Supplementary material A1 Calculation R^2, C_j', and F_{ij}

The R^2 of a nutrient is obtained from the linear regression of total nutrient intake on the nutrient intakes of all food items in the food list:

$$Z_i = \beta_0 + \beta_1 F_{i1} + \beta_2 F_{i2} + \cdots + \beta_j F_{ij} + \varepsilon_i \text{ ,}$$

where Z_i is the total nutrient intake of individual i (i = 1,…, I) and F_{ij} is individual i's intake of the nutrient from food list item j (j = 1,…, J). F_{ij} is calculated as the multiplication of the amount of the item that is reported with the nutrient content of that item. However, since a food list will contain single foods at level 5 and aggregated items at the other levels, F_{ij} may be either the nutrient intake of individual i from a single food or the nutrient intake from an aggregated item.

For single foods the F_{ij} is calculated as the multiplication of the reported amount of the food (A_{ij}) and C_j, which is the nutrient content (gram per gram) as reported in the food composition table: $F_{ij} = A_{ij} \times C_j$.

The nutrient content of an aggregated item is not reported in the food composition table, but has to be calculated via the single foods that are taken into account by this aggregated item j. Therefore for aggregated food items the weighted composition C_j' is used, with weights derived from the observed consumption:

$$C_j' = \sum_{k \in K_j} \left(C_k \times \frac{\sum_{i=1}^{I} A_{ik}}{\sum_{\kappa \in K_j} \sum_{i=1}^{I} A_{i\kappa}} \right)$$

where K_j is the set of single foods that are taken into account by aggregated item j, and $\sum_{i=1}^{I} A_{ik} / \sum_{\kappa \in K_j} \sum_{i=1}^{I} A_{i\kappa}$ is the fraction of the amount of aggregated item j attributable to single food k in the whole population. Then for an aggregated item j the F_{ij} is calculated as

$$F_{ij} = A_{ij} \times C_j'$$

The R^2 is then defined in the usual way:

$$R^2 = \frac{\sum_{i=1}^{I} \left(\hat{Z}_i - \overline{Z}\right)^2}{\sum_{i=1}^{I} \left(Z_i - \overline{Z}\right)^2} \text{ ,}$$

where $\hat{Z}_i$ is the predicted nutrient intake of individual i, $\overline{Z}$ is the mean observed nutrient intake in the population and Z_i is the observed nutrient intake of individual i.

Example of calculation of C_j' and F_{ij} of an aggregated food item

Suppose we have a food tree that contains an aggregated food item NON-CITRUS (referred to as j), constructed by aggregation of four single food items APPLE, PEAR, BANANA, KIWI (referred to as k = 1,...,4), see Figure A1.

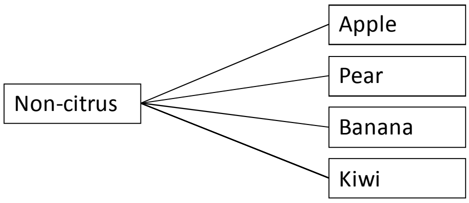

Figure A1. Example of a food tree that contains one aggregated item NON-CITRUS, constructed by aggregation of four single foods APPLE, PEAR, BANANA, KIWI.

Suppose that we have a population of 3 individuals i. Table A1 shows how much individual i has consumed of (single) food item k (A_{ik}), and it shows how much the food items k contain of the nutrient of interest (C_k). For example, individual i=1 consumed 300g of pear, and pear contains 30% of the nutrient of interest.

Table A1. Example of calculation of the nutrient content of aggregated food item NON-CITRUS, where A_{ik} denotes how much individual i has consumed of (single) food item k, and C_k is the nutrient content of food item k.

		A_{ik}	i=1	i=2	i=3	$\sum_{i=1}^{I} A_{ik}$	C_k
NON-CITRUS:	k=1	APPLE	150	250	0	400	0.40
	k=2	PEAR	300	200	100	600	0.30
	k=3	BANANA	350	250	400	1000	0.25
	k=4	KIWI	200	100	200	500	0.20
Consumption of NON-CITRUS:			1000	800	700	2500	

Total consumption of APPLE in the population can be calculated as

$$\sum_{i=1}^{I} A_{i1} = 150 + 250 + 0 = 400 .$$

Total consumption of NON-CITRUS in the population can be calculated as

$$\sum_{k \in Kj} \sum_{i=1}^{I} A_{ik} = (150 + 250 + 0) + ... + (200 + 100 + 200) = 400 + ... + 500 = 2500$$

and in this population the fraction of NON-CITRUS consumption attributable to APPLE is

$$\frac{\sum_{i=1}^{I} A_{i1}}{\sum_{k \in K_j} \sum_{i=1}^{I} A_{ik}} = \frac{400}{2500} .$$

So in this population the nutrient content of aggregated food item NON-CITRUS is

$$C_j' = \sum_{k \in K_j} C_k \frac{\sum_{i=1}^{I} A_{ik}}{\sum_{\kappa \in K_j} \sum_{i=1}^{I} A_{i\kappa}} = 0.40 \cdot \tfrac{400}{2500} + 0.30 \cdot \tfrac{600}{2500} + 0.25 \cdot \tfrac{1000}{2500} + 0.20 \cdot \tfrac{500}{2500} = 0.276 ,$$

and the nutrient intake of individual i via NON-CITRUS has to be calculated as

$$F_{ij} = A_{ij} \times C_j'.$$

For example, for individual 1 this is F_{ij} = (150+300+350+200)×0.276 = 276g.