

### RESEARCH ARTICLE

## Reproducibility of dietary intakes of macronutrients, specific food groups, and dietary patterns in 211 050 adults in the UK Biobank study

Jennifer L. Carter<sup>1\*</sup> , Sarah Lewington<sup>1,2</sup>, Carmen Piernas<sup>3</sup>, Kathryn Bradbury<sup>4</sup>, Timothy J. Key<sup>5</sup>, Susan A. Jebb<sup>2,3</sup>, Matthew Arnold<sup>6</sup>, Derrick Bennett<sup>1</sup> and Robert Clarke<sup>1</sup>

<sup>1</sup>Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>2</sup>Medical Research Council Population Health Research Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>3</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

<sup>4</sup>National Institute for Health Innovation, School of Population Health, The University of Auckland, Auckland, New Zealand

<sup>5</sup>Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>6</sup>Department of Public Health and Primary Care, University of Cambridge, Cambridge, Cambridgeshire, UK

(Received 22 May 2019 – Final revision received 19 September 2019 – Accepted 20 September 2019)

*Journal of Nutritional Science* (2019), vol. 8, e34, page 1 of 9

doi:10.1017/jns.2019.31

### Abstract

To detect modest associations of dietary intake with disease risk, observational studies need to be large and control for moderate measurement errors. The reproducibility of dietary intakes of macronutrients, food groups and dietary patterns (vegetarian and Mediterranean) was assessed in adults in the UK Biobank study on up to five occasions using a web-based 24-h dietary assessment ( $n$  211 050), and using short FFQ recorded at baseline ( $n$  502 655) and after 4 years ( $n$  20 346). When the means of two 24-h assessments were used, the intra-class correlation coefficients (ICC) for macronutrients varied from 0.63 for alcohol to 0.36 for polyunsaturated fat. The ICC for food groups also varied from 0.68 for fruit to 0.18 for fish. The ICC for the FFQ varied from 0.66 for meat and fruit to 0.48 for bread and cereals. The reproducibility was higher for vegetarian status ( $\kappa > 0.80$ ) than for the Mediterranean dietary pattern (ICC = 0.45). Overall, the reproducibility of pairs of 24-h dietary assessments and single FFQ used in the UK Biobank were comparable with results of previous prospective studies using conventional methods. Analyses of diet–disease relationships need to correct for both measurement error and within-person variability in dietary intake in order to reliably assess any such associations with disease in the UK Biobank.

**Key words:** 24-h recall: Dietary assessment: Dietary patterns: Food groups: Macronutrients: Reproducibility

Diet has been estimated to account for about one-third of all deaths from CVD and cancer<sup>(1,2)</sup>. Recent research has continued to highlight the ongoing limitations in the methods used to reliably assess dietary intake in population studies using either 24-h dietary recalls or FFQ<sup>(3,4)</sup>. Epidemiological studies need to be large enough to detect modest differences in risk, and failure to appreciate the importance of within-person variability in dietary intake will result in underestimation of associations with disease in population studies. It is important to obtain repeat measurements in individuals of dietary intake

at intervals of several months or years to control for random measurement error, within-person variability and true changes in diet, as all of these sources of variability contribute to misclassification of an individual's 'usual' dietary intake over a long period of time<sup>(5)</sup>. By controlling for these sources of variability, reliable estimates of the true associations between the 'usual' levels of dietary intake over a particular period and incident disease at the same or later time can be obtained<sup>(5,6)</sup>. Recent advances in technology now permit serial assessments of self-completed dietary questionnaires using web-based

**Abbreviation:** ICC, intra-class correlation coefficient.

\* **Corresponding author:** Jennifer L. Carter, fax +44 1865 743985, email [jennifer.carter@ndph.ox.ac.uk](mailto:jennifer.carter@ndph.ox.ac.uk)



platforms in large-scale cohort studies. The potential for these digital technologies to improve the estimation of dietary intake, however, is ongoing and requires quantification prior to their use in population studies of diet and incident disease.

The UK Biobank is a prospective cohort study of 0.5 million adults and is one of the largest prospective studies with repeat 24-h assessments of dietary intake and FFQ. Measuring the reproducibility of these measurements, or the degree of consistency between the repeated 24-h assessments, will quantify the magnitude of within-person variability (i.e. change or variation in an individual's dietary intake and chance fluctuations in a person's recorded dietary intake that average to 'true' intake over repeat administrations)<sup>(5,7,8)</sup>. While reproducibility does not indicate how valid these 24-h assessments are (i.e. how accurately the measurements reflect true dietary intake), the estimates of reproducibility can still be used to correct for the underestimation of relative risks arising from random measurement errors in exposures (i.e. 'regression dilution bias') in order to calculate more reliable associations between 'usual' dietary intake and incident disease outcomes in the UK Biobank<sup>(5,6)</sup>. The aims of the present report were to estimate the reproducibility of dietary intakes of macronutrients, food groups and selected dietary patterns using serial web-based 24-h dietary assessments in 211 050 participants from the UK Biobank study who completed such assessments; and to compare measurements of intakes from food groups from the mean of 24-h dietary assessments with those estimated using a single short FFQ on all 0.5 million participants.

## Methods

The UK Biobank is a population-based prospective cohort study of 502 655 participants aged 40–69 years at recruitment between 2006 and 2010. Participants living within 25 miles (40.2 km) of twenty-two assessment centres across England, Wales and Scotland were identified through National Health Service (NHS) central registers (containing 98 % of the UK population). All eligible adults (about 9.2 million) were sent an invitation letter, with a response rate of about 5.5 %. Details of the study protocol have been published elsewhere<sup>(9,10)</sup>. UK Biobank participants completed two types of dietary assessments on several occasions between 2006 and 2013: a short FFQ and a 24-h dietary assessment (see Fig. 1). At baseline assessment, all participants completed a short FFQ (described below). Towards the end of the recruitment period an enhancement was added to the baseline assessment protocol so that, between April 2009 and September 2010, 70 724 participants attending their first assessment centre visit at recruitment completed a modified 24-h dietary assessment capturing the foods and beverages consumed during the previous 24 h (detailed below; see Fig. 1). Between February 2011 and April 2012, identical web-based 24-h dietary assessments were emailed on four separate occasions (cycles) over a 16-month period to 331 013 participants who provided a valid email address. Invitations were scheduled to maximise variation by season and day of the week. The response rate for the web-based assessment was 53 % (*n* 176 012), with a total of 211 050 participants completing at

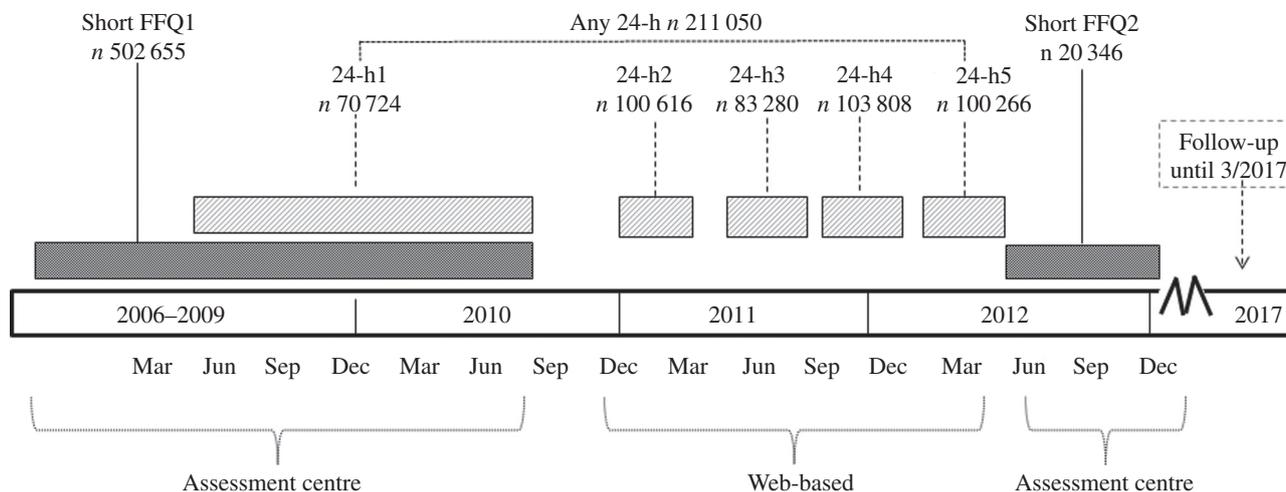
least one 24-h dietary assessment online or at the assessment centre. Of these, 84 175 (40 %) completed a single assessment; 48 129 (23 %) completed two assessments; 42 492 (20 %) completed three assessments; 30 488 (15 %) completed four assessments and 5766 (3 %) completed five assessments. The response rate varied between 26 and 33 % in each cycle, with some participants missing one cycle and then returning to complete a later cycle.

### *Dietary measures from the 24-h dietary assessment*

Unlike standard 24-h dietary recalls that ask respondents to report everything they ate in the previous day, the 24-h dietary assessments in the UK Biobank presented participants with a list of commonly consumed foods on which they indicated their consumption in the previous 24 h<sup>(11–13)</sup>. Consumption of up to 206 widely consumed foods and thirty-two types of drinks were recorded using the 24-h dietary assessment, which took approximately 15 min to complete (median 15 (interquartile range 11–19) min). Standard categories and descriptions were used for portion sizes of each food (for example, slices of bread), and participants were asked to enter the ingredients individually for composite dishes (for example, so spaghetti bolognaise would be recorded as separate items for pasta, meat and tomato-based sauce)<sup>(12)</sup>. An open text box was included at the end for additional items that were not included on the food list. Participants also indicated if their food consumption in the previous day was 'typical' or if they followed any special diets (for example, vegetarian or gluten-free). Total energy intake (kJ) and macronutrient (g) values were estimated for protein, total fat, saturated fat, polyunsaturated fat, carbohydrates, total sugars, starch, fibre and alcohol using previously validated methods<sup>(12)</sup>. Nutrient intakes were calculated using associated portion sizes for each food/beverage item and by multiplying the amount consumed by the nutrient composition, using standard food consumption tables for the UK<sup>(14–23)</sup>. Reproducibility was also calculated for six commonly consumed food groups: (i) meat (beef, lamb, pork, bacon, ham, sausage, liver, or chicken), (ii) cheese, (iii) bread and breakfast cereals, (iv) fruit (fresh or dried), (v) vegetables (excluding potatoes) and (vi) fish (oily and not oily).

### *Dietary measures from the short FFQ*

All 0.5 million UK Biobank participants completed a short computer touchscreen questionnaire at their initial assessment centre visit that included twenty-nine questions about their average diet over the previous 12 months, most of which asked about the frequency of consumption of the six food groups specified above. In addition, a sample of 20 346 participants completed a repeat assessment about 4 years later (Fig. 1). Typically, responses to the short FFQ included: never, less than once per week, once per week, 2–4 times per week, 5–6 times per week, or once or more daily. For fruit and vegetables, participants were asked to directly enter the average number of servings consumed each day; for bread and breakfast cereals, participants entered the average number of servings per week. Daily intake (g) for the short



**Fig. 1.** Timeline of UK Biobank dietary assessment measurement 2006–2015, including the 24-h dietary assessment (24-h) and the short FFQ.

FFQ was calculated by multiplying the frequency of intake by standard portion size (g) in the UK (for example, eating chicken 5–6 times per week would have been (5.5/7) multiplied by 120 g)<sup>(24)</sup>. See Supplementary Methods for additional details.

Vegetarian and Mediterranean dietary patterns were also assessed as these dietary patterns have been more strongly associated with incident disease than individual nutrients. Participants were classified as consuming a vegetarian diet if they self-reported a vegetarian diet (24-h dietary assessment) or if they reported never eating any meat or fish (short FFQ). A Mediterranean-style diet score, as defined by previous research, was estimated for each cycle of the 24-h dietary assessment by summing the indicators for each of the following (diet score range 0–9): if participants consumed greater than the median servings per d for vegetables, legumes, fruit, nuts, fish, wholegrains and the ratio of dietary monounsaturated to saturated fat intake; less than the median intake for red and processed meat; and between 5 and 25 g per d for ethanol<sup>(25,26)</sup>.

### Statistical methods

The mean intakes of individual macronutrients, food groups and dietary patterns were calculated across 24-h dietary assessments. The Mediterranean diet score and macronutrient intakes were normally distributed, except for alcohol, which was transformed with a rank-based inverse normal transformation; distributions of food group intakes were similarly transformed. Reproducibility was assessed using intra-class correlation coefficients (ICC; i.e. the proportion of between-subject variance to the total variance) from a one-way ANOVA for all dietary intakes, except for a vegetarian dietary pattern, which was assessed with a weighted  $\kappa$  statistic because it was a dichotomous variable<sup>(27)</sup>. Since variation in daily dietary intake is expected to be high, the average of an increasing number of 24-h dietary assessments should reduce random error and provide a better estimate of the true 'usual' intake<sup>(7,28)</sup>. In the main analyses, the mean of the first two

cycles of the 24-h dietary assessment was compared with the mean of the subsequent two cycles of the 24-h dietary assessment for participants who completed at least four 24-h assessments ( $n$  36 254, 17.2 % of the total sample). To assess the impact of bias due to systematically under- or over-reporting energy intake, the ICC were estimated before and after adjustment for total energy intake using the nutrient residuals method (where residuals from a regression of energy on the nutrient of interest are used when calculating reproducibility in order to determine the variation in nutrient intake due only to the composition of the diet)<sup>(29)</sup>. Participants were excluded from analyses if they reported implausible values of total energy intake that were approximately >99 % of the sex-specific distribution (>20 000 kJ for men and >18 000 kJ for women; approximately 700 participants from each cycle). A minimum threshold of energy intake was not imposed since participants could have had a special dietary restriction prior to their 24-h dietary assessment.

Reproducibility was calculated using the 24-h dietary assessment separately for subgroups classified by season, age group, socio-economic group, BMI and if collected on weekdays or weekends. Ignoring the first cycle of the 24-h dietary assessment, which was administered over a 17-month period in the assessment centre, seasonality was analysed by comparing measurements from the online cycles 2–5 as these were each recorded over 3-month periods corresponding to different seasons. Cycle 2 was recorded in winter (February–April 2011), cycle 3 in summer (June–August 2011), cycle 4 in autumn (October–December 2011), and cycle 5 in spring (April–June 2012). Age at the first 24-h dietary assessment was classified into 5-year age bands (<45, 45–49, 50–54, 55–59, 60–64, 65+ years); and socio-economic group (assessed by the Townsend index of deprivation based on the participant's postcode at recruitment) was classified into quintiles<sup>(30)</sup>. BMI at recruitment was classified according to WHO standard categories (<18.5 kg/m<sup>2</sup>, underweight; 18.5–24.9 kg/m<sup>2</sup>, normal weight; 25–29.9 kg/m<sup>2</sup>, overweight; 30–34.9 kg/m<sup>2</sup>, obese class I; 35 kg/m<sup>2</sup> +, obese class II/III). Analyses of day of the week compared assessments recorded



**Table 1.** Selected characteristics of all UK Biobank participants and the subset who completed at least one 24-h dietary assessment\* (Numbers of participants and percentages; mean values and standard deviations)

	Overall (2006–2010)		24-h Assessment (2011–2012)	
	<i>n</i>	%	<i>n</i>	%
Participants ( <i>n</i> )	502 655	–	211 050	–
Demographics				
Women	273 472	54	116 263	55
Men	229 183	46	94 787	45
Age (years)				
Mean		57		56
SD		8		8
White	472 837	95	201 319	96
Higher degree	161 215	33	89 779	43
Townsend index (% highest quintile)	100 691	20	44 940	21
Medical history				
Prior CVD	149 355	30	55 772	26
Prior diabetes	26 408	5	8869	4
Prior cancer	38 623	8	15 973	8
Lifestyle factors				
Alcohol intake				
Daily	101 792	20	48 155	23
1–4/week	244 791	49	105 092	50
<1/week	154 569	31	57 626	27
Days of moderate/vigorous exercise per week				
Mean		2.6		2.6
SD		1.9		1.8
Current smoking	52 990	11	16 555	8
Clinical measurements				
BMI (kg/m <sup>2</sup> )				
Mean		27		27
SD		4.8		4.7
Systolic blood pressure (mmHg)				
Mean		138		137
SD		19		18

\* Number missing: Townsend index (overall *n* 627, 0.1 %; 24-h *n* 265, 0.1 %), exercise (overall *n* 12 227, 2.4 %; 24-h *n* 2083, 0.9 %), BMI (overall *n* 3106, 0.6 %; 24-h *n* 597, 0.3 %), systolic blood pressure (overall *n* 30135, 6.0 %; 24-h *n* 7845, 3.7 %).

concordantly on weekdays (Monday–Friday) or weekends (Saturday/Sunday), with those recorded discordantly on both. All analyses were stratified by sex, and overall estimates were calculated as an inverse-variance weighted average of the sex-specific estimates.

In addition to comparing estimates of reproducibility for food group intake measured on the 24-h assessment and short FFQ, correlation coefficients between the FFQ and 24-h dietary assessment were estimated to compare the rankings of usual food group consumption. This was done using a method that deattenuated the coefficients for random error in the 24-h dietary assessments (by accounting for the ICC of the variable number of 24-h assessments) in the subset of those completing at least two 24-h dietary assessments<sup>(7,31)</sup>. All analyses were conducted using Stata version 13<sup>(32)</sup>.

### Ethics

Research in the UK Biobank study was conducted according to the Declaration of Helsinki, and was approved by the North West Multi-Centre Research Ethics Committee (reference number 06/MRE08/65). During the touchscreen assessment at baseline, all participants provided informed consent to

participate in the UK Biobank and be followed up, using a signature capture device.

### Results

#### Population characteristics and intake of macronutrients

Selected baseline characteristics of the study participants completing at least one 24-h dietary assessment and those in the overall UK Biobank study population are shown in Table 1. Similar to the overall cohort, participants who completed a 24-h dietary assessment had a mean age of 56 (SD 8) years, 55 % were women, 96 % reported having White British ancestry and 21.3 % were in the most affluent quintile of the Townsend index. BMI and blood pressure were also similar in both groups, but the participants who completed the 24-h assessment were more likely to have a university degree (43 *v.* 33 %) and less likely to be current smokers (8 *v.* 11 %) compared with those in the overall cohort.

The five cycles of the 24-h dietary assessment were recorded at intervals over a 28-month period (mean interval 5.9 (SD 2.0) months), and the absolute mean intake of macronutrients was similar across different cycles for both sexes (Supplementary Table S1). Although men consumed higher absolute amounts of energy and of each macronutrient than women, both sexes



**Table 2.** Dietary intake of macronutrients assessed by 24-h dietary assessment in the UK Biobank (Mean values and standard deviations)

	Overall (n 210 143)*				Men (n 94 311)				Women (n 115 832)			
	% Energy		g/d		% Energy		g/d		% Energy		g/d	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Energy (kJ)	–		8733	2452	–		9464	2570	–		8136	2177
Total fat	32	7	77.1	29.2	32	7	82.8	31.0	32	7	72.4	26.8
Saturated fat	12	4	29.5	12.4	12	4	31.9	13.3	12	4	27.6	11.3
Polyunsaturated fat	6	2	14.2	7.1	6	2	15.2	7.5	6	2	13.5	6.7
Monounsaturated fat	14	3	33.3	13.0	14	3	35.7	13.8	14	3	31.3	12.0
Protein	16	4	81.6	24.7	16	4	86.4	26.2	17	4	77.8	22.5
Carbohydrates	46	8	251.3	80.4	46	8	269.2	84.3	47	8	236.7	73.9
Total sugars	22	7	119.3	47.8	21	7	124.5	50.1	23	7	115.1	45.4
Starch	22	6	121.8	46.1	23	6	134.6	48.7	22	6	111.4	41.1
Fibre density†	1.9	0.7	16.3	6.5	1.8	0.7	16.6	6.8	2.0	0.7	16.1	6.3
Alcohol	5	7	16.1	21.1	6	7	21.4	21.2	4	6	11.8	15.9

\* For any participant with >1 measurement, the average over all 24-h cycles has been used.

† Fibre % energy is calculated as fibre density (g/MJ).

consumed similar proportions (% from energy) of macronutrients (Table 2).

### Reproducibility of macronutrient intake

The reproducibility of single 24-h dietary assessment was modest, with ICC ranging from 0.23 for polyunsaturated fat to 0.46 for sugars (Table 3). As expected, the estimates of reproducibility were higher when the means of two cycles of 24-h assessments were compared; for most macronutrients reproducibility estimates were moderate and varied between 0.50 and 0.60 (range 0.36 for polyunsaturated fat to 0.63 for alcohol).

Reproducibility was typically slightly higher in men than in women, and in participants aged  $\geq 65$  years than in younger participants, but did not vary substantially by season, day of the week, age, or index of deprivation. The reproducibility was also unaltered by adjustment for total energy intake so the nutrient residuals method is reported in the Supplement, and the main results present the unadjusted estimates (Supplementary Tables S2–S4). The only exception was

alcohol intake where reproducibility was greater if both measurements were recorded at weekends rather than on weekdays (for example, 0.60 *v.* 0.47 for men). The reproducibility for alcohol also increased with age (for example, from 0.38 in men <45 years to 0.52 in those aged  $\geq 65$  years). Participants who were underweight reported the highest reproducibility for macronutrient intake, with small differences seen across other groups categorised by BMI (Supplementary Table S5).

### Reproducibility of selected food groups and dietary patterns

The average dietary intake of food groups and dietary patterns from the short FFQ and 24-h assessments can be seen in Supplementary Table S6. Overall, the reproducibility of consumption of food groups assessed using a single 24-h dietary assessment was modest and varied between 0.10 (fish) to 0.52 (fruit) (Table 4). Reproducibility improved when comparing the means of two 24-h dietary assessments (range 0.18 for fish to 0.68 for fruit), but was still modest for several food groups like meat, cheese and fish. The reproducibility for

**Table 3.** Intra-class correlations of dietary intake of macronutrients estimated using single 24-h assessments and the average of two 24-h assessments, by sex\*

Nutrient (g)	Single measurements			Cycles (1 + 2) <i>v.</i> (3 + 4)		
	Overall	Men	Women	Overall	Men	Women
<i>n</i>	210 143	94 311	115 832	36 140	15 852	20 288
Energy (kJ)	0.35	0.36	0.34	0.50	0.51	0.48
Total fat	0.33	0.35	0.31	0.48	0.50	0.46
Saturated fat	0.35	0.38	0.33	0.51	0.54	0.49
Polyunsaturated fat	0.23	0.24	0.22	0.36	0.37	0.35
Monounsaturated fat	0.31	0.33	0.30	0.46	0.48	0.44
Protein	0.30	0.30	0.29	0.43	0.43	0.42
Carbohydrates	0.40	0.42	0.38	0.55	0.57	0.53
Total sugars	0.46	0.48	0.44	0.61	0.63	0.59
Starch	0.34	0.34	0.33	0.48	0.49	0.47
Fibre	0.44	0.45	0.44	0.60	0.62	0.59
Alcohol	0.44	0.45	0.44	0.63	0.65	0.62

\* Cycle (1 + 2) *v.* (3 + 4) intra-class correlations are computed from means of cycles 1 and 2, correlated with means of cycles 3 and 4.



**Table 4.** Intra-class correlations of food groups and dietary patterns in UK Biobank 24-h dietary assessment (*n* 211 050; '24-h') and short FFQ (*n* 502 655; 'FFQ')\*

	Overall						Men						Women					
	24-h			FFQ1			24-h			FFQ1			24-h			FFQ1		
	v. 24-h	(1 + 2) v. (3 + 4)	<i>n</i>	v. FFQ2	v. 24-h††	<i>n</i>	v. 24-h	(1 + 2) v. (3 + 4)	<i>n</i>	v. FFQ2	v. 24-h††	<i>n</i>	v. 24-h	(1 + 2) v. (3 + 4)	<i>n</i>	v. FFQ2	v. 24-h††	
<b>Food groups</b>																		
Total meat	0.24	0.38	126 875	0.66	0.50	20 346	0.23	0.37	15 915	0.65	0.46	56 142	0.24	0.40	20 339	0.68	0.53	
Cheese	0.22	0.38		0.63	0.51		0.24	0.39		0.64	0.53		0.21	0.36		0.61	0.49	
Fish	0.10	0.18		0.65	0.47		0.10	0.18		0.64	0.50		0.09	0.18		0.66	0.45	
Bread and cereals	0.37	0.55		0.48	0.52		0.37	0.55		0.49	0.51		0.37	0.54		0.47	0.54	
Fruit	0.52	0.68		0.66	0.63		0.53	0.70		0.67	0.64		0.51	0.66		0.64	0.61	
Vegetables	0.31	0.46		0.59	0.38		0.31	0.46		0.59	0.38		0.31	0.47		0.59	0.38	
<b>Dietary patterns</b>																		
Mediterranean‡	0.30	0.45		0.92	0.69		0.30	0.47		0.93	0.69		0.29	0.44		0.91	0.68	
Vegetarian§	0.81	0.84					0.81	0.84					0.81	0.84				

\* *n* is the maximum number of participants in each column, with each food group reporting <1% missing.

† FFQ1 v. 24-h is the deattenuated correlation coefficient between baseline FFQ measurements and mean intake on 24-h assessments for those with 2+ cycles.

‡ Greater than median intake (servings/d) for vegetables, legumes, fruit, nuts, fish, wholegrains, ratio of monounsaturated to saturated fat; less than median intake for red and processed meat; and between 5 and 25 g/d for ethanol.

§ Reproducibility coefficients for the vegetarian dietary pattern are  $\kappa$  statistics.

food groups assessed using the short FFQ was generally higher, with most ICC around 0.60, although two food groups reported higher reproducibility on the 24-h assessments (bread/cereal and fruit). Comparisons of the intakes of food groups assessed using the short FFQ and the averages of all 24-h dietary assessments (if  $\geq 2$  cycles completed) demonstrated moderate agreement, with deattenuated correlation coefficients varying from 0.38 for vegetables to 0.63 for fruit.

Similar to estimates for food groups, the reproducibility of the Mediterranean dietary pattern was 0.45 when the means of two pairs of 24-h dietary assessments were compared. There was very good agreement when a vegetarian dietary pattern was reported using either the short FFQ or the 24-h dietary assessment ( $\kappa > 0.80$ ).

As a sensitivity analysis, reproducibility was estimated after omitting participants reporting an 'atypical' diet in the previous 24 h, and the levels of reproducibility on 24-h dietary assessments improved by 0.03–0.04 for all types of dietary intake.

## Discussion

The reproducibility of the self-completed 24-h dietary assessments and the short FFQ was assessed in the large prospective UK Biobank study. The reproducibility of dietary intake assessed using the mean of two 24-h dietary assessments was moderate and generally varied between 0.50 and 0.60 for macronutrients. However, the reproducibility estimates of selected food groups using the mean of two 24-h dietary assessments were variable, and generally lower (albeit not consistently) than estimates reported using the short FFQ. The reproducibility of the vegetarian dietary pattern was high irrespective of the type of dietary assessment, although the reproducibility of the Mediterranean dietary pattern was lower and comparable with those for food groups and macronutrients.

As expected, the reproducibility estimates of all measures of dietary intake assessed solely using a pair of single 24-h dietary assessment in the UK Biobank were generally modest (between 0.30 and 0.40) due to measurement error and the day-to-day variation in dietary intake, consistent with estimates from previous studies<sup>(28,33–39)</sup>. While a single 24-h dietary assessment does not provide an adequate indication of usual long-term dietary intake for individuals, such measurements have still been used in some observational studies of diet and disease<sup>(40)</sup>, despite clear evidence from previous research that the means of several 24-h dietary assessments for individuals are required for reliable estimation of absolute values for usual intake of most macronutrients<sup>(28,41–43)</sup>. In the present study, the reproducibility of intake of macronutrients using the mean of two 24-h dietary assessments was higher, and approximately comparable with those reported by FFQ widely used in previous studies of diet and disease, with reproducibility estimates of around 0.50–0.60<sup>(5)</sup>. This suggests that averaging measurements from at least two 24-h assessments may be similar to FFQ for capturing the consistency of dietary intake over longer periods of time. The reproducibility coefficients for the mean of two 24-h dietary assessments were similar to those for blood pressure and total cholesterol, which are typically in the range 0.60–0.70<sup>(5,6)</sup>. Hence, the mean of at



least two 24-h dietary assessments in the UK Biobank have a similar magnitude of random measurement error or within-person variability and will be informative for analyses of associations with disease in observational studies. However, while methods used to correct for within-person random error are well established, the sources of error in dietary intake are likely to be more complex than for biochemical measurements, and validity studies with recovery or concentration biomarkers of dietary intake may be necessary to fully correct for such errors<sup>(8)</sup>.

While the within-person variability in a single 24-h dietary assessment was substantial in the UK Biobank, it did not appear to vary by subgroups such as age, sex, season, day of the week or index of deprivation. However, previous studies addressing these potential sources of variation reported mixed findings<sup>(34,43–47)</sup>. Previous studies have also reported that reproducibility of daily intake of food groups on 24-h assessments was generally lower than for macronutrients<sup>(5,34)</sup>. Similar findings were observed in the present study for some of the food groups, although certain foods that were eaten regularly on a daily basis, like fruit, were more reproducible on the 24-h assessments than most macronutrients. Estimates of reproducibility for intake of most food groups were also slightly higher in the short FFQ than in the 24-h assessments in this study. However, only a small number of food groups (*n* 6) were measured in both types of assessments and could be compared, and it would be expected that, for example, fish would be better captured with a FFQ than two 24-h assessments. Overall, agreement between the two types of dietary assessment was moderate, and consistent with findings of a previous report that suggested the short FFQ adequately discriminated between high and low intakes on the 24-h dietary assessment in the UK Biobank for selected food groups<sup>(48)</sup>. The level of agreement reported between dietary assessments used in this study was also comparable with estimates reported between traditional 24-h recalls and FFQ reported in previous studies<sup>(49–51)</sup>, indicating that measurements of food group intake in the non-traditional UK Biobank dietary assessments are acceptable and suitable for use in prospective analyses. However, some of the level of agreement between the FFQ and 24-h assessment may be due to correlated errors, and since the dietary assessments in the UK Biobank have not yet been compared with objective measures of dietary intake such as recovery biomarkers, the validity of these assessments has not yet been established.

Reproducibility of a vegetarian dietary pattern was high if self-identified (24-h assessment) or if indicated by non-consumption of meat and fish in the short FFQ. While no other estimates of reproducibility of the Mediterranean diet could be found, previous research using FFQ to define multifaceted 'prudent' or 'Western' dietary patterns have reported reproducibility estimates around 0.70<sup>(52)</sup>. However, the estimate of reproducibility for the mean of two 24-h dietary assessments of the Mediterranean dietary pattern in this study was lower, and similar to the less reproducible macronutrients with an ICC of 0.45. Components of the Mediterranean diet, like nuts, fish or olive oil, may not be consumed frequently enough in the UK for a small number of 24-h dietary assessments to record reliably, which should be considered if

researchers plan to assess associations of this dietary pattern with incident disease.

The findings in the present study indicating that the 24-h dietary assessments have acceptable reproducibility, provided at least two assessments are used, have important implications for analysis of UK Biobank dietary intake data. The UK Biobank is one of the largest studies with multiple 24-h dietary assessments, with over 78 000 participants completing at least three dietary assessments on different occasions. A recent analysis suggested that use of at least three 24-h dietary recalls had the best calibration with true intakes measured by recovery biomarkers when comparing serial measurements using FFQ and 24-h recalls<sup>(4)</sup>. However, the strength of the associations between usual dietary intake and disease risk in observational studies like the UK Biobank will still be underestimated if uncorrected for within-person variability in dietary assessments. Correction for such random within-person error (but not systematic error, which is likely to be still present in dietary assessments) can be done, as in the present report, using reproducibility measurements from a sample of the original cohort and does not require a validation study with a 'gold standard'<sup>(6)</sup>. The large size of the UK Biobank substudy is an additional strength as it allows reliable estimates of these correction factors<sup>(27)</sup>.

One strength of the 24-h dietary assessments used in the UK Biobank was that measurements were recorded using a web-based platform in a short period of time, with automatic coding of nutrients. This web-based platform limited the burden to participants and researchers (and hence improved the acceptability), and therefore permitted assessments to be administered to a large number of participants<sup>(13)</sup>. However, the web-based format may discourage or prohibit certain types of participants from completing it, particularly participants with ill health or poor computer literacy<sup>(53–55)</sup>. Previous analyses of 24-h assessments in the UK Biobank reported that participants completing multiple assessments tended to be older and more educated, but that there were small differences by sex and deprivation<sup>(13)</sup>. Furthermore, without the assistance of trained interviewers, participants may have omitted some food items and required assistance for portion sizes<sup>(28)</sup>. Collectively, this may have increased the measurement error in the 24-h dietary assessments. However, studies indicated comparable results between a web-based 24-h dietary assessment method and an interviewer-administered 24-h dietary assessment completed on the same day<sup>(12,55)</sup>. The web-based 24-h dietary assessment also suffers from the standard measurement issues common to many dietary assessments, all of which may have had an impact on reproducibility, such as incomplete information on ingredients in composite dishes, discrepancies in how participants reported portion sizes, and general issues with memory recall<sup>(28)</sup>. Since 24-h dietary assessments in the UK Biobank differed from standard '24-h recalls' that are commonly used in nutritional epidemiology, some foods may have been missed, as only 206 foods and 32 beverages were surveyed. However, total energy intake as computed from the 24-h assessments was not notably low, suggesting that few important foods were missed.

While the mean of two 24-h dietary assessments in the UK Biobank had acceptable reproducibility, this could still lead to



underestimation of the strength of associations with disease by up to 50 % (as correction factors from the ICC are still around 0.50–0.60). Previous reports on the recommended methods to correct for measurement error by Willett<sup>(7)</sup> and Bennett *et al.*<sup>(8)</sup> include several approaches for minimising the effects of regression dilution bias that can arise from both measurement error and within-person variability assessed in the present study. Therefore, it is important that analyses of diet–disease relationships take account of both measurement error and within-person variability as diet may be an even more important determinant of chronic disease risk than has previously been realised.

### Acknowledgements

This research has been conducted using the UK Biobank Resource under application number 13055.

This work was supported by core grants to CTSU (Clinical Trial Service Unit) from the Medical Research Council (J. L. C., S. L. and R. C., Clinical Trial Service Unit A310) and the British Heart Foundation (J. C., S. L. and R. C., CH/1996001/9454). Funding was received from the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research (S. A. J. and C. P.) and Care Oxford at Oxford Health NHS Foundation Trust (S. A. J. and C. P.). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The funders had no role in the design, analysis or writing of this article.

J. L. C. designed and conducted all the analyses, and was responsible for preparing the initial draft report. S. L., C. P. S., K. B., T. J. K., S. A. J., M. A., D. B. and R. C. all contributed to the design of the analysis, interpretation of data and results. All authors also suggested revisions to the draft report, and approved the final manuscript prior to submission for publication.

There are no conflicts of interest.

### Supplementary material

The supplementary material for this article can be found at <https://doi.org/10.1017/jns.2019.31>

### References

- Newton JN, Briggs ADM, Murray CJL, *et al.* (2015) Changes in health in England, with analysis by English regions and areas of deprivation, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **386**, 2257–2274.
- Forouzanfar MH, Alexander L, Anderson HR, *et al.* (2015) Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **386**, 2287–2323.
- Trepanowski JF & Ioannidis JP (2018) Perspective: limiting dependence on nonrandomized studies and improving randomized trials in human nutrition research: why and how. *Adv Nutr* **9**, 367–377.
- Freedman LS, Midthune D, Arab L, *et al.* (2018) Combining a food frequency questionnaire with 24-hour recalls to increase the precision of estimating usual dietary intakes – evidence from the Validation Studies Pooling Project. *Am J Epidemiol* **187**, 2227–2232.
- Willett W & Lenart E (2013) Reproducibility and validity of food frequency questionnaires. In *Nutritional Epidemiology, Monographs in Epidemiology and Biostatistics*, 3rd ed., vol. 40, pp. 96–141 [W Willett, editor]. New York: Oxford University Press.
- Clarke R, Shipley M, Lewington S, *et al.* (1999) Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *Am J Epidemiol* **150**, 341–353.
- Willett W (2013) Correction for the effects of measurement error. In *Nutritional Epidemiology, Monographs in Epidemiology and Biostatistics*, 3rd ed., vol. 40, pp. 287–304 [W Willett, editor]. New York: Oxford University Press.
- Bennett DA, Landry D, Little J, *et al.* (2017) Systematic review of statistical approaches to quantify, or correct for, measurement error in a continuous exposure in nutritional epidemiology. *BMC Med Res Methodol* **17**, 146.
- Allen N, Sudlow C, Downey P, *et al.* (2012) UK Biobank: current status and what it means for epidemiology. *Health Policy Technol* **1**, 123–126.
- UK Biobank (2007) UK Biobank: protocol for a large-scale prospective epidemiological resource. <http://www.ukbiobank.ac.uk/resources/> (accessed July 2016).
- Davey GK, Spencer EA, Appleby PN, *et al.* (2003) EPIC–Oxford: lifestyle characteristics and nutrient intakes in a cohort of 33 883 meat-eaters and 31 546 non meat-eaters in the UK. *Public Health Nutr* **6**, 259–268.
- Liu B, Young H, Crowe FL, *et al.* (2011) Development and evaluation of the Oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies. *Public Health Nutr* **14**, 1998–2005.
- Galante J, Adamska L, Young A, *et al.* (2015) The acceptability of repeat Internet-based hybrid diet assessment of previous 24-h dietary intake: administration of the Oxford WebQ in UK Biobank. *Br J Nutr* **115**, 681–686.
- Holland B, Welch A, Unwin I, *et al.* (1991) *McCance and Widdowson's The Composition of Foods*. Cambridge, UK: Royal Society of Chemistry, and Ministry of Agriculture, Fisheries and Food.
- Holland B, Unwin I & Buss D (1988) *Cereals and Cereal Products. Third Supplement to McCance and Widdowson's The Composition of Foods*, 4th ed. Royal Society of Chemistry, and Ministry of Agriculture, Fisheries and Food. London: HMSO.
- Holland B, Brown J & Buss D (1993) *Fish and Fish Products. Third Supplement to McCance and Widdowson's The Composition of Foods*, 5th ed. Royal Society of Chemistry, and Ministry of Agriculture, Fisheries and Food. London: HMSO.
- Holland B, Unwin ID & Buss DH (1992) *Fruit and Nuts. First Supplement to McCance and Widdowson's The Composition of Foods*, 5th ed. Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food. London: HMSO.
- Chan W, Brown J, Lee SM, *et al.* (1995) *Meat, Poultry and Game. Fifth Supplement to McCance and Widdowson's The Composition of Foods*, 5th ed. Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food. London: HMSO.
- Chan W, Brown J, Church S, *et al.* (1996) *Meat Products and Dishes. Sixth Supplement to McCance and Widdowson's The Composition of Foods*, 5th ed. Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food. London: HMSO.
- Holland B, Unwin I & Buss D (1989) *Milk Products and Eggs. Fourth Supplement to McCance and Widdowson's The Composition of Foods*, 4th ed. Royal Society of Chemistry, and Ministry of Agriculture, Fisheries and Food. London: HMSO.
- Chan W, Brown J & Buss D (1994) *Miscellaneous Foods. Fourth Supplement to McCance and Widdowson's The Composition of Foods*. Royal Society of Chemistry and the Ministry of Agriculture, Fisheries and Food. London: HMSO.
- Holland B, Welch AA & Buss DH (1996) *Vegetable Dishes. Second Supplement to McCance and Widdowson's The Composition of Foods*, 5th ed. Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food. London: HMSO.



23. Holland B, Unwin I & Buss D (1991) *Vegetables, Herbs and Spices. Fifth Supplement to McCance and Widdowson's The Composition of Foods*. Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food. London: HMSO.
24. Crawley H (1993) *Food Portion Sizes*, 2nd ed. Ministry of Agriculture, Fisheries and Food. London: HMSO.
25. Fung TT, McCullough ML, Newby P, *et al.* (2005) Diet-quality scores and plasma concentrations of markers of inflammation and endothelial dysfunction. *Am J Clin Nutr* **82**, 163–173.
26. Liese AD, Krebs-Smith SM, Subar AF, *et al.* (2015) The Dietary Patterns Methods Project: synthesis of findings across cohorts and relevance to dietary guidance. *J Nutr* **145**, 393–402.
27. Van Dam R & Hunter D (2012) Biochemical indicators of dietary intake. In *Nutritional Epidemiology, Monographs in Epidemiology and Biostatistics*, vol. **3**, pp. 150–212 [W Willett, editor]. New York: Oxford University Press.
28. Baranowski T (2013) 24-h Recall and diet record methods. In *Nutritional Epidemiology, Monographs in Epidemiology and Biostatistics*, 3rd ed., vol. **40**, pp. 49–69 [W Willett, editor]. New York: Oxford University Press.
29. Willett W (2013) Implications of total energy intake for epidemiologic analyses. In *Nutritional Epidemiology, Monographs in Epidemiology and Biostatistics*, 3rd ed., vol. **40**, pp. 260–286 [W Willett, editor]. New York: Oxford University Press.
30. Townsend P, Phillimore P & Beattie A (1988) *Health and Deprivation: Inequality and the North*. London: Croom Helm.
31. Chavarro JE, Rosner BA, Sampson L *et al.* (2009) Validity of adolescent diet recall 48 years later. *Am J Epidemiol* **170**, 1563–1570.
32. StataCorp (2013) *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.
33. Stein AD, Shea S, Basch CE, *et al.* (1991) Variability and tracking of nutrient intakes of preschool children based on multiple administrations of the 24-hour dietary recall. *Am J Epidemiol* **134**, 1427–1437.
34. Palaniappan U, Cue R, Payette H, *et al.* (2003) Implications of day-to-day variability on measurements of usual food and nutrient intakes. *J Nutr* **133**, 232–235.
35. Rush D & Kristal AR (1982) Methodologic studies during pregnancy: the reliability of the 24-hour dietary recall. *Am J Clin Nutr* **35**, 1259–1268.
36. Kubota Y, Maruyama K, Sato S, *et al.* (2010) Reproducibility of 24-hour dietary recall for vitamin intakes by middle-aged Japanese men and women. *J Nutr Health Aging* **14**, 196–200.
37. Liu T, Wilson NP, Craig CB, *et al.* (1992) Evaluation of three nutritional assessment methods in a group of women. *Epidemiology* **3**, 496–502.
38. Yuan C, Spiegelman D, Rimm EB, *et al.* (2017) Validity of a dietary questionnaire assessed by comparison with multiple weighed dietary records or 24-hour recalls. *Am J Epidemiol* **185**, 570–584.
39. Beaton GH, Milner J, McGuire V, *et al.* (1983) Source of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. *Am J Clin Nutr* **37**, 986–995.
40. Siri-Tarino PW, Sun Q, Hu FB, *et al.* (2010) Meta-analysis of prospective cohort studies evaluating the association of saturated fat with cardiovascular disease. *Am J Clin Nutr* **91**, 535–546.
41. Nelson M, Black AE, Morris JA, *et al.* (1989) Between- and within-subject variation in nutrient intake from infancy to old age: estimating the number of days required to rank dietary intakes with desired precision. *Am J Clin Nutr* **50**, 155–167.
42. Sempos CT, Liu K & Ernst ND (1999) Food and nutrient exposures: what to consider when evaluating epidemiologic evidence. *Am J Clin Nutr* **69**, 1330s–1338s.
43. Stote KS, Radecki SV, Moshfegh AJ, *et al.* (2011) The number of 24 h dietary recalls using the US Department of Agriculture's automated multiple-pass method required to estimate nutrient intake in overweight and obese adults. *Public Health Nutr* **14**, 1736–1742.
44. Sempos CT, Johnson NE, Smith EL, *et al.* (1985) Effects of intraindividual and interindividual variation in repeated dietary records. *Am J Epidemiol* **121**, 120–130.
45. Van Staveren WA, Deurenberg P, Burema J, *et al.* (1985) Seasonal variation in food intake, pattern of physical activity and change in body weight in a group of young adult Dutch women consuming self-selected diets. *Int J Obes* **10**, 133–145.
46. Beaton GH (1994) Approaches to analysis of dietary data: relationship between planned analyses and choice of methodology. *Am J Clin Nutr* **59**, 253S–261S.
47. Kipnis V, Subar AF, Midthune D, *et al.* (2003) Structure of dietary measurement error: results of the OPEN biomarker study. *Am J Epidemiol* **158**, 14–21; discussion 22–26.
48. Bradbury KE, Young H, Guo W, *et al.* (2017) Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire. *J Nutr Sci* **7**, e6.
49. Ocké MC, Bueno-de-Mesquita HB, Goddijn HE, *et al.* (1997) The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int J Epidemiol* **26**, S37–S48.
50. Kaaks R, Slimani N & Riboli E (1997) Pilot phase studies on the accuracy of dietary intake measurements in the EPIC project: overall evaluation of results. *Int J Epidemiol* **26**, S26–S36.
51. Bohlscheid-Thomas S, Hoting I, Boeing H, *et al.* (1997) Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. *Int J Epidemiol* **26**, S59–S70.
52. Hu FB, Rimm E, Smith-Warner SA, *et al.* (1999) Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr* **69**, 243–249.
53. Thompson FE, Dixit-Joshi S, Potischman N, *et al.* (2015) Comparison of interviewer-administered and automated self-administered 24-hour dietary recalls in 3 diverse integrated health systems. *Am J Epidemiol* **181**, 970–978.
54. Tabacchi G, Amodio E, Di Pasquale M, *et al.* (2014) Validation and reproducibility of dietary assessment methods in adolescents: a systematic literature review. *Public Health Nutr* **17**, 2700–2714.
55. Kirkpatrick SI, Subar AF, Douglass D, *et al.* (2014) Performance of the automated self-administered 24-hour recall relative to a measure of true intakes and to an interviewer-administered 24-h recall. *Am J Clin Nutr* **100**, 233–240.