Santos-Báez *et al.*
BMC Medical Research Methodology (2025) 25:115
https://doi.org/10.1186/s12874-025-02568-4

BMC Medical Research Methodology

RESEARCH

Open Access

Dietary misreporting: a comparative study of recalls vs energy expenditure and energy intake by doubly-labeled water in older adults with overweight or obesity

Leinys S. Santos-Báez[1*], Michele N. Ravelli[2], Diana A. Díaz-Rizzolo[1,3], Collin J. Popp[4], Dympna Gallagher[1], Bin Cheng[5], Dale Schoeller[2] and Blandine Laferrère[1]

Abstract

Background Self-report methods are widely used to assess energy intake but are prone to measurement errors. We aimed to identify under-reported, over-reported, and plausible self-reported energy intake by dietary recalls (rEI) using a standard method (Method 1) that calculates the rEI ratio against measured energy expenditure (mEE) by doubly-labeled water (DLW), and compare it to a novel method (Method 2), which calculates the rEI ratio against measured energy intake (mEI) by the principle of energy balance (EB = mEE + changes in energy stores).

Methods The rEI:mEE and rEI:mEI ratios were assessed for each subject. Group cut-offs were calculated for both methods, using the coefficient of variations of rEI, mEE, and mEI. Entries within ± 1SD of the cutoffs were categorized as plausible, < 1SD as under-reported, and > 1SD as over-reported. Kappa statistics was calculated to assess the agreement between both methods. Percentage bias (bβ) was estimated by linear regression. Remaining bias (dβ) was calculated after applying each method cutoffs.

Results The percentage of under-reporting was 50% using both methods. Using Method 1, 40.3% of recalls were categorized as plausible, and 10.2% as over-reported. With Method 2, 26.3% and 23.7% recalls were plausible and over-reported, respectively. There was a significant positive relationship between mEI with weight (ß = 21.7, $p < 0.01$) and BMI (ß = 48.8, $p = 0.04$), but not between rEI with weight (ß = 13.1, $p = 0.06$) and BMI (ß = 41.8, $p = 0.11$). The rEI relationships were significant when only plausible entries were included using Method 1 (weight: ß = 17.4, $p < 0.01$, remaining bias = 49.5%; BMI: ß = 44.6, $p = 0.01$, remaining bias = 60.2%) and Method 2 (weight: ß = 19.5, $p < 0.01$, remaining bias = 24.9%; BMI: ß = 44.8, $p = 0.03$, remaining bias = 56.9%).

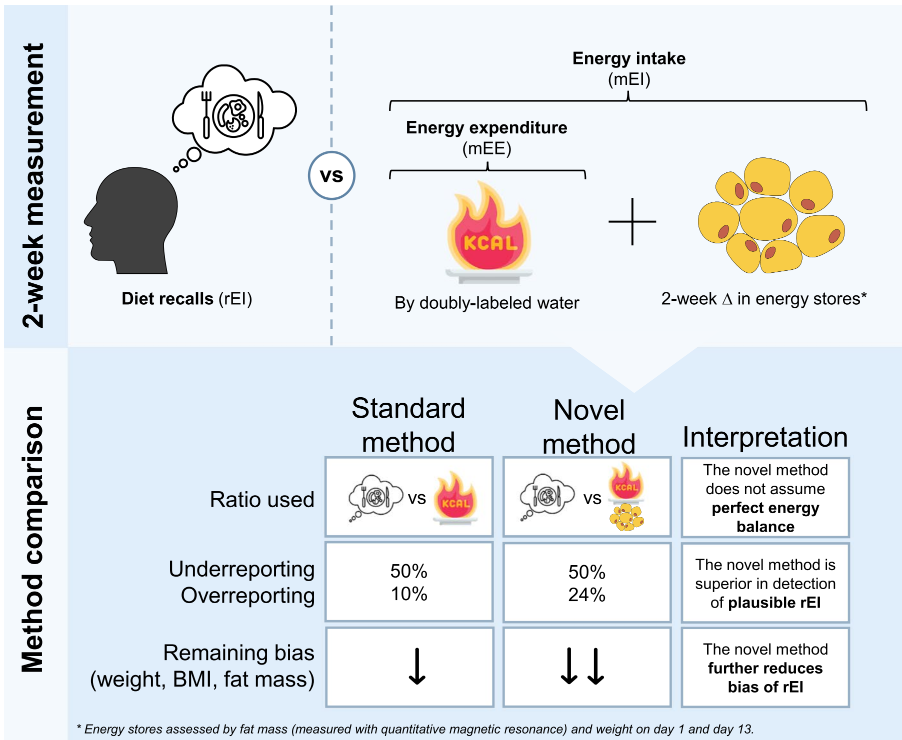
Conclusions The choice of method significantly impacts plausible and over-reported classification, with the novel method identifying more over-reported entries. While rEI showed no relationships with anthropometric measurements, applying both methods reduced bias. The novel method showed greater bias reduction, suggesting that it may have superior performance when identifying plausible rEI.

Clinical trials registration NCT04465721.

Keywords Nutrition assessment, Doubly labeled water, Energy intake, Self-report, Dietary recall, Dietary misreporting, Bias, Goldberg cutoff

*Correspondence:
Leinys S. Santos-Báez
lss2181@cumc.columbia.edu
Full list of author information is available at the end of the article

Graphical Abstract



Background

Despite their widespread use in clinical and research settings, dietary recalls and other retrospective and prospective dietary assessment methods of energy intake (EI) have long been scrutinized for their accuracy and validity [1] due to deliberate or inadvertent misreporting [2]. Although under-reporting dietary intake is well-documented [2–19], over-reporting dietary intake receives less attention [20, 21]. Neglecting over-reporting risks an incomplete understanding of the misreporting spectrum's dual nature. Under-reporting can obscure true associations between dietary intakes and outcomes of interest, while over-reporting can mask genuine deficiencies and exaggerate the effects of dietary patterns. This leads to skewed study findings that result in misleading interpretations. Nevertheless, although self-report methods are often viewed as too flawed for reliable scientific measurement [22], they remain a cost-effective and convenient tool in nutrition and clinical research [23]. Therefore, it is paramount to identify plausible dietary recalls, as measurement errors and discrepancies between actual and reported caloric intake could also be accompanied with inaccuracies in reporting nutrient composition [24].

To characterize self-reported EI (rEI) as plausible, a common approach excludes participants with rEI outside a pre-set range (e.g. 500–3,500 for women, and 800–4000 kcals/day for men) [25–29]. This one-size-fits-all method might overlook inaccurate reporting in individuals with overweight and higher energy requirements [30], or aging population and adults who struggle with progressive declines in energy expenditure (EE) [31, 32]. Goldberg et al. [33–36] proposed activity-based cut-offs using the ratio between rEI and basal metabolic rate (rEI:BMR) plus assignment of a physical activity level. This method requires weight stability and correct assignment of physical activity levels. As an alternative, predicted EE (pEE) and measured EE (mEE) obtained through the doubly-labeled water (DLW) method [37–41] have been widely used for rEI plausibility assessment.

The rEI:mEE ratio method considers the within-subject errors in both factors, including mEE measurement error and normal day-to-day variation [42, 43]. The mEE method has been shown to have the highest specificity to identify plausible reports [44], however, it assumes energy balance, and rEI is often based on only 1–2 recalls [37, 38, 44], which may not represent typical dietary intake [45]. Furthermore, it neglects to consider the influence of self-monitoring [3, 46, 47], and may wrongly classify valid entries as under-reported (e.g. during weight loss or illness). A recent equation that estimates pEE using body weight, height, age, sex, race, ethnicity, and elevation above sea level, was shown to highly correspond with mEE by DLW, using 95% predictive limits to identify plausible reports [48]. However, similar to the rEI:mEE ratio, it assumes energy balance during the measurement period. Regardless of the method used, most studies have shown that BMI, female sex, and older age predict the prevalence and magnitude of dietary misreporting [7, 48–53].

Measured EI (mEI) using the principle of energy balance can be calculated by mEE combined with changes in body energy stores (ΔES) [54]. Although more difficult to measure, using mEI can better represent a direct comparison against rEI. To our knowledge, rEI plausibility has not been compared with mEI. Therefore, using a well-characterized cohort, we aimed to compare a known method that uses the ratio of rEI to mEE, and a novel approach using the ratio of rEI to mEI to identify implausible rEI derived by dietary recalls across 3 to 6 non-consecutive days within a 2-week period. We hypothesized that this novel approach would provide a more accurate assessment of dietary plausibility. Furthermore, we aimed to examine how both methods influence the relationship between known predictors and dietary misreporting.

Methods

Study population

The study was completed using the baseline data collected between June 2021 and February 2024 in a sample cohort ($n = 39$) from the NY-TREAT Study [55]. This cohort consisted of male and female adults of any racial or ethnic group, aged 50 to 75 years, with overweight or obesity (BMI ≥ 25 and ≤ 45 kg/m^2), and a habitual long eating window ($\geq$ 14 h). Participants with were recruited from the New York City Metropolitan area by flyers and referrals. Additional details of inclusion and exclusion criteria have been previously published [55].

Study design

The parent study is a randomized controlled trial of 12-month duration at Columbia University. After informed consent, participants completed a 2-week baseline assessment, prior to being randomized to a 10-h time-restricted eating intervention or control group with habitual diet. The 2-week assessment was repeated at the end of the third month [55]. For this ancillary project, our goal was to assess dietary reporting by multiple 24-h recalls against mEE and mEI, with data obtained during the 2-week baseline assessments which occurred in ambulatory conditions except for in-person visits on days 1 and 13 of the 2-week period (Supplemental Fig. 1). During the baseline assessment, participants were advised to continue with their usual diet and physical activity routine and were blinded to the data collected.

Measurements

Anthropometric measurements

On days 1 and 13, body weight was measured to the nearest 0.1 kilogram (kg) using a calibrated scale (Ohaus Champ General Purpose Bench Scale, Ohaus Corp., Pine Brook, NJ, USA), while height was measured to the nearest 1 millimeter (mm) using a stadiometer (Holtain Ltd., Crymych, UK). Participants were instructed to empty their bladders, remove clothing and jewelry, and wear a provided hospital gown and slippers immediately before anthropometric measurements.

Body composition

Quantitative magnetic resonance (QMR, EchoMRI 2020, Echo Medical Systems, Houston TX, USA) is a noninvasive technique that employs proton nuclear magnetic resonance to measure body composition [56]. The system can accommodate individuals up to 250 kg and is standardized to detect changes in fat mass (FM) with a precision (replicated measurements CV) of $<0.5\%$ [57]. This technique is conducted by a trained technologist on days 1 and 13 of the 2-week ambulatory period. Participants were required to abstain from caloric and water intake for 12 h before each measurement, which took approximately three minutes and was done in duplicate. The system provides estimates of FM, lean mass, free water, and total body water. We analyzed FM and fat-free mass (FFM), which was calculated by subtracting FM from measured body weight.

EE assessment by doubly-labeled water

EE assessment by doubly-labeled water

The mEE was determined utilizing the gold-standard DLW method [42, 58–60]. Each participant orally received a dose comprising 1.68 g per kg of body water of oxygen- 18 water (10.8 APE) and 0.12 g per kg of body water of deuterium oxide water (99.8 APE). Urine samples were collected before dosing, within 3- and 4-h post-dose, and twice 12 days following ingestion using the two-point protocol for sample collection [61]. The samples were analyzed using isotope ratio mass spectrometers (Delta V IRMS and Delta Plus IRMS Thermo Fisher®) at the Isotope Ratio Mass Spectrometry Laboratory, University of Wisconsin-Madison. For mEE calculation, the carbon dioxide production (rCO_2) equation [59] was applied, considering a respiratory quotient of 0.86. The rCO_2 was then converted to total daily energy expenditure using the Weir equation [62].

EI assessment by principle of energy balance

The mEI was determined using the principle of energy balance [54]. This method considers the measurement of mEE and ΔES:

$$mEI_{(kcal/day)} = mEE_{(kcal/day)} + \Delta ES_{(kcal/day)}$$

In our study, the ΔES were computed based on changes in FM and FFM observed between days 1 and 13 QMR measurements. This computation involved multiplying the changes in FM and FFM by the respective energy density coefficients for each tissue (9.5 kcal/g for FM and 1.02 kcal/g for FFM), and dividing by the number of days between the measurements [60]. To address the significant within-individual SD seen in FM measurements seen in our cohort, a linear regression equation derived from baseline data for males and females was computed. Sex was considered separately due to the inherent differences in body composition between males and females. In this equation, the changes in FM (y-axis) served as the dependent variable, while the change in body weight (x-axis) was the independent variable. By utilizing this linear regression approach, the predicted change in FM (ΔFM_{adj}) was calculated for each participant based on their sex. Subsequently, the changes in FFM (ΔFFM_{adj}) were determined as the difference between the changes in body weight and ΔFM_{adj}. The formula was:

$$\Delta ES_{(kcal/day)} = \frac{\left(\Delta FM_{adj\,(kg)} \times 9500_{(kcal/kg)}\right) + \left(\Delta FFM_{adj\,(kg)} \times 1020_{(kcal/kg)}\right)}{days}$$

Self-reported energy intake (rEI)

Self-reported energy intake (rEI) was assessed by 24-h dietary recalls via the web-based, Automated Self-Administered 24-h® (ASA24®) Dietary Assessment Tool, a web-based tool modeled on the United States Department of Agriculture's Automated Multiple-Pass Method [63] in which participants recorded meals ingested in the previous 24-h period. Participants completed up to six recalls on non-consecutive weekdays and at least one weekend day over the 2-week period. Participants were requested to complete an additional rEI if the recall was submitted erroneously and incomplete (≤ 2 entries and ≤ 100kals in a 24-h diet recall). To assess different settings of rEI, the mean caloric intake considered the average 24-h calories in all reported recalls, the average 24-h calories in all in-clinic recalls (assisted by study staff), and the average 24-h calories in all free-living recalls (completed at home and not assisted by study staff), to address patterns in rEI.

Classification of plausible and misreported rEI

Using the following 2 methods, a cutoff of 1SD was calculated for the entire group, as it has been shown to exclude implausible rEI in previous work [37]. Recall entries that were within 1SD were categorized as plausible report. Recall entries < 1SD were categorized as under-reported, and entries > 1SD were categorized as over-reported.

Method 1: the ratio between average rEI and mEE (rEI:mEE) was assessed for each participant. Based on previously defined formulas [37], the cutoff used for diet recall categorization was calculated as:

$$1SD\ (Method\ 1) = \sqrt{\left(\frac{CVrEI^2}{d}\right) + \left(CVpEE^2 + CVmEE^2\right)}$$

where CV_{rEI} is the pooled within-subject variation in rEI, d is the average number of diet recalls, CV_{pEE} is the pooled CV of predicted EE (pEE), and CV_{mEE} is the within-subject variation of mEE. The pEE was computed with the equation developed by Vinken et al. [40], using the following calculation: pEE = 7.377–0.076 × age (years) + 0.0806 × weight (kg) + 0.0135 × height (cm) – 1.363 × sex (0 for males, 1 for females).

Method 2: the ratio between rEI and EI measured by the principle of energy balance (rEI:mEI) was assessed for each participant, developed from the principles described in the formulas used in Method 1. The cutoff used for categorization was calculated as:

$$1SD\ (Method\ 2) = \sqrt{\left(\frac{CVrEI^2}{d}\right) + \left(CVmEI^2\right)}$$

where CV_{rEI} and d were calculated as described in Method 1, and CV_{mEI} is the geometric mean of the within-subject variation of mEI.

Although all data analyzed as part of this report were completed during baseline prior to randomization, the data at baseline and 3-month period was used to analyze the CV of repeated measurements for mEI, using the control group only. This approach was adopted to minimize the influence of behavioral modification that might confound the true CV of the mEI approach.

Statistical analyses

Categorical variables, including sex, age, race, and ethnicity compared with chi-squared test. Point bi-serial correlation used to assess relationships between dichotomous variables and continuous variables. Pearson's and Spearman's correlations performed to assess relationships between parametric and non-parametric variables, respectively, which were determined by Shapiro–Wilk test. The sensitivity and specificity for under-reporting were calculated using the Method 1 as the reference test. Kappa statistics was calculated to assess the agreement between the Method 1 and Method 2. Systematic bias for method comparison was assessed with Bland–Altman analysis. Linear mixed models were used to evaluate the effect of call sequence on rEI and reporting ratios while adjusting for whether the call was completed in-clinic, sex, and an interaction between call sequence and sex. Separate linear regressions evaluated the linear relationships for continuous variables between mEI and rEI before and after cutoffs were applied ($rEI_{[raw]}$, and $rEI_{[METHOD1]}$ and $rEI_{[METHOD2]}$) against anthropometric outcomes (weight, BMI, and FM in kg) using participants with available rEI data after both method cutoffs were applied. These regressions were used on the basis that weight and body composition are associated with higher energy requirements to maintain energy balance [30, 64, 65]. The EI variables were used as independent variables, and the anthropometric outcomes were used as dependent variables. To assess bias, the estimated linear regression coefficients from each model are described as β_{mEI}, $\beta_{rEI[raw]}$, $\beta_{rEI[METHOD1]}$, and $\beta_{rEI[METHOD2]}$. To compute the degree of bias by $rEI_{[raw]}$, we calculated the percentage bias (b_β) [66], using the estimated linear regression coefficient:

$$\mathrm{b}\beta = \frac{\beta \mathrm{rEI[raw]} - \beta \mathrm{mEI}}{\beta \mathrm{mEI}} \times 100(\%)$$

To test whether rEI bias was reduced after Method 1 and Method 2 were applied, we computed the percentage remaining bias ($d\beta$), using the following formulas:

$$\text{Method 1 } \mathrm{b}\beta = \frac{\beta \mathrm{rEI[METHOD1]} - \beta \mathrm{mEI}}{\beta \mathrm{rEI[raw]} - \beta \mathrm{mEI}} \times 100(\%)$$

$$\text{Method 2 } \mathrm{b}\beta = \frac{\beta \mathrm{rEI[METHOD2]} - \beta \mathrm{mEI}}{\beta \mathrm{rEI[raw]} - \beta \mathrm{mEI}} \times 100(\%)$$

The degree of bias reduction using Method 1 and Method 2 are quantified by the remaining bias after subtracting the method cutoff bias from the raw rEI bias. A $d\beta = 0$ implies complete bias elimination, while a non-zero implies that bias remains. All analyses were performed using SAS version 9.4 (Cary, North Carolina, USA), IBM-SPSS 29.0 (Armonk, NY, USA), and GraphpadPrism 10.1.0 (Boston, Massachusetts, USA). Significance level was set at $\alpha = 0.05$.

Results

Overview

A total of 39 healthy adults (Supplemental Fig. 2) completed at least 2 dietary recalls for a total of 189 dietary recalls, of which 186 (4.8 ±1.0 per participant) were included in analyses after exclusion of erroneous or incomplete entries. Most participants were females (67%), aged 61 ±7 years. As expected, height ($p < 0.01$) and weight ($p = 0.01$) were higher in males, while the percentage of FM ($p = 0.02$) was higher in females, however, BMI was similar for both sexes at an average of 33.1 ±6.4 kg/m^2 (Table 1). A total of 31 (79.5%) participants completed 1 recall in-clinic on the first study visit (day 1), and all participants completed in-clinic recall on the second study visit (day 13). The remainder of the recalls were completed in a free-living setting.

There were no significant correlations between the mean rEI and mEI ($\beta = 0.221$, $p = 0.2$) or mEE ($\beta = 0.163$, $p = 0.3$). Over the course of two weeks, participants did experience some change in weight (range: − 2.2 kg to +2.2 kg) and FM (range: − 1.8 kg to +1.6 kg). This resulted in a lower estimated average mEI compared to mEE ($p = 0.03$). The rEI was 1885 ±633 kcals/day, while mEI was 2241 ±685 kcals/day, representing a non-significant ($p = 0.20$) mean underestimation of mEI by 10%, with a range between underestimation of 71% to overestimation of 89% (Fig. 1).

There were no patterns of increase or decrease of rEI with each additional dietary recall entry, and there was a non-significant continuous decline in rEI for each additional dietary recall when entries completed under free-living conditions were assessed separately (Supplemental Fig. 3). There were no significant differences in rEI completed in-clinic vs free-living conditions (Supplemental Fig. 3). Similarly, there were no consistent trends in reporting ratios (rEI:mEE and rEI:mEI) across repeated entries (Supplemental Table 1). Therefore, rEI entries and reporting ratios analyses were not stratified by in-clinic versus free-living setting nor by the ordinal number of recalls.

Assessment of implausible rEI

Crossing entries, Kappa statistics ranged from 0.49 to 0.79 for the first 5 entries (p-value <0.0001, data shown in Supplemental Table 2), indicating moderate to substantial agreement between the two methods. The sensitivity and specificity for the first 5 entries were consistently above 0.85 (Supplemental Table 3), indicating excellent discriminating ability of Method 2. Both reporting ratio methods showed a significant negative correlation with body weight (Table 2). Although the rEI:mEE and rEI:mEI were strongly correlated with each other ($r = 0.837$, $p < 0.01$), both were significantly different ($p =$

Table 1 Participant characteristics

Variable		All ($n=39$) N (%) ± SD	Male ($n=13$) N (%) ± SD	Female ($n=26$) N (%) ± SD	p-value
Age	Age	60.6 ± 6.9	59.9 ± 1.5	61 ± 7	0.653
Race	White	12 ± (30.8)	5 (38.5)	7 (26.9)	0.515
	Black	13 (33.3)	2 (15.4)	11 (42.3)	
	Asian	8 (20.5)	4 (30.8)	4 (15.4)	
	More than 1 race	3 ± (7.7)	1 (7.7)	2 (7.7)	
	Unknown	3 (7.7)	1 (7.7)	3 (7.7)	
Ethnicity	Hispanic	8 (21)	1 (7.7)	7 (26.9)	0.264
	Not Hispanic	31 (79)	12 (92.3)	19 (73.1)	
Anthropometrics	Height (cm)	166.3 ± 9.8	176.0 ± 7.4	161.5 ± 6.9	**< 0.01**
	Weight (kg)	92.1 ± 21.4	102.7 ± 18	86.8 ± 21.3	**0.014**
	BMI (kg/m^2)	33.1 ± 6.4	33 ± 4	33.2 ± 7.5	0.551
	Waist circumference (cm)	107.9 ± 15.2	114.5 ± 13.3	104.6 ± 15.2	0.055
	Fat mass (kg)	38.4 ± 14.3	34.9 ± 11.6	40.1 ± 15.4	0.283
	Fat mass (%)	41.3 ± 8.7	33.4 ± 6.2	45.2 ± 6.9	**0.015**
	2-week weight change	− 0.2 ± 1.0	− 0.1 ± 1.2	− 0.2 ± 0.9	0.740
	2-week fat mass change	− 0.1 ± 0.8	0.1 ± 0.6	− 0.2 ± 0.9	0.289
Dietary recall	rEI (kcals)	1884.9 ± 632.6	2039.3 ± 630.3	1807.7 ± 619.3	0.101
DLW measurements	mEE (kcals)	2407.22 ± 524.9	2786.2 ± 689.9	2217.7 ± 279.7	**0.006**
	mEI (kcals)	2240.6 ± 685.2	2705.4 ± 757.4	2008.2 ± 519.4	**0.001**

Nonparametric data was compared using the Mann U Whitney test, and parametric data was compared with the student t-test. The significance is shown in bold. Fat mas (kg) measured by quantitative magnetic resonance (QMR), and fat mass (%) measured as: (Fat Mass/Body Weight) × 100

Abbreviations: *BMI* Body mass index, *DLW* Doubly-labeled water, *mEE* Measured energy expenditure, *mEI* Measured energy intake, *rEI* Reported energy intake

0.02). The Bland–Altman plot demonstrated a systematic bias between rEI:mEE and rEI:mEI ratios, with a mean difference of 0.09 (95% CI: − 0.25 and 0.42). The difference between both ratios increased with higher measurement values, a pattern that suggests lower agreement between the two methods at higher ratios (Fig. 2).

With Method 1, the CV_{rEI} in our dataset was 0.34, d was 4.77, CV_{pEE} was 0.19, CV_{mEE} was 0.03, and the resulting 1SD cutoff was 0.25. With Method 2, the CV_{rEI} and d were the same as defined in Method 1, CV_{mEI} was 0.07, and the resulting 1SD cutoff was 0.17. The percentage of participants in plausible and over-reporting categories differed depending on which method cutoff was used. When Method 1 was applied, 40.3% (75 entries) were categorized as plausible, and 10.2% (19 entries) as over-reported, while these percentages changed to 26.3% (49 entries) and 23.7% (44 entries) when Method 2 was applied. The percentage of under-reported entries did not vary significantly: 49.5% (92 entries) and 50.0% (93 entries) for Method 1 and Method 2, respectively.

Relationship of average EI data with anthropometric measures before and after exclusion of implausible recalls

After cutoffs were applied, plausible entries data was available for further analyses in 27 participants (8 males, 19 females). There was a consistent significant positive relationship between mEI with weight and BMI in all participants (males and females combined). These relationships remained significant for weight, but not BMI, when males and females were analyzed separately. In contrast, there were no significant relationships between $rEI_{[raw]}$ and anthropometric measurements. After the cutoffs were applied, the slopes of the $rEI_{[METHOD1]}$ and $rEI_{[METHOD2]}$ were closer to the fit line found between mEI and anthropometric measures. A significant positive relationship with weight was maintained using the $rEI_{[METHOD1]}$ for all participants combined, as well as for males and females. A significant positive relationship with weight using the $rEI_{[METHOD2]}$ remained significant when all participants and females were analyzed, but not in males. A significant positive relationship with BMI was maintained using the $rEI_{[METHOD1]}$ for all participants and females. A significant positive relationship with BMI using the $rEI_{[METHOD2]}$ remained significant when all participants were analyzed, but not in analyses stratified by sex. Lastly, although the relationship between mEI and FM was not significant, the relationship between $rEI_{[METHOD1]}$ and FM was significant when all participants combined and females were analyzed, and the

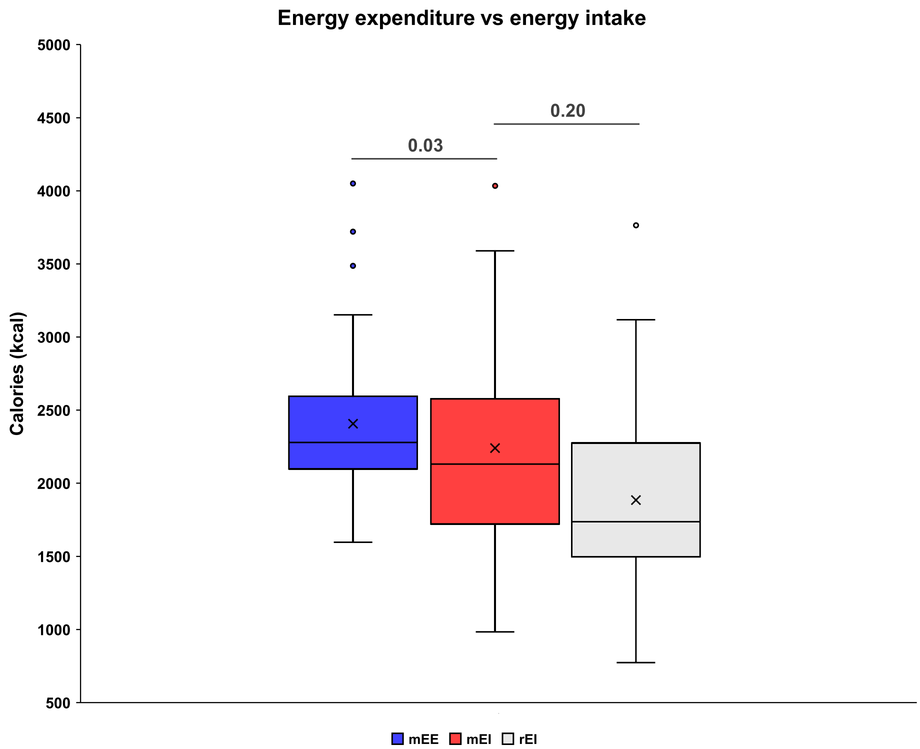


Fig. 1 Box plots of measured energy expenditure (mEE), measured energy intake (mEI), and average reported energy intake (rEI). The median value is indicated by the horizontal line, and the mean value is marked with an "x" within each box. The whiskers represent the minimum and maximum values excluding outliers. Outliers are shown as individual points outside of the boxes. The sign test was used to compare non-parametric paired variables

Table 2 Correlations between reporting ratios with sex, age and anthropometric measurements

Variable	rEI:mEE ratio		rEI:mEI ratio	
	Coeff	*p* value	Coeff	*p* value
Sex	0.089	0.588	0.176	0.283
Age (years)	− 0.018	0.915	0.043	0.793
Weight (kg)	− 0.332	**0.039**	− 0.378	**0.018**
BMI (kg/m^2)	− 0.204	0.212	− 0.210	0.200
Fat mass (kg)	− 0.156	0.342	− 0.129	0.433
Fat mass (%)	0.037	0.823	0.075	0.649

Point-biserial correlation is used to assess relationships between dichotomous variables and continuous variables. Pearson's and Spearman's correlations were performed to assess relationships between parametric and non-parametric variables. Fat mas (kg) measured by quantitative magnetic resonance (QMR), and fat mass (%) measured as: (Fat Mass/Body Weight) × 100

Abbreviations: *BMI* Body mass index, *kg* Kilograms, *mEI* Measured energy intake, *mEI* Measured energy expenditure, *rEI* Reported energy intake

relationship between rEI$_{[METHOD2]}$ and FM was significant for all participants combined only (See Fig. 3 for *p* values).

To assess the degree of bias elimination after the exclusion of implausible recalls using Method 1 and Method 2, we calculated the percentage of remaining bias. A bias reduction was observed in corrected rEI for weight, BMI and FM using Method 1 ($d\beta$= 49.5%, 60.2% and 51.0%, respectively) and Method 2 ($d\beta$= 24.9%, 56.9% and 24.7%, respectively). These results reveal a higher reduction using Method 2 for all measures, except BMI, in all participants combined as well as separate analyses for men. Although we did not expect a complete elimination of bias using either method, the bias remained for all participants when using both methods, and an overcorrection was seen for females in BMI and FM using Method 1 ($d\beta$=− 10.1% and − 108.6%, respectively), and FM using Method 2 ($d\beta$=− 36.3%). Moreover, bias was greater for weight in females using Method 1 ($d\beta$= 294.6%), however, a higher bias was also present, albeit lower, with Method 2 ($d\beta$= 152.7%) (Table 3).

Discussion

In this study, our data highlight the inherent value in evaluating self-reported EI plausibility against objective mEI (Method 2), instead of mEE (Method 1), and reinforces the existing research that consistently demonstrates the inaccuracies of EI assessments using self-report methods [3, 7, 52, 67–71]. Small changes in weight and FM

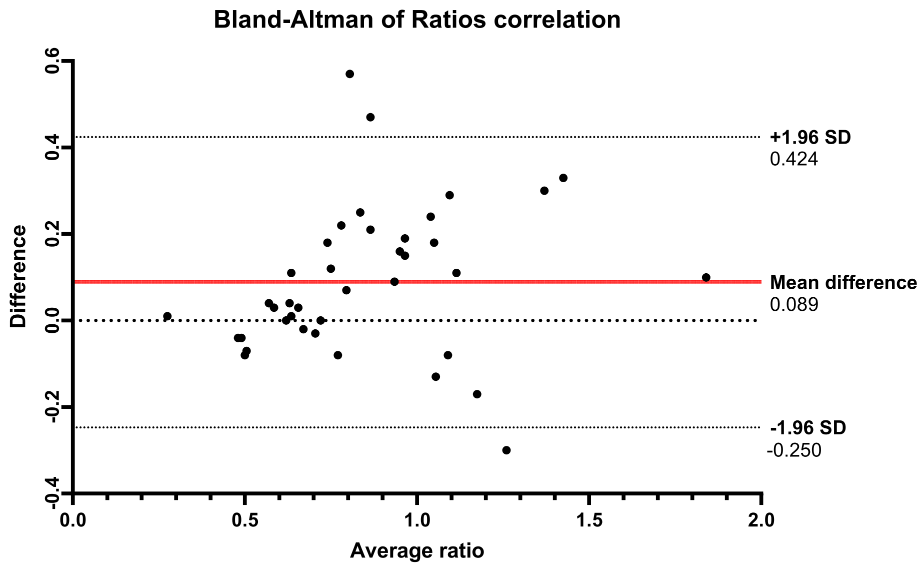


Fig. 2 Bland–Altman plot of ratios correlation. Bland–Altman plot of ratio comparing Method 1 and Method 2 ratios. The x-axis represents the average of the measurements of the Method 1 and Method 2 ratios, and the y-axis represents the difference between the measurements. The mean difference between both methods was 0.089, and the 95% limits of agreement were − 0.25 and 0.424, suggesting that 95% of the differences between the two methods fell within this range. The plot exhibited a heteroscedastic pattern, suggesting lower agreement between the two methods at higher ratio values, consistent with differing identification of over-reporting. Abbreviations: SD, standard deviation

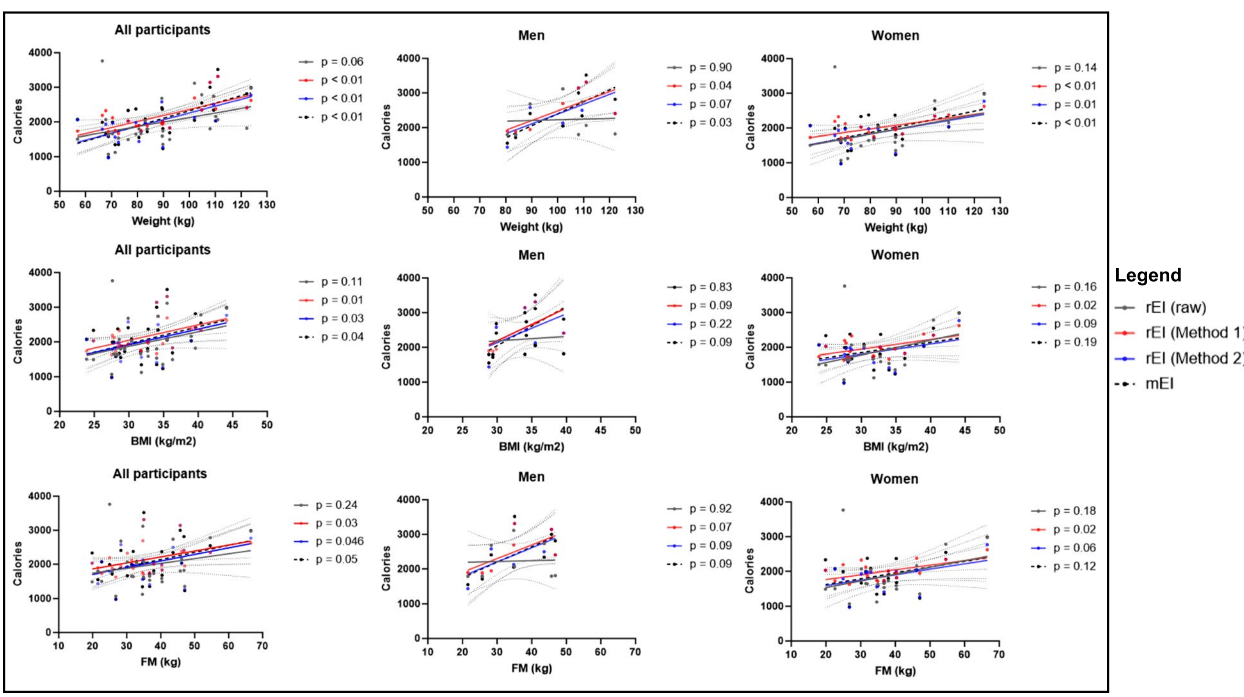


Fig. 3 Measured energy intake (mEI) and average reported energy intake (rEI) data before and after the Method 1 and Method 2 application. Linear regression models with mEI and rEI data before and after the Method 1 and Method 2 were applied in the entire group, males and females. Abbreviations: BMI, body mass index; cm, centimeters; FM, fat mass; kg, kilograms; mEI, measured energy intake; rEI, reported energy intake

were evident for all participants, despite instructions to maintain habitual dietary intake during the 2-week assessment, therefore a significantly lower mEI than mEE was seen in this group. While the baseline assessment involved no intervention, participants were monitoring their diet with dietary recalls and real-time meal

Table 3 Linear regression coefficients of the measured energy intake (mEI) and reported energy intake (rEI) before and after cutoffs were applied

Group	Predictor	mEI	rEI		rEI (Method 1)			rEI (Method 2)		
		Coeff	Coeff	Bias	Coeff	Bias	Remaining bias, %	Coeff	Bias	Remaining bias, %
All	Weight (kg)	21.65	13.05	− 8.60	17.39	− 4.26	49.53	19.51	− 2.14	24.88
	BMI (kg/m^2)	48.78	41.84	− 6.94	44.60	− 4.18	60.23	44.83	− 3.95	56.92
	Fat mass (kg)	21.01	14.36	− 6.65	17.62	− 3.39	50.98	19.37	− 1.64	24.66
Males	Weight (kg)	34.40	1.88	− 32.52	27.77	− 6.63	20.39	28.34	− 6.06	18.63
	BMI (kg/m^2)	113.30	13.02	− 100.28	94.93	− 18.37	18.32	83.22	− 30.08	30.00
	Fat mass (kg)	44.46	2.42	− 42.04	39.50	− 4.96	11.80	41.03	− 3.43	8.16
Females	Weight (kg)	15.42	13.77	− 1.65	10.56	− 4.86	294.55	12.90	− 2.52	152.73
	BMI (kg/m^2)	29.00	42.23	13.23	27.67	− 1.33	− 10.05	29.33	0.33	2.49
	Fat mass (kg)	16.39	19.17	2.78	13.37	− 3.02	− 108.63	15.38	− 1.01	− 36.33

Estimated regression coefficients from linear regressions of mEI and rEI before cutoffs and after cutoffs were applied. The EI variables were used as independent variables, and the anthropometric outcomes were used as dependent variables. To assess bias reduction, the coefficient of rEI after cutoffs were applied, were subtracted from the mEI coefficient. The bias reduction was computed subtracting the rEI (before and after cutoffs were applied) coefficient from the mEI coefficient. The percentage of remaining bias was computed dividing the methods bias by the raw rEI bias and multiplied by 100

Abbreviations: *BMI* Body mass index, *kg* Kilograms, *mEI* Measured energy intake, *mEI* Measured energy expenditure, *rEI* Reported energy intake

tracking of meal timing. The weight and FM changes may be attributed to day-to-day variability in body water at different measurement periods [72], or the effects of self-monitoring on behavior, as previous research suggests that lower rEI can reflect a genuine reduction in energy consumption to a certain extent, rather than intentional misrepresentation [3].

There was moderate to substantial agreement between the two methods using the Kappa statistics, and sensitivity and specificity for the first 5 entries (the number of entries completed by most participants), were consistently above 85%, indicating excellent discriminating ability of Method 2. While the ratios derived from Method 1 (rEI:mEE) and Method 2 (rEI:mEI) were strongly correlated with each other, discrepancies were evident, particularly at higher ratios. Approximately 50% of the entries were under-reported with the use of both cutoffs, however, only 10% were classified as over-reported with Method 1, and 24% were classified as over-reported with Method 2. Therefore, the application of Method 2 was comparable to Method 1 in the identification of under-reported entries, yet Method 2 detected more over-reported recalls that would otherwise have been classified as plausible by Method 1. We compared rEI by ASA24, a method that outperforms other self-report tools [71, 73], against objective measures of EE, shown to have the highest negative and positive predictive value [37, 38, 44]. With a 25% cutoff for the rEI:mEE ratio, Method 1 aligns with previous findings [37, 38, 74], and the rates of under- and over-reporting were within the published ranges of 20–70% and 2–10%, respectively [37, 38, 74, 75]. The choice of either mEE or mEI to validate rEI impacts the classification of recalls categorized as plausible, under- or over-reported, and the novel method may have identified more over-reported entries as it accounts for changes in body weight and FM during the measurement period. The systematic bias observed between methods at higher ratios underscores the need for using objective methods to measure EI when interpreting reporting ratio data. Energy requirements can be underestimated [76], therefore, methods that compare rEI against mEE [77] overlook inadvertent reductions in EI during self-monitoring periods [3].

To calculate the ratios in Method 1, a value specific to the dataset being analyzed can be used; however, previous studies often employed the standard CV_{rEI} of 23% and CV_{mEE} of 8.2%, or a combination of standard values and the studied dataset [66, 78]. A strength of our study is the use of group-specific values, thereby eliminating arbitrary assumptions. Moreover, we introduced the use of the mEI cutoff that uses the CV_{mEI}. Our CV_{rEI} of 34% is comparable to previous work where CV_{rEI} is approximately 23–30% [52, 74, 79]. Our CV_{mEE} was 3%, and while this is approximately half of what previous works have used for EI assessment [66, 80, 81], our CV is within the general reports of variation for EE by DLW [42, 82]. This disparity emphasizes the utility of a dataset-specific approach and the incorporation of updated DLW variability as methodological advancements with this measurement occur.

Our findings also show a consistent and significant association between mEI with weight and BMI, supporting the notion that higher energy requirements are needed to maintain energy balance in individuals with greater body mass and composition [30, 64, 65]. Significant relationships were also observed between mEI and weight for men and women. In contrast, $rEI_{[raw]}$ showed no significant relationships with anthropometric measures. These findings support prior literature suggesting that self-reporting energy intake without adjustments cannot adequately capture true EI [24, 50, 51]. However, the application of both cutoff methods resulted in significant associations between plausible recalls and anthropometric measurements when men and women were evaluated in combined analyses. For body weight, significant associations between plausible recalls and weight were present in men when Method 1 cutoff was used, but not with Method 2 cutoff. In contrast, significant associations were present for women when both Method 1 cutoff and Method 2 cutoff were applied. The improved associations between body weight and rEI indicate that the measurement error associated with self-report energy intake is attenuated when both method cutoffs are applied. Therefore, to measure the degree of bias reduction, we quantified the remaining bias from the $rEI_{[METHOD1]}$ and $rEI_{[METHOD2]}$ by subtracting its estimates from the $rEI_{[raw]}$.

The use of Method 2 resulted in a larger reduction of bias for most measures compared to Method 1. Method 2 only resulted in significant associations between $rEI_{[METHOD2]}$ and anthropometric measurements when all participants were considered. However, after stratifying by sex, there were associations with $rEI_{[METHOD2]}$ and weight in women only. This may be attributed to a reduction of statistical power due to higher sample losses. Nonetheless, a trend to significance was present in all cases, except BMI associations in men. Moreover, the remaining bias in weight was the highest in the entire group and stratified data, and while both methods effectively reduced bias, neither eliminated it entirely. Furthermore, we observed instances of increase or overcorrection of bias using both methods, and while Method 2 mitigated those instances to some extent, bias persisted. Ejima et. al 2019 had previously demonstrated that the use of Goldberg cutoffs does not always eliminate bias [66], with a remaining bias in weight of 56.1%. In our study, weight bias was reduced by nearly half (49.5%) with the use of a high-performance standard method [44], and further reduced (24.9%) with the use of a novel cutoff method.

There was a significant association between lower reporting ratios and higher weight with the use of both cutoff methods. However, we found no associations with BMI, sex, nor age. This is contrary to previous work [41, 48, 51, 53, 71], which reported associations between under-reporting and BMI, female sex, and older age. Our small sample size, a skewed sex distribution (female sex comprised two-thirds), and well-defined cohort with a narrow age and BMI ranges could explain these discrepancies, as previous work often included larger populations with greater heterogeneity in both BMI and age, as well as a more even distribution of sexes.

This study introduces methodological advancements and offers new insights into what we know about the plausibility of self-reported caloric intake. We addressed the limitations of previous research, in which the BMR or mEE were used, and introduced the potential to use mEI as a personalized option to interpret self-reported dietary intake data. Important strengths of this study include multiple ambulatory rEI, use of QMR and consideration of sex differences to assess small changes in body composition. Previous work validated energy balance assessment by DLW in controlled feeding using body weight or dual-energy x-ray absorptiometry [54]. There are limitations to acknowledge, the small sample size with specific inclusion criteria of individuals with overweight or obesity may not be generalizable to different populations. This limitation was exacerbated when the cutoffs were applied and sample losses probably led to reduced statistical power, however, this weakness was minimized by using dataset-specific CV. The measurements of mEE and mEI using DLW and QMR (or other proxies of FM change, such as DXA) is costly and not feasible in large scale studies, however, for studies that do have measurements of DLW, weight change could be considered to estimate mEI for assessment of dietary plausibility, instead of relying on mEE and assuming energy balance. Similarly, the identification of dietary plausibility can occur throughout all levels of reporting, even in plausible reports, therefore, nutrient intake cannot be assumed to be correct in plausible recalls identified with either method. Furthermore, although DLW is the gold standard to measure free-living EE, the method is not entirely resistant to small measurement errors, and given that the mEI is in part calculated by mEE, any measurement errors of mEE be incorporated in overall mEI as well. While QMR offers a precise measure of FM, the use of FFM for mEI calculations can introduce inaccuracies, as FFM includes all non-adipose tissue, including free, intra-cellular and extra-cellular water, which contributes to BMR [83]. Day-to-day variability in body water [72] could also explain some of the changes in weight and FM over the measurement period. To reduce errors due to variability in total body water during QMR measurements, participants were asked to avoid water intake for 12 h prior to the QMR measurements.

Conclusion

In conclusion, the implementation of cutoffs based on objectively measured EI demonstrated greater accuracy and provided a more reliable estimate of plausible dietary recalls by eliminating under- and over-reported entries, though this method, while reducing bias, does not completely eliminate it. Nevertheless, even without the application of data removal by this approach, reporting the rEI:mEI ratio remains a valuable option for a systematic assessment of the degree of misreporting and a comprehensive interpretation of published diet and health outcomes data. Further research is necessary to validate this new approach in larger populations of varying age, body weight and physical activity. This new approach could be used to identify plausible intakes using 95% predictive limits, similar to recent work using a newly-developed pEE equation [48], and assess how both of these novel methods compare. While using pEE and mEE has advantages, identifying energy intake with DLW, if feasible, remains a more sensible option to reduce assumptions about energy balance.

Abbreviations

ASA24®	Automated Self-Administered 24-h® Dietary Assessment Tool
BMR	Basal metabolic rate
CV	Coefficient of variation
DLW	Doubly-labeled water
EB	Energy balance
ES	Energy stores
ΔES	Changes in energy stores
FM	Fat mass
FFM	Fat-free mass
HbA1 C	Hemoglobin A1C
mEE	Measured energy expenditure
mEI	Measured energy intake
pEE	Predicted energy expenditure
NY-TREAT	NY- Time-Restricted EATing study
QMR	Quantitative magnetic resonance
rEI	Reported energy intake
T2D	Type 2 diabetes

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12874-025-02568-4.

Supplementary Material 1.

Acknowledgements
The authors express their sincere gratitude to the study participants for their generous participation, Thauany Nantes Guirao for the technical support in the DLW measurements, and the NY-TREAT trial staff for the recruitment, administrative and technical support.

Authors' contributions
L-S.S-B. conceived the study, performed the cutoff calculations, derived equations, performed data analyses, and wrote the first draft of the manuscript. M-N.R. performed the energy balance calculations. B.C. performed data analyses. All authors contributed to the interpretation of the results and critical revision of the manuscript for important intellectual content. All authors approved the final version of the manuscript.

Funding
This study was supported by the National Institute on Aging (NIA) R01 AG065569 -03, the National Center for Advancing Translational Sciences (NCATS) UL1 TR001873, and P30 DK026687. Santos-Báez was supported by T32 HL007343 -44, R01 AG065569 -03S1, and the Naomi Russ Berrie Fellowship. Díaz-Rizzolo was supported by Fundación Alfonso Martin Escudero and the Naomi Russ Berrie Fellowship. Popp was supported by R01 NR018916 and K99HL163474.

Data availability
The datasets used in this study are not readily available at time of publication as the data are part of an ongoing study. Requests to access the datasets should be directed to the corresponding author: Leinys S. Santos-Báez, MD (lss2181@cumc.columbia.edu).

Declarations

Ethics approval and consent to participate
This study was conducted according to the guidelines of the Declaration of Helsinki. Study was approved by the Columbia University Institutional Review Board (AAAS7791) and informed consent was obtained from all participants.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Division of Endocrinology, Nutrition Obesity Research Center, Columbia University Irving Medical Center, New York, NY, USA. [2]Isotope Ratio Mass Spectrometry Core Laboratory, Biotechnology Center, University of Wisconsin, Madison, WI, USA. [3]Health Science Faculty, Universitat Oberta de Catalunya (UOC), Barcelona, Spain. [4]Department of Population Health, New York University Langone Health Grossman School of Medicine, New York, NY, USA. [5]Department of Biostatistics, Columbia University Irving Medical Center, New York, NY, USA.

Received: 3 March 2025 Accepted: 15 April 2025
Published online: 26 April 2025

References

1. Archer E, Marlow ML, Lavie CJ. Controversy and debate: Memory-Based Methods Paper 1: the fatal flaws of food frequency questionnaires and other memory-based dietary assessment methods. J Clin Epidemiol. 2018;104:113–24.
2. Kirkpatrick SI, Troiano RP, Barrett B, Cunningham C, Subar AF, Park Y, et al. Measurement Error Affecting Web- and Paper-Based Dietary Assessment Instruments: Insights From the Multi-Cohort Eating and Activity Study for Understanding Reporting Error. Am J Epidemiol. 2022;191(6):1125–39.
3. Goris AH, Westerterp-Plantenga MS, Westerterp KR. Undereating and underrecording of habitual food intake in obese men: selective underreporting of fat intake. Am J Clin Nutr. 2000;71(1):130–4.
4. Lafay L, Mennen L, Basdevant A, Charles MA, Borys JM, Eschwège E, et al. Does energy intake underreporting involve all kinds of food or only specific food items? Results from the Fleurbaix Laventie Ville Santé (FLVS) study. International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity. 2000;24(11):1500–6.
5. Gemming L, Ni MC. Dietary under-reporting: what foods and which meals are typically under-reported? Eur J Clin Nutr. 2016;70(5):640–1.
6. Macdiarmid J, Blundell J. Assessing dietary intake: Who, what and why of under-reporting. Nutr Res Rev. 1998;11(2):231–53.
7. Johansson G, Wikman A, Ahrén AM, Hallmans G, Johansson I. Underreporting of energy intake in repeated 24-hour recalls related to gender, age, weight status, day of interview, educational level, reported food intake, smoking habits and area of living. Public Health Nutr. 2001;4(4):919–27
8. Gibson RS, Charrondiere UR, Bell W. Measurement Errors in Dietary Assessment Using Self-Reported 24-Hour Recalls in Low-Income Countries and Strategies for Their Prevention. Adv Nutr. 2017;8(6):980–91.
9. Nelson M, Atkinson M, Darbyshire S. Food photography. I: The perception of food portion size from photographs. Br J Nutr. 1994;72(5):649–63.

10. Young LR, Nestle MS. Portion sizes in dietary assessment: issues and policy implications. Nutr Rev. 1995;53(6):149–58.
11. Almiron-Roig E, Solis-Trapala I, Dodd J, Jebb SA. Estimating food portions. Influence of unit number, meal type and energy density. Appetite. 2013;71:95–103.
12. Amoutzopoulos B, Page P, Roberts C, Roe M, Cade J, Steer T, et al. Portion size estimation in dietary assessment: a systematic review of existing tools, their strengths and limitations. Nutr Rev. 2020;78(11):885–900.
13. Brogden N, Almiron-Roig E. Estimated portion sizes of snacks and beverages differ from reference amounts and are affected by appetite status in non-obese men. Public Health Nutr. 2011;14(10):1743–51.
14. Thompson FE, Subar AF, Loria CM, Reedy JL, Baranowski T. Need for technological innovation in dietary assessment. J Am Diet Assoc. 2010;110(1):48–51.
15. Dahle JH, Ostendorf DM, Zaman A, Pan Z, Melanson EL, Catenacci VA. Underreporting of energy intake in weight loss maintainers. Am J Clin Nutr. 2021;114(1):257–66.
16. Burrows TL, Ho YY, Rollo ME, Collins CE. Validity of Dietary Assessment Methods When Compared to the Method of Doubly Labeled Water: A Systematic Review in Adults. Front Endocrinol (Lausanne). 2019;10:850.
17. Markussen MS, Veierød MB, Ursin G, Andersen LF. 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. Br J Nutr. 2016;116(3):547–58.
18. Malinowska AM, Mlodzik-Czyzewska MA, Chmurzynska A. Dietary patterns associated with obesity and overweight: When should misreporters be included in analysis? Nutrition. 2020;70: 110605.
19. Hörnell A, Winkvist A, Hallmans G, Weinehall L, Johansson I. Mis-reporting, previous health status and health status of family may seriously bias the association between food patterns and disease. Nutr J. 2010;9:48.
20. Poslusna K, Ruprich J, de Vries JH, Jakubikova M, van't Veer P. Misreporting of energy and micronutrient intake estimated by food records and 24 hour recalls, control and adjustment methods in practice. Br J Nutr. 2009;101(Suppl2):S73-85.
21. Johansson L, Solvoll K, Bjørneboe GE, Drevon CA. Under- and overreporting of energy intake related to weight status and lifestyle in a nationwide sample. Am J Clin Nutr. 1998;68(2):266–74.
22. Ravelli MN, Schoeller DA. Traditional Self-Reported Dietary Instruments Are Prone to Inaccuracies and New Approaches Are Needed. Front Nutr. 2020;7:90.
23. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. J Nutr. 2015;145(12):2639–45.
24. Livingstone MB, Black AE. Markers of the validity of reported energy intake. J Nutr. 2003;133(Suppl3):895s–920s.
25. Nielsen SJ, Adair L. An alternative to dietary data exclusions. J Am Diet Assoc. 2007;107(5):792–9.
26. Fung TT, Malik V, Rexrode KM, Manson JE, Willett WC, Hu FB. Sweetened beverage consumption and risk of coronary heart disease in women. Am J Clin Nutr. 2009;89(4):1037–42.
27. Schulze MB, Manson JE, Ludwig DS, Colditz GA, Stampfer MJ, Willett WC, et al. Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. JAMA. 2004;292(8):927–34.
28. Fung TT, Rexrode KM, Mantzoros CS, Manson JE, Willett WC, Hu FB. Mediterranean diet and incidence of and mortality from coronary heart disease and stroke in women. Circulation. 2009;119(8):1093–100.
29. Turner-McGrievy GM, Davidson CR, Wilcox S. Does the type of weight loss diet affect who participates in a behavioral weight loss intervention? A comparison of participants for a plant-based diet versus a standard diet trial. Appetite. 2014;73:156–62.
30. de Boer JO, van Es AJ, van Raaij JM, Hautvast JG. Energy requirements and energy expenditure of lean and overweight women, measured by indirect calorimetry. Am J Clin Nutr. 1987;46(1):13–21.
31. Roberts SB, Dallal GE. Energy requirements and aging. Public Health Nutr. 2005;8(7a):1028–36.
32. Pontzer H, Yamada Y, Sagayama H, Ainslie PN, Andersen LF, Anderson LJ, et al. Daily energy expenditure through the human life course. Science. 2021;373(6556):808–12.
33. Goldberg GR, Black AE, Jebb SA, Cole TJ, Murgatroyd PR, Coward WA, et al. Critical evaluation of energy intake data using fundamental principles of energy physiology: Derivation of cut-off limits to identify under-recording. Eur J Clin Nutr. 1991;45(12):569–81.
34. Black AE, Goldberg GR, Jebb SA, Livingstone MB, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: Evaluating the results of published surveys. Eur J Clin Nutr. 1991;45(12):583–99.
35. Tooze JA, Krebs-Smith SM, Troiano RP, Subar AF. The accuracy of the Goldberg method for classifying misreporters of energy intake on a food frequency questionnaire and 24-h recalls: comparison with doubly labeled water. Eur J Clin Nutr. 2012;66(5):569–76.
36. Black AE. 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 International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity. 2000;24(9):1119–30.
37. McCrory MA, McCrory MA, Hajduk CL, Roberts SB. Procedures for screening out inaccurate reports of dietary energy intake. Public Health Nutr. 2002;5(6a):873–82.
38. Huang TT, Roberts SB, Howarth NC, McCrory MA. Effect of screening out implausible energy intake reports on relationships between diet and BMI. Obes Res. 2005;13(7):1205–17.
39. Schoeller DA, Bandini LG, Dietz WH. Inaccuracies in self-reported intake identified by comparison with the doubly labelled water method. Can J Physiol Pharmacol. 1990;68(7):941–9.
40. Vinken AG, Bathalon GP, Sawaya AL, Dallal GE, Tucker KL, Roberts SB. Equations for predicting the energy requirements of healthy adults aged 18–81 y. Am J Clin Nutr. 1999;69(5):920–6.
41. Subar AF, Kipnis V, Troiano RP, Midthune D, Schoeller DA, Bingham S, et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. Am J Epidemiol. 2003;158(1):1–13.
42. Trabulsi J, Troiano RP, Subar AF, Sharbaugh C, Kipnis V, Schatzkin A, et al. Precision of the doubly labeled water method in a large-scale application: evaluation of a streamlined-dosing protocol in the Observing Protein and Energy Nutrition (OPEN) study. Eur J Clin Nutr. 2003;57(11):1370–7.
43. Schoeller DA, Hnilicka JM. Reliability of the doubly labeled water method for the measurement of total daily energy expenditure in free-living subjects. J Nutr. 1996;126(1):348s-s354.
44. Batista LD, de França NAG, Fontanelli MM, Martinez-Arroyo AG, Fisberg RM. Misreporting of dietary energy intake obtained by 24 h recalls in older adults: a comparison of five previous methods using doubly labeled water. Eur J Clin Nutr. 2022;76(4):535–43.
45. Chi SA, Lee H, Lee JE, Lee HS, Kim K, Yeo IK. An ensemble method based on marginal-effect models (EMM) for estimating usual food intake from single-day dietary data and internal/external two-day dietary data. Eur J Clin Nutr. 2023;77(3):325–34.
46. Harvey J, Krukowski R, Priest J, West D. Log Often, Lose More: Electronic Dietary Self-Monitoring for Weight Loss. Obesity (Silver Spring). 2019;27(3):380–4.
47. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc. 2011;111(1):92–102.
48. Bajunaid R, Niu C, Hambly C, Liu Z, Yamada Y, Aleman-Mateo H, et al. Predictive equation derived from 6,497 doubly labelled water measurements enables the detection of erroneous self-reported energy intake. Nat Food. 2025;6(1):58–71.
49. Lichtman SW, Pisarska K, Berman ER, Pestone M, Dowling H, Offenbacher E, et al. Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. N Engl J Med. 1992;327(27):1893–8.
50. Nunes CL, Jesus F, Oliveira MV, Thomas DM, Sardinha LB, Martins P, et al. The impact of body composition on the degree of misreporting of food diaries. Eur J Clin Nutr. 2024;78(3):209–16.
51. Lissner L, Troiano RP, Midthune D, Heitmann BL, Kipnis V, Subar AF, et al. OPEN about obesity: recovery biomarkers, dietary reporting errors and BMI. Int J Obes (Lond). 2007;31(6):956–61.
52. Price GM, Paul AA, Cole TJ, Wadsworth ME. Characteristics of the low-energy reporters in a longitudinal national dietary survey. Br J Nutr. 1997;77(6):833–51.
53. Freedman LS, Commins JM, Moler JE, Arab L, Baer DJ, Kipnis V, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. Am J Epidemiol. 2014;180(2):172–88.

54. Ravelli MN, Schoeller DA. An objective measure of energy intake using the principle of energy balance. Int J Obes (Lond). 2021;45(4):725–32.
55. Santos-Báez LS, Garbarini A, Shaw D, Cheng B, Popp CJ, Manoogian ENC, et al. Time-restricted eating to improve cardiometabolic health: The New York Time-Restricted EATing randomized clinical trial - Protocol overview. Contemp Clin Trials. 2022;120:106872.
56. Napolitano A, Miller SR, Murgatroyd PR, Coward WA, Wright A, Finer N, et al. Validation of a quantitative magnetic resonance method for measuring human body composition. Obesity (Silver Spring). 2008;16(1):191–8.
57. Gallagher D, Thornton JC, He Q, Wang J, Yu W, Bradstreet TE, et al. Quantitative magnetic resonance fat measurements in humans correlate with established methods but are biased. Obesity (Silver Spring). 2010;18(10):2047–54.
58. Westerterp KR. Doubly labelled water assessment of energy expenditure: principle, practice, and promise. Eur J Appl Physiol. 2017;117(7):1277–85.
59. Speakman JR, Yamada Y, Sagayama H, Berman ESF, Ainslie PN, Andersen LF, et al. A standard calculation methodology for human doubly labeled water studies. Cell Rep Med. 2021;2(2): 100203.
60. Thomas DM, Martin CK, Redman LM, Heymsfield SB, Lettieri S, Levine JA, et al. Effect of dietary adherence on the body weight plateau: a mathematical model incorporating intermittent compliance with energy intake prescription. Am J Clin Nutr. 2014;100(3):787–95.
61. Welle S. Two-point vs multipoint sample collection for the analysis of energy expenditure by use of the doubly labeled water method. Am J Clin Nutr. 1990;52(6):1134–8.
62. Weir JB. New methods for calculating metabolic rate with special reference to protein metabolism. J Physiol. 1949;109(1–2):1–9.
63. Subar AF, Kirkpatrick SI, Mittl B, Zimmerman TP, Thompson FE, Bingley C, et al. The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. J Acad Nutr Diet. 2012;112(8):1134–7.
64. Westerterp KR. Control of energy expenditure in humans. Eur J Clin Nutr. 2017;71(3):340–4.
65. Prentice AM, Black AE, Coward WA, Cole TJ. Energy expenditure in overweight and obese adults in affluent societies: an analysis of 319 doubly-labelled water measurements. Eur J Clin Nutr. 1996;50(2):93–7.
66. Ejima K, Brown AW, Schoeller DA, Heymsfield SB, Nelson EJ, Allison DB. Does exclusion of extreme reporters of energy intake (the "Goldberg cutoffs") reliably reduce or eliminate bias in nutrition studies? Analysis with illustrative associations of energy intake with health outcomes. Am J Clin Nutr. 2019;110(5):1231–9.
67. Ferrari P, Slimani N, Ciampi A, Trichopoulou A, Naska A, Lauria C, et al. 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. 2002;5(6b):1329–45.
68. Orcholski L, Luke A, Plange-Rhule J, Bovet P, Forrester TE, Lambert EV, et al. Under-reporting of dietary energy intake in five populations of the African diaspora. Br J Nutr. 2015;113(3):464–72.
69. Scagliusi FB, Ferriolli E, Pfrimer K, Laureano C, Cunha CS, Gualano B, et al. Characteristics of women who frequently under report their energy intake: a doubly labelled water study. Eur J Clin Nutr. 2009;63(10):1192–9.
70. Tooze JA, Vitolins MZ, Smith SL, Arcury TA, Davis CC, Bell RA, et al. High levels of low energy reporting on 24-hour recalls and three questionnaires in an elderly low-socioeconomic status population. J Nutr. 2007;137(5):1286–93.
71. Park Y, Dodd KW, Kipnis V, Thompson FE, Potischman N, Schoeller DA, et al. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. Am J Clin Nutr. 2018;107(1):80–93.
72. Bunt JC, Lohman TG, Boileau RA. Impact of total body water fluctuations on estimation of body fat from body density. Med Sci Sports Exerc. 1989;21(1):96–100.
73. Subar AF, Potischman N, Dodd KW, Thompson FE, Baer DJ, Schoeller DA, et al. Performance and Feasibility of Recalls Completed Using the Automated Self-Administered 24-Hour Dietary Assessment Tool in Relation to Other Self-Report Tools and Biomarkers in the Interactive Diet and Activity Tracking in AARP (IDATA) Study. J Acad Nutr Diet. 2020;120(11):1805–20.
74. Black AE. The sensitivity and specificity of the Goldberg cut-off for EI:BMR for identifying diet reports of poor validity. Eur J Clin Nutr. 2000;54(5):395–404.
75. Tooze JA, Subar AF, Thompson FE, Troiano R, Schatzkin A, Kipnis V. Psychosocial predictors of energy underreporting in a large doubly labeled water study. Am J Clin Nutr. 2004;79(5):795–804.
76. Tooze JA, Schoeller DA, Subar AF, Kipnis V, Schatzkin A, Troiano RP. Total daily energy expenditure among middle-aged men and women: the OPEN Study. Am J Clin Nutr. 2007;86(2):382–7.
77. Mendez MA, Popkin BM, Buckland G, Schroder H, Amiano P, Barricarte A, et al. Alternative methods of accounting for underreporting and overreporting when measuring dietary intake-obesity relations. Am J Epidemiol. 2011;173(4):448–58.
78. Berta Vanrullen I, Volatier JL, Bertaut A, Dufour A, Dallongeville J. Characteristics of energy intake under-reporting in French adults. Br J Nutr. 2014;111(7):1292–302.
79. Livingstone MB, Robson PJ, Black AE, Coward WA, Wallace JM, McKinley MC, et al. An evaluation of the sensitivity and specificity of energy expenditure measured by heart rate and the Goldberg cut-off for energy intake: basal metabolic rate for identifying mis-reporting of energy intake by adults and children: a retrospective analysis. Eur J Clin Nutr. 2003;57(3):455–63.
80. Andersen LF, Tomten H, Haggarty P, Løvø A, Hustvedt BE. Validation of energy intake estimated from a food frequency questionnaire: a doubly labelled water study. Eur J Clin Nutr. 2003;57(2):279–84.
81. Black AE, Cole TJ. Within- and between-subject variation in energy expenditure measured by the doubly-labelled water technique: implications for validating reported dietary energy intake. Eur J Clin Nutr. 2000;54(5):386–94.
82. Schoeller DA. Measurement of energy expenditure in free-living humans by using doubly labeled water. J Nutr. 1988;118(11):1278–89.
83. Boschmann M, Steiniger J, Hille U, Tank J, Adams F, Sharma AM, et al. Water-induced thermogenesis. J Clin Endocrinol Metab. 2003;88(12):6015–9.

Publisher's Note