**Process for including new food items in price analysis**

**1. Food price data**

**A: *Identify foods to include***

* Download item indices for each of last 10 years from [ONS Website](https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/consumerpriceindicescpiandretailpricesindexrpiitemindicesandpricequotes)
* Identify items that are available in all years
* Remove items that include an element of service, eg, takeaway or hot meal in a pub
* Remove items that contain no nutrients meaningful to the analysis, eg, tea, coffee, mineral water.

***B: Calculate food price***

* Remove prices with a validity < 3 as these prices are unreliable (especially during the COVID-19 pandemic)
* Calculate median price of each item for each quarter and use to produce a mean price for each item per year. **🡪 See document ‘CPI data for original items’**
* Steps described in ‘Syntax calculating median price 2021’

**C: Identify purchased weight**

The CPI data provides a price per unit, with no field for weight. To establish a price per 100g, we must identify the weight of the product purchased in one of the following ways:

* CPI item description - ‘*potatoes-new-per-kg’* as 1000g 🡪 for foods with a range (e.g. potatoes 500g-1000g) use the mean weight (i.e. 750g)
* Online supermarket aggregator – for items such as *individual pizza.* The original analysis used *MySupermarket* to identify items captured by the CPI description. Where more than one is available, choose the one closest to the median price calculated above and use its weight. *MySupermarket* ceased to exist on 01 March 2020. As a replacement, you can use <https://www.trolley.co.uk/>
* USDA National Nutrient database (<https://fdc.nal.usda.gov/>) – for items with variable weights, for example, pineapple, peaches. This database provides a standard reference weight for such items.

**2. Nutrition data**

**A: Obtain National Diet and Nutrition Survey data**

The National Diet and Nutrition Survey (NDNS) collects survey data on foods as consumed by adults in the survey. The nutrient content of foods consumed are reported in detail per portion.

The NDNS data is available to download (<https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=6533&type=Data>) for year 1-9 of the survey. The original analysis used data available from year 1-3.

**B: Process NDNS data**

* Use nutrient bank database to get nutritional information per 100g (i.e. energy in kcal) 🡪 e.g. ndns\_yr11\_nutrientdatabank\_2021-03-19

**3. Linkage process**

The CPI price data and nutritional data are two separate datasets and need to be linked.

**A: Identify appropriate NDNS items**

* Use the description from each CPI item and match to suitable NDNS item(s).
* Merge data on price per 100g and nutritional data

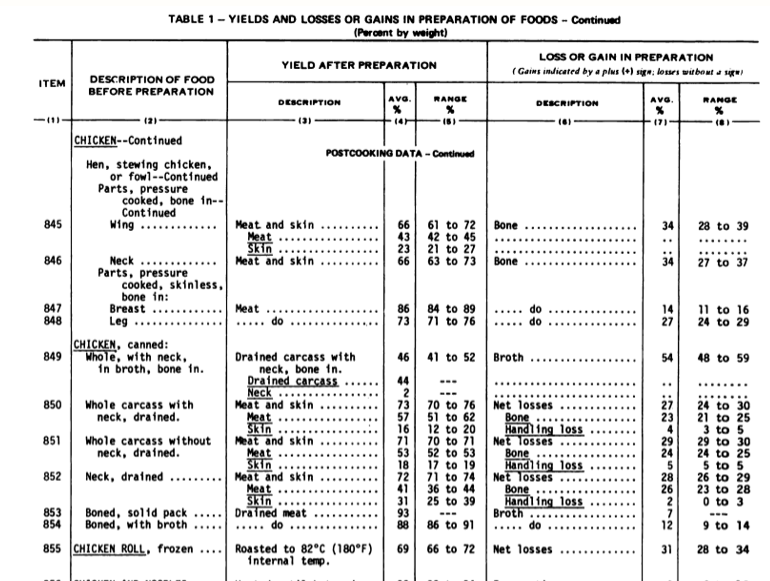
**B: Calculate weighted mean for items with multiple NDNS items**

* The CPI item descriptions can be broad, eg, tinned fruit, and at times, multiple NDNS items, eg, tinned pineapple and tinned peaches can match. If CPI item matches more than 1 item on the NDNS:
  + Identify frequency it was consumed in NDNS dataset and calculate the mean nutrition data weighted by consumption frequency 🡪 Use food consumption data to calculate the number of times a specific food has been selected (e.g. ndns\_rp\_yr11a\_foodleveldietarydata\_uk\_20210831). Run tabs on food\_item\_number to see the number of times items were selected

**Step 3: Adjust price for food yields**

The food price data relates to food items as purchased for example, 100g of raw chicken breast. The nutritional data relates to food as consumed, for example 100g grilled chicken breast and therefore the data needs to be adjusted to account for this difference.

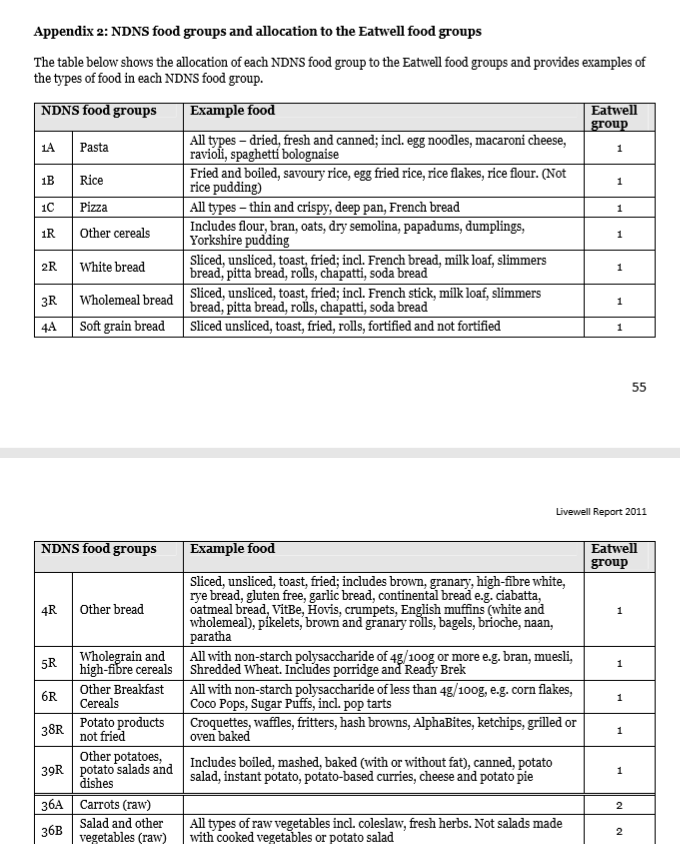
* Use *US Department of Agriculture handbook 102: Food yields (*<https://babel.hathitrust.org/cgi/pt?id=uc1.b3614369&view=1up&seq=1> - *extract below)*
* For each food, identify the food’s yield % after preparation from the handbook.
* Use % food yield to calculate the edible purchased mass. Using the example below, the yield percentage of a chicken breast is 86%, therefore when cooking 1000g of raw chicken, the cooked chicken breast is 860g.
* The edible purchased mass is used to calculate the price per 1000kcal.



**4. Food group-based classification of foods**

To generate a graph according to the five categories of the eatwell plate, each food must be assigned to an Eatwell category:

* Using reference table on page 55 of livewell report (extract below and attached) assign each food to an eatwell food group based on the NDNS data
* Report can be found here: <http://assets.wwf.org.uk/downloads/livewell_report_corrected.pdf>



**5. Nutrient based classification of foods**

To generate a graph according to the Food Standards Agency categorisation of more and less healthy, each food must be assigned to a category. Use the nutrient profiling technical guidance (attached) and the nutritional data for each item to calculate the score and assign to the more/less healthy categories. See the report ‘Nutrient Profiling Technical Guidance’ from 2011.