European Journal of Clinical Nutrition (2011) 65, S92–S101

0954-3007/11
www.nature.com/ejcn

ORIGINAL ARTICLE

Comparing four methods to estimate usual intake distributions

OW Souverein[1], AL Dekkers[2], A Geelen[1], J Haubrock[3], JH de Vries[1], MC Ocké[2], U Harttig[3], H Boeing[3] and P van 't Veer[1], on behalf of the EFCOVAL Consortium

[1]*Division of Human Nutrition, Wageningen University, Wageningen, The Netherlands;* [2]*National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands and* [3]*Department of Epidemiology, German Institute of Human Nutrition, Potsdam-Rehbruecke, Germany*

Background/Objectives: The aim of this paper was to compare methods to estimate usual intake distributions of nutrients and foods. As 'true' usual intake distributions are not known in practice, the comparison was carried out through a simulation study, as well as empirically, by application to data from the European Food Consumption Validation (EFCOVAL) Study in which two 24-h dietary recalls (24-HDRs) and food frequency data were collected. The methods being compared were the Iowa State University Method (ISU), National Cancer Institute Method (NCI), Multiple Source Method (MSM) and Statistical Program for Age-adjusted Dietary Assessment (SPADE).
Subjects/Methods: Simulation data were constructed with varying numbers of subjects (n), different values for the Box–Cox transformation parameter (λ_{BC}) and different values for the ratio of the within- and between-person variance (r_{var}). All data were analyzed with the four different methods and the estimated usual mean intake and selected percentiles were obtained. Moreover, the 2-day within-person mean was estimated as an additional 'method'. These five methods were compared in terms of the mean bias, which was calculated as the mean of the differences between the estimated value and the known true value. The application of data from the EFCOVAL Project included calculations of nutrients (that is, protein, potassium, protein density) and foods (that is, vegetables, fruit and fish).
Results: Overall, the mean bias of the ISU, NCI, MSM and SPADE Methods was small. However, for all methods, the mean bias and the variation of the bias increased with smaller sample size, higher variance ratios and with more pronounced departures from normality. Serious mean bias (especially in the 95th percentile) was seen using the NCI Method when $r_{var}=9$, $\lambda_{BC}=0$ and $n=1000$. The ISU Method and MSM showed a somewhat higher s.d. of the bias compared with NCI and SPADE Methods, indicating a larger method uncertainty. Furthermore, whereas the ISU, NCI and SPADE Methods produced unimodal density functions by definition, MSM produced distributions with 'peaks', when sample size was small, because of the fact that the population's usual intake distribution was based on estimated individual usual intakes. The application to the EFCOVAL data showed that all estimates of the percentiles and mean were within 5% of each other for the three nutrients analyzed. For vegetables, fruit and fish, the differences were larger than that for nutrients, but overall the sample mean was estimated reasonably.
Conclusions: The four methods that were compared seem to provide good estimates of the usual intake distribution of nutrients. Nevertheless, care needs to be taken when a nutrient has a high within-person variation or has a highly skewed distribution, and when the sample size is small. As the methods offer different features, practical reasons may exist to prefer one method over the other.
European Journal of Clinical Nutrition (2011) **65,** S92–S101; doi:10.1038/ejcn.2011.93

Keywords: diet surveys; statistical models; humans

Introduction

The European Food Consumption Validation (EFCOVAL) Project aimed at the further development and validation of the 24-h dietary recall (24-HDR) software EPIC-Soft (the software developed to conduct 24-HDRs in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study) (Slimani *et al*., 1999), in order to be used for estimation of the intake of foods, nutrients and potentially hazardous chemicals within the European adult population. The 24-HDR was selected by the European Food Cosumption Survey Method (EFCOSUM) Project as the most suitable method to obtain comparable data in pan-European dietary

Correspondence: Dr OW Souverein, Division of Human Nutrition, Wageningen University, PO Box 8129, 6700 EV Wageningen, The Netherlands.
E-mail: Olga.Souverein@wur.nl

surveys (Brussaard et al., 2002). A single 24-HDR per individual can be used to characterize the average consumption of a group of individuals fairly well, provided the sample is representative of the underlying population and all days of the week and all seasons are equally represented (Biro et al., 2002; Kroes et al., 2002). However, the variance of the usual group intake is inflated by day-to-day variation in individual intake, resulting in misleading estimates of the prevalence of low or high intakes (Mackerras and Rutishauser, 2005). If repeated 24-HDRs are collected, it is possible to eliminate the intra-individual variability of the data and thereby to obtain an estimate of the usual intake distribution of the population. Several statistical procedures for estimating the usual intake distribution from repeated 24-HDRs are available (National Research Council, 1986; Slob, 1993; Wallace et al., 1994; Buck et al., 1995; Nusser et al., 1996a; Hoffmann et al., 2002), including a number of recently developed methods (Haubrock et al., 2011; Slob, 2006; Tooze et al., 2006; Waijers et al., 2006).

As the estimates of usual intake distributions are one of the main outcomes of food consumption surveys, it is important to know how the different methods for estimating usual intake distributions perform. In 2002, Hoffmann et al. (2002) published a comparison of six different methods to estimate usual intake distributions. As then new methods have been developed, which have not been compared so far. Only the National Cancer Institute Method (NCI) and Multiple Source Method (MSM) were compared on a number of aspects by Haubrock et al. (personal communication). Nevertheless, there are differences between the methods in, for instance, the statistical models, the assumptions and the back-transformations used. Therefore, the objective of the present study was to compare recently developed methods. The NCI Method (Tooze et al., 2006), MSM (Haubrock et al., 2011) and Statistical Program for Age-adjusted Dietary Assessment (SPADE) (Dekkers et al., in preparation) were included, and for reference, the methods were also compared with the 2-day within-person mean (WPM) and the Iowa State University Method (ISU) Method (Nusser et al., 1996a), which is one of the most commonly used methods. The comparison was made through a simulation study, which allows comparison of the estimated distributions to a known theoretical true distribution, and empirically by application of the different methods to data from the EFCOVAL Study. The simulation study was limited in scope, only investigating whether the methods were at least able to perform adequately in simple situations.

Materials and methods

Methods to estimate usual intake distributions

All methods for estimating usual intakes have the following global approach for nutrients and frequently (almost everyday) consumed foods. First, a transformation step is used to obtain more or less symmetrically distributed data. In a second step, the mean usual intake on the transformed scale is estimated (sometimes as a function of age or for different age classes), as well as the within- and between-person variance. In the last step, the within-person variance is eliminated and the results are back-transformed to the original scale, resulting in the usual intake distribution. Furthermore, these methods assume that 24-HDR intake is an unbiased estimator of usual intake in the untransformed scale (see also the review of Dodd et al., 2006).

To be able to estimate the usual intake distribution for episodically consumed foods, the consumption frequency distribution needs to be estimated. Episodically consumed foods are foods that are not consumed everyday, and therefore short-term measurements of intake of these foods contain many zero consumptions. The final usual intake distribution of these foods is obtained by combining the estimated consumption frequency distribution and the usual intake distribution based on positive observations (that is, the amounts consumed).

The Box–Cox transformation is used by most methods in the transformation step, but with different definitions. In this paper, we use the following definition: let x represent the original intake data, with obviously $x \geqslant 0$, and let y represent the Box–Cox transformed data, then $y = (x^{\lambda} - 1)/\lambda$, for $\lambda \neq 0$ and $y = \log(x)$, for $\lambda = 0$.

Iowa State University Method

The ISU Method was implemented in the SIDE (Nusser et al., 1996a; for nutrients and foods consumed daily) and C-SIDE programs (Nusser et al., 1996 b; for episodically consumed foods). The method allows for initial adjustments for nuisance effect, such as day of the week, month, interview mode or interview sequence, and uses a power transformation to make the distribution of the observed data nearly symmetric.

National Cancer Institute Method

The NCI Method (Tooze et al., 2006), which was developed specifically to deal with episodically consumed foods, consists of a two-part model with correlated person-specific effects. The first part consists of estimating the probability of consuming a food, whereas the second part specifies the consumption-day amount. This method can also be used to analyze nutrients and foods that are consumed daily by running only the second part of the model (the amount model). The back-transformation used is an approximation of the exact formula. The NCI Method has been implemented in SAS macros, which were downloaded from the website (http://riskfactor.cancer.gov/diet/usualintakes/macros.html; 9 September 2008). Covariates such as age and body mass index maybe included in the model. Moreover, additional information from a food propensity questionnaire (FPQ) maybe included in the model as a covariate. Since 2009, the estimation of individual usual intakes is possible with an additional SAS macro that is available on the website (Kipnis et al., 2009).

Multiple Source Method

MSM (https://nugo.dife.de/msm; user guide available) aims at estimating the population's usual intake distribution by estimating the individual usual intakes. First, a consumption probability is estimated using logistic regression (step 1), and consumption-day amount is estimated using linear regression (step 2) for each individual. In both steps, a Box–Cox transformation is used to achieve a symmetric residual distribution. The back-transformation consists of an exact integral solution for non-negative integers of the Box–Cox parameters (that is, $\lambda = 1/n$, $n = 1, 2, \ldots$). Second, estimates of the individual usual intakes are obtained by multiplying consumption probability and consumption-day amount (step 3). Parameters of the usual intake distribution of the population are calculated directly from the distribution, which consists of the estimated individual usual intakes. Again, this method can also be used to analyze nutrients and foods that are consumed daily by running only the part of the model for consumption-day amount.

MSM was developed within the framework of the EFCOVAL Project and implemented at a website (https://nugo.dife.de/msm). Covariates maybe included in the regression models of steps 1 and 2. By definition, MSM provides estimates of individual usual intakes. Additional information from an FPQ maybe used in two ways. First, FPQ information can be used to identify consumers among those reported as non-consumers in the 24-HDRs. For those consumers, consumption amount is estimated through simulation based on covariate information (see user guide for detailed information). Second, FPQ information may be added as a covariate in the regression models. As an additional feature, external information on the proportion of non-consumers of a food can also be handled by MSM.

Statistical program to assess dietary exposure

The first part of SPADE estimates the consumption frequency distribution using a logistic regression with an individual-specific random effect. The second part of SPADE uses only positive intakes and is a new version of Agemode (Waijers *et al*., 2006). This part of SPADE starts with a one-parameter Box–Cox transformation to obtain symmetrically distributed data. Thereafter, the mean usual intake is modeled as a constant or as a function of age by fractional polynomial regression, and also the within- and between-individual variations are estimated (Waijers *et al*., 2006). Once more, nutrients and foods that are consumed daily may be analyzed by running only the second part. The back-transformation is based on a Gaussian Quadrature (Dekkers *et al*., personal communication) for integrating out the within-individual variance, which is much faster than the back-transformation by simulation in Agemode (Waijers *et al*., 2006). In the third step, the distribution of the usual intakes is obtained by joining the two distributions obtained in the first two steps together by simulation. The model is implemented in R. More detailed information can be found in the user guide (Dekkers *et al*., in preparation).

Data of the simulation study

A simulation study was conducted to investigate the extent to which the estimated usual intake distributions obtained with the different methods deviated from a theoretical true usual intake distribution in different situations. To create the simulated data, the overall mean (μ), the between-person standard deviation (s.d.$_b$), the within-person standard deviation (s.d.$_w$), the lambda for the inverse Box–Cox transformation (λ_{BC}, where $\lambda_{BC} = 0$ is a log-normal distribution and $\lambda_{BC} = 1$ is a normal distribution), and the number of subjects (n) were set. The ranges of the λ_{BC} and variance ratio ($r_{var} = s_w^2/s_b^2$) were based on what was found in the EFCOVAL Study. Eight different scenarios were investigated: (1) to compare the effect of study size, the number of individuals was set to either 1000, 500 or 150 (with $\lambda_{BC} = 0.0$ and $r_{var} = 1.0$); (2) to compare the effect of the 'skewness' of the intake data, λ_{BC} was set to either 0.0, 0.2 or 0.5 (with $n = 1000$ and $r_{var} = 1.0$); and (3) to compare the effect of the variance ratio, this ratio was set to 0.25, 1, 4 or 9 (with $n = 1000$ and $\lambda_{BC} = 0.0$). For each scenario, a total of 100 simulations were analyzed.

To obtain the simulated data, a mean intake on the transformed scale for each of n individuals was obtained from a normal distribution with μ and s.d.$_b$. Given this individual mean intake, 2 days of intakeon the transformed scale were obtained for each individual by drawing twice from a normal distribution with s.d.$_w$. The s.d.$_w$ was assumed to be equal for each individual. Next, λ_{BC} was used to obtain the intake of 2 days for each individual on the original scale. These data were then analyzed using each of the methods. Given μ, s.d.$_b^2$, s.d.$_w^2$ and λ_{BC}, estimates of the percentiles and mean of the true usual intake distribution on the 'original'/untransformed scale were obtained by applying Gaussian Quadrature. The simulation data were created using R version 2.8.1.

The ISU, NCI, MSM and SPADE Methods were applied to the 100 data sets per scenario, and for each the estimated mean and percentiles of the usual intake distribution were obtained. Moreover, the mean of the 2-day WPM was obtained for the simulated data, which represents the sample mean. For each of the eight different scenarios, these means and the 5th, 50th and 95th percentiles were compared in terms of mean bias, which is calculated as the mean of 100 differences between the estimated value and the true value. Moreover, the 95% confidence interval (CI) for the bias was calculated. To illustrate the estimated usual intake distributions, the density plots were drawn. For the ISU Method, density estimates were provided as output of the analysis, and therefore these were plotted directly. For NCI and SPADE, the 100 percentiles of the usual intake distribution were obtained. For successive percentiles P_x, P_{x+1} ($x = 1, 2, \ldots 98$), 100 values were simulated from the uniform distribution on the interval $[P_x, P_{x+1}]$. Density plots were drawn from these simulated individual values. For MSM, the density plot was based on the estimated individual usual intakes. These plots were constructed using R version 2.8.1.

Data of the EFCOVAL Study
To illustrate the different methods in a real-life setting, the usual intake distributions of several nutrients and foods were calculated using data from the male participants of the EFCOVAL Study (Crispim *et al.*, 2011). This study was designed to evaluate the validity of two independent days of 24-HDRs using EPIC-Soft for assessing food and nutrient intake within countries in Europe, and for comparisons between these countries. A total of 600 healthy volunteers were recruited from five European countries, namely the Netherlands, Belgium, France, Norway and Czech Republic. Subjects were men and women aged between 45 and 65 years. Each participant was scheduled to complete two computer-assisted 24-HDR interviews and an FPQ. The 24-HDRs were conducted in each country by trained dieticians with use of the EPIC-Soft Program (Slimani *et al.*, 1999). The two recalls were scheduled to take place at least 1 month apart. The FPQ consisted of four multiple-choice questions on the frequency of usual consumption of vegetables, fruit, fish and alcohol. The study was approved by the ethics committees of each participating institution, and written informed consent was obtained from each participant.

For each method and the WPM, the mean and the 5th, 10th, 25th 50th, 75th, 90th and 95th percentile of the estimated usual intake distribution, as well as the variance ratio and the λ_{BC} were obtained for protein (g/day), potassium (mg/day), protein density (that is, protein/energy; g/MJ/day, based on the ratio of daily intakes (Freedman *et al.*, 2010)), vegetables (g/day), fruit (g/day) and fish (g/day). No covariates were included in these analyses. Analyses were performed for men from the five countries together. A total of 297 men were included in the study, of which 293 completed both 24-HDRs. As the incorporation of FPQ data have been suggested to lead to an improvement of the estimated distribution (Subar *et al.*, 2006; Tooze *et al.*, 2006), the usual intake distributions of vegetables, fruit and fish were estimated with and without incorporation of the FPQ data for the NCI method and MSM. Without the individual FPQ information, both these methods consider all participants as consumers by default. The categorical data of the FPQ were recoded into consumption frequency per day (that is, never = 0; 1 day per month or less = 0.033; 1–3 days per month = 0.066; 1 day per week = 0.143; 2–3 days per week = 0.357; 4–5 days per week = 0.643; 6–7 days per week = 0.929). One subject had not completed the FPQ, and therefore 292 men were available for the analysis of vegetable and fish intake.

Results

Comparison for nutrients
Figure 1 shows the mean bias with 95% CI of the 5th, 50th and 95th percentiles, and the mean as estimated with the different methods for the scenarios with $n = 1000$, $n = 500$ and $n = 150$. In addition, the mean bias (95% CI) of the mean calculated with the 2-day WPM is shown. This mean bias in the estimated usual mean intake for the 2-day WPM is −0.11,

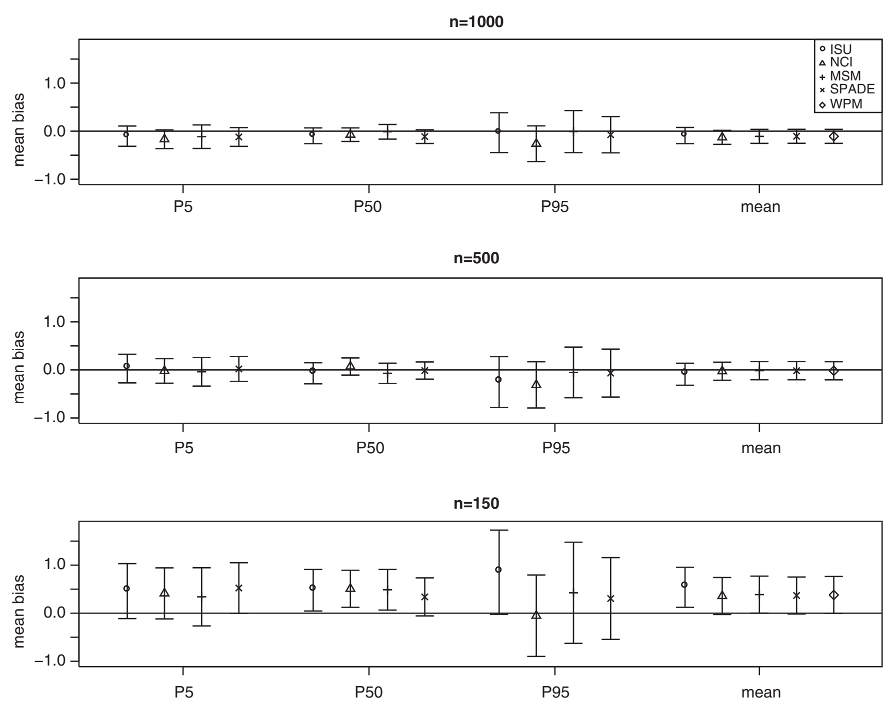

Figure 1 Mean bias for different sample sizes as obtained with the methods. The error bars indicate the 95% confidence interval for the bias. $\lambda_{BC} = 0$ and $r_{var} = 1$ for all scenarios in this figure.

−0.02 and 0.38, which is the average deviation from the true mean for this sample of 100 simulations with $n=1000$, $n=500$ and $n=150$, respectively. As expected, the distribution estimated with WPM is much wider than the true usual intake distribution, and therefore, this is not shown in the figure. For all sample sizes, NCI, MSM and SPADE performed similarly with mean biases close to the average deviation in the sample. The mean bias of ISU seems somewhat higher for $n=500$ and $n=150$. For $n=1000$ and $n=500$, none of the mean biases differed from zero, as indicated by the 95% CI. However, for ISU, the mean bias was significantly different from zero (0.54 (0.12–0.96)) for $n=150$.

The width of the CI indicates the variation in bias and a wider CI, thus, indicates that a method is more uncertain compared with another method. Overall, the CI is largest in the P95 (mean s.d. for $n=1000$:2.10) and smallest in the mean (mean s.d. for $n=1000$:0.77) and median (mean s.d. for $n=1000$:0.78). Moreover, the s.d. is larger for $n=150$ compared with $n=500$, and the s.d. is smallest for $n=1000$ (mean s.d.:2.14, 1.04 and 0.78, respectively). The s.d. was somewhat larger for the ISU Method and MSM compared with NCI and SPADE (for example, for $n=1000$, P50:0.84 (ISU), 0.78 (MSM), 0.72 (NCI) and 0.73 (SPADE)).

The performance of the methods for different variance ratios is plotted in Figure 2. The results of the ISU, MSM and SPADE Methods do not indicate serious bias for any of the investigated ratios. Nevertheless, the s.d. of the bias increases when the variance ratio increases (for example, average s.d. for P50:0.69 ($r_{var}=0.25$), 0.78 ($r_{var}=1$), 1.08 ($r_{var}=4$) and 1.80 ($r_{var}=9$)). However, the mean bias for the NCI Method with $r_{var}=9$, $\lambda_{BC}=0$ and $n=1000$ was −1.34 (95% CI:−1.65; −1.04) for P50, and −5.00 (95% CI:−7.58; −2.43) for P95. In addition, the mean usual intake of the population estimated with the NCI Method was biased for this scenario (mean bias:−1.51, 95% CI:−1.79 to −1.23), as well as for the scenario where $r_{var}=4$, $\lambda_{BC}=0$ and $n=1000$ (mean bias:−0.27, 95% CI:−0.44 to −0.09).

For $\lambda_{BC}=0.5$, mean bias was 0.00 in almost all cases, except for P5 and P50 for the ISU Method that showed a very small mean bias of −0.01. The s.d. of the bias was also small, ranging from 0.06 for P95 to 0.03 for P50 and the mean. For $\lambda_{BC}=0.2$, the mean bias was also small, ranging from 0.01 to 0.03 for P5 and from −0.01 to −0.03 for P95. In this scenario, the sample deviation was 0.01. The s.d. of the bias was somewhat larger than those found for $\lambda_{BC}=0.5$, ranging from 0.08 for NCI and SPADE to 0.10 for ISU and MSM (P50). Again, the s.d. was larger for P5 and P95 (average s.d., 0.16 for P5 and 0.21 for P95), but ISU and MSM showed consistently larger s.d's than NCI and SPADE.

Figure 3 shows the usual intake distributions as estimated with each method as well as the intake distribution of the original data. In the plot entitled 'simulation', the dashed–dotted line represents the true usual intake distribution with $\lambda_{BC}=0$ and $r_{var}=1$, the dotted line represents the density function of all the simulated 'individuals' that were used in the simulations with $n=1000$ and $n=150$ put together

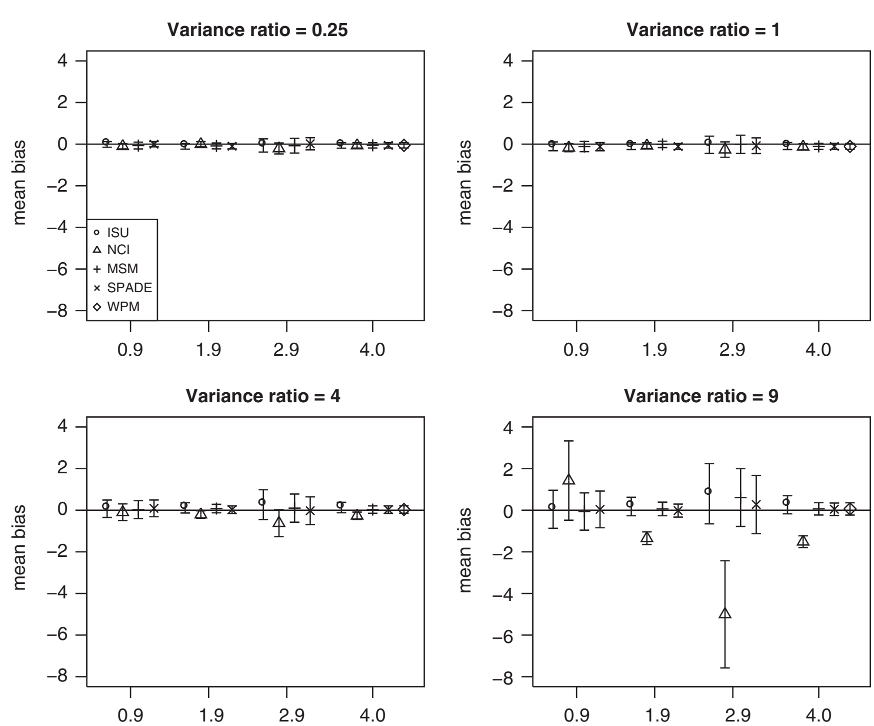

Figure 2 Mean bias for different variance ratios as obtained with the methods. The error bars indicate the 95% confidence interval for the bias. $n=1000$ and $\lambda_{BC}=0$ for all scenarios in this figure.

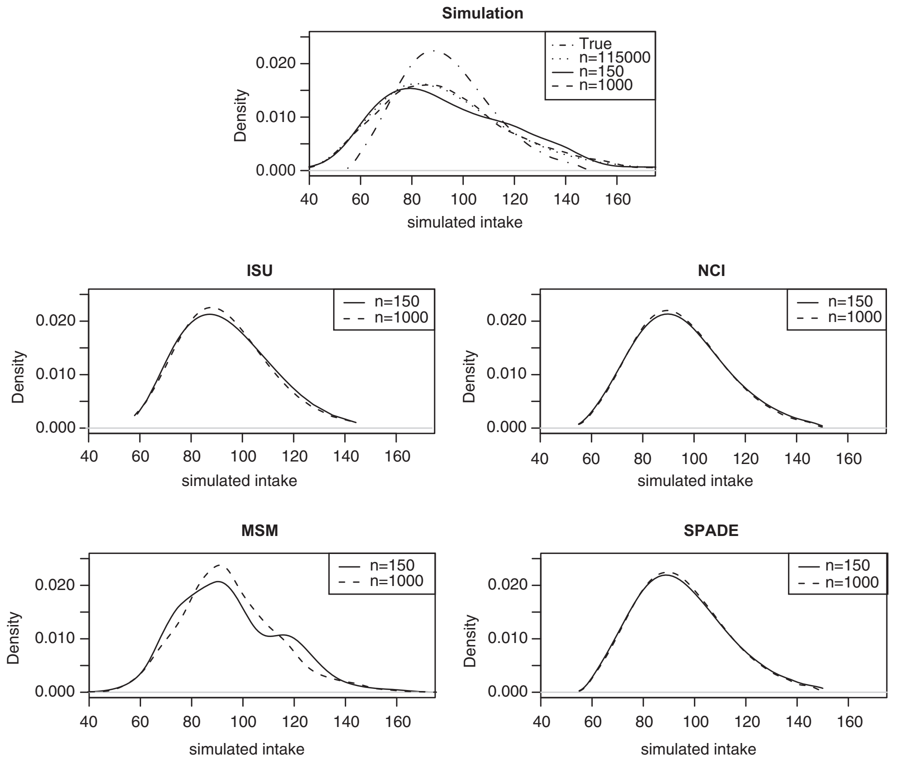

Figure 3 Smoothed usual intake distributions of the first simulation set of scenarios with parameters $\lambda_{BC}=0$, $r_{var}=1$ and $n=150$ or $n=1000$ estimated with the different methods. The usual intake distribution of MSM is based on estimated individual usual intakes. In the graph 'Simulation', the dotted line represents the density function of all the simulated 'individuals' that were used in the simulations with $n=1000$ and $n=150$ put together ($n=(1000+150)\times 100=115\,000$).

($n=(1000+150)\times 100=115\,000$), the dashed line is the first original sample of $n=1000$ and the solid line is the first original sample of $n=150$. The distribution of the original samples is much wider than the true, as these samples contain within-person variation as well as between-person variation. In addition, the original sample of $n=150$ is not as unimodal as those of $n=1000$ or $n=115\,000$. The density plots obtained with ISU, NCI and SPADE Methods are unimodal curves that closely resemble the true usual intake distribution. This is also true for MSM when $n=1000$, although this density function seems somewhat wider than the true usual intake distribution. However, the curve from MSM, based on estimated individual intakes for $n=150$, shows the same 'peaks' that are also present in the original sample.

Table 1 presents the percentiles and means of the estimated usual intake distributions as well as the estimated variance ratios and λ_{BC} of protein, protein density and potassium obtained with the different methods using data of the EFCOVAL Study. The estimated usual intake distributions are very similar, with all estimates of the percentiles within 5% of each other. The largest relative differences are found in potassium intake between the ISU Method and MSM (P5:4.6%) and between the NCI method and MSM (P10:4.4%). The ISU Method estimated the mean protein (104.0 g/day) and potassium (4126 mg/day) intake is somewhat higher than that expected, based on the 2-day WPM mean (103.3 g/day for protein and 4084 mg/day for potassium). On the other hand, the mean protein density as estimated by ISU is 9.24 g/MJ/day, which is somewhat smaller than the mean intake of 9.34 g/MJ/day as estimated by 2-day WPM. For the other methods, the estimated mean intake of the three nutrients is (almost) identical to the mean intake estimated with WPM. The differences in the λ_{BC} are caused by different implementations of the Box–Cox transformation in the different methods.

Comparison for episodically consumed foods

Table 2 presents the estimated usual intake distributions for vegetables, fruit and fish consumption. For vegetables, 2.1% (6 out of 293) of men did not consume vegetables on both recall days and 10.6% (31 out of 293) of men did not consume vegetables on one of the recall days. None of the men reported on the FPQ that they had never consumed vegetables in the past year. Although it is not possible to state which of the methods approximates the true usual intake of these foods more closely, it is reasonable to assume that the estimated mean should be approximately equal to the mean of the population (the mean of 2-day WPM). All of

Table 1 Results of the estimated usual intake distributions for men ($n=293$) from the EFCOVAL Study of protein, protein density and potassium

Nutrient	*Method*	*P5*	*P10*	*P25*	*P50*	*P75*	*P90*	*P95*	*Mean*	r_{var}	λ_{BC}
Protein (g/day)	2-Day WPM	62.1	69.0	81.6	98.7	121.1	143.3	154.3	103.3	—	—
	ISU	72.1	78.1	90.0	102.4	117.2	131.9	141.4	104.0	1.63	—
	NCI	71.3	77.4	88.3	101.7	116.5	131.1	140.4	103.2	1.62	0.295
	MSM	71.4	77.3	88.1	101.1	116.8	133.8	139.8	103.3	1.61	0.100
	SPADE	71.4	77.4	88.3	101.7	116.5	131.1	140.5	103.3	1.63	0.264
Protein density (g/MJ/day)	2-Day WPM	6.75	7.14	8.02	9.10	10.38	11.72	12.38	9.34	—	—
	ISU	7.54	7.87	8.45	9.16	9.94	10.71	11.20	9.24	2.80	—
	NCI	7.53	7.88	8.51	9.26	10.09	10.89	11.40	9.34	2.79	0.010
	MSM	7.56	7.93	8.50	9.26	10.08	10.88	11.25	9.34	2.83	0
	SPADE	7.54	7.89	8.50	9.25	10.08	10.90	11.43	9.34	2.83	−0.152
Potassium (mg/day)	2-day WPM	2306	2747	3301	3907	4804	5747	6288	4084	—	—
	ISU	2647	2919	3417	4040	4741	5443	5898	4126	0.96	—
	NCI	2584	2862	3372	4003	4704	5407	5857	4083	0.96	0.335
	MSM	2530	2988	3438	3968	4665	5461	5911	4085	0.96	0.125
	SPADE	2591	2866	3369	3998	4705	5413	5871	4084	0.96	0.300

Abbreviations: λ_{BC}, Box–Cox transformation parameter; EFCOVAL, European Food Consumption Validation; ISU, Iowa State University; NCI, National Cancer Institute Method; MSM, Multiple Source Method; r_{var}, within- and between-person variance; SPADE; Statistical Program for Age-adjusted Dietary Assessment; WPM, within-person mean.

Table 2 The percentiles and mean of the estimated usual intake distributions for men ($n=292$) from the EFCOVAL Study of vegetables, fruit and fish together with the ratio of the within-subject and between-subject variance and the estimate value of the Box–Cox transformation

Food	*Method*	*P5*	*P10*	*P25*	*P50*	*P75*	*P90*	*P95*	*Mean*	r_{var}	λ_{BC}
Vegetables	2-Day WPM	32.6	56.1	116.6	176.5	258.5	340.1	401.9	193.3	—	—
	ISU Foods	99.0	119.0	154.0	195.0	243.0	291.0	322.0	200.6	3.7	—
	NCI without FPQ[a]	112.2	129.3	161.3	202.0	248.3	294.9	325.1	208.2	3.7	0.448
	NCI with FPQ[a]	106.1	125.1	159.0	201.1	248.2	294.7	324.1	205.8	4.7	0.448
	MSM without FPQ	92.0	109.4	149.5	185.9	236.9	281.7	311.4	194.1	3.5	0.333
	MSM with FPQ	79.2	105.5	140.4	191.2	241.3	289.9	314.8	193.7	4.7	0.333
	SPADE	87.8	105.2	139.5	185.6	240.8	298.8	337.4	195.6	3.8	0.408
Fruit	2-Day WPM	0	1.9	70.0	158.5	273.0	403.1	502.8	193.3	—	—
	ISU Foods	37	57	102	169	249	339	402	186.8	0.9	—
	NCI without FPQ	26.2	44.6	92.1	168.5	266.5	373.2	444.9	192.9	0.9	0.458
	NCI with FPQ	25.9	43.8	91.9	171.6	269.8	371.1	436.7	193.1	1.3	0.450
	MSM without FPQ	20.1	40.4	99.9	169.7	260.8	349.5	418.0	190.2	0.9	0.333
	MSM with FPQ	27.1	45.3	91.7	170.1	265.5	373.4	441.3	192.6	1.4	0
	SPADE	49.0	70.5	116.9	184.5	270.5	356.7	417.7	202.8	1.0	0.444
Fish	2-Day WPM	0	0	0	10	83.8	129.5	157.8	46.0	—	—
	ISU Foods	—	—	—	—	—	—	—	—	—	—
	NCI without FPQ	7.3	11.2	22.5	43.1	68.6	87.3	94.4	46.3	87.4	0.352
	NCI with FPQ	9.8	13.7	23.3	40.7	65.9	87.5	91.5	46.1	29.6	0.355
	MSM without FPQ	11.0	12.1	14.6	43.3	72.8	97.8	108.8	47.1	15.1	0.333
	MSM with FPQ	9.2	10.1	13.7	41.9	75.0	96.8	114.4	47.3	15.1	0.333
	SPADE	10.7	14.7	25.1	41.1	61.5	79.8	89.7	44.5	−(5.9/0)	0.216

Abbreviations: λ_{BC}, Box–Cox transformation parameter; EFCOVAL, European Food Consumption Validation; FPQ, food propensity questionnaire; ISU, Iowa State University; NCI, National Cancer Institute Method; MSM, Multiple Source Method; r_{var}, within- and between-person variance; SPADE; Statistical Program for Age-adjusted Dietary Assessment; WPM, within-person mean.
[a]Owing to the limited number of zero consumptions on the recall the model for episodically consumed could not be used. Therefore, the model for nutrients was used instead. When zero consumptions are replaced by small numbers (0.001), the mean intake is estimated as 194.9 g/day without FPQ and as 195.0 g/day with FPQ.

the methods estimated the mean vegetable consumption too high compared with the mean intake based on 2-day WPM (WPM: 193.3 g/day, MSM: 194.1 g/day, NCI: 208.2 g/day, SPADE: 195.6 g/day, ISU: 200.6 g/day). Owing to the relatively low number of zero consumptions on the recall, the model for episodically consumed foods of the NCI Method did not converge. Therefore, the model for nutrients and daily consumed foods was used. This resulted in a high estimated mean compared with the 2-day WPM. When the zero consumptions were replaced by a small number (0.001), the estimated mean intake was 194.9 g/day without FPQ information and 195.0 g/day with FPQ information.

The inclusion of the FPQ information resulted in only a marginal change of the estimated mean intake from 208.2

to 205.8 g/day for the NCI Method and from 194.1 to 193.7 g/day for MSM. The estimated variance ratio (s_w^2/s_b^2) was similar for all methods without additional FPQ information (that is, 3.5 for MSM, 3.7 for ISU Foods and NCI and 3.8 for SPADE). Both the NCI Method and MSM showed an increase in the variance ratio when the FPQ information was included (that is, 4.7 for NCI and 4.7 for MSM).

For fruit, 9.6% of men did not report fruit consumption on both recall days and 25.3% of men reported no fruit consumption on one of the recall days. Out of 292 men, 2 reported never to eat fruit on the FPQ. The mean fruit intake according to 2-day WPM was 193.3 g/day, and all methods estimated the mean fruit intake somewhat lower (ISU: 186.8 g/day, NCI: 192.9 g/day, MSM: 190.2 g/day), except SPADE (202.8 g/day). The inclusion of the FPQ information resulted in a somewhat higher estimate of the mean fruit intake (NCI: 193.1 g/day, MSM: 192.6 g/day). The estimated s_w^2/s_b^2 ratio was ~ 1 for all methods. The estimate of the SPADE Method of the P5 is almost twofold the estimates of the NCI Method and MSM.

Only 58 out of 293 men (19.8%) consumed fish on both recall days; 101 men (34.5%) consumed fish on one of the recall days and 134 men (45.7%) did not consume fish on both days. None of the men reported on the FPQ that they had never consumed fish in the past year. The estimated mean fish intake was higher compared with 2-day WPM (46.0 g/day) for NCI (46.3 g/day) and MSM (47.1 g/day), and lower for SPADE (44.5 g/day). No results are available for the ISU Method, because the large number of subjects that did not consume fish caused convergence problems. The inclusion of the FPQ information did not alter the estimated mean intake markedly (NCI, from 46.3 to 46.1 g/day, MSM, from 47.1 to 47.3 g/day). The estimated variance ratio was very high (87.4 for NCI without FPQ, 15.1 for MSM and infinite for SPADE). For NCI, the variance ratio decreased from 87.4 to 29.6 when the additional FPQ information was included, whereas the variance ratio did not change for MSM.

Discussion

The objective of this study was to compare methods to estimate usual intake distributions through a simulation study as well as empirically by applying the different methods to intake data collected during the EFCOVAL Study. Overall, the simulation study showed no indication that any of the methods provide biased estimates of the intake distribution of the simulated 'nutrient', with the exception of the NCI Method when the within-person variation was much higher than the between-person variation, and the intake data were log-normally distributed. Although s_w^2/s_b^2 ratios of 9 or higher are rare, they have been reported for some nutrients (McAvay and Rodin, 1988; Willett, 1998). However, the NCI Method was recently updated. Tooze *et al.* (2010) advise the use of the back-transformation used by the ISU Method instead of the Taylor series approximation, when the within-person variability is much larger than the between-person variability and the data are highly skewed.

The s.d. of the bias was somewhat higher for the ISU Method and MSM for all percentiles in all scenarios, indicating that the uncertainty of these methods is somewhat larger than that of the NCI and SPADE Methods. Furthermore, results from the simulation study showed that the estimated usual intake distributions became more variable when study size decreased, when the intake data were less normally distributed and when the variance ratio increased. However, this was the case for all four methods. Nevertheless, it is important to consider this when planning, analyzing and interpreting a food consumption survey.

Although the mean bias is similar for the different methods, it is important to realize that the ISU, NCI and SPADE Methods estimate smooth intake distributions, whereas MSM may provide uneven intake distributions. This is because of the fact that the ISU, NCI and SPADE Methods estimate parameters of the distribution (namely, the mean, s.d.$_w$, s.d.$_b$ and λ_{BC}), whereas MSM produces usual intake distributions for the population, based on estimates of the individual usual intakes. Therefore, for MSM, any unevenness in the original intake data will show in the final estimated density function. Of course, this is mostly a problem in small sample sizes in which chance findings are much more likely to occur.

A limitation of this simulation study is that for all scenarios only 100 simulation sets were used. Therefore, deviations of the sample mean from the true mean usual intake have occurred by chance, which can be seen when looking at the bias of the mean of the simulated sample, which is the mean usual intake estimated with the 2-day WPM. When, for instance, 10 000 simulations were used to calculate the mean usual intake with 2-day WPM in the scenario with $n = 1000$, $\lambda_{BC} = 0.0$ and $r_{var} = 1$, the mean usual intake is identical to the true mean (resulting in a mean bias of 0.00; data not shown). However, with eight different scenarios and five different methods, more than 100 simulations would not have been feasible in the given time. Furthermore, the scope of the simulation study was limited. Simulations were conducted mostly with log-transformed data, whereas distributions of nutrients and foods are typically less skewed in reality and difficult to transform into symmetrical distributions. However, the log-transformation was used because in this case, the Gaussian Quadrature Method used to obtain the true distribution does not bias the results in favor of the SPADE Method. Nevertheless, it would be of interest to perform a simulation study with nutrient data that present more noise, and therefore are more representative of real nutrient data.

Overall, the results from the empirical comparison show that, for nutrients, the methods produce comparable usual intake distributions. As expected, the distributions estimated by 2-day WPM are much wider than those obtained with any

Table 3 Summary of availability, costs and some important attributes of methods

	Software	*Costs*	*Covariates can be included*	*Survey weights can be included*	*Proportion of population above/below cutoff*
ISU	Available as a SAS/IML version and a stand-alone version	$300	Initial adjustments	Yes	Yes
NCI	NCI method was implemented in SAS macros	Requires the SAS package	Yes	Yes	Yes
MSM	MSM was implemented on a website	None	Yes	No	No
SPADE	SPADE was implemented in R	None	Age	Yes	Yes

Abbreviations: ISU, Iowa State University; NCI, National Cancer Institute Method; MSM, Multiple Source Method; SPADE; Statistical Program for Age-adjusted Dietary Assessment.

of the statistical methods. However, for the three foods analyzed in this study (that is, vegetables, fruit and fish), the usual intake distributions estimated with the methods differed markedly, especially without the inclusion of the FPQ in the models. The ISU Method was even unable to produce results for the usual intake of fish, because of the low numbers of consumers on the 24-HDRs. Furthermore, the inclusion of individual frequency information from the FPQ into the model changed the estimated percentiles somewhat for both MSM and the NCI Method. Without the individual FPQ information, the NCI Method considers all participants as consumers. By default, MSM also assumes this, but it has an option to change the percentage habitual non-consumers, for instance, when information regarding the population is available from another survey. The effect of the FPQ information was small for vegetables and fruit for both MSM and NCI. The simulation study of Tooze *et al*. (2006) showed no striking effect of inclusion of FPQ information on the mean bias in the correlated model, but the mean bias was lower when including the FPQ than without the FPQ for the uncorrelated model. However, it should be noted here that the EFCOVAL sample (293 men and 2 days) was too small to reliably estimate the usual intake distributions of foods with such high variance. Because of the small sample size, the estimated usual intake distributions are highly uncertain and, therefore, these empirical results should be regarded carefully. Given these results, a simulation study, based on foods with different percentages of non-consumers, which also investigates the effect of FPQ information, is recommended.

The methods offer a number of different features (Table 3). MSM and SPADE are freely available, and ISU is available for $300 from ISU. However, the NCI Method requires the SAS software package (SAS Institute Inc., Cary, NC, USA), which is expensive and requires some experience before it can be used. Another important aspect in 'real life' is the inclusion of covariates, as this offers the opportunity to decompose the variance into within-person, explained between-person and unexplained between-person variation. NCI and MSM allow covariates to be included, and ISU allows for initial adjustments of covariates. However, SPADE only allows the inclusion of age at this point. The inclusion of survey weights and the estimation of the proportion of the population above or below a specific cutoff value is possible for ISU, NCI and SPADE, but not for MSM. Furthermore, MSM and the NCI Method with additional macros provide estimates of individual usual intakes, which might be used in an epidemiological study.

In conclusion, the methods that were compared in the present paper seem to provide good estimates of the usual intake distribution of nutrients. Nevertheless, the choice of the most suitable method depends on available nutrient or food intake data; care needs to be taken particularly with regard to a high within-person variation in intake, a highly skewed distribution, a high proportion of zero intakes and small sample size.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The content of the paper reflects only our views and the Community is not liable for any use that may be made of the information contained therein. OS designed the study, performed the statistical analysis and drafted the paper. ALD participated in the design and analysis of the study and helped to draft the paper. AG, JHdV, MCO, JH and PvtV participated in the design of the study. We were involved in the interpretation of the results and the review of the manuscript. The EFCOVAL Study takes place with the financial support of the European Community Sixth RTD Framework Program (contract FOOD-CT-2006-022895).

References

Biro G, Hulshof KF, Ovesen L, Amorim Cruz JA (2002). Selection of methodology to assess food intake. *Eur J Clin Nutr* **56** (Suppl 2), S25–S32.

Brussaard JH, Lowik MR, Steingrimsdottir L, Moller A, Kearney J, De Henauw S *et al*. (2002). A European food consumption survey method–conclusions and recommendations. *Eur J Clin Nutr* **56** (Suppl 2), S89–S94.

Buck RJ, Hammerstrom KA, Ryan PB (1995). Estimating long-term exposures from short-term measurements. *J Expo Anal Environ Epidemiol* **5**, 359–373.

Crispim SP, de Vries JH, Geelen A, Souverein OW, Hulshof PJ, Lafay L *et al*. (2011). Two non-consecutive 24 h recalls using EPIC-Soft software are sufficiently valid for comparing protein and potassium intake between five European centers—results from the European Food Consumption Validation (EFCOVAL) study. *Br J Nutr* **105**, 447–458.

Dekkers ALM, Verkaik J, Van Rossum CTM, Slob W, Ocké MC.*SPADE: Statistical Program To Assess Dietary Exposure—User's Manual*. National Institute for Public Health and Environment: Bilthoven (in preparation).

Dodd KW, Guenther PM, Freedman LS, Subar AF, Kipnis V, Midthune D *et al*. (2006). Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. *J Am Diet Assoc* **106**, 1640–1650.

Freedman LS, Guenther PM, Dodd KW, Krebs-Smith SM, Midthune D (2010). The population distribution of ratios of usual intakes of dietary components that are consumed every day can be estimated from repeated 24-h recalls. *J Nutr* **140**, 111–116.

Haubrock J, Nöthlings U, Volatier JL, Dekkers A, Ocké M, Harttig U *et al*. (2011). Estimating usual food intake distributions by using the Multiple Source Method in the EPIC-Potsdam Calibration Study. *J Nutr* **141**, 914–920.

Hoffmann K, Boeing H, Dufour A, Volatier JL, Telman J, Virtanen M *et al*. (2002). Estimating the distribution of usual dietary intake by short-term measurements. *Eur J Clin Nutr* **56** (Suppl 2), S53–S62.

Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM *et al*. (2009). Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes. *Biometrics* **65**, 1003–1010.

Kroes R, Muller D, Lambe J, Lowik MR, van Klaveren J, Kleiner J *et al*. (2002). Assessment of intake from the diet. *Food Chem Toxicol* **40**, 327–385.

Mackerras D, Rutishauser I (2005). 24-h national dietary survey data: how do we interpret them most effectively? *Public Health Nutr* **8**, 657–665.

McAvay G, Rodin J (1988). Interindividual and intraindividual variation in repeated measures of 24-h dietary recall in the elderly. *Appetite* **11**, 97–110.

National Research Council SiCfDE (1986). *Nutrient Adequacy: Assessment Using Food Consumption Surveys*. National Academy Press: Washington, DC.

Nusser SM, Carriquiry AL, Dodd KW, Fuller WA (1996a). A semi-parametric transformation approach to estimating usual daily intake distributions. *J Am Stat Assoc* **91**, 1440–1449.

Nusser SM, Fuller WA, Guenther P (1996b). Estimating usual dietary intake distributions: adjusting for measurement error and nonnormality in 24-h food intake data. In: Lyberg L, Biemer P, Collins M, De Leeuw E, Dippo C, Schwartz N *et al*. (eds). *Survey Measurement and Process Quality*. Wiley: New York.

Slimani N, Deharveng G, Charrondiere RU, van Kappel AL, Ocke MC, Welch A *et al*. (1999). Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Comput Methods Programs Biomed* **58**, 251–266.

Slob W (1993). Modeling long-term exposure of the whole population to chemicals in food. *Risk Anal* **13**, 525–530.

Slob W (2006). Probabilistic dietary exposure assessment taking into account variability in both amount and frequency of consumption. *Food Chem Toxicol* **44**, 933–951.

Subar AF, Dodd KW, Guenther PM, Kipnis V, Midthune D, McDowell M *et al*. (2006). The food propensity questionnaire: concept, development, and validation for use as a covariate in a model to estimate usual food intake. *J Am Diet Assoc* **106**, 1556–1563.

Tooze JA, Kipnis V, Buckman DW, Carroll RJ, Freedman LS, Guenther PM *et al*. (2010). A mixed-effects model approach for estimating the distribution of usual intake of nutrients: the NCI method. *Stat Med* **29**, 2857–2868.

Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF *et al*. (2006). A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J Am Diet Assoc* **106**, 1575–1587.

Waijers PM, Dekkers AL, Boer JM, Boshuizen HC, van Rossum CT (2006). the potential of AGE MODE, an age-dependent model, to estimate usual intakes and prevalences of inadequate intakes in a population. *J Nutr* **136**, 2916–2920.

Wallace LA, Duan N, Ziegenfus R (1994). Can long-term exposure distributions be predicted from short-term measurements? *Risk Anal* **14**, 75–85.

Willett WC (1998). *Nutritional Epidemiology*, 2nd edn. Oxford University Press: New York.