Supplemental Examples

TITLE AND ABSTRACT

Nut-1 State the dietary/nutritional assessment method(s) used in the title, abstract or keywords

Example 2.
Dietary arachidonic and oleic acid intake in ulcerative colitis etiology: a prospective cohort study using 7-day food diaries. (1)

METHODS

Nut-5. Settings: Describe any characteristics of study settings that might affect the dietary intake or nutritional status of the participants, if applicable

Example 2.
From April 1993 through December 1998, a total of 63,257 Chinese women and men aged 45–74 years were enrolled in the study. Study subjects were restricted to the 2 major dialect groups of Chinese in Singapore, that is, the Hokkiens and the Cantonese, who originated from the contiguous provinces of Fujian and Guangdong, respectively, in the southern part of China. Participants were residents of government-built housing estates, where 86% of the Singapore population resided during the enrolment period. (2)

Example 3.
The women were invited to information meetings before the study began, and detailed sampling instructions were given. The importance of careful sampling and not changing dietary habits during the study period was emphasized. We collected duplicate diets to determine the dietary intake of cadmium. Duplicate portions of all foods and beverages, including drinking water, consumed during 4 consecutive days were collected (...). To cover both the inter-seasonal and intraweekly variations in food consumption, samples were collected during four study periods between December 1991 and October 1992. During each (4-day) study period, half of the subjects started sampling on a Sunday and the other half on a Wednesday. (3)

Nut-7.1. Variables: Clearly define foods, food groups, nutrients or other food components

Example 2.
In the Sidama Zone of Southern Ethiopia, maize (Zea mays L.) and fermented enset (Enset ventricosum) products are the major staple foods, contributing up to 90% of energy. (...)
Samples of each of the 2 fermented enset products, kocho and bulla, were purchased from local markets and analyzed for folate and vitamin B-12 content. (4)
Nut-7.2 Variables: When using dietary patterns or indices, describe the methods to obtain them and their nutritional properties

Example 2.
Cluster analysis was performed using a k-means method, an iterative partitioning procedure that attempts to group the data into k clusters in such a way as to maximize the overall R^2 value, defined as R^2=1-W/T, where W is the sum of squared Euclidean distances between each data point and its within-cluster mean (or center), and T is the sum of squared distances between each data point and the overall mean. The k-means methodology is recommended when working with large data sets, and have previously been used in a large number of diet-chronic disease studies. Clustering was based on the 181 energy-adjusted and standardized food frequency variables for k =3–12 clusters. A final number of clusters was chosen based on the stability of large clusters (n>10 000) that were formed, and on the overall R^2 values. When plotting the R^2 values against the number of clusters, six clusters for men and nine clusters for women accounted for most of the increase in R^2 and ensured three stable large clusters for each gender. Four clusters in men and three in women were used in subsequent analyses. The distributions of relative food frequencies and the medians of total energy and energy-adjusted nutrient variables were examined across clusters. The distribution of common risk factors for colorectal cancer was examined by Chi^2 analysis. The Cox proportional hazards regression, with time since entry into the study as the time scale, was used to examine the association between clusters and incidence of colorectal cancer, colon cancer and rectal cancer. The largest cluster (labeled ‘Many foods’ in both men and women) was used as the reference category. (5)

Example 3.
Dietary patterns were derived from 47 food groups with the use of 2 different empirical methods: principal component analysis (PCA) and reduced rank regression (RRR) (SAS PROC FACTOR with varimax rotation and SAS PROC PLS with the RRR method option, respectively). In the RRR analysis, the specified response variables were CRP, IL-6, fibrinogen, and homocysteine (natural log values due to non-normal distribution of original values). For the PCA analysis, a 4-pattern solution was chosen after the evaluation of a scree plot of the eigenvalues and after the interpretability of the final solution was considered. For the RRR analysis, a 4-pattern solution was used in accordance with the number of response variables specified. The 4-factor PCA solution explained 26.2% of the variation in food group intake and 1.5% of the variation in response variables. The 4-factor RRR solution explained 12.6% of the variation in food group intake and 5.0% of the variation in response variables. To present data concisely and because other dietary patterns were not significantly associated with the outcomes, the results are presented for only the primary factor (dietary pattern) derived by the PCA and the primary factor derived by RRR. A score was calculated for each participant for each PCA and RRR dietary pattern as a sum of the 47 food groups, each weighted according to the factor loadings. Participant scores were then categorized into quintiles separately for each dietary pattern. Thus, for each dietary pattern, quintile 5 was composed of persons whose diets conformed most closely to that particular pattern. (6)

Example 4.
The level of adherence to the Mediterranean Diet (MD) was assessed using the adapted relative Mediterranean diet (arMED), an adapted version of the relative MED score, which excludes alcoholic beverages as they are known breast cancer risk factors. The arMED score is based on the score designed by Trichopoulou et al. but consists of a 16-point scale that incorporates eight key components of the MD. For the six components presumed to fit the MD; fruits (including nuts and seeds), vegetables (excluding potatoes), legumes, fish (fresh or frozen, excluding fish products and preserved fish), olive oil and cereals, a score of 0–2 was
assigned to the country-specific tertiles of intake. The scoring was inverted for the components presumed to not fit MD; meat and dairy products. The tertile scoring for olive oil was adapted for non-Mediterranean countries owing to their low consumption, by assigning 0 to nonconsumers, 1 for subjects below the median and 2 for subjects equal or above this median. The points were summed to define the arMED score that ranges from 0 (no adherence) to 16 (maximum adherence) and represents the level of adherence to the MD. (7)

Nut-8.1. Data sources/measurements: Describe the dietary assessment method(s) (e.g., portion size estimation, number of days and items recorded, how it was developed and administered, and how quality was assured). Report if and how supplement intake was assessed.

Example 2.
All examinees were eligible for two 24 hour dietary recall interviews. The first dietary recall interview was collected in-person in the mobile examination centre (MEC) while the second was collected by telephone 3–10 days later but never on the same day of the week as the MEC interview. Dietary interviews were conducted by trained interviewers using the validated United States Department of Agriculture Automated Multiple-Pass Method. For children under 9 years of age, the interview was conducted with a proxy; for children between 6 and 8 years of age, in the presence of the child. Children aged 9–11 years provided their own data assisted by an adult household member (assistant). The preferred proxy/assistant was the most knowledgeable person about the child's consumption on the day before the interview. If the child had more than one caregiver, several individuals could contribute to the intake data. (8)

Nut-8.2. Data sources/measurements: Describe and justify food composition data used. Explain the procedure to match food composition with consumption data. Describe the use of conversion factors, if applicable

Example 2.
In order to account for time-related changes in polychlorinated biphenyls (PCB) concentrations in food, two large recipe-based databases were developed to estimate the dietary exposure to CB-153 (2,2',4,4',5,5'-hexachlorobiphenyl) for the 1997 and the 2004–2006 exposure assessment. Concentration data on PCBs in foods sold on the Swedish market was obtained from regular control and monitoring programs run by the Swedish National Food Agency. When concentration data for the specific time periods was not available, concentrations retrieved for the time period between 1992 and 2009 (over 1200 samples) were used with extrapolation backwards or forward to reflect the concentrations of PCBs in food during the years 1997 and 2004. The extrapolation was based on an average PCB concentration decline of 8% per year, which mirrors the temporal trend of CB-153 (...) To account for changes in PCB concentrations of fried and boiled fish or of cooked meat products, the PCB concentrations (expressed on a lipid weight basis of raw food) were multiplied by the lipid weight of the cooked/processed food. For stews and similar dishes and cooking methods, no loss of fat was assumed and, instead, a factor for the loss of water during cooking was applied to the cooked/processed food. (9)

Example 3.
In order to attribute country-specific nutrient values to each food occurrence, they had to be linked to local Food Composition Table (FCT) items. (...) No new (analysed) data were generated, only data readily available in national FCT were considered. Using information on nutrient values provided by the national compilers, the nutrient values were standardized
as far as possible using references in terms of unit, mode of expression, definition and chemical analysis. To assign nutrient values to occurrences, there could be a direct link with an item of the FCT, or values could be derived from a recipe or by calculating a weighted average. Some algorithms could also be applied to adjust for raw-to-cooked gains and losses and to adjust the fat, sugar and skin contents. Indeed the European Prospective Investigation into Cancer and Nutrition (EPIC) Nutrient Manager (ENMan) software developed for the EPIC Nutrient Database (ENDB) project has been used to document and compute folate values based on the matching information, the recipe and generic item databases, the folate values assigned to the FCT items and a set of algorithms and retention and yield factors in line with European Food Information Resource Network (EuroFIR) guidelines to calculate folate values for cooked foods and for recipes (...). The ENDB folate database can be obtained on request. (10)

**Nut-8.3. Data sources/measurements:** Describe the nutrient requirements, recommendations or dietary guidelines and the evaluation approach used to compare intake with the dietary reference values, if applicable

**Example 2.**
The estimated average requirement (EAR) for adolescents is derived from a formula based on the EAR for adults, weight of the adolescents, weight of adults and a growth factor:

\[
\text{EAR}_{\text{adolescent}} = \text{EAR}_{\text{adult}} \times \left( \frac{\text{weight}_{\text{adolescent}}}{\text{weight}_{\text{adult}}} \right)^{0.75} \times (1 - \text{growth factor}).
\]

The EAR for adults is equal to the recommended dietary allowance (RDA) divided by 1.3 with a coefficient of variation of 15%. Weight\text{adolescent} is derived from the present study; 45.0 kg for boys and 51.3 kg for girls. The World Health Organization reference weight of adults, 65 kg for men and 55 kg for women, is used as weight\text{adult} in the formula, and the growth factor is 15%. (11)

**Nut-8.4. Data sources/measurements:** When using nutritional biomarkers, additionally use the STROBE-ME. Report the type of biomarkers used and usefulness as dietary exposure markers

**Example 2.**
Alkylresorcinols are phenolic lipids present in the bran fraction of wheat, rye, and barley kernels. About half (45%–71%) of alkylresorcinols are absorbed, and intact alkylresorcinols and its metabolites can be measured in plasma. The two main metabolites of alkylresorcinol, 3,5-dihydroxybenzoic acid and 3,5-dihydroxyphenylpropanoic acid, have been suggested as biomarkers of whole-grain intake in Nordic populations. (12)

**Nut-8.5. Data sources/measurements:** Describe the assessment of non-dietary data (e.g., nutritional status and influencing factors) and timing of the assessment of these variables in relation to dietary assessment

**Example 2.**
Self-reported physical activity is assessed using the Short Questionnaire to ASsess Health-enhancing physical activity (SQUASH). The general purpose of this questionnaire is to assess habitual physical activity, with a reference period of a normal week in the past months. Participants are asked to report their average time spend on the following pre-structured types of activities: commuting activities, activity at work, household activities and leisure time activities (walking, bicycling, gardening, odd jobs and up to four sports). The SQUASH consists of three main queries: days per week, average time per day, and intensity. The recorded activity will be converted into Metabolic Equivalent (MET)-scores using the
Compendium of Physical Activities. Validation studies showed that the SQUASH-questionnaire is fairly reliable and reasonably valid in an adult population and may be used to rank participants based on their physical activity level and to categorize them according to the Dutch physical activity guideline (30 minutes or more of at least moderate intense physical activity for a minimum of 5 days per week. (13)

Nut-8.6. Data sources/measurements: Report on the validity of the dietary or nutritional assessment methods and any internal or external validation used in the study, if applicable

Example 3.
Energy and food intake estimated from the 270-item food frequency questionnaire (FFQ) was validated. Energy intake was evaluated against independent measures of energy expenditure using the ActiReg system (PreMed AS) (motion detection), whereas 7 days weighed food records were used to study the relative validity of food and nutrient intake. The correlation coefficient between energy intake and energy expenditure was 0·54. Correlation coefficients between the FFQ and the weighed food records were 0·41 for berries, 0·61 for fruit and 0·38 for vegetables. This FFQ was also validated for ranking individuals according to their usual intake of fruit, juices and vegetables by using the method of triads with two independent and specific biomarkers of fruit and vegetables, the FFQ and 7 days weighed food records. The validity coefficients ranged from 0·60 to 0·94. (14)

Nut-8.6. Data sources/measurements: Report on the validity of the dietary or nutritional assessment methods and any internal or external validation used in the study, if applicable

Example 2.
Correlations for total gram of food and total energy on all plates for each of the five days were all high (r=0.853-0.977, p<0.001 and r=0.874-9.966, p<0.001, for Iceland and Sweden respectively). (...) The modified Bland-Altman plots indicated a slight bias towards underestimation (...) Larger deviations from the mean were observed with increasing amounts (kcal) on the plates. (...) Inter-rater reliability between estimators in each country showed high intra-class correlation coefficients (ICC) (>0.90) for all foods except tomato salsa (ICC = 0.80; 95% CI 0.70–0.87) and bean salad (ICC = 0.74; 95% CI 0.61–0.83) in Sweden. (15)

Nut-9. Bias: Report how bias in dietary or nutritional assessment was addressed (e.g., misreporting, changes in habits as a result of being measured, data imputation from other sources)

Example 2.
The questionnaire also included questions about the health of parents and siblings and a dichotomous variable Health of parents and siblings (ill/healthy) was constructed including high blood pressure, diabetes, and stroke. (...) The variables Previously Ill (ill/healthy) as well as Health of parents and siblings (ill/healthy) were also included as covariates in the final regression analysis. (16)
Nut-11. Quantitative variables: Explain categorization of dietary/nutritional data (e.g., use of N-tiles and handling of non-consumers) and the choice of reference category, if applicable

Example 2.

Table 1
Year 0 characteristics of study sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Year 0 (1985-1986)</th>
<th>Whole-fat milk</th>
<th>SSBs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>Consumer</td>
<td>Nonconsumer</td>
</tr>
<tr>
<td>1865 (51.9)</td>
<td>1731 (48.1)</td>
<td>3425 (95.3)</td>
<td>171 (4.8)</td>
</tr>
<tr>
<td>Demographic characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female (%)</td>
<td>53.5 ± 0.8</td>
<td>50.0 ± 1.2</td>
<td>57.4 ± 1.2*</td>
</tr>
<tr>
<td>Black (%)</td>
<td>47.4 ± 0.8</td>
<td>65.7 ± 1.0</td>
<td>27.7 ± 1.1*</td>
</tr>
<tr>
<td>Age (y)</td>
<td>25.0 ± 3.6</td>
<td>24.5 ± 0.09</td>
<td>25.5 ± 0.1*</td>
</tr>
<tr>
<td>Smoking (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>28.1 ± 0.8</td>
<td>32.9 ± 1.1</td>
<td>23.1 ± 1.0*</td>
</tr>
<tr>
<td>Former</td>
<td>13.1 ± 0.6</td>
<td>10.3 ± 0.7</td>
<td>16.2 ± 0.9*</td>
</tr>
<tr>
<td>Never</td>
<td>58.7 ± 0.8</td>
<td>56.8 ± 1.1</td>
<td>60.7 ± 1.2*</td>
</tr>
<tr>
<td>Physical activity (EU/wk)</td>
<td>429 ± 302</td>
<td>408.8 ± 7.1</td>
<td>451.2 ± 7.1*</td>
</tr>
<tr>
<td>Energy from food (kcal)</td>
<td>2347 ± 1373</td>
<td>2576.6 ± 36.2</td>
<td>2100.5 ± 25.9*</td>
</tr>
<tr>
<td>Anthropometric and blood chemistry measurements</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>24.5 ± 5.0</td>
<td>24.2 ± 0.11</td>
<td>24.3 ± 0.11</td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>77.3 ± 10.9</td>
<td>77.2 ± 0.3</td>
<td>77.3 ± 0.3</td>
</tr>
<tr>
<td>Glucose (mg/dL)</td>
<td>82.2 ± 12.9</td>
<td>81.8 ± 0.3</td>
<td>82.6 ± 0.2</td>
</tr>
<tr>
<td>LDL cholesterol (mg/dL)</td>
<td>108.7 ± 30.7</td>
<td>108.1 ± 0.7</td>
<td>109.3 ± 0.7</td>
</tr>
<tr>
<td>HDL cholesterol (mg/dL)</td>
<td>53.7 ± 13.0</td>
<td>53.9 ± 0.3</td>
<td>53.6 ± 0.3</td>
</tr>
<tr>
<td>Triglycerides (mg/dL)</td>
<td>70.9 ± 45.2</td>
<td>69.5 ± 1.0</td>
<td>72.4 ± 1.2</td>
</tr>
<tr>
<td>Blood pressure (mm Hg)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic</td>
<td>110.3 ± 10.9</td>
<td>111.0 ± 0.3</td>
<td>109.6 ± 0.3*</td>
</tr>
<tr>
<td>Diastolic</td>
<td>68.6 ± 9.6</td>
<td>68.5 ± 0.2</td>
<td>68.6 ± 0.2</td>
</tr>
</tbody>
</table>

1 Values are means ± SDs or percentages ± SEs unless otherwise indicated; n = 3596 (sample used to examine incident metabolic syndrome). SSBs, sugar-sweetened beverages; EU, exercise units.

2 Estimates within beverage group (ie, SSB or whole-fat milk consumer compared with nonconsumer) are significantly different from one another, $P < 0.05$ (chi-square test for dichotomous variables or Student’s $t$ test for continuous variables).

[Reproduced with permission from reference (17)]
Nut-12.2. Statistical methods (b): Describe and justify the method for energy adjustments, intake modelling and use of weighting factors, if applicable

Example 3.
The long-term dietary exposure to phthalates was calculated using the 'Monte Carlo risk assessment' programme, release 7.0, available for registered users at the Netherlands National Institute for Public Health and the Environment website. (...) The resulting distribution of daily phthalate exposures includes both the variation between individuals and between days within individuals. However, to assess the long-term intake within a population only the former type of variation is of interest, since in the long run the intake between different days of one individual will level out. Therefore, the distributions of daily exposures were corrected for the within-person (between days) variation using the betabinomial-normal model. (18)

Nut-17. Other analyses: Report any sensitivity analysis (e.g., exclusion of misreporters or outliers) and data imputation, if applicable

Example 2.
Associations between food pattern groups (FPG) and background factors, nutrient intakes, and results of the health check-up were performed on the whole sample as well as separately for the four groups: (i) Previously Healthy Adequate Reporters (9,376 women, 8,691 men), (ii) Previously Healthy Low Energy Reporters (14,300 women, 14,061 men), (iii) Previously Ill Adequate Reporters (2,311 women, 1,862 men), and (iv) Previously Ill Low Energy Reporters (4,895 women, 3,943 men). Most differences among FPG for Previously Healthy were significant regardless of reporting accuracy, while most differences among FPG for Previously Ill were non-significant. However, in almost all analyses the order with respect to macro- and micronutrient intake between the FPGs within the sexes was identical regardless of subgroup. (16)

Example 3.
Multivariate analyses performed only on complete data cases often lead to the exclusion of a substantial proportion of the original sample, thus limiting the generalization of the models. In our study, the answer ‘Don’t know’ was responsible for all missing data except for blood mercury levels. Therefore, we assumed that our data were missing at random and we used the multiple imputation procedure to cope with missing values. This procedure uses auxiliary variables (i.e. additional variables that are not part of the main model of interest but that might predict missing values) along with variables in the model of interest to replace each missing value with a set of plausible values. The percentage of missing data for variables included in the multiple imputation procedure ranged from 0% to 8.9% (...). For all variables to be included in the regression model, the minimum value was set to zero, whereas the maximum value was set to the maximum observed value plus one standard deviation. (19)

Nut-20. Interpretation: Report the nutritional relevance of the findings, given the complexity of diet or nutrition as an exposure

Example 2.
In this study, we observed a U-shaped association between fish consumption and all-cause mortality, with both high and low levels of fish consumption being associated with higher mortality, especially in women. Similar results were observed for cardiovascular disease specific mortality. The mortality risk associated with low and high fish consumption, although significantly increased, may only translate into a reduction in survival of a few months. (20)
COMPLEMENTARY MATERIAL

Nut-22.1. Ethics: Describe the procedure for consent and study approval from ethics committee(s)

Example 2.
Written consent from parents/carers and assent from participants was obtained prior to entering the study. Ethical approval was granted by the National Health Service Ethics Service (10/H0808/122) and the University of Surrey Ethics Committee (EC/2010/115/FHMS). (21)

Nut-22.2. Online material: Provide data collection tools and data as online material or explain how they can be accessed or why they cannot be provided

Example 2.
Supporting Information Table S1 Individual Data. Excel file containing individual data for adults in the United Kingdom National Diet and Nutrition Survey (men aged 19–64 y and women aged 20–49 y): serum ferritin (mg/L), daily iron intake (mg), calculated quantity of iron absorbed (mg/d) at efficiencies of absorption ranging from 1–25%, estimated physiological requirements of iron to replace obligatory losses (mg/d), and predicted prevalence of inadequate intake at each % absorption efficiency. (22)