Assessment of the validity of reported energy intakes – review and recent developments

By Gail R. Goldberg and Alison E. Black

Abstract

Studies of dietary intake are based on the assumption that the methods used are valid and provide measures of habitual intake. It is only recently that data from whole-body calorimetry and doubly-labelled water studies of energy expenditure have yielded insights into the validity and interpretation of energy intake data. In this paper we review the evidence from these and other studies that have led to researchers' current level of understanding of how best to analyse and interpret reported energy intake data. We also discuss what we consider should be the focus of future research needed to refine experimental protocols and evaluate data.

Introduction

In the 60 years since Widdowson and McCance's pioneering studies measuring the food intake of individuals, a variety of methods have been used routinely in a wide range of subjects and nutritional settings. Most studies of dietary intake are based on the assumption that the methods used are valid and provide measures of habitual intake. However, as Garrow stated almost 25 years ago "the measurement of the habitual food intake of an individual must be among the most difficult tasks a physiologist can undertake" (1). A major problem has been the lack of independent markers to validate measurements of intake. It is only recently that measurements of energy expenditure by whole-body calorimetry and doublylabelled water (DLW) have yielded insights into the validity and interpretation of energy intake data. In this paper we review the evidence from these and other studies that have led to our current level of understanding of the mis-reporting of energy and macronutrient intake and future research needed to further refine experimental protocols and evaluate data.

The doubly-labelled water method

It is not an exaggeration to state that doubly-labelled water (DLW, ${}^{2}H_{2}{}^{18}O$) has revolutionised studies of human energy expenditure (2,3). The method was originally developed for use in small animals over 40 years ago, but more recently advances in technology have made its application feasible in humans. It enables total

energy expenditure (TEE) to be measured under genuinely free-living conditions with the minimum of inconvenience or hindrance to the subject. After oral dosing with the labelled water, subjects are required only to collect a single urine sample daily for 10-15 days. The difficulties of the method are confined to the laboratory – the sample analysis and interpretation of data is complex. The main drawback to widespread use of DLW is the price of one of the stable isotopes, oxygen-18. It costs approximately £500 (SEK 6000) to dose a 70 kg subject.

Over the past 15 years DLW has been used throughout the world by many investigators in a wide variety of subjects and circumstances. The studies are far too numerous to list here, but a comprehensive review by *Black* et al. contains an extensive list of original citations (4). Because the measurement of TEE is integrated over a period of 10-15 days, it is likely, in the majority of subjects, to be representative of their habitual (long term average) energy expenditure. DLW therefore also gives an independent measure of energy requirements and can be used as a bio-marker of energy intake. A number of studies have now been conducted in which DLW and dietary intake assessments have been made concurrently in the same individuals. The first of these in 1986 (5) demonstrated conclusively the answer to a paradox that many researchers in energy physiology had been trying to answer, i.e. that obese subjects apparently existed on energy intakes that were the same as, or less than, their lean peers. Despite decades of research no defects in energy expenditure were identified which were of sufficient magnitude to explain these differences. Prentice et al. (5) found that reported energy intakes (EI_{rep}) in obese subjects were a mean of 3.5 MJ/day less than measurements of TEE. In contrast a group

of lean women demonstrated excellent agreement between EI_{rep} and TEE. Since then, other investigators have also observed that mean energy intake measurements are lower than simultaneous measurements of TEE in populations of obese adults and obese adolescents (6,7). Subsequently groups of post obese subjects were found to be more likely than their never-obese peers to underreport energy intake (8). These findings in relatively small groups of volunteer subjects confirmed what energy physiologists and some dietitians had long suspected; that obese persons tend to underreport their food intake. However, in 1990 Livingstone et al. (9) demonstrated underreporting also in a randomly selected sample of men and women taking part in a large study of diet and health. This was clearlyan issue of potentially far greater concern with serious implications for previous and ongoing studies, both large and small. Since DLW is too expensive and technically demanding to use routinely to validate measurements of food intake, we proposed evaluating energy intake by comparing it with presumed energy requirements, both expressed as multiples of BMR (10). This method depends on the fundamental principle of energy physiology, that, if body weight is stable then total energy expenditure equals total energy intake.

Evaluation of energy intake – physical activity levels

The 1985 FAO/WHO/UNU report on energy requirements expressed total daily energy expenditure (and therefore energy requirements) as a multiple of basal metabolic rate (BMR) (11). The use of this ratio with BMR as the denominator removes much of the inter-individual variability due to weight, height, age and sex. Thus

Gail R. Goldberg*, BSc, Alison E. Black, BSc SRD, Medical Research Council, Dunn Clinical Nutrition Centre, Hills Road, Cambridge CB2 2DH, United Kingdom. *Correspondence**

E-mail: Gail.Goldberg@mrc-dunn.cam.ac.uk

physical activity levels (TEE/BMR or PAL), provide a means by which the activity levels of individuals can be compared directly. We proposed that EI_{rep} data be evaluated by comparing EI:BMR with the presumed PAL of the population under investigation. This is not quite as straightforward as first appears, since when evaluating EI_{rep} data the investigator needs to ask two questions (1) Can the EI_{rep} be representative of *habitual* (long term average) intake or (2) can the EI_{rep} be a valid estimate of *actual* intake during the short period of investigation when an intake either higher or lower than average could obtained by chance.

Group mean energy intake over the long term

The criterion for answering this question was obtained from studies using wholebody calorimeters. These enable minuteby-minute changes in energy expenditure to be measured and 24 hour energy expenditure (24 hr EE) can be broken down into its major components (basal metabolic rate, BMR; diet induced thermogenesis DIT; physical activity). Calorimeters provide a comfortable, although artificial, environment for subjects. Investigators often use strictly standardised protocols in order to remove as much behavioural noise as possible enabling the underlying physiology to be studied (12). In the context of assessing dietary intakes, data from whole-body calorimeters have been vital for yielding 24 hr EE and PAL values at or below the minimum which might reasonably be expected in healthy, free-living people. The mean \pm SD of 207 individual measurements from 9 studies in 5 different laboratories was 1.35±0.58 (10). In all these studies subjects were required to be very sedentary, were measured in the thermoneutral range (to exclude thermoregulatory thermogenesis) and were close to energy balance (to exclude excess DIT). Because of the physical constraints of a calorimeter compared with normal patterns of physical activity, it is highly unlikely that a healthy free-living adult would have a habitual PAL lower than 1.35. Therefore, for EI_{rep} to be representative of habitual or average intake over the long term, it must be greater than 1.35 x BMR. Values less than 1.35 cannot represent habitual intake. Values above 1.35 cannot be rejected as invalid, but are not necessarily valid as physical activity may be greater.

Group mean energy intake over the short period

Again, the basic premise is, if weight is stable then energy expenditure equals

energy intake. Both EI_{rep} and presumed energy expenditure can be expressed as multiples of BMR. However, the question of validity is much more complex. Although in groups of subjects average TEE should equal average EI_{rep}, errors on all the measurements mean that absolute agreement is not necessarily obtained. For evaluating whether mean EI_{ren} is a valid estimate of actual intake during the period of investigation we need to estimate the confidence limits of the agreement between EI:BMR and PAL. Statistically if mean reported EI:BMR is less than the lower 95% confidence limit or cut-off, then it is highly improbable that the reported intake could represent genuinely low intakes obtained by chance. Equally, if reported mean EI:BMR is greater than the upper 95% confidence limit or cut-off, then it is highly improbable that the reported intake could represent genuinely high intakes obtained by chance.

A number of factors need to be considered. These include the within-subject day-to-day variation in energy intake; the number of days of dietary assessment (more days reduces the day-to-day variation); the number of subjects in the sample population (more subjects improves the precision of the measurement); the variation in repeated measurement of BMR or the precision of estimated versus measured BMR; the between-subject variability in PAL; and the expected mean freeliving PAL. Our formula to calculate the 95% confidence limits of agreement between EI:BMR and PAL taking all these variables into account is shown in Figure 1.

The values for CVs we used in these equations can be found in the original paper (10). It is important to note that investigators can, of course, use their own data if appropriate. In particular, they should choose a PAL for comparison with EI_{rep} that is appropriate to their own study population.

Underreporting

Using this formula we calculated study specific lower cut-off limits and evaluated mean and individual EI data from both our own studies and a number of published studies (13). It is important to note here that for these evaluations we chose to derive cut-offs using a PAL of 1.55 for the assumed physical activity of the populations (the yardstick for comparison). This is the value defined by WHO as a sedentary level of energy expenditure (11). We chose this value because we had no knowledge of the physical activity in the studies we examined. For this same reason we made no attempt to calculate an upper cut-off. Our analyses showed conclusively that underreporting of habitual and Figure 1. Formula to calculate the 95% confidence limits of agreement between EI:BMR and PAL taking all these variables into account.

Cut-off value for EI:BMR
= PAL x exp [SD x
$$(S/100)$$
]
 \sqrt{n}

Where:

- *PAL* is assumed physical activity level for the population under study
- SD is -2 for lower or +2 for upper 95% confidence limits or -3 for lower or +3 for upper 99% confidence limits
- *n* is the number of subjects
- *S* is the overall CV (coefficient variation) for PAL taking into account the variability in energy intake and BMR

	S	is given by the equation
	$S = \sqrt{[}$	$[CV_{IW}^2/k + CV_B^2 + CV_P^2]$
	Where	· · · · · · · · · · · · · · · · · · ·
	CV^2_{IW}	is within individual variation
	1.0	in energy intake
	k	is the number of days of diet
		assessment
	CV_{R}^{2}	is the variation in repeated
	D	BMR measurements or
		precision of estimated vs
		measured BMR
	CV^2_{P}	is the between-subject variation
	1	in PAL
_		

actual EI was both widespread and serious even in the most carefully conducted studies. When entire data-sets were examined, improbably low intakes from individuals were not balanced by high intakes at the upper end of the distribution and thus even mean intake data was not valid. In many instances mean or individual EI_{rep} were less than BMR. Subsequent analysis of nearly 600 DLW measurements has shown a mean PAL greater than 1.55 in both men and women in all age groups except those over 75 years old (4,14). Thus the extent of underreporting in our original analysis which used a PAL of 1.55 was almost certainly an underestimate of the problem.

Recent developments

The use of cut-offs has heightened awareness about the prevalence of patently incorrect estimates of energy (and hence nutrient) intakes and it has taken only a few years for it to be almost universally accepted that underreporting is a problem. Dietary underreporting is a difficult area in which to work, because our demonstration that it was a feature of almost all studies seriously undermined the dietary

Figure 2. Evaluation of energy intake.

In order for energy intake to be evaluated properly dietary studies should routinely include:

- * Measurements of height and weight (to estimate BMR)
- * Some assessment of subjects' occupation and leisure activities so that an informed choice of PAL is used for comparison.
- * Questions to identify weight-conscious individuals
- * Psychological assessments if appropriate and external validation if possible

Then:

- * Calculate individual EI:BMR
- * Compare mean EI:BMR with appropriate PAL
- * Examine mean and distribution of EI:BMR in sub-groups
- * Look for bias linked to
- parameters under study* Look at characteristics of those
- giving biased records
- * Possibly reject individual records

intake data from many very large and also smaller scale published and ongoing studies. From the outset we emphasised that highlighting the problem of underreporting must not be interpreted as suggesting that dietary surveys are a waste of time. Drawing the scientific community's attention to the huge implications was the first step. Investigators were urged to routinely include measurements of height and weight (to estimate BMR) and also some assessment of subjects' occupation and leisure activities to permit an informed choice of 'yardstick' PAL (13,15) (Figure 2). It was also a matter of priority to focus on the identification of the sources of bias (Figure 2). This in turn should lead to better study designs and strategies for interpreting data. In the past seven years various investigators have indeed begun to address these issues by reanalysing their own data, particularly those obtained from large cohorts, and by designing specific studies to answer particular questions.

Food intake data is now being examined and interpreted far more critically. Many researchers have used our cut-off values, and the three most commonly raised questions are how best to

- *l*) evaluate the reported intakes of
- individuals vs groups of subjects
- 2) choose an appropriate PAL
- 3) identify overreporting.

Dealing with individual data

If we are to understand underreporting and devise new methods either for collecting data or for analysing the data that we do collect, then we need to understand who are the underreporters and why they underreport. To do this, we need to be able to identify them. Our formula to evaluate EI_{rep} by comparison with the mean expected PAL of the population incorporates a factor for the number of subjects (n). A logical extension to evaluating the mean EI_{ren} is to argue that, if n=1, then the formula calculates the cut-off value to identify invalid data from individuals. This approach has been taken by several authors in re-examining large national databases (16), and this has led to significant advances in our understanding. These authors defined underreporters as those subjects with an EI:BMR below a certain level, usually 1.1 or 1.2 i.e. a level defined using a 'yardstick' PAL of 1.55. There are two problem with this approach. First, this cut-off only identifies underreporters by comparison with a sedentary PAL. DLW data show that there is underreporting at all levels of energy expenditure and that a cut-off around 1.2 identifies only about 50% of them (17). The situation is slightly improved if a more appropriate higher mean PAL is used for the comparison in groups that are more active. It is improved further if each individual is evaluated against a PAL appropriate to him/her self (18). This is the basis for our recommendation that all dietary studies should incorporate assessments of physical activity as routine. This improves the specificity of the cut-off. The second problem is that, for n=1, the confidence limits are wide and the sensitivity of the technique to identify all underreporters is limited.

The choice of appropriate PAL

The choice of PAL can be made at three levels of specificity depending on information available:

l. A suitable mean PAL for the group for comparison with mean EI_{rep} to calculate a cut-off at the group level and identify the presence or absence of bias to underreporting. As explained above, this can also be used to calculate the cut-off for n=1 and identify some of the underreporters at the individual level. The meta-analysis of nearly 600 DLW data has been used to determine mean PAL for different age and sex groups (14).

2. If there is information on activity sufficient to classify subjects into low medium or high levels of activity, then three different cut-offs for n=1 can be calculated and used to identify individual underreporters in each of three broad bands of activity. From the meta-analysis of DLW data in free-living subjects, the mean PAL in the low, medium and high

thirds of the distribution was 1.4, 1.65 and 1.95/2.0 (women/men) respectively. These values are close to the three WHO levels of energy expenditure for light (1.55), moderate (1.64/1.78) and heavy (1.82/2.10) levels of energy expenditure (11). Either set of figures could be used.

3. Where detailed information on occupational and leisure activity has been obtained, then subject-specific PAL could be devised and used to evaluate the EI of each subject individually. The meta-analysis showed that individual PAL in normally active free-living individuals (not selected for specific occupations and activities such as soldiers and athletes) can range from 1.2 to >2.2 x BMR at all ages (4,14).

Overreporting

The initial focus of the evaluation of energy intake was on undereporting. This was for two reasons. Firstly, there is a bias towards underreporting. As discussed above, implausibly low intakes from individuals are not balanced by high intakes at the upper end of the distribution. Secondly, until very recently it was not possible to define an appropriate PAL for calculating upper 95% confidence limits or cut-offs and thus identifying overreporters. DLW data from individuals with very high levels of physical activity are now available and can be used as the basis for assessing individuals or populations known to have high levels of energy expenditure and for identifying possible overreporters (4). It is of course important to make the distinction between the maximum achievable PAL over a defined period of time and the maximum sustainable habitual PAL. DLW data obtained from Polar explorers, mountaineers and elite athletes in competition show that under conditions which elicit extremely high levels of energy expenditure over relatively short periods of time PALs can be as high as 5.0. Data obtained from athletes and soldiers during extended periods of training and subsistence farmers, suggest that a PAL of 2.5 is indicative of an extremely physically active lifestyle, and is probably the maximum sustainable in the long term. A value of 2.0 is likely among manual workers in the West where activity may vary widely from day-to-day or week-to-week. As an indication of overreporting, if PALs of 1.55 or 2.0 and n=1 are incorporated into our formula (while retaining the original factors for CV_{IW} , CV_B and CV_P), the upper cut-off limits for 7 day energy intake measurements are 2.19 x BMR and 2.82 x BMR, respectively. Thus an individual reporting EI:BMR >2.82 x BMR is likely to be overreporting actual food intake.

The future

The initial focus of the evaluation of dietary intake data was on undereporting of energy intake. This is not the place for a full review of recent developments since a number of investigators, including ourselves, are now addressing more specific issues. What is clear is that it is not only obese subjects who underreport their energy intake, and that not all obese subjects underreport. However, the pro-

References

- Goldberg GR: From individual variation in energy intakes to variations in energy require-ments and adaptations to them. Brit J Nutr 997;78 (suppl 2):S81-S94.
- IDECG. International Dietary Energy Consultancy Group. The doubly-labelled water method for measuring energy expenditure: Technical recommendations for use in humans.
- Cambridge: IDECG/IAEA, Vienna, 1990. Speakman JR: Doubly-labelled water: theory and practice. London: Chapman and Hall, 1997 3
- Black AE, Coward WA, Cole TJ, Prentice AM: 4 Human energy expenditure in affluent socie-ties: an analysis of 574 doubly-labelled water measurements. Eur J Clin Nutr 1996;50:72-92. Prentice AM, Black AE, Coward WA, Davies HL, Goldberg GR, Murgatroyd PR, et al: High
- 5. levels of energy expenditure in obese women. Brit Med J 1986;292:983-87.
- Bandini LG, Schoeller DA, Cyr HN, Dietz WH: Validity of reported energy intake in obese and nonobese adolescents. Amer J Clin Nutr 1990; 52:421-5
- 7. Lichtman SW, Pisarska K, Berman ER, Pestone M, Dowling H, Offenbacher E, et al: Dis-crepancy between self-reported and actual
- crepancy between sen-reported and actual caloric intake and exercise in obese subjects. New Engl J Med 1992;327:1893-8. Black AE, Jebb SA, Bingham SA, Runswick SA, Poppitt SD: The validation of energy and protein intakes by doubly-labelled water and 24-hour urinary excretion in post-obese sub-iacts. J Hum Nutr Diet 1005:8:51-64 jects. J Hum Nutr Diet 1995;8:51-64
- Livingstone MBE, Prentice AM, Strain JJ, Coward WA, Black AE, Barker ME, et al: Accuracy of weighed dietary records in studies of diet and health. Brit Med J 1990;300:708-12.

bability of underreporting rises as BMI rises. Underreporting of energy intake occurs at all levels of energy expenditure and in all groups of people (16,17). Characterisation of the sub-groups most likely to mis-report their food intake requires some form of psychological assessment (eg. dietary restraint, attitudes to food and other concerns) and identification of the reasons for doing so (16, 19-22). The issue of differential reporting of

- 10. 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. 1. Deri-
- vation of cut-off limits to identify under-recording. Eur J Clin Nutr 1991;45:569-81. FAO/WHO/UNU. Report of a joint expert consultation: Energy and protein requirements. Technical Report Series 724. Geneva: WHO, 11. 1985
- 12. Murgatroyd PR, Shetty PS, Prentice AM: Techniques for the measurement of human energy expenditure: a practical guide. Internat J Obes 1993;17:549-68. Black AE, Goldberg GR, Jebb SA, Livingstone
- 13. MBE, Cole TJ, Prentice AM: Critical evaluation of energy intake data using fundamental principles of energy physiology. 2. Evaluating the results of dietary surveys. Eur J Clin Nutr 1991; 45.583-99
- 14. Black AE: Physical activity levels from a metaanalysis of doubly-labelled water studies for validating energy intake as measured by dietary assessment. Nutr Rev 1996;54:170-4
- 15. Black AE, Prentice AM, Goldberg GR, Jebb SA Bingham SA, Livingstone MBE, et al: Measurements of total energy expenditure pro-vide insights into the validity of dietary mea-surements of energy intake. J Amer Diet Assoc 1993;93:572-9
- 16. Price GM, Paul AA, Cole TJ, Wadsworth MEJ: Characteristics of the low-energy reporters in a longitudinal national dietary survey. Brit J Nutr 1997:77:833-51.
- 17. Black AE: Underreporting of energy intake at all levels of energy expenditure: evidence from doubly labelled water studies. Proc Nutr Soc

macronutrients and whether particular food types, meals or snack foods are more likely to be misreported and the possible reasons for this (22-25) need study. When greater understanding of what is happening, and why, has been achieved, then perhaps techniques for reporting food intake can be improved. Certainly, analysis and interpretation of the data that has been collected can be improved, and conclusions drawn with greater certainty.

1997;56(1A):121A.

- 18 Black AE, Bingham SA, Johansson G, Coward WA: Validation of dietary intakes of protein and energy against 24 hour urinary N and DLW energy expenditure in middle-aged women, retired men and post-obese subjects: com-parisions with validation against presumed energy requirements. Eur J Clin Nutr 1997;51: 405-13.
- Lindroos A-K, Lissner L, Mathiassen ME, 19 Karlsson J, Sullivan M, Bengtsson C, et al: Dietary intake in relation to restrained eating, disinhibition, and hunger in obese and nonobese
- Similation, and nunger in obese and nonobese Swedish women. Obes Res 1997;5:175-82. Macdiarmid JI, Blundell JE: Dietary underreporting: what people say about recording their food intake. Eur J Clin Nutr 1997;51:199-200.
- 21. Mela DJ, Aaron JI: Honest but invalid: what subjects say about recording their food intake. J Amer Diet Assoc 1997;97:791-3.
- Lafay L, Basdevant A, Charles M-A, Vray M, Balkau B, Borys J-M, et al: Determinants and nature of dietary underreporting in a free-living population: the Fleurbaix Laventie Dante (FLVS) study. Intern J Obes 1997;21:567-73.
- Lissner L, Lindroos A-K: Is dietary under-reporting macronutrient specific? Eur J Clin Nutr 1994;48 (letter):453-4. 23.
- 24. Heitmann BL, Lissner L: Dietary underreporting by obese individuals - is it specific or non-specific ? Brit Med Journal 1995;311:986-9.
- 25 Poppitt SD, Swann D, Black AE, Prentice AM: Assessment of selective underreporting of food intake by both obese and non-obese women in a metabolic facility. Internat J Obes 1998;22:303-