

The effect of under-reporting of energy intake on dietary patterns and on the associations between dietary patterns and self-reported chronic disease in women aged 50–69 years

Marianne S. Markussen¹, Marit B. Veierød^{1,2}, Giske Ursin^{1,3,4*} and Lene F. Andersen^{1,5}

¹Department of Nutrition, Institute of Basic Medical Sciences, University of Oslo, PO Box 1046 Blindern, 0317 Oslo, Norway

²Department of Biostatistics, Center for Biostatistics and Epidemiology, Institute of Basic Medical Sciences, University of Oslo, PO Box 1122 Blindern, 0317 Oslo, Norway

³Cancer Registry of Norway, PO Box 5313 Majorstuen, 0304 Oslo, Norway

⁴Department of Preventive Medicine, University of Southern California, Soto Street Building, 2001 N Soto Street, Los Angeles, CA 90032-3628, USA

⁵Division of Cancer, Transplantation and Surgery, Norwegian Advisory Unit on Disease-Related Malnutrition, Oslo University Hospital, PO Box 4950 Nydalen, 0424 Oslo, Norway

(Submitted 21 October 2015 – Final revision received 13 April 2016 – Accepted 9 May 2016 – First published online 6 June 2016)

Abstract

The aim of this cross-sectional study was to investigate whether under-reporting of energy intake affects derived dietary patterns and the association between dietary patterns and self-reported chronic disease. Diets of 6204 women aged 50–69 years participating in the Norwegian Breast Cancer Screening Program were assessed using a 253-item FFQ. We identified dietary patterns using principal component analysis. According to the revised Goldberg cut-off method, women with a ratio of reported energy intake:estimated BMR < 1·10 were classified as low energy reporters (*n* 1133, 18%). We examined the associations between dietary patterns and self-reported chronic diseases by log-binomial regression, and the results are presented as prevalence ratios (PR) and CI. 'Prudent', 'Western' and 'Continental' dietary patterns were identified among all reporters and plausible reporters. The PR expressing the associations between the 'Western' and 'Prudent' dietary pattern scores and self-reported chronic diseases were consistently highest among plausible reporters except for joint/muscle/skeletal disorders. The largest difference in PR among plausible *v.* all reporters was found for the association between the 'Prudent' pattern and diabetes (PR for highest *v.* lowest tertile: PR_{all reporters} 2·16; 95% CI 1·50, 3·13; *P*_{trend} < 0·001; PR_{plausible reporters} 2·86; 95% CI 1·81, 4·51; *P*_{trend} < 0·001). In conclusion, our results suggest that under-reporting can result in systematic error that can affect the association between dietary pattern and disease. In studies of dietary patterns, investigators ought to consider reporting effect estimates both for all individuals and for plausible reporters.

Key words: Dietary patterns: Principal component analyses: Misreporting of energy: Low energy reporters: Goldberg cut-off method

The relationship between diet and chronic disease is complex. We consume foods and nutrients in different combinations and as part of meals. Thus, evaluating diet as a whole on the basis of dietary patterns is a complementary approach to the study of single nutrients or foods to understand the relationship between diet and disease⁽¹⁾. When investigating associations between diet and disease, the findings are in most cases based on self-reported dietary intake. Previous research has revealed extensive misreporting, especially under-reporting, of self-reported dietary intake^(2–5). The misreporting can be general under-reporting of food intake, or under- or over-reporting of

certain food groups related to social desirability^(6–8). Under- or over-reporting of certain food groups may distort dietary patterns, and such distortion could result in erroneous conclusions regarding the associations between dietary patterns and disease.

The doubly labelled water (DLW) technique has been looked upon as a gold standard in the evaluation of reported energy intake (EI). Unfortunately, the DLW method is technically challenging and extremely expensive, and therefore not possible to implement in most studies. The more simple method developed by Goldberg *et al.*⁽⁹⁾ and later revised by Black⁽¹⁰⁾

Abbreviations: DLW, doubly labelled water; EE, energy expenditure; EI, energy intake; PAL, physical activity level; PCA, principal component analysis; PR, prevalence ratio.

* **Corresponding author:** G. Ursin, email giske.ursin@krefregisteret.no

has been proposed as an alternative to identify potential misreporters of EI. By using the level of discrepancy between the ratio of EI:estimated BMR and the presumed physical activity level (PAL) of the population, individuals can be classified as likely to be low energy, plausible or high energy reporters^(9,10).

A few studies have investigated the effect of under-reporting on empirically derived dietary patterns^(11–18); four studies have reported that the composition of food groups that significantly contributed to the dietary patterns remained relatively unchanged after removal of low energy reporters^(11,13–15). One study found that the number of dietary patterns differed between plausible reporters and all reporters⁽¹²⁾. In all these studies, cluster analysis was used to define dietary patterns. The distribution of low energy reporters across clusters was not uniform, and whether the highest proportion of low energy reporters were found in the healthy or unhealthy clusters differed between the studies^(11,12,14–18). The cluster analysis assigns the study subjects to one of a number of discrete clusters or dietary patterns. When deriving dietary patterns by principal component analysis (PCA) an individual's diet is characterised using a continuous score for each of the derived patterns; thus, this method has the advantage that it looks at more than one dimension of variation in the diet⁽¹⁹⁾. Recently, a study among Swedish adults investigated the effect of excluding low energy reporters on dietary patterns derived by PCA⁽²⁰⁾ and found that the patterns were largely consistent. That study is, to the best of our knowledge, the only study that has investigated the effect of under-reporting of EI on dietary patterns derived by PCA.

The aims of the present study were to investigate the effect of under-reporting of EI, by excluding low energy reporters from the study sample, on (a) the dietary patterns derived by PCA and (b) the association between the dietary patterns and self-reported chronic disease.

Methods

Study sample

The Norwegian Breast Cancer Screening Program is a government-funded national screening programme administered by the Cancer Registry of Norway⁽²¹⁾. All Norwegian women aged 50–69 years are invited to a bilateral two-view mammogram biennially. The participation rate is 77%⁽²²⁾, with about 250 000 women invited/year. In 2006/2007, the Norwegian Breast Cancer Screening Program's invitation letter for mammographic screening included a question on willingness to complete a dietary questionnaire. A total of 67 527 women agreed to participate. In 2008, a consent form and a FFQ were sent to a random sample of 10 000 of these women living all over Norway. A total of 6974 returned the FFQ, and 676 women were excluded because of the following reasons: the FFQ were not filled in (n 46); missing data on height and/or weight (n 158), age (n 5), smoking status (n 41), education (n 79), physical activity (n 104); height < 125 cm (n 7) and weight < 30 or > 170 kg (n 13); age not within the range 50–69 years (n 15); BMI < 18.5 or \geq 40 kg/m² (n 98); or EI < 2100 or > 15 000 kJ/d (n 204). This left us with a total sample of 6204 women.

This study was conducted according to the guidelines laid down in the Declaration of Helsinki, and all procedures involving human subjects were approved by the regional ethics committee and the Norwegian Data Protection Authority. Written informed consent was obtained from all subjects.

Dietary assessment

The 16-page, 253-item FFQ was designed to measure the habitual food intake among Norwegian adults the preceding year. The questionnaire had an extra focus on fruits, vegetables and other antioxidant-rich foods and beverages, and thus the foods accounting for the variation in antioxidant intake in a population could be investigated⁽²³⁾. The 253-item FFQ has been described in detail earlier⁽²⁴⁾. In short, it was based on a previously validated 180-item FFQ designed to measure total EI in the Norwegian population⁽²⁵⁾, which later was expanded to a 270-item FFQ to cover most of the antioxidant-rich foods and beverages in Norway⁽²⁶⁾. The energy and food intakes estimated from the 270-item FFQ have been validated^(26,27). The EI was compared with independent measures of energy expenditure (EE) using the ActiReg[®] system (motion detection)⁽²⁸⁾, whereas 7-d weighed food records were used to study the relative validity of food and nutrient intakes. The correlation coefficient between EI and EE was 0.54⁽²⁶⁾. Correlations between FFQ and the weighed food records were 0.41 for berries, 0.61 for fruits and 0.38 for vegetables⁽²⁶⁾. This FFQ has also been validated for ranking individuals according to their usual intakes of fruits, juices and vegetables using the method of triads with two independent and specific biomarkers of fruits and vegetables and 7-d weighed food records⁽²⁷⁾. Using the method of triads, the validity correlation was found to range from 0.60 to 0.94. The 253-item FFQ used in this study was revised from the original 270-item FFQ by removing seventeen items that were seldom or never eaten (curly kale, red cabbage, mushroom, globe artichoke, sundried tomatoes, tofu, cumin, turmeric, ginger powder, caraway, cloves, piri piri, sage, rosehip tea, organic blueberry juice, organic blackcurrant juice and crowberry juice). The questionnaire also collected information about dietary supplements, age, height, weight, smoking status, physical activity, chronic diseases (present or previous) and use of medication. Daily intake of energy, nutrients and foods were computed using the food database AE-07 and KBS software system (KBS, version 4.9, 2008) developed at the Department of Nutrition, University of Oslo, Norway. The food database AE-07 is based on the 2006 edition of the *Norwegian Food Composition Table* (www.norwegianfoodcomp.no). Intakes from dietary supplements were included in the calculations.

The 253 food items were categorised into forty-nine food groups on the basis of similarity in ingredients, nutrient profile or culinary usage (online Supplementary Table S1).

Disease assessment

In the FFQ, the participants were asked whether they had currently or previously been diagnosed with one or more of the following diseases: asthma, joint inflammation, muscle or



skeletal disorder, chronic gastrointestinal disease, chronic respiratory disease, depression or psychiatric disorder, stroke, heart attack or angina, hypertension and diabetes (type 1 or type 2). We defined six disease groups: total chronic disease (composed of all of the following disease groups), CVD (stroke, heart attack, angina and hypertension), diabetes (type 1 and 2), chronic respiratory disease (asthma and chronic respiratory disease), cancer and joint/muscle/skeletal disorders (joint inflammation and muscle and skeletal disorders). A participant was identified to belong to a disease group if she had been diagnosed with at least one of the diseases in the group.

Physical activity assessment

Physical activity was assessed using a modified version⁽²⁹⁾ of the physical activity questionnaire used in the California Teachers Study⁽³⁰⁾. Subjects were asked to assess habitual weekly physical activity and report all physical activity lasting at least 10 min/session. They were provided examples of light activities (defined as walking or cross-country skiing at a slow pace), moderate activities (defined as activities where some effort is required and that cause increased breathing, such as bicycling, swimming or cross-country skiing at a moderate pace, jogging at a slow pace, dancing) and strenuous activities (defined as activities that require hard effort and causes substantial increased breathing, such as aerobics, running, cross-country skiing or bicycling at a brisk pace). The subjects were asked to estimate their mean hours per week (none, <0.5, 0.5–1, 1.5–2, 2.5–3.5, 4–6, ≥7 h) of participation at each level of activity. We created separate light, moderate and strenuous activity variables in minutes per week by summing up hours per week for each level of activity multiplied with 60. We calculated EE as the number of hours of each physical activity multiplied by its estimated metabolic cost⁽³¹⁾ and expressed this variable in metabolic equivalent task (h/week).

Definition of low energy reporters

Low energy reporters were determined using the revised Goldberg cut-off method⁽¹⁰⁾. This method is based on the principle that EI equals EE when weight is stable (equation (1): EI = EE). EE can also be expressed as multiples of BMR and PAL, and replacing EE in equation (1) with BMR × PAL gives equation (2): EI/BMR = PAL. The idea by Goldberg *et al.* was that the ratio EI:BMR can be derived from a dietary assessment method and then be evaluated against an expected PAL for a population.

The revised Goldberg cut-off values⁽¹⁰⁾ used in the present study are based on estimated 95% confidence limits (cut-offs) for the plausible EI. The values of these cut-off points vary according to PAL, number of days of food recording and whether the evaluation of EI/BMR is at the individual or group level. Subjects are defined as plausible, low energy or high energy reporters from their ratio of EI:BMR according to whether this ratio is within, below or above the 95% confidence limits calculated, respectively.

A PAL for sedentary lifestyle (1.55)^(32,33) was applied for all participants, because of lack of an objective measure of total physical activity in the present study. We have used the lower

95% confidence limits published by Black⁽¹⁰⁾ to identify low energy reporters, which is based on a PAL of 1.55 and infinity number of days of food recording (habitual intake measured by a FFQ) at the individual level (*n* 1). Black calculated the value of this cut-off point to be 1.10, and all women with EI/BMR < 1.10 were classified as low energy reporters in this study.

In the present study, BMR was calculated from the following equations⁽³⁴⁾:

$$\text{BMR women 31–60 years: } 0.0433W + 2.57H - 1.180$$

and

$$\text{BMR women 61–70 years: } 0.0342W + 2.10H - 0.0486.$$

Statistical methods

We divided the study sample into all and plausible reporters, and each of these subsamples was stratified by BMI: 18.5 kg/m² ≤ BMI < 25 kg/m² (normal weight) and 25 kg/m² ≤ BMI < 40 kg/m² (overweight/obese). Differences between under-reporters and plausible reporters and between normal weight and overweight/obese subjects were studied by the two-sample *t* test for continuous variables and the χ^2 test for categorical variables. Physical activity and alcohol intake data were log_e transformed in these analyses.

PCA was used in order to reduce the dimensionality of the dietary data. Before the PCA, the 253 food items were categorised into forty-nine food groups (g/d) according to similarity in ingredients, nutrient profile or culinary usage. In the PCA, the food groups are aggregated in linear combinations called principal components, according to the degree to which the food groups are correlated to each other⁽¹⁾. Before extracting components, the suitability for using PCA was assessed by the Kaiser–Meyer–Olkin measure of sampling adequacy and the Bartlett's test of sphericity, which tests whether our correlation matrix is significantly different from an identity matrix⁽³⁵⁾. The Kaiser–Meyer–Olkin value was 0.76 for both all and plausible reporters, which is above the suggested minimum of 0.50⁽³⁶⁾, and Bartlett's test of sphericity was statistically significant ($P < 0.001$), supporting the suitability of the data for PCA. The input variables were standardised using the correlation matrix of the forty-nine food group variables in the PCA, and not the covariance matrix. To determine the number of meaningful components or dietary patterns to retain, we considered the eigenvalue-one criterion, the scree test, the proportion of variance accounted for and the interpretability of the patterns⁽³⁷⁾. After extraction of the dietary patterns, a rotation method is usually applied to improve interpretation⁽¹⁾. We investigated unrotated, orthogonal (varimax) rotation and oblique (oblimin) rotation. The unrotated dietary patterns were less interpretable than the rotated ones. The orthogonal (varimax) and oblique (oblimin) rotation methods gave the same result, and we chose to apply the orthogonal (varimax) rotation method in order to achieve a simpler structure of the dietary patterns with greater interpretability. The varimax rotation is the most common rotation method applied in dietary pattern analysis^(1,38–54) and leads to uncorrelated dietary patterns. Food groups with a factor loading ≥ 0.3 (absolute value) were considered to be important contributors to a



component. Factor loadings can be interpreted as correlation coefficients between food groups and dietary patterns⁽⁵⁵⁾. We labelled the dietary patterns according to the more or less healthy combinations of food groups and according to the influence of international cuisines. Finally, individual scores were calculated for each of the retained components. The overall dietary pattern of a participant is represented by her factor scores on all the identified dietary patterns. A high factor score for a given dietary pattern indicated high intake of food groups constituting that dietary pattern, whereas a low score indicated low intake of those food groups.

Owing to the high prevalence of outcomes, a generalised linear regression with a log link and binomial distribution (log-binomial regression)⁽⁵⁶⁾ was used to estimate the association between dietary pattern scores and the prevalence of self-reported chronic diseases among all and plausible reporters. Women with self-reported chronic diseases were compared with those reporting not having a disease. The dietary pattern scores were categorised into tertiles, and we estimated adjusted prevalence ratios (PR) and 95% CI for each tertile compared with the lowest tertile of each dietary pattern. We analysed trends across tertiles of dietary pattern scores by treating the variable as a continuous variable in the regression analysis. We consider the succeeding variables as potential

confounders and adjusted for them as follows: age (50–60, 61–69 years), education (\leq upper secondary school, academy/college/university ≥ 4 years), smoking status (yes, no), BMI (18.5–24.9, 25–39.9 kg/m²), physical activity (continuous) and EI (continuous). Wald's test was used to test for interaction between BMI and dietary pattern scores. No significant interactions were found.

All tests were two sided, and $P < 0.05$ was considered to be statistically significant. The analyses were conducted using SPSS version 20.0 (IBM Corp.).

Results

A total of 1133 (18.3%) of the 6204 women were defined as low energy reporters (Table 1). Low energy reporters had significantly lower EI, higher BMI, lower physical activity, lower alcohol intake and lower educational level than plausible reporters ($P \leq 0.02$). Moreover, there was an indication of a higher proportion of smokers among low energy reporters compared with plausible reporters ($P = 0.09$). The prevalence of self-reported total chronic disease, CVD, diabetes and joint/muscle/skeletal disorders was significantly higher ($P \leq 0.04$) among low energy reporters compared with plausible reporters.

Table 1. Selected characteristics and prevalence of disease of all reporters, under-reporters and plausible reporters (Mean values and standard deviations; numbers; percentages; medians and 25th and 75th percentiles (P))

Characteristics	All reporters		Low energy reporters		Plausible reporters		P*
	Mean	SD	Mean	SD	Mean	SD	
n	6204		1133		5071		
Age (years)	57.9	4.5	57.6	4.5	57.9	4.5	0.08
Energy intake (kJ/d)	8698	2207	5850	972	9334	1879	<0.001
Height (m)	1.67	0.06	1.67	0.06	1.66	0.06	<0.001
Weight (kg)	70.8	11.2	75.5	12.4	69.7	10.7	<0.001
BMI (kg/m ²)	25.5	3.8	27.0	4.2	25.2	3.6	<0.001
18.5 \leq BMI < 25 (%)	51		37.2		54.2		<0.001
25 \leq BMI < 40 (%)	49		62.8		45.8		<0.001
Physical activity (MET-h/week)							<0.001†
Median	14.3		12.0		15.4		
P25, P75	7.5, 24.2		4.9, 20.0		7.6, 24.2		
Alcohol intake (g/d)							<0.001†
Median	4.9		3.7		5.2		
P25, P75	1.1, 11.3		0.8, 8.5		1.3, 11.7		
Smoking status (%)	19.3		21.1		18.9		0.09
Education (%)							0.02
Primary and secondary school	19.9		22.3		19.4		
Upper secondary school	39.8		41.3		39.5		
Academy/college/university (≤ 4 years)	24.4		22.5		24.8		
Academy/college/university (> 4 years)	15.9		13.9		16.3		
Diseases (%)							
No disease	39.8		36.2		40.6		0.005
Total chronic disease‡	55.5		59.2		54.7		0.005
CVD§	20.4		23.8		19.6		0.001
Diabetes	3.0		4.8		2.6		<0.001
Chronic respiratory disease ¶	11.0		12.5		10.6		0.06
Cancer	6.2		6.9		6.0		0.29
Joint/muscle/skeletal disorder	36.4		39.1		35.8		0.04

MET, metabolic equivalent task.

* Comparison of under-reporters and plausible reporters: two-sample *t* test for continuous variables and χ^2 test for categorical variables.

† Physical activity and alcohol were log_e transformed for the comparison of BMI groups.

‡ Disease group composed of CVD, diabetes, chronic respiratory disease, cancer and joint inflammation and muscle and skeletal disorder.

§ Disease group composed of stroke, heart attack, angina and hypertension.

|| Disease group composed of diabetes type 1 and type 2.

¶ Disease group composed of asthma and chronic respiratory inflammation.

Overweight/obesity was more common in low energy reporters than plausible reporters (63 and 46 %, respectively).

Among all reporters, EI was significantly higher for the normal weight group than for the overweight/obese group ($P=0.02$) (Table 2). By removal of the low energy reporters, this changed to the opposite ($P=0.001$). In both all and plausible reporters, the overweight/obese group was slightly older, had lower physical activity, lower alcohol intake, were less likely to smoke, were less educated and had a higher prevalence of chronic diseases than the normal weight group ($P<0.001$).

We identified three major dietary patterns for both all and plausible reporters, all with eigenvalues ≥ 2.0 . The point at which the slope of the graph in the scree plot showed a change, and the interpretation of the components, justified retaining three dietary patterns. Table 3 presents the three dietary patterns for all and plausible reporters, with food groups having factor loadings with absolute values ≥ 0.30 . The three dietary patterns accounted for 17.4 and 16.7% of the total variance among all and plausible reporters, respectively. Among all reporters, the dietary pattern labelled 'Prudent' was characterised by high positive loadings for vegetables, fish as dinner, fruits, herbs and spices, berries, nuts and seeds, legumes, meat dishes, salad dressings, poultry, vegetarian food, soup and tea. Although the 'Prudent' pattern derived for the plausible reporters was substantially similar to that of all reporters, differences were noted for three food groups: vegetarian food, tea and salad dressings had factor loadings <0.30 among plausible reporters. Furthermore, among all reporters the 'Prudent' pattern was the pattern explaining most of the variance in the dietary data. After excluding the low energy reporters from the study sample, the 'Prudent' pattern explained the lowest amount of variance in the dietary data. Among all reporters, the 'Western' dietary pattern was characterised by high loadings for potatoes, sauce, refined grains, processed meat, cakes and desserts, margarine, sweet spreads, red meat and game as well as high negative loadings for wine and herbs and spices. For plausible reporters, a similar 'Western' pattern was found, but this pattern also showed a high negative loading for vegetarian food. The 'Western' pattern explained the highest total variance among the plausible reporters. The third pattern was labelled 'Continental', because several of the food groups contributing significantly to this pattern were influenced by international cuisine. Among all reporters, it was characterised by high loadings for tomato sauce, pasta, processed meat, fat-rich potatoes, salty snacks, pizza, salad dressings, rice, poultry, mustard, sweets and wine. We found a similar 'Continental' pattern among plausible reporters, but here soya sauce was also among the highly loaded food groups (0.31), and wine had a slightly lower loading (0.29) than in the 'Continental' pattern among all reporters.

Table 4 presents the adjusted PR of self-reported chronic disease by tertiles of the dietary pattern scores among all and plausible reporters. Self-reported CVD was significantly positively associated with 'Western' pattern scores among plausible reporters but not among all reporters (PR for highest *v.* lowest tertile: PR_{all reporters} 1.05; 95 % CI 0.94, 1.18; $P_{\text{trend}}=0.40$; PR_{plausible reporters} 1.15; 95 % CI 1.02, 1.31; $P_{\text{trend}}=0.03$). Self-reported CVD was also significantly positively associated with

'Prudent' pattern scores among both all and plausible reporters, but the association was slightly stronger among plausible reporters (PR for highest *v.* lowest tertile: PR_{all reporters} 1.27; 95 % CI 1.14, 1.43; $P_{\text{trend}}<0.001$; PR_{plausible reporters} 1.37; 95 % CI 1.20, 1.56; $P_{\text{trend}}<0.001$). The largest differences between all and plausible reporters were found for the association between self-reported diabetes and the 'Prudent' pattern, where a stronger association was found among the plausible reporters, although with wide CI (PR for highest *v.* lowest tertile: PR_{all reporters} 2.16; 95 % CI 1.50, 3.13; $P_{\text{trend}}<0.001$; PR_{plausible reporters} 2.86; 95 % CI 1.81, 4.51; $P_{\text{trend}}<0.001$). In addition, a significant positive association between self-reported chronic respiratory disease and 'Prudent' pattern scores was found among plausible reporters, but not among all reporters (PR for highest *v.* lowest tertile: PR_{all reporters} 1.18; 95 % CI 0.98, 1.42; $P_{\text{trend}}=0.12$; PR_{plausible reporters} 1.33; 95 % CI 1.08, 1.63; $P_{\text{trend}}=0.007$). Finally, we found a significant inverse association between self-reported joint/muscle/skeletal disorders and the 'Continental' pattern and a significant positive association between these disorders and the 'Prudent' pattern. However, there were no differences in the effect estimates between all and plausible reporters.

The online Supplementary Table S2 shows the effect of including the covariates one by one in the log-binomial regression model of the relationship between the tertiles of dietary pattern score and self-reported total chronic disease among all and plausible reporters. Only minor differences were found between the effect estimates in the different models.

Discussion

We identified almost one-fifth of the women to be low energy reporters based on the revised Goldberg cut-off method⁽¹⁰⁾. The majority of food groups contributing significantly to the 'Prudent', 'Western' and 'Continental' patterns were consistently found for both all and plausible reporters, differing only with a few food groups. The PR expressing the associations between the 'Western' and 'Prudent' dietary pattern scores and self-reported chronic diseases were consistently highest among plausible reporters except for joint/muscle/skeletal disorders.

Studies using the DLW method have clearly shown that all dietary assessment methods tend to underestimate EI to various degrees⁽⁵⁷⁻⁵⁹⁾. Previous studies have reported prevalence of low energy reporting ranging from 10 to 60% depending on the dietary assessment method, the reference method used to identify low energy reporters and the characteristics of the study population^(4,11,58,60-72). In the revised Goldberg cut-off equations⁽¹⁰⁾, the individual's physical activity is taken into account. To increase sensitivity, Black⁽¹⁰⁾ recommended collecting more information about home or occupational and leisure-time physical activity, to be able to assign subjects to low, medium and high activity categories; three different cut-offs can then be calculated for the subjects belonging to the different activity categories. This would probably have resulted in a higher prevalence of low energy reporters in our study sample. Unfortunately, the physical activity questionnaires used in the present study did not give enough information about the

Table 2. Selected characteristics and prevalence of disease of all v. plausible reporters stratified by BMI (Mean values and standard deviations; numbers; percentages; medians and 25th and 75th percentiles (P))

Characteristics	All reporters (n 6204)					Plausible reporters (n 5071)				
	18.5 ≤ BMI < 25		25 ≤ BMI < 40		P*	18.5 ≤ BMI < 25		25 ≤ BMI < 40		P*
	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
n	3167		3037			2746		2325		
Low energy reporters (%)	6.8		11.5		<0.001	–		–		
Age (years)	57.6	4.5	58.1	4.5	<0.001	57.6	4.5	58.2	4.5	<0.001
Energy intake (kJ/d)	8763	2235	8631	2176	0.02	9255	1960	9428	1774	0.001
BMI (kg/m ²)	22.7	1.6	28.5	3.0	–	22.6	1.6	28.2	2.8	–
Physical activity (MET-h/week)					<0.001†					<0.001†
Median	16.9		12.5			17.5		12.5		
P25, P75	9.9, 26.6		7.2, 20.2			10.5, 27.5		7.2, 20.5		
Alcohol intake (g/d)					<0.001†					<0.001†
Median	5.7		4.0			5.9		4.3		
P25, P75	1.6, 12.1		0.8, 10.4			1.7, 12.4		0.8, 10.9		
Smoking status (%)	21.3		17.2		<0.001	20.6		16.8		<0.001
Education (%)					<0.001					<0.001
Primary and secondary school	15.8		24.2			15.4		24.1		
Upper secondary school	38.6		41.2			38.5		40.8		
Academy/college/university (≤4 years)	26.9		21.7			27.3		21.8		
Academy/college/university (>4 years)	18.7		12.9			18.9		13.3		
Diseases (%)										
No disease	47.6		31.7		<0.001	47.4		32.6		<0.001
Total chronic disease‡	20.8		38.1		<0.001	20.7		37.1		<0.001
CVD§	13.1		27.9		<0.001	13.1		27.3		<0.001
Diabetes	1.7		4.8		<0.001	1.3		4.2		<0.001
Chronic respiratory disease¶	8.6		13.1		<0.001	8.9		12.6		<0.001
Cancer	7.4		6.4		0.60	5.8		6.3		0.50
Joint/muscle/skeletal disorder	32.1		41.5		<0.001	31.5		40.9		<0.001

MET, metabolic equivalent task.

* Comparison of BMI groups: two-sample *t* test for continuous variables and χ^2 test for categorical variables.

† Physical activity and alcohol were log_e transformed for the comparison of BMI groups.

‡ Disease group composed of CVD, diabetes, chronic respiratory disease, cancer and joint inflammation and muscle and skeletal disorder.

§ Disease group composed of stroke, heart attack, angina and hypertension.

|| Disease group composed of diabetes type 1 and type 2.

¶ Disease group composed of asthma and chronic respiratory inflammation.

Table 3. Factor loadings for the three dietary patterns found in the principal component analysis for all (*n* 6204) and plausible (*n* 5071) reporters

Food groups	Prudent		Western		Continental	
	All reporters	Plausible reporters	All reporters	Plausible reporters	All reporters	Plausible reporters
Vegetables	0.65*	0.64*	-0.05	-0.16	0.04	0.06
Fish, dinner	0.53*	0.57*	0.18	0.08	-0.09	-0.09
Fruits	0.52*	0.47*	0.03	-0.14	-0.20	-0.18
Herbs and spices	0.50*	0.38*	-0.30*	-0.41*	0.19	0.25
Berries	0.48*	0.46*	0.19	0.06	-0.16	-0.13
Nuts and seeds	0.47*	0.36*	-0.14	-0.28	0.05	0.08
Legumes	0.40*	0.34*	-0.22	-0.29	0.07	0.12
Meat dishes	0.40*	0.47*	0.22	0.18	0.16	0.16
Vegetarian food	0.32*	0.20	-0.21	-0.32*	0.004	0.07
Soup	0.30*	0.30*	0.13	0.05	0.09	0.12
Tea	0.30*	0.22	-0.15	-0.23	0.01	0.05
Egg	0.24	0.24	-0.02	-0.05	0.12	0.12
Fish, breadspread	0.22	0.27	0.14	0.10	-0.04	-0.05
Water	0.21	0.24	-0.05	-0.06	-0.01	-0.01
Fruit juice	0.13	0.05	0.04	-0.04	0.02	0.03
Potatoes	-0.02	0.09	0.59*	0.57*	-0.11	-0.16
Sauce	-0.05	0.02	0.57*	0.58*	0.24	0.21
Refined grains	-0.04	-0.03	0.54*	0.50*	0.05	0.05
Processed meat	-0.02	0.03	0.47*	0.48*	0.44*	0.41*
Cakes and desserts	0.08	0.06	0.46*	0.40*	0.16	0.17
Margarine	-0.10	-0.10	0.41*	0.39*	0.06	0.04
Sweet spreads	0.06	0.07	0.38*	0.32*	-0.17	-0.16
Red meat and game	0.11	0.20	0.37*	0.40*	0.28	0.24
Wine	0.10	-0.01	-0.32*	-0.35*	0.30*	0.29
Cheese	0.08	0.04	0.26	0.19	0.03	0.01
Whole grains	0.25	0.20	0.26	0.11	-0.19	-0.20
Mayonnaise	0.02	0.04	0.25	0.23	0.15	0.13
Sugar-sweetened beverages	-0.09	-0.09	0.23	0.21	-0.01	-0.01
Coffee	-0.11	-0.03	0.23	0.27	0.07	0.04
Butter	-0.01	-0.03	0.17	0.13	0.04	0.04
Sugar	-0.001	-0.03	0.14	0.09	0.02	0.02
Tomato sauce	0.21	0.15	0.09	0.05	0.53*	0.55*
Pasta	0.11	-0.02	-0.04	-0.11	0.51*	0.54*
Fat-rich potatoes	0.03	-0.003	0.21	0.19	0.39*	0.39*
Salty snacks	-0.08	-0.11	0.12	0.13	0.38*	0.37*
Pizza	-0.05	-0.07	0.14	0.14	0.38*	0.37*
Salad dressings	0.36*	0.24	-0.15	-0.22	0.37*	0.42*
Rice	0.20	0.11	-0.05	-0.12	0.36*	0.40*
Poultry	0.33*	0.32*	-0.14	-0.16	0.34*	0.33*
Mustard	0.14	0.12	0.07	0.04	0.34*	0.34*
Sweets	0.02	-0.06	0.13	0.08	0.30*	0.30*
Soya sauce	0.25	0.17	-0.16	-0.22	0.28	0.31*
Barbecue and taco seasoning	0.06	0.07	0.07	0.09	0.25	0.24
Low-fat dairy products	0.06	-0.06	0.12	-0.02	-0.24	-0.24
Beer	0.05	-0.02	-0.03	-0.07	0.21	0.21
Sweeteners	-0.03	-0.01	0.05	0.09	0.19	0.17
High-fat dairy products	0.03	-0.09	0.17	0.06	-0.18	-0.17
Artificially sweetened beverages	-0.06	-0.03	0.02	0.06	0.16	0.14
Liquor	-0.05	-0.06	0.02	0.03	0.14	0.12
Total variance explained (%)	6.1	5.2	5.9	6.0	5.4	5.5

* Factor loadings with an absolute value ≥ 0.30 .

individuals' total amount of physical activity. Therefore, we used a PAL of 1.55, which is the value defined by FAO/WHO/United Nations University⁽³³⁾, representing a sedentary level of EE⁽⁵⁸⁾, in order not to overestimate the extent of under-reporting. Nevertheless, it could be criticised to be a very conservative PAL value for this population, and misclassifications of more active participants could exist. We found a prevalence of 18.3% low energy reporters in our study sample. Other studies using the revised Goldberg cut-off method⁽¹⁰⁾ have found a prevalence of low energy reporters of 21–33%^(73–78). The differences between

studies in the prevalence of low energy reporters could be due to differences in the accuracy of reporting or it might be due to the differences in criteria used to identify low energy reporters or the way the dietary data are collected and calculated. It is important to take into account that the confidence limits calculated by Black are wide, and only extreme degrees of misreporting can be identified⁽¹⁰⁾.

The low energy reporters in this study reported higher BMI, lower physical activity, lower alcohol intake and lower educational level than the plausible reporters (Table 1). This is

Table 4. Relationship between prevalence of self-reported chronic disease and tertiles (T) of dietary pattern scores among all and plausible reporters (Numbers; prevalence ratio (PR)* and 95 % confidence intervals)

Diseases	All reporters (n 6204)								Plausible reporters (n 5071)									
	n	T1	T2			T3			P _{trend}	n	T1	T2			T3			P _{trend}
			PR	95 % CI		PR	95 % CI					PR	95 % CI		PR	95 % CI		
No disease	2470								2060									
Total chronic disease†	3444								2773									
Western		1.00	0.99	0.94, 1.04		1.00	0.95, 1.06		0.99		1.00	1.01	0.96, 1.08		1.04	0.98, 1.11		0.21
Continental		1.00	0.96	0.92, 1.01		0.96	0.91, 1.01		0.07		1.00	0.98	0.93, 1.04		0.95	0.90, 1.01		0.09
Prudent		1.00	1.01	0.95, 1.06		1.14	1.08, 1.21		<0.001		1.00	1.08	1.02, 1.15		1.18	1.10, 1.25		<0.001
CVD‡	1264								994									
Western		1.00	1.03	0.93, 1.14		1.05	0.94, 1.18		0.40		1.00	1.09	0.97, 1.24		1.15	1.02, 1.31		0.03
Continental		1.00	0.93	0.84, 1.02		0.96	0.86, 1.06		0.31		1.00	0.97	0.88, 1.08		0.96	0.85, 1.09		0.53
Prudent		1.00	1.04	0.94, 1.16		1.27	1.14, 1.43		<0.001		1.00	1.19	1.05, 1.34		1.37	1.20, 1.56		<0.001
Diabetes§	188								134									
Western		1.00	0.85	0.62, 1.18		0.89	0.63, 1.26		0.48		1.00	0.97	0.64, 1.46		1.12	0.75, 1.67		0.57
Continental		1.00	0.89	0.65, 1.21		0.89	0.63, 1.26		0.49		1.00	1.17	0.79, 1.73		1.03	0.67, 1.58		0.85
Prudent		1.00	1.36	0.98, 1.89		2.16	1.50, 3.13		<0.001		1.00	1.95	1.27, 2.99		2.86	1.81, 4.51		<0.001
Chronic respiratory disease	680								538									
Western		1.00	0.95	0.80, 1.12		0.99	0.82, 1.19		0.86		1.00	1.08	0.89, 1.30		1.11	0.91, 1.36		0.31
Continental		1.00	0.95	0.81, 1.12		1.04	0.88, 1.22		0.70		1.00	1.02	0.85, 1.23		1.09	0.90, 1.32		0.37
Prudent		1.00	0.93	0.79, 1.09		1.18	0.98, 1.42		0.12		1.00	1.14	0.95, 1.37		1.33	1.08, 1.63		0.007
Cancer	384								306									
Western		1.00	1.01	0.80, 1.29		1.20	0.92, 1.57		0.19		1.00	1.00	0.76, 1.31		1.22	0.93, 1.61		0.15
Continental		1.00	0.95	0.76, 1.19		0.96	0.76, 1.21		0.72		1.00	1.00	0.78, 1.28		0.88	0.68, 1.15		0.37
Prudent		1.00	0.83	0.66, 1.05		1.00	0.76, 1.31		0.86		1.00	0.92	0.71, 1.19		1.13	0.84, 1.51		0.51
Joint/muscle/skeletal disorder	2260								1817									
Western		1.00	0.98	0.92, 1.06		1.01	0.93, 1.09		0.89		1.00	1.00	0.92, 1.08		1.04	0.95, 1.13		0.38
Continental		1.00	0.93	0.87, 1.00		0.90	0.84, 0.97		0.005		1.00	0.95	0.88, 1.02		0.89	0.82, 0.97		0.006
Prudent		1.00	1.02	0.94, 1.09		1.21	1.11, 1.31		<0.001		1.00	1.12	1.03, 1.21		1.24	1.13, 1.35		<0.001

* Adjusted for age (50–60, 61–69 years), education (\leq upper secondary school, academy/college/university ≥ 4 years), smoking status (yes, no), physical activity (continuous), BMI ($18.5 \leq \text{BMI} < 25$, $25 \leq \text{BMI} < 40 \text{ kg/m}^2$) and energy (continuous).

† Disease group composed of CVD, diabetes, chronic respiratory disease, cancer and joint inflammation and muscle and skeletal disorder.

‡ Disease group composed of stroke, heart attack, angina and hypertension.

§ Disease group composed of diabetes type 1 and type 2.

|| Disease group composed of asthma and chronic respiratory inflammation.

in line with previous studies investigating characteristics of low energy reporters^(13,58,79).

We have previously discussed the dietary patterns derived from this study in detail⁽⁸⁰⁾. In the current analyses, we wanted to investigate whether the measurement errors introduced by under-reporting distorted the food groups defining the dietary patterns. We found three major dietary patterns among both all and plausible reporters that were not vastly different from each other, differing only with a few food groups in each pattern. Other studies have also identified relatively similar patterns after removal of low energy reporters from the analysis compared with the total sample^(13–15). Interestingly, the dietary pattern explaining the highest extent of variance in the dietary intake differed between all and plausible reporters, with the ‘Prudent’ pattern explaining the highest extent of variance among all reporters and the ‘Western’ pattern explaining the highest extent of variance among plausible reporters. This may be related to the fact that low energy reporters tend to over-report foods perceived as healthy or/and under-report foods perceived as unhealthy^(8,63,81,82). The identification of the first principal component as a prudent dietary pattern among all reporters is comparable with other studies investigating dietary patterns derived by PCA^(19,46,48,83,84). Some, but not all⁽¹⁸⁾, studies investigating the

association between dietary patterns derived by cluster analysis and under-reporting observed more severe under-reporting among subjects in healthy dietary pattern clusters^(16,85,86).

The implications of under-reporting might be distortion of the associations between diet and disease. Most studies in nutritional epidemiology have excluded subjects with implausible high or low EI using cut-off values for plausible EI, usually < 2100 and $> 15\,000 \text{ kJ/d}$ ^(87–92); however, this does not account for all the misreporting. In our study, the associations between dietary patterns and self-reported chronic disease were somewhat stronger among plausible reporters (additional exclusion of low energy reporters as defined by the revised Goldberg cut-off method⁽¹⁰⁾) compared with all reporters (excluding those with implausible EI < 2100 and $> 15\,000 \text{ kJ/d}$). Specifically, the associations between the ‘Prudent’ pattern and self-reported chronic diseases were strengthened. The positive relationship between the ‘Prudent’ pattern and several of the chronic diseases indicated that the participants tried to eat healthy in order to reduce either the symptoms of their condition or reduce the likelihood of possible detrimental consequences. A positive relationship between a healthy dietary pattern and disease has also been reported in a Swedish study, where the highest prevalence of previously known

health problems was observed in the healthy 'fruit and vegetables' cluster among women⁽⁸⁵⁾.

Effects of under-reporting on diet–disease associations have been reported in some studies. A Swedish study investigated the effect of under-reporting on the association between risk of breast cancer and alcohol intake⁽⁹³⁾. The researchers reported an increased risk of breast cancer with high alcohol intakes, and the risk estimates were strengthened among the plausible reporters compared with all reporters. A study in the US⁽⁹⁴⁾ investigated the use of calibrated EI to account for under-reporting and the effect on the association between risk of breast, colon, endometrial and kidney cancer. They produced the calibrated consumption estimates based on calibration equations developed in a substudy among 544 women where DLW was used to estimate total EE and urinary N was used as the recovery biomarker for protein⁽⁹⁵⁾. The researchers found calibrated energy consumption to be positively associated with the risk of breast, colon, endometrial and kidney cancer, while uncalibrated energy was not. In a few studies, the investigators have adjusted for under-reporting of EI in their analyses in order to avoid biased conclusions^(96–98). Our results and those of other studies show that it is important to consider under-reporting in dietary studies and the effect this might have on associations between dietary patterns and health outcomes.

The extensive information on diet, lifestyle and self-reported chronic diseases and the large study sample from different parts of the country are important strengths of the present study. However, there are some limitations. First, it might be that the women responding to the FFQ were healthier and/or more health conscious than those not responding. Moreover, as the FFQ had an extra focus on fruits and vegetables, these food items may have been overestimated. Furthermore, by using PCA to derive dietary patterns, many subjective decisions were made that can impact the number and type of dietary patterns^(1,52,99,100). The DLW method is the best method to measure EE, and therefore the best method to evaluate the reported EI, but it is too expensive and impractical for application to large-scale epidemiological studies. Therefore, in this study, the reported EI was evaluated against presumed energy requirements as proposed by Black⁽¹⁰⁾. The questionnaire used in the present study was designed for a study on diet and breast cancer and focused on recreational physical activity. Therefore, recreational light, moderate and vigorous physical activities were assessed, but not occupational or household physical activities, which are important contributors to total EE. Owing to the lack of information about total physical activity, the PAL value for a sedentary lifestyle was used for our sample of middle-aged and old women. Using this PAL value may have resulted in misclassification of more physically active individuals as plausible reporters. In this cross-sectional study, the measurement of exposure and disease was made at the same time and it was impossible to determine which came first. In addition, we do not have information on whether the participant actually had the disease at the time the questionnaire was filled in. Therefore, we could not adjust for co-morbidity in the analyses. We attempted to adjust for potential confounding in the statistical analyses; however, as there is always a chance of measurement errors in the

confounders or unmeasured confounders, we cannot rule out the possibility of residual confounding.

In conclusion, in this large sample of women aged 50–69 years, we identified three dietary patterns: 'Prudent', 'Western' and 'Continental' for both all and plausible reporters. The food groups significantly contributing to the dietary patterns were quite similar for both all and plausible reporters; however, the pattern contributing most to the explanation of variances in the dietary data was the 'Prudent' pattern among all reporters and the 'Western' pattern among plausible reporters. We also found that under-reporting of EI attenuated the associations between dietary patterns and self-reported chronic diseases. Our findings suggest that under-reporting of food items can result in measurement errors of dietary patterns, which may affect the association between dietary patterns and disease. Investigators should consider reporting effect estimates for the associations between dietary patterns and disease for all individuals as well as restricting analyses to those with plausible intake.

Acknowledgements

This project was supported by the Research Council of Norway (G. U., grant number 196999). The Research Council of Norway had no role in the design, analysis or writing of this article.

M. S. M. carried out the calculations of the daily intakes of energy, nutrients and foods, the statistical analyses and drafted the manuscript. M. B. V. contributed to the statistical analysis, interpretation of the data and the revisions of the manuscript. G. U. designed the study, obtained the funding, planned and executed the data collection process and participated in the discussion of results and development of the manuscript. L. F. A. contributed to the interpretation of the data and the revisions of the manuscript. All authors read and approved the final version of the manuscript.

None of the authors have any financial or other interests concerning the outcomes of the investigation or any conflicts of interest to declare.

Supplementary material

For supplementary material/s referred to in this article, please visit <http://dx.doi.org/doi:10.1017/S000711451600218X>

References

1. Newby PK & Tucker KL (2004) Empirically derived eating patterns using factor or cluster analysis: a review. *Nutr Rev* **62**, 177–203.
2. Schoeller DA (1995) Limitations in the assessment of dietary energy intake by self-report. *Metabolism* **44**, Suppl. 2, 18–22.
3. Schoeller DA (1990) How accurate is self-reported dietary energy intake? *Nutr Rev* **48**, 373–379.
4. Subar AF, Kipnis V, Troiano RP, *et al.* (2003) Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol* **158**, 1–13.
5. Jonnalagadda SS, Mitchell DC, Smiciklas-Wright H, *et al.* (2000) Accuracy of energy intake data estimated by a multiple-pass, 24-hour dietary recall technique. *J Am Diet Assoc* **100**, 303–308 quiz 309–311.

6. Connor Gorber S & Tremblay MS (2010) The bias in self-reported obesity from 1976 to 2005: a Canada-US comparison. *Obesity (Silver Spring)* **18**, 354–361.
7. Hebert JR, Hurley TG, Peterson KE, *et al.* (2008) Social desirability trait influences on self-reported dietary measures among diverse participants in a multicenter multiple risk factor trial. *J Nutr* **138**, 226s–234s.
8. Scagliusi FB, Polacow VO, Artioli GG, *et al.* (2003) Selective underreporting of energy intake in women: magnitude, determinants, and effect of training. *J Am Diet Assoc* **103**, 1306–1313.
9. Goldberg GR, Black AE, Jebb SA, *et al.* (1991) Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-reporting. *Eur J Clin Nutr* **45**, 569–581.
10. Black AE (2000) Critical evaluation of energy intake using the Goldberg cut-off for energy intake: basal metabolic rate. A practical guide to its calculation, use and limitations. *Int J Obes Relat Metab Disord* **24**, 1119–1130.
11. Funtikova AN, Gomez SF, Fito M, *et al.* (2015) Effect of energy under-reporting on secular trends of dietary patterns in a Mediterranean population. *PLOS ONE* **10**, e0127647.
12. Shaneshin M, Jessri M & Rashidkhani B (2014) Validity of energy intake reports in relation to dietary patterns. *J Health Popul Nutr* **32**, 36–45.
13. Bailey RL, Mitchell DC, Miller C, *et al.* (2007) Assessing the effect of underreporting energy intake on dietary patterns and weight status. *J Am Diet Assoc* **107**, 64–71.
14. Winkvist A, Homell A, Hallmans G, *et al.* (2009) More distinct food intake patterns among women than men in northern Sweden: a population-based survey. *Nutr J* **8**, 12.
15. Martikainen P, Brunner E & Marmot M (2003) Socioeconomic differences in dietary patterns among middle-aged men and women. *Soc Sci Med* **56**, 1397–1410.
16. Scagliusi FB, Ferriolli E, Pfrimer K, *et al.* (2008) Under-reporting of energy intake is more prevalent in a healthy dietary pattern cluster. *Br J Nutr* **100**, 1060–1068.
17. Holmback I, Ericson U, Gullberg B, *et al.* (2009) Five meal patterns are differently associated with nutrient intakes, lifestyle factors and energy misreporting in a sub-sample of the Malmo Diet and Cancer Cohort. *Food Nutr Res* **53**, 10.3402/fnr.v53i0.1970.
18. Pryer JA, Nichols R, Elliott P, *et al.* (2001) Dietary patterns among a national random sample of British adults. *J Epidemiol Community Health* **55**, 29–37.
19. Crozier SR, Robinson SM, Borland SE, *et al.* (2006) Dietary patterns in the Southampton Women's Survey. *Eur J Clin Nutr* **60**, 1391–1399.
20. Ax E, Warensjo Lemming E, Becker W, *et al.* (2016) Dietary patterns in Swedish adults; results from a national dietary survey. *Br J Nutr* **115**, 95–104.
21. Hofvind SS, Wang H & Thoresen S (2003) The Norwegian Breast Cancer Screening Program: re-attendance related to the women's experiences, intentions and previous screening result. *Cancer Causes Control* **14**, 391–398.
22. Hofvind S, Geller B, Vacek PM, *et al.* (2007) Using the European guidelines to evaluate the Norwegian Breast Cancer Screening Program. *Eur J Epidemiol* **22**, 447–455.
23. Qureshi SA, Lund AC, Veierod MB, *et al.* (2014) Food items contributing most to variation in antioxidant intake; a cross-sectional study among Norwegian women. *BMC Public Health* **14**, 45.
24. Markussen MS, Veierod MB, Sakhi AK, *et al.* (2015) Evaluation of dietary patterns among Norwegian postmenopausal women using plasma carotenoids as biomarkers. *Br J Nutr* **113**, 672–682.
25. Andersen LF, Solvoll K, Johansson LR, *et al.* (1999) Evaluation of a food frequency questionnaire with weighed records, fatty acids, and alpha-tocopherol in adipose tissue and serum. *Am J Epidemiol* **150**, 75–87.
26. Carlsen MH, Lillegaard IT, Karlsen A, *et al.* (2010) Evaluation of energy and dietary intake estimates from a food frequency questionnaire using independent energy expenditure measurement and weighed food records. *Nutr J* **9**, 37.
27. Carlsen MH, Karlsen A, Lillegaard IT, *et al.* (2011) Relative validity of fruit and vegetable intake estimated from an FFQ, using carotenoid and flavonoid biomarkers and the method of triads. *Br J Nutr* **105**, 1530–1538.
28. Berntsen S, Hageberg R, Aandstad A, *et al.* (2010) Validity of physical activity monitors in adults participating in free-living activities. *Br J Sports Med* **44**, 657–664.
29. Qureshi SA, Ellingjord-Dale M, Hofvind S, *et al.* (2012) Physical activity and mammographic density in a cohort of postmenopausal Norwegian women; a cross-sectional study. *Springerplus* **1**, 75.
30. Dallal CM, Sullivan-Halley J, Ross RK, *et al.* (2007) Long-term recreational physical activity and risk of invasive and *in situ* breast cancer: the California Teachers Study. *Arch Intern Med* **167**, 408–415.
31. Ainsworth BE, Haskell WL, Whitt MC, *et al.* (2000) Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc* **32**, Suppl., S498–S504.
32. Food and Agriculture Organization, World Health Organization & UNU (1985) *Energy and protein requirements. Report of a joint FAO/WHO/UNU expert consultation. World Health Organization Technical Report Series*, no. 724. Geneva: WHO.
33. Food and Agriculture Organization, World Health Organization & UNU (2001) *Human Energy Requirements. Report of a Joint FAO/WHO/UNU Expert Consultation*. Geneva: WHO.
34. Henry CJ (2005) Basal metabolic rate studies in humans: measurement and development of new equations. *Public Health Nutr* **8**, 1133–1152.
35. Tabachnick BG & Fidell LS (2007) *Using Multivariate Statistics*, 5th ed. Boston, MA: Allyn & Bacon.
36. Hucheson G & Sofroniou N (1999) *The Multivariate Social Scientist*. London: Sage.
37. Michels KB & Schulze MB (2005) Can dietary patterns help us detect diet-disease associations? *Nutr Res Rev* **18**, 241–248.
38. Dekker LH, van Dam RM, Snijder MB, *et al.* (2015) Comparable dietary patterns describe dietary behavior across ethnic groups in the Netherlands, but different elements in the diet are associated with glycated hemoglobin and fasting glucose concentrations. *J Nutr* **145**, 1884–1891.
39. Gadgil MD, Anderson CA, Kandula NR, *et al.* (2015) Dietary patterns are associated with metabolic risk factors in South Asians living in the United States. *J Nutr* **145**, 1211–1217.
40. Niu K, Momma H, Kobayashi Y, *et al.* (2016) The traditional Japanese dietary pattern and longitudinal changes in cardiovascular disease risk factors in apparently healthy Japanese adults. *Eur J Nutr* **55**, 267–279.
41. Barbaresko J, Siegert S, Koch M, *et al.* (2014) Comparison of two exploratory dietary patterns in association with the metabolic syndrome in a Northern German population. *Br J Nutr* **112**, 1364–1372.

42. Odegaard AO, Koh WP, Yuan JM, *et al.* (2014) Dietary patterns and mortality in a Chinese population. *Am J Clin Nutr* **100**, 877–883.
43. Mourouti N, Papavagelis C, Plytzanopoulou P, *et al.* (2015) Dietary patterns and breast cancer: a case-control study in women. *Eur J Nutr* **54**, 609–617.
44. Engeset D, Hofoss D, Nilsson LM, *et al.* (2015) Dietary patterns and whole grain cereals in the Scandinavian countries – differences and similarities. The HELGA project. *Public Health Nutr* **18**, 905–915.
45. Fransen HP, May AM, Stricker MD, *et al.* (2014) *A posteriori* dietary patterns: how many patterns to retain? *J Nutr* **144**, 1274–1282.
46. Link LB, Canchola AJ, Bernstein L, *et al.* (2013) Dietary patterns and breast cancer risk in the California Teachers Study cohort. *Am J Clin Nutr* **98**, 1524–1532.
47. Santaliestra-Pasias AM, Mouratidou T, Huybrechts I, *et al.* (2014) Increased sedentary behaviour is associated with unhealthy dietary patterns in European adolescents participating in the HELENA study. *Eur J Clin Nutr* **68**, 300–308.
48. Chocano-Bedoya PO, O'Reilly EJ, Lucas M, *et al.* (2013) Prospective study on long-term dietary patterns and incident depression in middle-aged and older women. *Am J Clin Nutr* **98**, 813–820.
49. Northstone K, Smith AD, Cribb VL, *et al.* (2014) Dietary patterns in UK adolescents obtained from a dual-source FFQ and their associations with socio-economic position, nutrient intake and modes of eating. *Public Health Nutr* **17**, 1476–1485.
50. Kristiansen AL, Lande B, Sexton JA, *et al.* (2013) Dietary patterns among Norwegian 2-year-olds in 1999 and in 2007 and associations with child and parent characteristics. *Br J Nutr* **110**, 135–144.
51. Charreire H, Kesse-Guyot E, Bertrais S, *et al.* (2011) Associations between dietary patterns, physical activity (leisure-time and occupational) and television viewing in middle-aged French adults. *Br J Nutr* **105**, 902–910.
52. Smith AD, Emmett PM, Newby PK, *et al.* (2013) Dietary patterns obtained through principal components analysis: the effect of input variable quantification. *Br J Nutr* **109**, 1881–1891.
53. Balder HF, Virtanen M, Brants HA, *et al.* (2003) Common and country-specific dietary patterns in four European cohort studies. *J Nutr* **133**, 4246–4251.
54. Hearty AP & Gibney MJ (2009) Comparison of cluster and principal component analysis techniques to derive dietary patterns in Irish adults. *Br J Nutr* **101**, 598–608.
55. Schulze MB, Hoffmann K, Kroke A, *et al.* (2001) Dietary patterns and their association with food and nutrient intake in the European prospective investigation into cancer and nutrition (EPIC)-Potsdam study. *Br J Nutr* **85**, 363–373.
56. Barros AJ & Hirakata VN (2003) Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. *BMC Med Res Methodol* **3**, 21.
57. Hill RJ & Davies PS (2001) The validity of self-reported energy intake as determined using the doubly labelled water technique. *Br J Nutr* **85**, 415–430.
58. Livingstone MB & Black AE (2003) Markers of the validity of reported energy intake. *J Nutr* **133**, Suppl. 3, 895s–920s.
59. Schoeller DA & Schoeller DA (2002) Validation of habitual energy intake. *Public Health Nutr* **5**, 883–888.
60. Black AE, Goldberg GR, Jebb SA, *et al.* (1991) Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Evaluating the results of published surveys. *Eur J Clin Nutr* **45**, 583–599.
61. Salle A, Ryan M & Ritz P (2006) Underreporting of food intake in obese diabetic and nondiabetic patients. *Diabetes Care* **29**, 2726–2727.
62. Yannakoulia M, Panagiotakos DB, Pitsavos C, *et al.* (2007) Low energy reporting related to lifestyle, clinical, and psychosocial factors in a randomly selected population sample of Greek adults: the ATTICA study. *J Am Coll Nutr* **26**, 327–333.
63. Rasmussen LB, Matthiessen J, Biloft-Jensen A, *et al.* (2007) Characteristics of misreporters of dietary intake and physical activity. *Public Health Nutr* **10**, 230–237.
64. Poslusna K, Ruprich J, de Vries JH, *et al.* (2009) Misreporting of energy and micronutrient intake estimated by food records and 24 hour recalls, control and adjustment methods in practice. *Br J Nutr* **101**, Suppl. 2, S73–S85.
65. Mendez MA, Popkin BM, Buckland G, *et al.* (2011) Alternative methods of accounting for underreporting and overreporting when measuring dietary intake-obesity relations. *Am J Epidemiol* **173**, 448–458.
66. Ferrari P, Slimani N, Ciampi A, *et al.* (2002) Evaluation of under- and overreporting of energy intake in the 24-hour diet recalls in the European Prospective Investigation into cancer and nutrition (EPIC). *Public Health Nutr* **5**, 1329–1345.
67. Rennie KL, Coward A & Jebb SA (2007) Estimating under-reporting of energy intake in dietary surveys using an individualised method. *Br J Nutr* **97**, 1169–1176.
68. Becker W, Foley S, Shelley E, *et al.* (1999) Energy under-reporting in Swedish and Irish dietary surveys: implications for food-based dietary guidelines. *Br J Nutr* **81**, Suppl. 2, S127–S131.
69. Murakami K, Sasaki S & Okubo H (2012) Characteristics of under- and over-reporters of energy intake among young Japanese women. *J Nutr Sci Vitaminol (Tokyo)* **58**, 253–262.
70. Freisling H, van Bakel MM, Biessy C, *et al.* (2012) Dietary reporting errors on 24 h recalls and dietary questionnaires are associated with BMI across six European countries as evaluated with recovery biomarkers for protein and potassium intake. *Br J Nutr* **107**, 910–920.
71. Koebnick C, Wagner K, Thielecke F, *et al.* (2005) An easy-to-use semiquantitative food record validated for energy intake by using doubly labelled water technique. *Eur J Clin Nutr* **59**, 989–995.
72. Luhrmann PM, Herbert BM & Neuhauser-Berthold M (2001) Underreporting of energy intake in an elderly German population. *Nutrition* **17**, 912–916.
73. Murakami K & Livingstone MB (2015) Prevalence and characteristics of misreporting of energy intake in US adults: NHANES 2003–2012. *Br J Nutr* **114**, 1294–1303.
74. Berta Vanrullen I, Volatier JL, Bertaut A, *et al.* (2014) Characteristics of energy intake under-reporting in French adults. *Br J Nutr* **111**, 1292–1302.
75. Gemming L, Jiang Y, Swinburn B, *et al.* (2014) Under-reporting remains a key limitation of self-reported dietary intake: an analysis of the 2008/09 New Zealand Adult Nutrition Survey. *Eur J Clin Nutr* **68**, 259–264.
76. Kye S, Kwon SO, Lee SY, *et al.* (2014) Under-reporting of energy intake from 24-hour dietary recalls in the Korean National Health and Nutrition Examination Survey. *Osong Public Health Res Perspect* **5**, 85–91.
77. Lutomski JE, van den Broeck J, Harrington J, *et al.* (2011) Sociodemographic, lifestyle, mental health and dietary factors associated with direction of misreporting of energy intake. *Public Health Nutr* **14**, 532–541.
78. Johansson L, Solvoll K, Bjorneboe GE, *et al.* (1998) Under- and overreporting of energy intake related to weight

status and lifestyle in a nationwide sample. *Am J Clin Nutr* **68**, 266–274.

79. Park HA, Lee JS & Kuller LH (2007) Underreporting of dietary intake by body mass index in premenopausal women participating in the healthy women study. *Nutr Res Pract* **1**, 231–236.
80. Markussen MS, Veierod MB, Kristiansen AL, *et al.* (2016) Dietary patterns of postmenopausal women and associations with nutrient intake, sociodemographic factors and key risk factors for noncommunicable diseases. *Public Health Nutr* (epublication ahead of print version 19 January 2016).
81. Lafay L, Mennen L, Basdevant A, *et al.* (2000) Does energy intake underreporting involve all kinds of food or only specific food items? Results from the Fleurbaix Laventie Ville Sante (FLVS) study. *Int J Obes Relat Metab Disord* **24**, 1500–1506.
82. Krebs-Smith SM, Graubard BI, Kahle LL, *et al.* (2000) Low energy reporters vs others: a comparison of reported food intakes. *Eur J Clin Nutr* **54**, 281–287.
83. Williams DE, Prevost AT, Whichelow MJ, *et al.* (2000) A cross-sectional study of dietary patterns with glucose intolerance and other features of the metabolic syndrome. *Br J Nutr* **83**, 257–266.
84. Herber-Gast GC & Mishra GD (2013) Fruit, Mediterranean-style, and high-fat and -sugar diets are associated with the risk of night sweats and hot flushes in midlife: results from a prospective cohort study. *Am J Clin Nutr* **97**, 1092–1099.
85. Wirfalt E, Mattisson I, Gullberg B, *et al.* (2000) Food patterns defined by cluster analysis and their utility as dietary exposure variables: a report from the Malmo Diet and Cancer Study. *Public Health Nutr* **3**, 159–173.
86. Hornell A, Winkvist A, Hallmans G, *et al.* (2010) Misreporting, previous health status and health status of family may seriously bias the association between food patterns and disease. *Nutr J* **9**, 48.
87. Anderson AL, Harris TB, Houston DK, *et al.* (2010) Relationships of dietary patterns with body composition in older adults differ by gender and PPAR-gamma Pro12Ala genotype. *Eur J Nutr* **49**, 385–394.
88. Flores M, Macias N, Rivera M, *et al.* (2010) Dietary patterns in Mexican adults are associated with risk of being overweight or obese. *J Nutr* **140**, 1869–1873.
89. Villegas R, Yang G, Gao YT, *et al.* (2010) Dietary patterns are associated with lower incidence of type 2 diabetes in middle-aged women: the Shanghai Women's Health Study. *Int J Epidemiol* **39**, 889–899.
90. Sanchez-Villegas A, Delgado-Rodriguez M, Martinez-Gonzalez MA, *et al.* (2003) Gender, age, socio-demographic and lifestyle factors associated with major dietary patterns in the Spanish Project SUN (Seguimiento Universidad de Navarra). *Eur J Clin Nutr* **57**, 285–292.
91. Song Y & Joung H (2012) A traditional Korean dietary pattern and metabolic syndrome abnormalities. *Nutr Metab Cardiovasc Dis* **22**, 456–462.
92. Tucker KL, Chen H, Hannan MT, *et al.* (2002) Bone mineral density and dietary patterns in older adults: the Framingham Osteoporosis Study. *Am J Clin Nutr* **76**, 245–252.
93. Mattisson I, Wirfalt E, Aronsson CA, *et al.* (2005) Misreporting of energy: prevalence, characteristics of misreporters and influence on observed risk estimates in the Malmo Diet and Cancer cohort. *Br J Nutr* **94**, 832–842.
94. Prentice RL, Shaw PA, Bingham SA, *et al.* (2009) Biomarker-calibrated energy and protein consumption and increased cancer risk among postmenopausal women. *Am J Epidemiol* **169**, 977–989.
95. Neuhouser ML, Tinker L, Shaw PA, *et al.* (2008) Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. *Am J Epidemiol* **167**, 1247–1259.
96. McNaughton SA, Mishra GD & Brunner EJ (2009) Food patterns associated with blood lipids are predictive of coronary heart disease: the Whitehall II study. *Br J Nutr* **102**, 619–624.
97. McNaughton SA, Mishra GD & Brunner EJ (2008) Dietary patterns, insulin resistance, and incidence of type 2 diabetes in the Whitehall II Study. *Diabetes Care* **31**, 1343–1348.
98. Brunner EJ, Mosdol A, Witte DR, *et al.* (2008) Dietary patterns and 15-y risks of major coronary events, diabetes, and mortality. *Am J Clin Nutr* **87**, 1414–1421.
99. Hu FB (2002) Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* **13**, 3–9.
100. Northstone K, Ness AR, Emmett PM, *et al.* (2008) Adjusting for energy intake in dietary pattern investigations using principal components analysis. *Eur J Clin Nutr* **62**, 931–938.