CHAPTER 3.

Can energy intake and expenditure (energy balance) be measured accurately in epidemiological studies? Is this important?

Walter C. Willett and Changzheng Yuan

CHAPTER 3

The roles of energy intake and expenditure are extremely important in human health and disease, for many reasons. Thus, the assessment and interpretation of energy intake and expenditure are major issues in epidemiological studies. Overweight and obesity have been recognized to be major risk factors for cancer, cardiovascular disease, diabetes, and many other health conditions. Therefore, the difference between energy intake and expenditure, frequently referred to as energy balance, has become of great interest, because of its direct relationship to long-term gain or loss of adipose tissue.

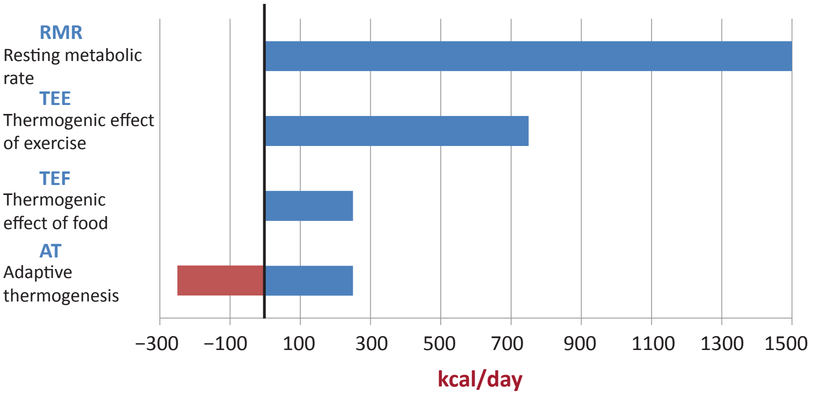
For this reason, questions have arisen about whether energy intake and expenditure can be measured adequately in epidemiological studies, to enable energy balance to be assessed adequately. Some [1–3] have argued that self-reported energy intake has no value and should be abandoned, and have extended this argument to all self-reported information. Others [4–8] have suggested that it is not realistically possible to measure energy intake and expenditure with sufficient precision in epidemiological studies to assess energy balance, but that this is not a serious problem because other means can be used to evaluate the effects (e.g. on disease incidence or mortality) of energy balance as an exposure and to study the determinants (e.g. dietary factors and physical activity) of energy balance as an outcome.

In this chapter, the factors contributing to energy balance and the measurement of these factors are reviewed. Notably, energy intake and expenditure have important roles in human health, and in epidemiological studies, independent of their contribution to energy balance; although these other applications are not the focus of this chapter, they are also mentioned.

Components of total energy expenditure

Total energy expenditure has traditionally been partitioned into several components: resting metabolic rate (RMR), physical activity, thermogenic effect of food, and adaptive thermogenesis (Fig. 3.1) [9]. RMR is quantitatively the most important, making up approximately 60% of total energy expenditure in an individual with moderate physical activity. In a moderately active individual, physical activity accounts for approximately 30% of total energy expenditure. The thermogenic effect of food (i.e. the metabolic cost of

Fig. 3.1. Components of energy expenditure during weight maintenance for a 70 kg man consuming 2500 kcal/day, and the potential modifying effect of adaptive thermogenesis. Adapted with permission from Horton (1983) [9].



absorbing and processing macronutrients) accounts for only about 10% of total energy expenditure. Adaptive thermogenesis (i.e. the compensatory capacity of an individual to conserve or expend energy in response to variable intake of food or temperature extremes) has been estimated to be less than ± 10% of total energy expenditure [9].

In epidemiological studies, the thermogenic effect of food is not likely to vary appreciably, because this becomes important only on extreme diets, and this can generally be assumed to be constant. Adaptive thermogenesis is practically important, because it can account for resistance to weight loss in the face of moderate restriction of energy intake by downregulating metabolic processes to become more energy-efficient. These differences in metabolic efficiency are difficult to measure even under highly controlled conditions as well as in epidemiological studies, so this needs to be recognized as a source of modest unmeasured variation in energy expenditure. RMR is determined mainly by body weight, although this is primarily a function of lean body mass. Because measurement of RMR requires metabolic facilities and is therefore not feasible in epidemiological studies and most clinical investigations, a series of prediction models based on age, weight, and sex have been developed [4]. In principle, height should also be added to the prediction models because, for the same weight, a taller person would be leaner, but height appears to add minimal variability. Because age, weight (or body mass index [BMI] plus height), and sex are routinely covariates in epidemiological studies, RMR is reasonably controlled for in most epidemiological analyses.

After age, weight, and sex have been controlled for, physical activity assumes a relatively large role in determining the variation in energy expenditure among free-living individuals. The true proportion of variation in energy expenditure accounted for by physical activity differs substantially among populations and is likely to be underestimated in most studies because of imperfect measurements of physical activity. Ravussin et al. [10] have demonstrated that even motor activity within the confines of a respiratory chamber (“fidgeting”) varies dramatically between individuals and can account for hundreds of kilocalories per day. Such differences in activity would not be detected by typical questionnaires. Thus, physical activity, which includes both fine motor and major muscle movement, is a major determinant of between-person variation in energy expenditure in many populations. Indeed, in most instances total energy intake can be interpreted as a crude measure of physical activity, especially after controlling for body size, age, and sex.

Energy balance and deviations from energy balance

Energy balance exists when weight is constant because energy expenditure equals energy intake. This can happen when no components change or when a change in one component is compensated for by changes in other components.

In adults, deviations from energy balance are a critical concern because these underlie weight gain and ultimately obesity. Of particular concern to public health and epidemiologists are small increments in weight, such as 0.5–1 kg per year, which are typical of many high-income populations [11] and which over a period of 20–30 years lead to large changes in weight and major morbidity and mortality [12]. The deviations from energy balance needed to produce this change in weight are very modest. For example, simply on the basis of the energy content of adipose tissue [4], if an adult man who consumes 2500 kcal/day (10 460 kJ/day) increases his energy intake by only 1% while other factors remain constant, over a 10-year period a theoretical weight gain of 10 kg would result. In reality, the increase in weight will be considerably less, because the additional energy cost of maintaining and moving the added body mass eventually equals the increment in energy intake and a new steady state in weight, i.e. balance, is reached.

By combining data on energy intake and weight gain and on the compensatory effects of added body mass on energy expenditure, Hall et al. [13] estimated that for each 10 kcal/day (42 kJ/day) increase in energy intake, the new mean steady-state weight will be about 0.5 kg higher (equivalently, each change of 100 kJ/day will lead to a weight change of 1 kg) and that the new steady-state weight will be reached in about 3 years. Their model, based on demographic models and repeated measures of weight, is available interactively at http://bwsimulator.niddk.nih.gov and was recently validated by comparison with repeated dual-energy X-ray absorptiometry (DEXA) measures of body composition and the doubly labelled water (DLW) method [14]. Using a similar approach, Wang et al. [15] estimated that an average increase in energy intake of 110–165 kcal/day (460–690 kJ/day) accounted for the large increase in weight gain among children in the USA between 1990 and 2000. Similarly, Hall et al. [13] estimated that a difference of only 7 kcal/day (30 kJ/day) between energy intake and expenditure could explain the average increase in weight (about 10 kg) in adults in the USA between 1978 and 2005. However, they estimated that the average additional energy intake accumulated over that time was about 220 kcal/day (920 kJ/day); this is the amount by which energy intake would have to be decreased in order to return to the distribution of body weight in 1978.

The important conclusion of these analyses using different approaches is that very small deviations from energy balance, on the order of 1–2% of daily energy intake, can result in large long-term changes in body weight, with major individual and public health implications.

Accuracy and precision of measures of energy intake

Because the calculation of energy balance is based on the difference between energy intake and energy expenditure, and differences of approximately 1% are of great importance, extremely accurate and precise measurements of both intake and expenditure would be needed. Energy intake represents a unique challenge for dietary assessment because, unlike any other nutrient, energy intake is tightly regulated by physiological controls, and thus between-person variation is low after taking into account weight and demographic variables.

A large literature exists on the accuracy of measures of energy intake, specifically comparing mean values obtained by different methods [4, 16]. Progress has been hampered by the lack of a perfect reference method; the closest to that would be the 24-hour whole-body calorimeter, but it artificially constrains physical activity. The DLW method, which measures the relative turnover of hydrogen and oxygen during a period of days or weeks, has become the operational gold standard for assessing energy expenditure, because it is unobtrusive and provides similar mean values to whole-body and respiratory calorimetry. However, the DLW method is extremely expensive, and therefore not practical in epidemiological studies, and is not robust across laboratories, because values have ranged widely in blinded testing [17]. Compared with this standard, most dietary intake assessment methods, including 24-hour dietary recalls, dietary records, and many food frequency questionnaires, underestimate energy intake by 10–20%, although this varies with the population, the details of the specific method, and BMI [4, 18].

Precision is also critical. Precision is difficult to quantify because it is hard to separate true changes in energy intake from measurement errors, although in epidemiological applications, within-person variation due to both sources will have similar implications. A large validation study has recently been completed in which all four methods of dietary assessment were used, which allowed assessment of mean intakes and within-person variation over a 1-year period (see Table 3.1). For all measures of total energy expenditure or intake, the within-person variability, expressed as the within-person coefficient of variation (CV%) and the intraclass correlation coefficient (ICC), is considerable. This includes the DLW method (CV%, 9%; ICC, 0.73) even though its major determinant, weight, has low variability (ICC, 0.98). This degree of variation is similar to what has been seen in other populations; for example, among 111 women the ICC for repeated DLW measurements over 6 months was 0.72 [19]. In another recent evaluation, measurements of energy expenditure using DLW were reproducible over a period of several years, but precision (CV%, ~5%) was still not sufficient for reliably detecting individual changes of 1–2% per year [20]. These analyses of within-person variation underestimate the measurement errors because they assume that each measurement is an unbiased estimate of the true value for individuals, i.e. that there is no systematic within-person error, also described as person-specific bias. This assumption would not apply to food frequency questionnaires, because of their structured nature [16], but is very likely to affect all measurements to some degree, including DLW assessments.

Assessment of energy expenditure due to physical activity in epidemiological studies has been less well developed than assessment of energy intake. Most questionnaires have been focused on discretionary activities or moderate to vigorous activities, assuming that other activities are less important for health or relatively constant in modern lifestyles [21]. Energy expenditure due to physical activity is usually not calculated, in recognition

Table 3.1. Distribution, within-person coefficient of variation (CV%), and intraclass correlation coefficient (ICC) for different measures of energy intake and expenditure in the Women's Lifestyle Validation Study [25] (data provided by 622 female nurses in the USA aged 45–80 years)[a]

Method (*n* = 622)	Time interval	Mean (SD)	Within-person SD	Within-person CV%	ICC
FFQ (kcal/day)	1 year	1901 (480)	286	15	0.70
Dietary records (kcal/day)	~6 months	1745 (334)	226	13	0.63
ASA24 (kcal/day)	Every 3 months	1825 (475)	507	28	0.29
DLW (kcal/day)	6–12 months	2195 (360)	190	9	0.73
Weight (lb)	Every 3 months	157.6 (33.4)	5.1	3	0.98
PAQ (MET-h/day)	1 year	16.5 (6.8)	5.3	32	0.54
Accelerometer (min/day)[b]	~6 months	19.5 (16.6)	8.9	46	0.75
Accelerometer (counts/day)	~6 months	243 056 (94 356)	45 827	19	0.79
PAEE (kcal/day)	6–12 months	708 (239)	166	23	0.53
PAEE (kcal/day)[c]	6–12 months	708 (237)	165	23	0.53
PAEE (kcal/day)[d]	6–12 months	708 (230)	165	23	0.51

ASA24, automated self-administered 24-hour dietary recall; CV, coefficient of variation; DLW, doubly labelled water; FFQ, food frequency questionnaire; ICC, intraclass correlation coefficient; METs, metabolic equivalents; PAEE, physical activity energy expenditure; PAQ, physical activity questionnaire; SD, standard deviation.

[a] ICC and CV% were calculated based on the original value.

[b] Moderate and vigorous activity (min/day), 1-minute bouts.

[c] Measure of activity assessed by DLW with weight regressed out.

[d] Measure of activity assessed by DLW with weight, age, and height regressed out.

of the fact that the data do not capture many activities of daily living and fine motor movements. Physical activity records and 24-hour physical activity recalls, analogous to their corresponding dietary assessment methods, have been minimally used in epidemiological studies thus far. Motion sensors – small devices for monitoring physical activity – are becoming sufficiently inexpensive to be used in epidemiological studies, but the best way to convert movement counts to energy expenditure is still being evaluated. The DLW measure, after subtracting energy expenditure due to RMR, is now often considered to be the gold standard for evaluation of other methods to assess energy expenditure due to physical activity. The variation of these methods over 1 year is also shown in Table 3.1. The most consistent measure appears to be by accelerometer, expressed as counts over 1 day (CV%, 19%; ICC, 0.79), and the DLW measure for physical activity was considerably more variable than that for total energy expenditure (CV%, 23%; ICC, 0.51).

As can be appreciated, the within-person CV% values both for total energy intake and for physical activity assessed by all methods are all far greater than the approximately 1% deviation from energy balance that would be needed to evaluate small but important long-term deviations from energy balance in individuals. Because the deviation would be calculated as the difference between the variables, its within-person error would be even greater because it would include variability from both measures of energy intake and physical activity. Thus, as has been noted earlier [4], it is clear that available methods for measuring energy intake and physical activity in epidemiological studies, as well as methods considered to be the gold standard, are far from adequate for assessing long-term deviations from energy balance in individuals. The relative absence of this approach in the epidemiological literature reflects this understanding. Further, it is unlikely that such methods will become available, because of inherent challenges in obtaining highly precise measurements of long-term behaviours of free-living individuals.

Alternative methods to assess energy balance in epidemiological studies

Fortunately, the study of energy balance does not require measurements of energy intake and expenditure, because attained weight and changes in weight are readily measured with high precision, even by self-report [4]. These measurements of weight provide a simple but precise time-integrated measure of changes in energy balance. Also, weight and changes in weight directly represent the primary health concern due to deviations from energy balance, which is adiposity. Thus, the inability to evaluate long-term deviations from energy balance in individuals by measuring energy intake and expenditure is not important.

The use of weight in epidemiological analyses, both as an exposure and as an outcome, is an important topic that has been discussed widely [22]. When adjusted for height, often expressed as BMI, weight is widely used as a surrogate for adiposity. Although conceptually imperfect because it does not separate lean mass and fat mass, BMI works remarkably well compared with gold-standard methods [23]. When it is used as an exposure, it is important for the analysis to address confounding by smoking, reverse causation due to underlying disease, and loss of lean mass due to frailty at older ages. When it is used as an outcome to study the effects of diet and activity, the study design needs careful consideration.

In cross-sectional studies, reverse causation can readily occur. In prospective studies with only a baseline measurement, the results can be misleading, because a change in diet or activity will often result in a change in weight for some period of time, and then a new steady-state weight will be reached. For example, if physical activity is increased, weight may decrease initially but does not continue to decrease to zero. If most study participants have already reached a steady-state weight at baseline, an effect of physical activity on weight could be missed. A better design will usually be to examine change in diet or activity in relation to change in weight [24], which more closely approximates the design of a clinical trial. Unlike most studies with disease outcomes, which can require many thousands of participants and many years of follow-up, the effects of changes in diet and activity on change in weight can be investigated with a few hundred subjects and 1 or 2 years of follow-up. Randomized trials should play a large role in addressing the effects of diet on weight, because they better control for confounding by variables that are hard to measure.

Other applications of data on total energy intake and expenditure

Although measurements of energy intake and expenditure will not be useful for assessing energy balance in epidemiological studies, they do play other important roles. For example, population trends in mean energy intake over time using 24-hour dietary recalls can provide useful information, because the effects of within-person variation over time can be dampened with large sample sizes. If the method remains standardized over time, temporal trends can still be valid even if there is some systematic underestimation or overestimation. Unfortunately, standardized methods for physical activity assessment over time do not seem to have been used, so there is less certainty about temporal trends in energy expenditure.

In nutritional epidemiological studies, assessment of energy intake is also important as an adjustment variable for nutrient intakes, because the focus is primarily on the composition of diet rather than on absolute intakes. This is because the composition of diet is what can most realistically be changed by individuals or a population [4]. Multiple aspects of dietary composition have been associated with weight changes [11], probably due to differences in satiety and possible hormonal effects that favour or inhibit accumulation of lean mass versus fat mass. Adjustment for total energy intake also has the benefit of cancelling correlated errors in nutrients, thus reducing measurement errors [4].

Assessment of physical activity, primarily by structured questionnaires, has documented the importance of moderate to vigorous activity in prevention of many diseases. Although these measures of physical activity have error, they have been validated by comparisons with more detailed assessments [21] and can thus provide useful information in prospective studies. The fact that these are based on self-reports rather than an objective measure is not important, because objective measures also have error and are subject to confounding. Even if a good measure of total energy expenditure from physical activity were available, this would not provide the important information on specific types of activity that can be obtained by questionnaires. Small motion sensors are now being incorporated into epidemiological studies; the structure of their measurement errors is now being investigated.

Conclusions

Deviations from energy balance are important in human health and disease. However, these cannot be assessed adequately in epidemiological studies by differences between energy intake and expenditure, because very small long-term deviations in energy intake or expenditure can have major effects on body weight. Neither the available methods nor the foreseeable future methods will be sufficiently precise and accurate to measure these small differences. However, body weight and change in weight provide precise indicators of long-term deviations from energy balance and are widely available for epidemiological studies. These simple and inexpensive measures of energy balance can be used as both exposure and outcome variables, taking into consideration their other determinants and confounding factors. Although they are not useful for assessing energy balance, which requires extreme accuracy and precision, measures of energy intake and physical activity will continue to play other important roles in epidemiological studies and in monitoring population trends.

Key points

- Very small differences between energy intake and expenditure can, over time, lead to important gains in weight. Measures of both energy intake and expenditure that are sufficiently precise to quantify these differences will not be available in the foreseeable future.
- Weight and changes in weight provide simple and inexpensive measures of deviations from energy balance.
- Measures of physical activity and energy intake are still valuable in epidemiological studies, even if they cannot be used to evaluate energy balance.

Research needs

- Continued work is needed to evaluate and improve the assessment of physical activity, both amount and type, in epidemiological studies.
- Further effort is needed to understand sources of error in measuring energy intake, including errors in the DLW method, the presumed gold standard.

References

1. Schoeller DA, Thomas D, Archer E, Heymsfield SB, Blair SN, Goran MI, et al. (2013). Self-report-based estimates of energy intake offer an inadequate basis for scientific conclusions. Am J Clin Nutr. 97(6):1413–5. http://dx.doi.org/10.3945/ajcn.113.062125 PMID:23689494

2. Archer E, Blair SN (2015). Implausible data, false memories, and the status quo in dietary assessment. Adv Nutr. 6(2):229–30. http://dx.doi.org/10.3945/an.114.007799 PMID:25770263

3. Dhurandhar NV, Schoeller D, Brown AW, Heymsfield SB, Thomas D, Sørensen TIA, et al.; Energy Balance Measurement Working Group (2015). Energy balance measurement: when something is not better than nothing. Int J Obes (Lond). 39(7):1109–13. http://dx.doi.org/10.1038/ijo.2014.199 PMID:25394308

4. Willett WC (2013). Implications of total energy intake for epidemiologic analyses. In: Willett WC, editor. Nutritional epidemiology. 3rd ed. New York (NY), USA: Oxford University Press; pp. 260–86.

5. Satija A, Yu E, Willett WC, Hu FB (2015). Objective measures are complementary to, rather than a replacement for, self-reported methods. Int J Obes (Lond). 39(7):1179. http://dx.doi.org/10.1038/ijo.2015.80 PMID:25924712

6. Freedman LS, Carroll RJ, Neuhouser ML, Prentice RL, Spiegelman D, Subar AF, et al. (2015). Reply to E Archer and SN Blair. Adv Nutr. 6(4):489. http://dx.doi.org/10.3945/an.115.009118 PMID:26178033

7. Hébert JR, Hurley TG, Steck SE, Miller DR, Tabung FK, Kushi LH, et al. (2015). Reply to E Archer and SN Blair. Adv Nutr. 6(2):230–3. http://dx.doi.org/10.3945/an.114.007831 PMID:25770264

8. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. (2015). Addressing current criticism regarding the value of self-report dietary data. J Nutr. 145(12):2639–45. http://dx.doi.org/10.3945/jn.115.219634 PMID:26468491

9. Horton ES (1983). Introduction: an overview of the assessment and regulation of energy balance in humans. Am J Clin Nutr. 38(6):972–7. PMID:6650453

10. Ravussin E, Lillioja S, Anderson TE, Christin L, Bogardus C (1986). Determinants of 24-hour energy expenditure in man. Methods and results using a respiratory chamber. J Clin Invest. 78(6):1568–78. http://dx.doi.org/10.1172/JCI112749 PMID:3782471

11. Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB (2011). Changes in diet and lifestyle and long-term weight gain in women and men. N Engl J Med. 364(25):2392–404. http://dx.doi.org/10.1056/NEJMoa1014296 PMID:21696306

12. Willett WC, Dietz WH, Colditz GA (1999). Guidelines for healthy weight. N Engl J Med. 341(6):427–34. http://dx.doi.org/10.1056/NEJM199908053410607 PMID:10432328

13. Hall KD, Sacks G, Chandramohan D, Chow CC, Wang YC, Gortmaker SL, et al. (2011). Quantification of the effect of energy imbalance on bodyweight. Lancet. 378(9793):826–37. http://dx.doi.org/10.1016/S0140-6736(11)60812-X PMID:21872751

14. Sanghvi A, Redman LM, Martin CK, Ravussin E, Hall KD (2015). Validation of an inexpensive and accurate mathematical method to measure long-term changes in free-living energy intake. Am J Clin Nutr. 102(2):353–8. http://dx.doi.org/10.3945/ajcn.115.111070 PMID:26040640

15. Wang YC, Gortmaker SL, Sobol AM, Kuntz KM (2006). Estimating the energy gap among US children: a counterfactual approach. Pediatrics. 118(6):e1721–33. http://dx.doi.org/10.1542/peds.2006-0682 PMID:17142497

16. Willett W, Lenart E (2013). Reproducibility and validity of food-frequency questionnaires. In: Willett W, editor. Nutritional epidemiology. 3rd ed. New York (NY), USA: Oxford University Press; pp. 96–141.

17. Roberts SB, Fuss P, Evans WJ, Heyman MB, Young VR (1993). Energy expenditure, aging and body composition. J Nutr. 123 (2 Suppl): 474–80. PMID:8429407

18. Freedman LS, Commins JM, Moler JE, Willett W, Tinker LF, Subar AF, et al. (2015). Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for potassium and sodium intake. Am J Epidemiol. 181(7):473–87. http://dx.doi.org/10.1093/aje/kwu325 PMID:25787264

19. Neuhouser ML, Tinker L, Shaw PA, Schoeller D, Bingham SA, Horn LV, et al. (2008). Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. Am J Epidemiol. 167(10):1247–59. http://dx.doi.org/10.1093/aje/kwn026 PMID:18344516

20. Wong WW, Roberts SB, Racette SB, Das SK, Redman LM, Rochon J, et al. (2014). The doubly labeled water method produces highly reproducible longitudinal results in nutrition studies. J Nutr. 144(5):777–83. http://dx.doi.org/10.3945/jn.113.187823 PMID:24523488

21. Hu FB (2013). Assessment of physical activity in nutritional epidemiology. In: Willett W, editor. Nutritional epidemiology. 3rd ed. New York (NY), USA: Oxford University Press; pp. 241–59.

22. Manson JE, Stampfer MJ, Hennekens CH, Willett WC (1987). Body weight and longevity. A reassessment. JAMA. 257(3):353–8. http://dx.doi.org/10.1001/jama.1987.03390030083026 PMID:3795418

23. Sun Q, van Dam RM, Spiegelman D, Heymsfield SB, Willett WC, Hu FB (2010). Comparison of dual-energy x-ray absorptiometric and anthropometric measures of adiposity in relation to adiposity-related biologic factors. Am J Epidemiol. 172(12):1442–54. http://dx.doi.org/10.1093/aje/kwq306 PMID:20952596

24. Smith JD, Hou T, Hu FB, Rimm EB, Spiegelman D, Willett WC, et al. (2015). A comparison of different methods for evaluating diet, physical activity, and long-term weight gain in 3 prospective cohort studies. J Nutr. 145(11):2527–34. http://dx.doi.org/10.3945/jn.115.214171 PMID:26377763

25. Yuan C, Spiegelman D, Rimm EB, Rosner BA, Stampfer MJ, Barnett JB, et al. (2017). Validity of a dietary questionnaire assessed by comparison with multiple weighed dietary records or 24-hour recalls. Am J Epidemiol. 185(7):570–84. PMID:28338828