classifications, but there were no obvious and coherent combinations that included a substantial proportion (8% or more) of the households. Comparisons of the three "extreme" dietary patterns may therefore be interesting, but the "non-specified" group will contain households with a wide diversity of dietary patterns, so that further analysis of the patterns associated with this group are unlikely to be particularly illuminating.

Figure 4 shows the mean proportion intakes of the five 'eatwell plate' components for the households in each of these four identified groups. Note how the groups with specified high levels of a particular component have mean proportion intakes for these components that are substantially higher than for the other diets, and that the mean values for the two components not included in the definitions (Meat and Fish, Bread and Cereals) have similar means for all four diet groups. In addition, the High Milk and Dairy diet tends to be associated with relatively low levels of Fats and Sugars (compared with the "non-specified" diet), and the High Fats and Sugars diet tends to be associated with relatively low levels of Milk and Dairy. The High Fruit and Vegetables diet tends to have low levels of each of the other four components (particular Fats and Sugars).

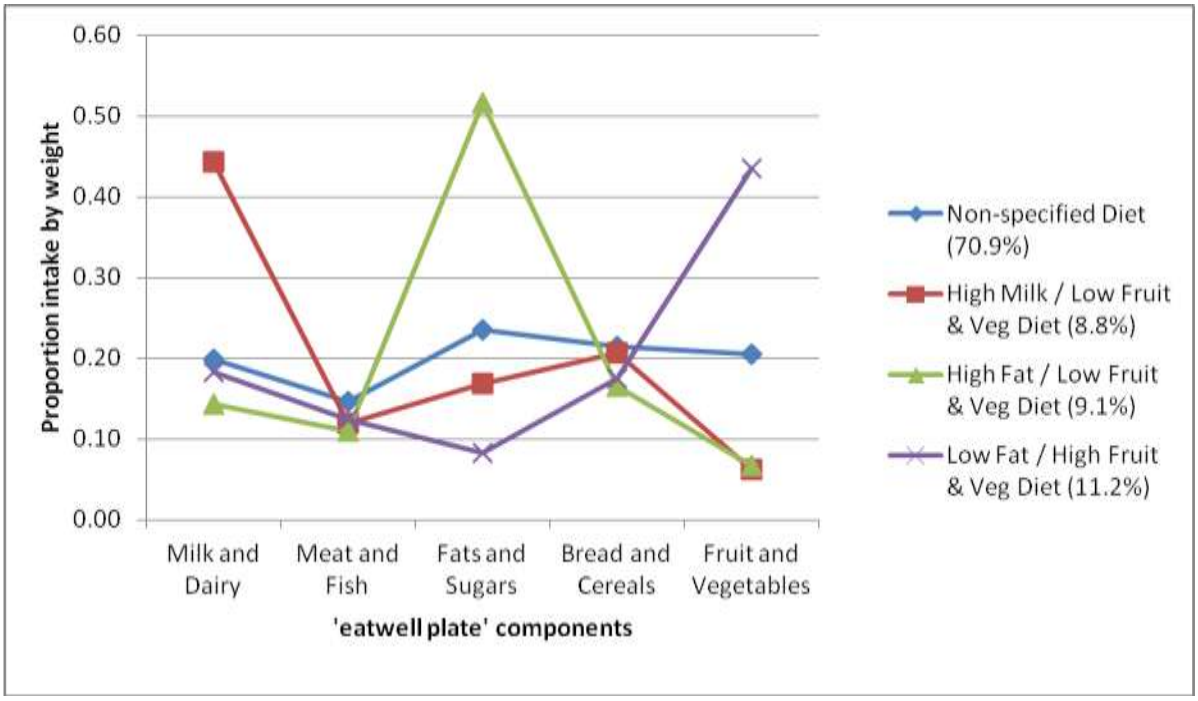


Figure 4 Mean proportions for each 'eatwell plate' component across the households allocated to each of four dietary pattern groups. The groups were defined based on three extreme pair-wise combinations of 'eatwell plate' category classifications, plus a fourth group containing the remaining majority of the households. Legend also indicates the percentage of households in the dataset falling into each diet group.

This approach was repeated for various combinations of the eight key nutritional variables – Saturates; Calcium; Iron; Vitamin C; Sodium; Non-Milk Extrinsic Sugars; Fibre; Fruit & Vegetable Portions – considered in the study, again identifying groups with common dietary patterns based on reasonably substantial numbers of households with particularly high or low levels for particular pairs of variables. However, while this approach worked for the 'eatwell plate' components, because of the percentages (proportions) having to sum to 100%, strong positive correlations among a number of the key nutritional variables made it difficult to identify differences in the dietary patterns between the groups, with mean values and distributions for the key nutritional variables used to identify the groups being quite similar between groups.

A final approach to identifying groups of households with similar dietary patterns was to use a series of hierarchical cluster analyses of the households, using either the 'eatwell plate' component intake proportions, or the values for the eight key nutritional components. Computational limitations meant that separate cluster analyses were first performed for the households surveyed in each financial year separately, with groups combined across years based on a cluster analysis of the group means. Figure 5 shows the mean proportion intakes of each of the 'eatwell plate' components for the groups identified through the application of this approach to the 'eatwell plate' data. Descriptions for each of the groups were generated based on the mean levels shown for each of the 'eatwell plate' components in Figure 5. Some of these groups are easy to identify and describe – for example, there was a small group (1.0% of households) with very high proportions of Milk and Dairy intake and consequently low levels of all other components (High Milk & Dairy 1, mean Milk and Dairy intake of 64%). There were two further groups (6.3%, 7.4% of households respectively) with relatively high levels of Milk and Dairy intake (High Milk & Dairy 2 and 3, mean Milk and Dairy intakes of 37% and 26% respectively), with slightly higher levels of each of the other four components. Three further groups had high levels of Fats and Sugars intake and relatively low levels of all other components (High Fats & Sugars 1, 5.2% of households, mean Fats and Sugars intake of 42%; High Fats & Sugars 2, 22.0% of households, mean Fats and Sugars intake of 34%; High Fats & Sugars 3, 5.5% of households, mean Fat and Sugars intake of 28%). Two further groups had high levels of Bread and Cereals intake (High Bread & Cereals, 5.6% of households, mean Bread and Cereals intake of 34%), and high levels of Fruit and Vegetables intake (High Fruit & Veg, 3.5% of households, mean Fruit and Vegetables intake of 45%), respectively. Other groups had relatively high levels of two components (High Milk & Dairy/High Fruit & Veg, 25.2% of households), or a high level of one and a low level of another (e.g. Low Milk & Dairy/High Meat & Fish, 1.7% of households). Finally, two groups (10.7%, 1.4% of households respectively) were quite distant from the ‘eatwell plate’ “ideal”, having roughly 20% intake for each of the five components (there are small differences between these two groups in the proportions of Meat and Fish intake, and Bread and Cereals intake). This approach to defining diets is not without flaws, as there are still, for example, households in the “High Fruit & Veg” dietary group with low percentage intakes (< 6%) of Fruit and Vegetables. But the methodology should ensure reasonable uniformity in the dietary patterns of households within each cluster (group).

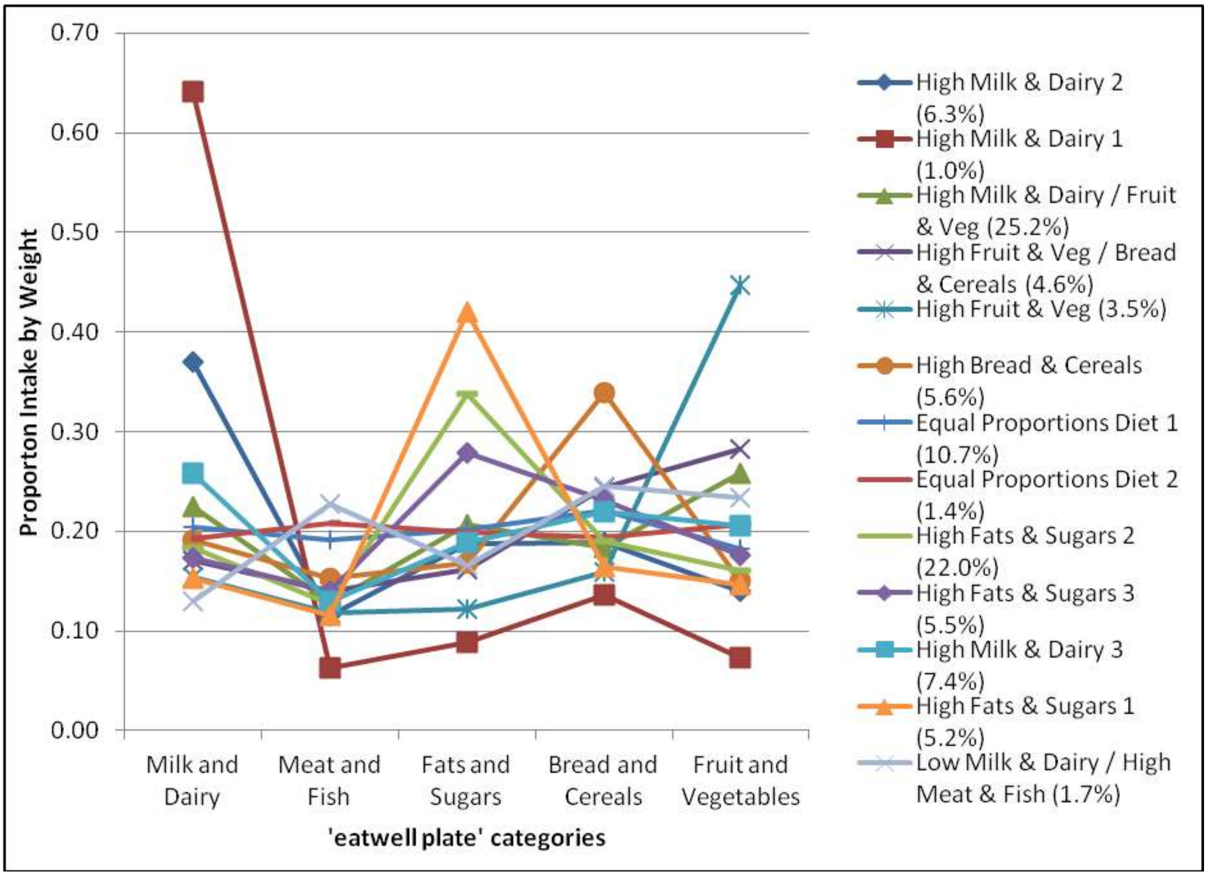


Figure 5 Mean proportions of each ‘eatwell plate’ component for the households allocated to each of the thirteen groups (clusters) defined based on the multiple hierarchical cluster analyses of the ‘eatwell plate’ component data, using a normalised Euclidean metric to calculate the similarities between households. The legend includes names for the groups based on the mean dietary patterns for each group, and also indicates the percentage of households in the dataset falling into each dietary group.

Thus we used three different empirical methods to approach the identification of patterns of dietary behaviour that could be regarded as reasonably similar within a group of households and sufficiently distinct from the patterns for other groups of households. Of these different statistical approaches, the last approach, multiple hierarchical cluster analysis of the ‘eatwell plate’ component data (repeated for the eight key nutritional variables), seemed to be the most promising in allowing the data to identify households with similar dietary patterns. Of course there is always a subjective element in any statistical analysis, in this instance in the definition of the clusters – we could have “cut” the dendrograms at different levels, which would have produced more or fewer clusters for each year, thus modifying the secondary clustering of groups across years. Also, the names given to groups were based on group mean values, and obviously do not entirely describe all individual households within each group. The disadvantage of the second approach described above, which constructed groups based on the categorised components, is that it is difficult to extract more than two or three independent groups (two of the groups identified were not, in fact, mutually exclusive), and the result was that the majority of households did not fit into any of the specific groups. Finally, although the classification of households using the ‘eatwell plate distance’ indicator, discussed first above, provided a valuable method for combining the large amount of data across the five ‘eatwell plate’ components, the ‘distance’ values of the indicator were arrived at through a statistical process, and arguably have no basis in epidemiology. In other words, the development of an indicator of a ‘healthy diet’ in terms of the distance of a households intake from the ‘eatwell plate’ “ideal”’ was explored using sampled data, and the nutritional or health significance of different degrees of ‘distance’ from the “ideal” is not yet known. Nevertheless, we do know that the further from the “ideal” the dietary pattern of a given household (or group of households) is (i.e. the larger the ‘eatwell plate distance’), the less likely it is that members of the household(s) are consuming diets conducive to good health. Conversely, the closer the dietary pattern of a household (or group of households) is to the ‘eatwell plate’ “ideal” (i.e. the smaller the ‘eatwell plate distance’) then the more likely it is that members of the household(s) are eating a healthy diet. It would be useful to explore further the relationship between such a ‘distance’ indicator, tested here as reasonably robust statistically, and other datasets which include biomedical markers of health.

Finally, one way to combine statistical and pattern information from the “cluster analysis” dietary groups and the ‘eatwell plate distance’ indicator is to examine the summaries of the ‘eatwell plate distance’ indicator for each of the “cluster analysis” dietary groups (Table 75). This clearly identified some dietary groups (e.g. “Equal Proportions Diet 1”) as having less within-group variability in ‘eatwell plate distance’, and others (e.g. High Fats and Sugars 1, 2 and 3) as generally being more distant from the ‘eatwell plate’ “ideal”. Further statistical observations about the identification of groups with common dietary patterns and the relationships with the ‘eatwell plate distance’ are explored in Annex 2.

Exploring the Assessment of Associations between Household Dietary Patterns and Household Socio-economic Characteristics

It is important to again emphasise that the groupings of households based on socio-economic characteristics and dietary patterns have been developed solely to illustrate the potential of the statistical analysis approach to use the FFS data to explore associations between these characteristics, so that any results described are for illustrative purposes only with further research on the appropriateness of these classifications needed before any firm inferences can be drawn.

Once these different statistical methods for exploring ways of grouping households with similar dietary patterns had been applied to the complete household dataset, and some reasonably consistent patterns established, such that households could be assigned to groups with more or less ‘healthy’ dietary patterns (e.g. based on distance from the ’eatwell plate’ ideal), the association between the dietary pattern groups and the socio-demographic and economic indicators could be explored. Illustrative results described below relate to the analysis of the associations with the groups defined using the ‘eatwell plate distance’ quintiles using the log-linear modelling approach described above, though equivalent analyses for the dietary pattern groups defined using other the methods, and the hierarchical analysis of variance of the calculated ‘eatwell plate distance’ values show similar patterns.

Of particular interest are the strong interactions involving the presence/absence of children and income quintile (Model 5), and overall household composition and income quintile (Model 6) (Table 76). Certain Government Regions (London, Eastern England, South East and South West) appear to be more associated with a diet that is closed to the ‘eatwell plate’ ideal, with others (North East, North West, Yorkshire and Humberside, Wales and Scotland) more associated with a larger distance from the ‘eatwell plate’ (based on the percentages of households from each Region in the different ‘eatwell plate distance’ categories – Table 77). Households with lower incomes also tend to have a larger distance from the ‘eatwell plate’ ideal (52.9% of households in the lowest equivalised income quintile had an ‘eatwell plate distance’ of more than 15, compared with 44.6% overall and 35.4% in the highest equivalised income quintile) and those with higher incomes tend to have dietary patterns that are closer to the ‘eatwell plate’ ideal (41.9% of households in the highest equivalised income quintile had an ‘eatwell plate distance’ of less than 12, compared with 34.0% overall and 27.4% in the lowest equivalised income quintile). Households with a younger minimum adult age (<50) also tend to have dietary patterns with a larger distance from the ‘eatwell plate’ ideal (40.8% of households with a minimum adult age between 18 and 30, and 31.8% of households with a minimum adult age between 31 and 50, had an ‘eatwell plate distance’ of more than 18, compared with 26.5% overall), while those comprising of older adults (minimum adult age > 50) tend to have better dietary patterns (38.7% of households with a minimum adult age between 51 and 65, and 40.1% of households with a minimum adult age between 66 and 85, had an ‘eatwell plate distance’ of less than 12, compared with 34.0% overall). There has been relatively little change in the distribution of households between the ‘eatwell plate distance’ quintiles across the 8 financial years covered by the analysed datasets (Table 79), though the percentage of households with an ‘eatwell plate distance’ of more than 18 increased to 28.2% and 27.9% in 2003/04 and 2004/05, respectively, and decreased to 25.0% in 2005/06, compared to an overall average of 26.5, with corresponding decreases and an increase in the percentage of households with an ‘eatwell plate distance’ of less than 12 in these same years. The distributions in the last three years of the study (2006/07, 2007/08 and 2008/09) have shown little variation from the overall distribution across all 8 years. There was no evidence for any interaction between the effects of financial year and equivalised income quintile in any of the analyses including both terms (Table 76), suggesting, in particular, that the quality of dietary pattern of households in the lowest equivalised income quintile was not influenced by the economic downturn in the last few years of the period studied.

Overall, households with children tend to have dietary patterns with a larger distance to the ‘eatwell plate’ than those without (53.9% of households with children had an ‘eatwell plate distance’ of more than 15 compared with 39.9% of households without children). In addition, single adult households with children tend to have dietary patterns that are further away from the ‘eatwell plate’ ideal than multiple adult households with children (63.3% of single adult household with children had an ‘eatwell plate distance of more than 15 compared 51.2% for multiple adult households with children), and single adult households without children also tend to have dietary patterns with a larger distance to the ‘eatwell plate’ ideal than households with more adults but without children (48.2% of single adult households without children had an ‘eatwell plate distance’ of more than 15 compared with 34.3% of multiple adult households without children). There is a strong interaction effect on the ‘eatwell plate distance’ between the presence or absence of children and equivalised income quintile (Table 81). So, while dietary pattern tends to improve (decreased percentages of households with larger ‘eatwell plate distances’) with increased income, the reduction is less dramatic for households without children (from 45.9% of households in the first equivalised income quintile with an ‘eatwell plate distance’ of more than 15, to 40.7%, 40.1% and 38.8% in the second, third and fourth quintiles, and 34.6% in the fifth quintile) than for those with children (66.3% of households in the first equivalised income quintile, reducing to 61.1%, 55.0% and 47.2% in the second, third and fourth quintiles, and to 37.3% in the fifth quintile). These patterns seem to be consistent across households with different numbers of adults. Finally, considering the range of 15 household compositions, households with children and all-male households tend to have dietary patterns with a larger distance to the ‘eatwell plate’ ideal (based on the percentages of households with an ‘eatwell plate distance’ of more than 18), while mixed gender households without children tend to have dietary patterns with a smaller distance.

All of these illustrative results depend on the acceptance of the concept of the ‘eatwell plate distance’ as a reliable indicator of the healthiness of a diet. Further research to develop this measure is needed before the ‘eatwell plate distance’ can be considered as a reliable indicator. However, these results do demonstrate the potential for using the log-linear analysis approach to assess for differences in dietary patterns between different socio-economic groups, once a reliable indicator of the healthiness of different diets can be established.

Assessing the Costs Associated with different Dietary patterns (and associations with Household Socio-economic Characteristics)

It is again important to emphasise that the classifications of households identified in this report have been developed solely to illustrate the potential of the statistical analysis approach to use the FFS data to assess associations between these classifications, so that all results described are for illustrative purposes only, with further research needed on the appropriateness of these classifications before any firm inferences can be drawn.

The previous sections have detailed the procedures used to explore ways of classifying households into different dietary pattern groupings, and to explore ways of examining how these dietary pattern groupings could be assessed in terms of their relative healthiness (in particular, based on quantitative assessment of their similarity to, or distance from, an ‘eatwell plate’ “ideal”). This section outlines the procedures for examining the expenditure associated with different dietary patterns, and the potential impact of household socio-economic characteristics on expenditures associated with different dietary patterns. The FFS records all household expenditure and thus the dataset contained records of expenditure on food, as well as on the food commodities purchased and the nutrient breakdown.

Two separate adjustments to the expenditure data were required to enable appropriate comparison of monetary expenditures on food across all households. The first was that the monetary values had to be adjusted for inflation: this was achieved by using monthly values of the Retail Price Index over the whole period of the survey to scale all expenditure records to have a cost equivalent to the values in December 2008 (the last month from which samples were included). The second adjustment was to make allowance for household composition using the McClement’s scale (McClement, 1977), which adjusts for both economies of scale and the ages of household members. The expenditure data are then expressed as equivalised to the standard household of two adults (of course, these figures could be adjusted for some other household composition if required). Further details are given in Annex 2. As well as considering the exact calculated adjusted cost for each household, households were then classified based on the observed quintiles of the household composition adjusted expenditure values (these were: <£65, £65 - £92, £92 - £119, £119 - £155, and >£155 per two week period, adjusted to a standard two-adult household composition). Adjusted expenditure values were also calculated for each of the five 'eatwell plate' components (analyses not presented).

A series of analyses were carried out using all the available data, therefore including household records from the different financial years, which allowed assessment of any changes in these adjusted and equivalised expenditure values over time. The first analysis explored the association between adjusted and equivalised expenditure (for household composition and RPI), classified into quintiles, and dietary patterns, using the 'eatwell plate distance' categories. One finding from the FFS data from 2002 to 2008 is that low expenditure was associated with a much higher percentage of households with larger 'eatwell plate distances' than for the higher expenditure categories – 40.6% of households in the lowest adjusted expenditure quintile (<£65) had an 'eatwell plate distance' of more than 18, compared with 27.6%, 24.3%, 21.5% and 18.9% of households, respectively, in the other four quintiles (Table 83). However, the percentages of households with an 'eatwell plate distance' of less than 9 (better dietary patterns) only ranged from 11.3% in the first adjusted expenditure quintile to 16.5% and 16.2% in the top two quintiles, with more variability in the 'eatwell plate distance' 9-12 category (from 13.3% for the first adjusted expenditure quintile to 23.9% for the fifth adjusted expenditure quintile). Therefore the evidence from this research was unable to demonstrate any clear association between diet quality and expenditure over the period from 2002 to 2008. A further analysis, including data from later time periods, would be needed to assess whether these patterns have changed more recently.

Log-linear analyses of the associations between the adjusted expenditure quintile classification and the various household socio-economic characteristics, using the eight models described above, showed statistically significant associations of adjusted expenditure with equivalised income quintile, season, government region, minimum adult age, economic status, ethnic origin, presence/absence of children, number of adults, general household composition, and income support/dwelling status. However, there was no clear association with financial year. Clear associations include that the North East, Yorkshire and Humberside, East Midlands and West Midlands had higher than average percentages of households in the lower adjusted expenditure quintiles, with the East of England, South East, Scotland and Northern Ireland having higher than average percentages of households in the higher adjusted expenditure quintiles, and with London having higher than expected percentages in both the first and fifth adjusted expenditure quintiles. Households with a minimum adult age in the 18-30 or >65 categories tended to have lower levels of adjusted expenditure (higher percentages than average in the lower two quintiles), with households with a minimum adult age between 31 and 65 tended to have higher levels of adjusted expenditure (higher percentages than average in the upper two quintiles). Single adult households, both with and without children, tended to have lower levels of adjusted expenditure, while two-adult households, again both with and without children, tended to have higher levels of adjusted expenditure. There was also a strong association between equivalised income and adjusted expenditure, with a tendency for households with lower incomes to have lower levels of expenditure and those with higher incomes to have higher levels of expenditure. Further, detailed patterns could be extracted from these analyses once appropriate particular questions are identified.

Two subsequent analyses examined the associations between expenditure, quality of dietary pattern (using the dietary groups defined based on the quintiles of the 'eatwell plate distance') and income (using the equivalised income quintiles). The first analysis calculated summary statistics for expenditure (mean, maximum, standard deviation) for each combination of 'eatwell plate distance' group and equivalised income quintile (Table 84). The increase of mean expenditure with increasing income (higher equivalised income quintile) is not surprising and reflects the purchase of food of higher social or culinary value (e.g. better cuts of meat, ready prepared vegetables, more exotic foods, etc) and cost as income rises. Those dietary patterns with a larger 'eatwell plate distance' were associated with lower average cost, which is also not surprising. However, what was also clear was that households with dietary patterns with small distances to the 'eatwell plate' "ideal" spent less on average than those with dietary patterns which were slightly further from the 'eatwell plate "ideal"' (comparing the "<9" 'eatwell plate distance category with the "9 – 12" category), with a much lower maximum expenditure for households in the lower equivalised income quintiles. Furthermore, it is interesting to note that there is less variability (shown by the smaller standard deviation) for the cost associated with the dietary patterns with small distances to the 'eatwell plate' "ideal" than with the dietary patterns with a larger 'eatwell plate distance'.

Final Discussion

The three key issues to be addressed in this part of this project were the development of statistical analysis methods for:

- Assessing the cost of a healthy diet (Objective 5)
- Assessing the diet patterns for those at risk of social exclusion (Objective 6), and
- Assessing the impact of the (*start of the*) economic downturn on diet patterns (Objective 7)

To assess the cost of a healthy diet (Objective 5), we first needed to characterise the healthiness of some different diets. While nutritional guidance is available on the characteristics of a healthy diet, the analysis of the FFS data required a clear quantitative characterization of the parameters associated with healthier and less healthy diets. The 'eatwell plate' provided one quantitative characterization of the healthiness of different diets, though recognises that there are many combinations of food choices that fit within the overall model of a healthy diet that it presents. A number of statistical approaches were applied to the observed data, alongside the 'eatwell plate' definition, to generate particular dietary groups. Having characterised these different dietary groups, an identification of more or less healthy diets is then required. With some dietary group definitions, such as that based on the 'eatwell plate distance' quintiles, this identification is relatively obvious, though it should be noted that most of the different dietary groups identified using all methods) included households with a wide range of dietary patterns, so that the mean value of each component per dietary group is not necessarily completely informative about the dietary patterns of those households allocated to each group. However, the 'eatwell plate' "ideal", from which these distances are calculated, is based on established nutritional guidelines, so that the 'eatwell plate distance' provides a potentially valuable tool to identify what a healthy diet looks like. However, while we can compare various attributes of different dietary groups, it is difficult to provide any definitive identification of those household characteristics associated with healthy or unhealthy diets. This is particularly true for the assessment of the costs of different diets, which may also be influenced by the other household characteristics, so that there is inevitably a distribution of costs associated with different diets, even having allowed for household composition and RPI.

While the 'eatwell plate distance' appears to provide a potential approach to defining better and poorer quality diets, it should be emphasised that it is just a statistical tool which has not yet been validated from a nutritional viewpoint. It is not yet entirely clear how to interpret the values of this metric, and therefore how to interpret the analyses presented in this report that are based on this metric. It should also be noted that different dietary patterns will be appropriate for different groups of people (even individuals), so that there is no single "healthy" diet, even if this could be identified using the statistical approaches used in this study. The methods used in this study provide potential approaches to identify groups of households with dietary patterns (nutrition and food intakes) that are similar, but, as described, there are a number of different ways of achieving this, so that a unique set of diet definitions cannot be obtained using statistical approaches. It is also important to note that we cannot draw sensible conclusions about the costs of a diet in isolation, but need to compare the costs between "alternative" diets. The costs of a particular diet will also depend on various household characteristics (particularly household composition and income) so that adjustment for these factors (using the McClement's scale and RPI) is necessary. As demonstrated for different income groups, we can then assess for the impacts of diet on expenditure within different household characteristic groups.

To assess the diet patterns of those at risk of social exclusion (Objective 6), we clearly need to be able to identify groups at risk. Some possibilities are obvious based on the household characteristic data collected, and included within one of the household characteristics models – particularly geographical regions (Model 1), pensioners (Model 2), low income (most models), non-workers (Model 3), single parent households (Model 5), young male households or single gender households (Model 6) and those on income support or non-home-owners (Model 7). However, it is not always easy to discriminate between possible "at risk" groups based on the available household characteristic data, and there is also the potential for confounding between the different household characteristics. In addition, to assess the patterns for those "at risk" we need to draw comparisons with other groups of households – the study again allows us to make comparisons between different groups, but not to identify those patterns as being related to particular causes of social exclusion.

Our analysis approaches potentially provide information on the factors associated with variation in individual dietary components, as well as those associated with different dietary patterns. Many of the household characteristic factors had significant impacts, and while there was potential for confounding between the factors, an assessment of this for a couple of models showed there to be minimal confounding. The initial analyses indicated that there are differences between groups of households, but not how the responses differ – assessment of particular contingency tables and other summary statistics provide the basis for such interpretation, but we consider it important to consider the variability of responses between households within groups rather than just the mean response, which can be achieved by considering the distributions (via contingency tables or more detailed approaches). It should be noted, however, that there may be additional factors influencing the observed distributions of values (see "Limitations" below). Again, it is important to note that we cannot draw conclusions about dietary patterns for groups at risk of social exclusion without comparing these patterns with other groups – it is the relative patterns that are informative, not any absolute patterns. Where clear definitions of "at risk" groups is possible, assessment of the associations with well-defined "unhealthy" diets could yield some valuable results, but both of these sets of definitions are challenging with the existing data.

To assess the potential of using this approach to consider any impact of the economic downturn (Objective 7), our analyses considered changes to mean values and distributions of values (for diets and individual components) for different financial years. It was not possible to draw any conclusions on the impact of the downturn, based on 2008/9 data with the data for the previous years. The recorded expenditure needed to be standardised, to take account of both household composition and inflation (using RPI), and then showed relatively little impact of financial year. Further analysis is needed for data over a longer time-scale to be able to detect any changes as a result of economic downturn, as economists suggest that there would be a considerably longer lag period than is possible given the time-period covered by this study, before evidence of an impact would be seen.

Statistical approaches

A conventional analysis of the data considered in this project would use some form of analysis of variance (such as the hierarchical approach described here) to assess for the significance of potential explanatory factors. However, as shown here, most effects will appear to be highly significant because such analyses are over-powered as a result of the size of the dataset. This means that even very small differences in group means will appear to be highly statistically significant, so providing little illumination on the importance of key factors. The alternative log-linear modelling approach used here considers the whole distribution of values, admittedly often based on a fairly crude classification of the observations, and therefore allows assessment of the more subtle differences that are likely to be needed to address questions of interest. As has been demonstrated, this method has the potential to identify large numbers of significant associations, and so careful interpretation (primarily nutritional and social, rather than statistical) is needed to extract the key information from this vast quantity of summary information. The presented analyses allow a broad exploration of patterns and associations in the data, so that clear identification of the precise questions to be addressed, in terms of both the levels of the household characteristics and the food component and nutrient variables being analysed, is important in applying these analysis techniques to the FFS data. Results from the log-linear analyses should be interpreted using the contingency tables associated with particular significant terms in the analysis model (as shown in various tables included in Annex 2), identifying changes in the distribution of households across the levels of the response variable (food component or nutrient) for different levels of the household characteristic variable. These changes are most clearly seen (see Annex 2) by presenting the numbers of households as percentages of the total number for each level of the household characteristic variable.

The 'eatwell plate distance' (or, more correctly, the distance to the 'eatwell plate' "ideal") has the potential to be a useful tool for measuring the quality of a diet, but further exploration is needed of how to interpret the different values together with a nutritional validation of the approach. It is obvious that there is considerable variation in the sizes of different components for similar values of this distance, and it is too simplistic to expect to be able to define the health of a dietary pattern based on a single index – different dietary patterns are certainly appropriate for different people (as recognised in the definition of the Dietary Reference Values (DRVs) used to categorise nutritional and food component intakes), and so further summaries of the 'eatwell plate' component values may be needed to capture this complexity.

Limitations

One immediate limitation of the available data was that it was impossible to consider all combinations of potentially interesting household characteristics – there were only 52,000 observations available, and just combining financial year (8), season (2), government region (12), income quintile (5), household composition (15), economic status (3), ethnic origin (2) and income support (2) comes to 172,800 combinations, without the further division needed for categories of the response (food component or nutrient) variable. To even consider fitting all of the main effects of the identified household characteristics of interest (the ideal approach for all three Objectives) means constructing the full contingency table, a sparse table, including lots of combinations containing no households, with the construction of this table and subsequent analysis proving too large a problem for the available statistical software (two software packages, GenStat and R, were used in the project, but while both provided access to the generalised linear model facilities necessary for fitting the log-linear models, neither had the data storage capacity to construct and analyse the full contingency table). Hence the approach taken for all three objectives was to consider a number of separate models, with the importance of different household characteristics determined based on the significance of terms fitted in different models. It is also worth noting that there is considerable confounding amongst the different household characteristic factors considered in the analyses, so that even if the full model could be fitted, it is almost certain that many interaction terms could not be fully estimated. The advantage of considering a number of separate sub-models is that these could be tailored to address particular questions about the impacts of household characteristics, though it is important to understand that the results from fitting these different models might be explaining similar sources of variation in the data, and so associations of dietary patterns with particular household characteristics cannot be assumed to be indicative of causal relationships.

As noted in the section above on “Using Statistical Approaches to Explore Ways of Grouping Different dietary Patterns” (relevant to all three Objectives, but primarily to Objective 5), software limitations of data storage capacity also influenced the approach that was able to be taken to use hierarchical cluster analysis to identify groups of households with similar diets. Households were first clustered within financial year, and the resulting clusters then combined on the basis of the mean nutrient/food component values for the clusters. It is highly likely that this will have introduced more variability among the dietary patterns for households included in the different clusters than would have been achieved based on a single hierarchical cluster analysis of all 52.000 households, which is then likely to bias the assessment of any associations between the dietary groups formed and other household characteristics. The extent of this bias is difficult to ascertain, though some overlap of household nutrient/food component profiles between clusters was noted in Annex 2.

A further limitation of the data is associated with the way in which the data were collected. By being based on purchases rather than consumption, it is possible that some household “baskets” will include “bulk buys” of some key components, while other “baskets” may omit key components that have been purchased in bulk in a previous period. This introduces a potential bias to the distributions of nutrient and food components associated with these particular purchases, as well as to the cost associated with the “basket”. It has been suggested that one approach to overcoming this limitation would be to apply the “infrequency of purchase” model to adjust for the potential bias introduced by these bulk buys. However, this would require a substantial number of additional assumptions to be made, possibly unsupported by the available data, which might introduce alternative biases. A better approach might be to modify the FFS data collection protocol to cover a longer period of term (maybe monthly to reflect what is probably a more common purchasing pattern), so that the collected data would be more appropriate for the developed analysis approaches. Without having comparable data collected over such a longer period, it is impossible to determine how much of a limitation this data collection protocol for the FFS is. However, for nutrients and food components associated with bulk-buy products, this bias is likely to result in more variability between households in the estimated levels of consumption. It is not clear how this extra variability would be realised for different socio-economic groups.

It should also be noted that the calculation of household requirements based on DRVs could result in misleading results, as the DRVs provide a description of the distribution of values required by the whole population, rather than providing the range of values that might be needed by an individual. However, we consider that our combination of these threshold values in this way still provides a potentially useful comparison of how the observed purchase values relate to the distribution of household requirements, particularly at the extremes of these requirement distributions. The calculation of the household requirement based on the DRVs of individual household members essentially assumes that all household members have access to the food components and nutrients via the household food supply (proportionate to their requirements?). Other models could have been considered, though more detailed data on individual requirements and food purchases would need to be taken into account (and these data were not available). A particular query was raised about the use of a particular smaller portion size for Fruit and Vegetable portions for children – the use of a half-size portion was based on comments about children having smaller requirements based on body mass, and alternative models could have been considered, particularly if more detailed data on body mass were available. But while these changes might alter the numerical household values for different components, and hence the distribution of households across categories and the significance and interpretation of associations, these changes should not alter the identification of the value provided by the statistical approaches applied to the FFS data.

The translation of food purchases to the different 'eatwell plate' components was made based on a list provided by those responsible for collecting the FFS data, so that any compromises and value judgements used mirror those made by the Defra group responsible for collecting and reporting on the FFS. We have assumed that these are agreed (published) associations, agreed from a nutritional viewpoint, but results relating to the 'eatwell plate' are clearly influenced by the accuracy of these allocations.

A further limitation of the available data relates to the definition of potential "at risk" groups (Objective 6), as these can only be defined based on the available household characteristics data. We found it fairly challenging to separate potential risk groups reliably based on these characteristics, and the potential for confounding between the household characteristics, and hence the definitions of "at risk" groups, means that any analysis can identify potential associations but cannot indicate any causal effects. Possibly the most promising approach to identify groups is based on the identification of "poor" or "unhealthy" diets, though this obviously depends on our ability to define such diets. Certainly the results presented here suggest the need for a more focussed data collection to address this issue associated with identified potential "at risk" factors, rather than using this existing data.

The diet characterisations (dietary groups) used in this study (Objective 5) were based purely on statistical ideas, without any consideration of nutritional aspects. While the multi-stage cluster analysis (another compromise caused by the limitations of the available statistical software, as noted above) provided an objective approach, it was still not informed by nutritional requirements. Given more nutritional information, such as the importance of different dietary components, the clustering could be weighted more appropriately, but this still does not solve the difficulty with providing a single definition of a "healthy" diet (Objective 5). As noted above, a better understanding is needed of the meaning of the 'eatwell plate distance', particularly related to the variability in dietary content for dietary patterns with similar distances, before this approach can be used as a reliable measure of diet quality.

General Conclusions

The presented analyses provide information about statistically significant patterns and trends, a number of which have been presented using contingency tables to illustrate the potential of a new statistical analysis approach to extract additional information from Defra's Family Food Survey/Food Expenditure Survey – but these tables require substantial nutritional/social input to adequately interpret these trends and patterns, and the clear identification of the questions of interest to identify the appropriate analyses to consider. This feasibility project was concerned with the development of statistical approaches that could be applied to the available data to address such issues, rather than with answering particular questions. Hence comparison of the results from our approach with those from other studies will depend on the particular questions being asked, and whether the available data are capable of addressing those questions. As has been emphasised throughout this report, the classifications of households have been developed solely to illustrate the potential of the statistical analysis approach to be applied to the FFS data, with further research needed to validate the appropriateness of these classifications before any firm inferences can be made.

The potential value of the information presented also depends on whether the characterisations and groupings of households are appropriate to the questions of interest, and on the appropriateness of the methods used to classify the dietary components ('eatwell plate', diets defined by the 'eatwell plate distance', classification of households based on an aggregation of individual DRVs). While the 'eatwell plate distance' is an interesting and potentially useful idea, further research is needed to consider the relationship between this index and other evidence on the health impacts of dietary patterns with different values of the 'eatwell plate distance'. Certainly, there is the need for further assessment of the nutritional values associated with dietary patterns with different values of this index.

We consider that our statistical analysis approaches, probably after further refinement and validation of the appropriateness of the household classifications used, have the potential to provide answers to specific questions about the associations between dietary patterns and household characteristics, but we need identification of these specific questions to identify the appropriate results to present, and the collection of appropriate quantities of data (and appropriate household characteristic information, via the FFS) to allow these questions to be addressed reliably. Whether the impacts and trends presented are important requires further nutritional and social interpretation, relative to these more specific questions. While such analyses might indicate statistically significant associations, the identification of causality will certianly require the collection of additional data focussed on particular social groupings.

Specific Conclusions for Objective 5: Explore possibilities of using Defra Family Food survey data (2002 to 2007) to assess the cost of a healthy diet.

Combining statistical approaches to identify groups of households with similar dietary patterns with a classification of the household expenditure (suitably adjusted for household composition) allowed the analysis of potential associations between diet and cost. However, further research into the appropriateness of these household classifications is needed before any firm inferences can be drawn. This statistical analysis approach enables some consideration for the variation in (or distribution of) expenditure for different dietary groups, where a conventional analysis of variance approach would only consider differences in mean values between dietary groups. However, healthy eating guidance recognises that a range of diets fits within the 'eatwell plate' description of a healthy, balanced diet, and does not attempt to define an 'unhealthy'diet. This means that these analyses can only indicate differences in the distributions of costs between different dietary groups, rather than identifying the cost of a healthy diet. Refining of the statistical clustering approach to incorporate additional nutritional information might produce better diet definitions, and the further interpretation of these diets from a nutritional perspective would allow a better assessment of the relative costs of healthy and unhealthy diets.

The analysis approach was further applied to identify potential associations between household characteristics and the costs of particular identified diets, though the size of the data set (and available statistical software) constrained the complexity of the models that could be considered. While these software constraints are likely to limit application of this analysis approach to similar data sets in the future, careful identification of the sub-models to address particular questions, assessment of the partial confounding amongst potential explanatory household characteristic factors, and careful interpretation of the trends indicated for models with partially confounded factors should still allow the approach to provide valuable insights into effects of socio-economic factors on the costs of healthy and unhealthy diets.

The 'eatwell plate distance' (distance from the 'eatwell plate' "ideal") provides a promising approach as an indicator of a healthy diet, but needs further research and refining, particularly as this radial measurement approach allocates similar distances (levels of "healthiness") to diets with very different compositional profiles. Additional summaries of the 'eatwell plate' data, linked to other evidence on the health impacts of different dietary patterns are probably needed alongside this initial approach.

While the Defra Family Food survey data allow the identification of households with different dietary patterns (assuming that purchase data is a good surrogate for consumption data), and hence to assess the costs of these different dietary patterns, further nutritional expertise is needed to interpret the healthiness of different diets, and therefore to assess the cost of a healthy diet.

Specific Conclusions for Objective 6: Explore possibilities of using Defra Family Food survey data (2002 to 2007) to examine patterns in the diets of those at risk of social exclusion.

The development of a log-linear analysis approach allows assessment of the impact of different socio-economic factors on the distribution of the levels of a range of nutrients and food components, where a conventional analysis of variance approach would only consider differences in mean values between levels of these socio-economic factors. However, further research into the appropriateness of these socio-economic classifications is needed before any firm inferences can be drawn. While the comparison of mean values for different groups might be interesting, the size of the data set means that even very small differences are very highly statistically significant, making it difficult to identify the really important differences. By considering impacts on the whole distribution of values, albeit through a classification of the values based on dietary reference values (DRVs) and related ideas, it is evident that there are large overlaps in the distributions of values for different socio-economic groups, as a result of the large number of different factors introducing variability. This analysis approach therefore provides a valuable tool for exploring the FFS to address more specific questions about patterns of dietary components for different socio-economic groups, and therefore of groups at risk of social exclusion, by comparison with other groups, should it be possible to identify such groups using the available socio-economic data collected in the FFS

Potential limitations include the reliability and validity of the socio-economic data collected in the FFS for defining groups at risk of social exclusion, the validity of the assumption that the purchase data collected in the FFS can be used as a surrogate for consumption data, and the validity of using population dietary reference values (DRVs) for individuals within a household to estimate the household requirement and hence identify the relative level of intake.

Combining statistical approaches to identify groups of households with similar dietary patterns with these socio-economic classifications allowed the same analysis approach to be used to assess the impact of a range of difference socio-economic factors on the distribution of households between these different dietary patterns, providing a broader assessment of the differences in dietary patterns between different socio-economic groups (and potentially those at risk of social exclusion). As noted above for Objective 5, the range of diets that may fit within descriptions of a healthier (and hence a less healthy) diet means that analyses can only indicate differences in dietary patterns rather than providing an indication of the absolute healthiness of the dietary patterns adopted by different socio-economic groups. Refining of the clustering approach to incorporate additional nutritional information might produce better diet definitions and hence more reliable conclusions about differences in the healthiness of diets.

The ‘eatwell plate distance’ (distance from the ‘eatwell plate’ “ideal”) provides a promising approach as an indicator of a healthy diet, but needs further research and refining, particularly as this radial measurement approach allocates similar distances (levels of “healthiness”) to diets with very different compositional profiles. Additional summaries of the ‘eatwell plate’ data, linked to other evidence on the health impacts of different dietary patterns are probably needed alongside this initial approach.

While the analysis approaches provide the potential to explore questions about differences in dietary patterns between socio-economic groups using the FFS data, assuming the limitations identified above are not considered critical, the software constraints are likely to similarly limit application of this analysis approach to related data sets in the future. Careful identification of the sub-models to address particular questions will then be required, together with assessment of the partial confounding amongst potential explanatory household characteristic factors and careful interpretation of the trends indicated for models with partially confounded factors. This should then still allow the approach to provide valuable insights into effects of socio-economic factors on dietary patterns.

Specific Conclusions for Objective 7: Make comparisons between Defra Family Food survey data (2002 to 2007) and data for 2008 to explore whether the economic downturn has affected food purchases and thereby the diet of different social groups.

All of the models considered in the log-linear model analysis of associations between socio-economic factors and nutrient and food component classifications included financial year, except for model 8, which was concerned with differences between urban and rural communities for which data were only collected in 2008. For most variables analysed, there was a statistically significant main effect of financial year (see Table 33 for details) with a few variables showing more complex interactions involving financial year and one of the other socio-economic factors (see Table 34 for details). However, further research is needed to determine whether these main effects are of nutritional significance, as well as to determine the appropriateness of the household classifications identified as having statistically significant interactions with financial year. Changes in the distribution patterns between years were explored in more detail for the four variables with the most significant effects (Sodium; Fruit and Vegetables, Bread, Cereals and Potatoes; Vitamin C), shown in Table 35. While the distribution for 2008/9 was often different from the previous financial year, this was either part of a continuing trend across the whole of the time period being studied (such as for Sodium), or the difference between these years was no larger than the differences between other successive years. Table 46 illustrates the interaction effect for financial year and government region on Iron intake, but again there is no clear evidence for 2008/9 being particularly different from all the previous years, just of variation between years. There is no evidence of the impact of the economic downturn from 2008/9. Data on later time-periods is needed to answer this question, which is outside the scope of this research.

Similar analyses assessed the impact of financial year on the distributions of households across the different overall dietary patterns. Again, there was little evidence for 2008/9 having a particularly different distribution across dietary groups (results for the 'eatwell plate distance' shown in Table 79) compared to the earlier financial years, with the last three financial years each showing relatively little difference from the overall distribution. So, again, there is no evidence from these data for a clear impact of the (*start of the*) economic downturn. Further research using later time-periods is needed to answer this question.

Of course, the limitations of the data and analysis methods noted above for Objectives 5 and 6 are also relevant for these analyses, with the concern about the impact of bulk buying possibly more important when considering the potential impacts of the economic downturn. Certainly further data on the impact of the economic downturn on shopping patterns could be important in determining whether the two-week data collection period used in the FFS is likely to provide a reliable representation of food purchase patterns. It is also possible that the models included in this study have missed some key interactions between financial year and other socio-economic factors, possibly because the relevant socio-economic data are not collected in the FFS. Finally, an improved identification of the healthiness of different diets, based on stronger nutritional information, would potentially provide stronger evidence for the relative impact of the economic downturn on both healthy and unhealthy dietary patterns for different social groups.

Key findings:

- This project has demonstrated the feasibility of analysing the Family Food Survey data to identify possible associations between diet and intake (purchase) levels of a range of nutrients and food components, socio-economic factors, and expenditure, based on household aggregated values and considering changes in the distributions of numbers of households between different categories.
- The 'distance' from the 'eatwell plate' "ideal" provides a potentially valuable approach as an indicator of a healthy diet, but needs further refining to better incorporate the nutritional importance of different dietary components, and possibly extension to understand and incorporate additional aspects of how household diets differ from the 'eatwell plate' "ideal".
- However, there are a number of limitations associated with the data used in the study. These include that the Family Food Survey data record the purchases for a household over a two-week period, so that analyses are based on assumptions about these data being representative of longer periods, and also reflecting the food intake of the whole household.
- A further limitation of the approach is the assumption that household food component and nutrient intake (purchase) data can be related to the aggregated requirements of the individuals within the household, based on the population-based thresholds of requirements (which relate to the requirements of particular proportions of the population).
- Applying the developed log-linear analysis approaches to the Family Food Survey data will be most effective if focussed on addressing specific questions (for well-defined socio-economic or dietary groups of households, related to particular dietary components). The hypothetical case studies considered in this project were only used to illustrate the potential of the analysis approach, with identified socio-economic groupings based on statistical features of the data. The potential value of the approach is therefore in the application to the FFS data to address clearly defined research questions.
- It is essential to develop practical quantitative characterisations of healthier and less healthy diets that allow the identification of parameters that can be used to categorise households based on a quantitative assessment of household intakes (purchases) to allow interpretation of associations between different diets and socio-economic factors that might be identified using this analysis method.
- Hierarchical cluster analysis of the households based on dietary intake patterns, to identify common dietary groups, has the potential for further development, and hence improvement of the groups formed, by incorporating further nutritional knowledge and information.
- Annex 2 to this report provides a more detailed exploration and illustration of how the statistical analysis methods developed in this project could be used to analyse the FFS data to address questions about associations between socio-economic classifications and aspects of diet. The project has demonstrated the potential value of the developed statistical analysis approaches, but further research is needed to identify appropriate socio-economic classifications of households and nutritionally meaningful classifications of dietary components before these statistical analysis approaches can be used to draw any firm conclusions.

Knowledge transfer

To date, two papers have been given at academic conferences (following competitive abstract submission), one paper has been published and another has been accepted for publication. All of these outputs are based on the first part of the project (Objectives 1-4).

<u>Conference papers</u>

‘Thinking about “food security”: engaging with UK consumers’ (interim results), Dowler, Collier, Kneafsey, Lambert & Inman, presented at the 2nd International Conference *Food, Society and Public Health,* BSA/Food Study Group, July 5-6th 2010.

‘Consumer perceptions of “food security”: perspectives from the UK’, Kneafsey, Lambie, Dowler, Collier & Inman, presented at the Royal Geographical Society/IBG Annual Conference, theme *Food Security*, Sept 1st -3rd 2010.

<u>Publications</u>

Dowler, E.A., Kneafsey, M., Lambie, H., Inman, A. & Collier, R. (2011). Thinking about ‘food security’: engaging with UK consumers. Critical Public Health 21 (4), 403-416.

Kneafsey, M., Dowler, E., Lambie-Mumford, H., Inman, A., Collier, R. (2012). Consumers and food security: Uncertain or empowered? Journal of Rural Studies, In press.

References to published material

9. This section should be used to record links (hypertext links where possible) or references to other published material generated by, or relating to this project.

Akin, J.S., Guilty, D.K., Popkin, B.M., & Faunally, M.T. (1986) Cluster analysis of food consumption patterns of older Americans. *Journal of the American Dietary Association*, 86, 616-624.

Braunsberger, K., Wybenga, H. & Gates, R. (2007) A comparison of reliability between telephone and web-based surveys, *Journal of Business Research*, 60, 758–764.

Chesher A. (1997) Diet revealed? Semiparametric estimation of nutrient intake-age relationships. *J R Stat Soc A;* 160:389-428.Defra (2008). A framework for pro-environmental behaviours. Report January 2008. http://www.defra.gov.uk/evidence/social/behaviour/documents/behaviours-jan08-report.pdf

DH (1991)Dietary Reference Values for Food Energy and Nutrients for the UK: Report of the Panel on DRVs of the Committee of Medical Aspects of Food Policy (COMA). London: TSO.

DFID/Defra (2010) Policy Narrative on Global Food Security and Sustainable Agriculture. March 2010.

Dowler, E. and Davis, L (2009) Potential for changes in food supply systems to shape consumer demand in different locations and circumstances: a rapid review. Contribution to Defra Project FO0416.

H.M. Treasury (2010) Comprehensive Spending Review. http://www.hm-treasury.gov.uk/spend_index.htm

Hulshof, K.F.A.M., Wedel, M., Lewik, M.R.H, Kok, F.J., Kistemaker, C., Hermus, R.J.J., ten Hoor, F & Ockhuizen, T.H. (1992) Clustering of dietary variables and other lifestyle factors (dutch National Surveillance System). *Journal of Epidemiology and Community Health*, 46, 417-424.IGD (2008) *Adapting to change: shoppers and the credit crunch*. Oct. (online summary and presentation to Council of Food Policy Advisers, Jan 2008)

IGD (2009), *Shopper Trends 2009: Food Shopping in a Recession*. Watford: Institute of Grocery Distribution.

McClements, L.D. (1977). Equivalence scales for children. *Journal of Public Economics, 8*, 191-210.

NatCen (2008) *An assessment of dietary sodium levels among adults (aged 19-64) in the UK general population in 2008, based on analysis of dietary sodium in 24 hour urine samples*. Report to FSA, June, http://www.food.gov.uk/multimedia/pdfs/sodiumreport08.pdf

Nelson, M. & Bingham. S. (1997) Assessment of food consumption and nutrient intake, ch6 in : Margetts, B. & Nelson, M. (eds) 2nd edn, Design concepts in nutritional epidemiology, pp 123-171, Oxford: OUP.

Pryer, J.A & Rogers, S. (2009) Dietary patterns among a national sample of British children aged 1½ - 4½ years. *Public Health Nutrition*, 12, 957-966.

Pryer, J.A., Cook, A. & Shetty, P. (2001) Identification of groups who report similar patterns of diet among a representative national sample of British adults aged 65 years of age or more. *Public Health Nutrition*, 4, 787-795.

Trichopoulou, A., Naska, A. & DAFNE iii group, (2003) European food availability databank based on household budget surveys, *Euro J Pub Hlth*, 13, 3 (suppl) 24-28.

Tucker, K.L., Dallal, G.E. & Rush, D. (1992) Dietary patterns of the elderly Boston-area residents defined by cluster analysis. *Journal of the American Dietary Association*, 92, 487-491.

Vasdekis VGS, Stylianou S, & Naska A. Estimation of age and gender-specific food availability from household budget survey data. *Public Health Nutr* 2001 ;4(5B): 1149-52.