British Journal of Nutrition (2001), **86**, 601–605
DOI: 10.1079/BJN2001449

Energy intake by multiple pass 24 h recall and total energy expenditure: a comparison in a representative sample of 3–4-year-olds

John J. Reilly[1]*, Colette Montgomery[1], Diane Jackson[1]†, Jane MacRitchie[2] and Julie Armstrong[2]

[1]*University of Glasgow Department of Human Nutrition, Yorkhill Hospitals, Glasgow G3 8SJ, Scotland, UK*
[2]*School of Biological and Biomedical Sciences, Glasgow Caledonian University, Cowcaddens Road, Glasgow, Scotland, UK*

(Received 27 November 2000 – Revised 18 June 2001 – Accepted 18 July 2001)

The accuracy of the multiple pass 24 h recall for assessment of habitual energy intake in pre-school children is unclear. The primary aim of this study was to assess its accuracy by comparison with measurement of total energy expenditure by doubly-labelled water in a representative sample of forty-one 3–4-year-olds. The recall method was well tolerated by subjects and was administered quickly and easily. However, it produced estimates of energy intake which significantly exceeded measures of total energy expenditure from doubly-labelled water, mean paired difference 660 kJ/d ($P<0{\cdot}01$). Agreement between the two methods was poor at the individual level: limits of agreement 660+/−3018 kJ/d. Error and imprecision in both methods contribute to individual differences, but the present study suggests that the multiple pass 24 h recall does not provide accurate estimates of dietary energy intake in individual children.

Doubly-labelled water: Energy expenditure: Children: Energy intake misreporting

Assessments of dietary energy intake (EI) in pre-school children are commonly made, in applications which range from clinical settings to experimental investigations of the causes of energy imbalance to population surveys. In older children and adolescents (Bandini *et al.* 1990; Livingstone *et al.* 1992; Maffeis *et al.* 1994; Champagne *et al.* 1996), and adults (Livingstone *et al.* 1990; Sawaya *et al.* 1996; Kroke *et al.* 1999) such assessments tend to be time consuming yet produce results which are prone to imprecision and inaccuracy (largely due to dietary under-reporting).

In 3–4-year-old children, the accuracy and practical utility of EI assessments are less clear because the evidence is somewhat limited: a recent review of dietary assessment methodology for use in children concluded that more validation studies were necessary (Livingstone & Robson, 2000). In subjects in energy balance EI should approximate total energy expenditure (TEE) and so measurement of TEE, with doubly-labelled water, is widely used as a criterion method against which dietary assessment techniques can be validated. This assumption is fair even in young, growing, children because the energy cost of growth is very small after infancy (typically <2 % of daily EI; Kuzawa, 1998). Our literature search revealed only one comparison of TEE and EI in 3–4-year-olds. Davies *et al.* (1994) found in twenty-seven 3·5–4·5-year-olds in England that average EI, measured by 4 d weighed record, slightly, but not significantly, exceeded average TEE, but that error at the individual level was large. In two studies of slightly older children in the USA the accuracy of a food-frequency questionnaire (Kaskoun *et al.* 1994; *n* 45, 4–7-year-olds) and the multiple pass 24 h recall (Johnson *et al.* 1996; *n* 24, 4–7-year-olds) was poor relative to TEE measured by doubly-labelled water, though bias in estimation by multiple pass 24 h recall was small, suggesting that it might be suitable for group estimates of EI (Johnson *et al.* 1996).

The multiple pass 24 h recall (Johnson *et al.* 1996; Guenther *et al.* 1998) was designed in part with the aim of reducing the problem of under-reporting of EI, and was intended to be very practical. It is widely used in the USA but its use has been very limited outside the USA. The primary aim of the present study was to assess the validity of the multiple pass 24 h recall for assessment of EI in a representative sample of 3–4-year-olds.

Subject and methods

Subjects and design

We aimed to recruit a sample of 3–4-year-olds representative of Glasgow, Scotland, UK, in terms of social class. This

Abbreviations: EI, energy intake; TEE, total energy expenditure.

* **Corresponding author:** Dr John J. Reilly, fax +44 141 201 9275, email jjr2y@clinmed.gla.ac.uk
† **Present address**: Aberdeen Centre for Energy Regulation and Obesity, Human Nutrition Unit, Rowett Research Institute, Aberdeen, Scotland, UK.

was achieved by recruiting children from nurseries in postcodes selected to approximate the social class distribution of Glasgow, one of the most socially deprived areas of the UK (General Register Office Scotland, 1993). The sample and sampling frame have been described elsewhere (Jackson *et al.* 2000).

Measurement of energy intake

EI was assessed using the multiple pass 24 h recall. This method, described in detail elsewhere (Johnson *et al.* 1996; Guenther *et al.* 1998), was originally designed to minimise under-reporting, while reducing the amount of effort required of the investigator and subject. To our knowledge the present study was its first use outside the USA.

In brief, the multiple pass recall consisted of three elements (passes): quick list; detailed description; review; as previously described by Johnson *et al.* (1996). Food portion sizes were estimated by a single dietitian using a combination of the photographic atlas of portion sizes for adults (Nelson *et al.* 1996, 1997), guide weights for portion sizes of selected foods consumed by children of this age in the UK based on a national survey (Gregory *et al.* 1995), and estimates provided in household measures by the carer. The recall interviews were conducted over the telephone after a face-to-face interview with the carer; in almost all cases the mother was the primary carer and was the target of the dietary recall. Photographs of food portion sizes and written instructions were provided to carers. Recalls were carried out over 3 d (two weekdays, one weekend day) during the 7 d period of the TEE measurements. It was felt that 3 d would be adequate to obtain a reasonably precise estimate of EI on the basis of earlier evidence in children of similar age (Davies *et al.* 1994, 4 d; Johnson *et al.* 1996, 3 d), though the standard error of the estimate for dietary EI was calculated and is discussed later. In addition, extension of the recall beyond a 3 d period was considered an unreasonable burden on carers.

Food intake records were converted to energy intake estimates using nutritional analysis software (Compeat version 4.0; Nutrition Systems, Grantham, Lincs., UK) based on the UK Food Composition Database (Royal Society of Chemistry, Cambridge, Cambs., UK). All interviews and nutrient analyses were carried out by the same investigator, a dietitian (CM).

Measurement of total energy expenditure

We measured TEE using the doubly-labelled water method. This is the most accurate method for measurement of TEE in free-living subjects, and is widely regarded as the only accurate method (Speakman, 1998). The method was used as previously described (Reilly *et al.* 1998, 1999). In brief, this involved giving each child an accurately weighed, sterilised, oral dose of ^{18}O (10 % enriched, 1·6 ml/kg body weight) mixed with 0·06 ml 99·9 % enriched $^{2}H_2O$/kg. Urine samples were obtained from each child before the dose, then on days 1 and 7 after the dose. A pilot study of four subjects suggested that 7 d was adequate in that it represented two half-lives of the ^{18}O isotope, though rate constants for both isotopes in the main study are given later. We used equation A6 of Schoeller *et al.* (1986): validation studies suggest that this equation is as accurate as the alternatives (Speakman, 1998). We converted estimates of CO_2 production to energy expenditure by using an average food quotient based on food records from the present study of 0·86, an energy equivalent of 23·7 kJ/l. Individualised estimates of food quotient from each child's food records were not used for three reasons. First, this would have involved the (doubtful) assumption that diet composition data were accurate for individuals. Second, individual differences in food quotient in the present and previous studies in Glasgow (e.g. Reilly *et al.* 1998, 1999) were small. Finally, the effect of small differences in food quotient on final measurements of TEE were negligible, as expected (Speakman, 1998).

Urine samples were prepared in triplicate for ^{2}H analysis using the method of Scrimgeour *et al.* (1993). In brief, aliquots of 400 ul urine were pipetted into 10 ml vials containing Pt catalyst (5 % Pt on alumina powder; Sigma Aldrich, Gillingham, Kent, UK), which were then placed on a twenty tube manifold and evacuated for 1 min. The manifold was brought to atmospheric pressure prior to flushing with reference gas for 5 min. Samples were filled with reference gas (H–He (5:95, v/v); Air Products Special Gases, Crewe, Ches., UK) for 10 s and left to equilibrate for at least 48 h at room temperature prior to analysis. Samples for analysis of ^{18}O were prepared using the method of McMillan *et al.* (1989). After ^{2}H analysis, samples were refrigerated until the analysis of ^{18}O. Vials were evacuated as described earlier and filled with reference gas (CO_2–N_2 (3:97, v/v); Air Products Special Gases) and samples left to equilibrate for 24 h at room temperature. The abundance of ^{2}H and ^{18}O in the gas phase was measured using a continuous flow isotope ratio MS (Hydra, Europa, Crewe, Ches., UK; Prosser & Scrimgeour, 1995). Reference samples of water were prepared and analysed with each batch of urine samples.

The energy cost of growth was expected to be small in these subjects (<2 % EI; Kuzawa, 1998) and so no correction was made for it and energy requirement was assumed to be equal to TEE. Measurements and calculations used for the doubly-labelled water method were made 'blind' to the assessments of EI and different staff measured the two variables. For descriptive purposes heights (to 0·1 cm) and weights (to 0·1 kg) were measured in light indoor clothing (without shoes) in all subjects and BMI values expressed as SD scores relative to UK 1990 reference data (Cole *et al.* 1995).

Statistical analysis and power

Differences between EI and TEE were tested for significance using paired *t* tests and calculation of 95 % CI. The study was intended to assess agreement at the individual level so we also calculated the 'biases and limits of agreement' between EI and TEE using the method of Bland & Altman (1986). This method was used as previously described (Davies *et al.* 1994; Kaskoun *et al.* 1994).

Differences in validity between boys and girls were formally tested using a *t* test. To estimate the power of the validation study we referred to previous studies which compared assessments of EI and TEE in young children (Davies *et al.* 1994; Kaskoun *et al.* 1994; Johnson *et al.* 1996). In all of these studies, sample sizes in the 3–4-year-old age range were less than thirty, yet this was adequate for identification of biases and determination of the degree of agreement at the individual level. We therefore aimed to recruit at least thirty children to the study.

Results

Characteristics of subjects

We recruited forty-one children (twenty-three boys, eighteen girls) to the study (mean age 3·7 (SD 0·4) years). Characteristics of subjects are given in Table 1.

Comparison of energy intake and total energy expenditure

Mean EI was 6·5 (SD 1·4) MJ/d and TEE 5·8 (SD 1·2) MJ/d (Table 1). We found no significant differences in the EI *v.* TEE comparison between the sexes, so data from both sexes were combined. EI significantly exceeded TEE: mean paired difference 660 kJ/d (paired t, $P<0·01$), with 95 % CI 183, 1137 kJ/d.

The mean difference ('bias') and limits of agreement (bias +/− t.SD of the differences) were 660+/−3018 kJ/d. The bias (Fig. 1) tended to increase with increasing EI but this did not reach significance (r 0·26, $P>0·05$).

Dilution space ratios (^2H:^{18}O) used in the Schoeller (1986) equation A6 are fixed, but our observed uncorrected mean was 1·022 (SD 0·019). Mean rate constants were 0·184 (SD 0·046) for ^{18}O and 0·135 (SD 0·044) for ^2H, with mean ^{18}O:^2H elimination rates of 1·36 (SD 0·12). These elimination rates corresponded to a mean study period of 1·9 biological half-lives for ^{18}O and 1·5 half-lives for ^2H.

Discussion

The present study found that use of the multiple pass 24 h recall in a representative sample of 3–4-year-olds was relatively quick for investigators and respondents, and well tolerated. In a pilot study of three families, used for training purposes, the initial recall interview averaged 33 (SD 4) min, but this reduced to 19 (SD 2) min for subsequent interviews. In another pilot study of practical utility involving seventeen children, the initial interview lasted an average of 17 (SD 5) min. The subsequent interviews, on second and third days, were significantly shorter at 8 (SD 4) and 7 (SD 4) min respectively, giving an average total interview time of 32 min over 3 d. A total of 98 % of recalls in this pilot study (50/51) were made successfully at either the first or second attempt to contact the mother by telephone.

Despite these important practical advantages, the multiple pass 24 h recall produced a significant overestimate of EI (by about 11 %). We found no evidence that bias was related to diet composition, or to weight status. Complete concordance of EI and TEE measurements would not be expected given the imprecision and error inherent in both methods. In addition, a small (approximately 2 %; Davies *et al.* 1994; Kaskoun *et al.* 1994; Kuzawa, 1998) excess of EI over TEE, representing the energy deposited in growth, might be expected. However, the observed bias in the present study was larger than the excess energy required for growth and might have increased with increasing energy requirement (Fig. 1), though this trend did not reach significance. Reasons for the bias are unclear, though one possibility might lie in the portion-size estimates used. There are currently two main tools in use for estimation of food portion sizes in the UK, both validated, but both based on adult data (Ministry of Agriculture, Fisheries and Food, 1993; Nelson *et al.* 1996, 1997). Errors in using these tools in children are likely and in our experience, mothers tended to report portions at the lower end of the adult size range. However, gross errors in portion size estimation in the present study were minimised in a number of ways. First, for many food items (e.g. biscuits, slices of bread; proportions of canned food and drinks consumed) average adult quantities were appropriate. Second, where possible estimated values from surveys of adults were compared with guide weights for typical portion sizes of young children (Gregory *et al.* 1995) and were usually in broad

Table 1. Characteristics of subjects†
(Mean values and standard deviations)

	Mean	SD
Age (years)	3·7	0·4
Weight (kg)	16·5	2·6
Height (m)	1·01	0·05
BMI (kg/m^2)	16·1	1·8
BMI SD score	0·13	1·19
Energy intake (MJ/d)	6·5**	1·1
Total energy expenditure (MJ/d)	5·8	1·2

† n 41, twenty-three boys, eighteen girls.
Mean value was significantly different from that of total energy expenditure (**$P<0·01$)

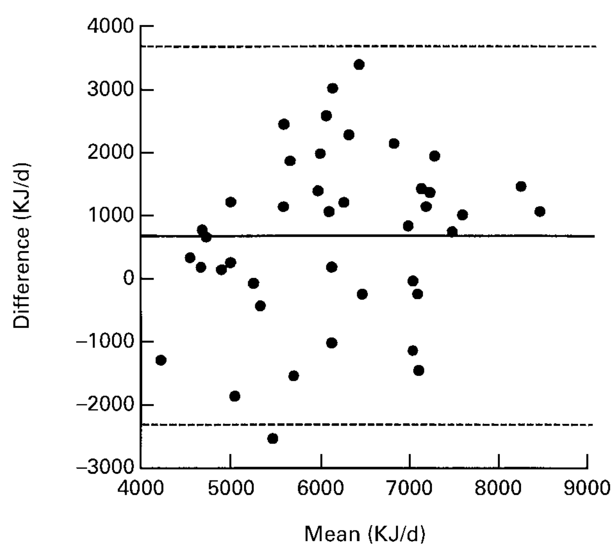

Fig. 1. Plot of individual errors in energy intake assessment (kJ/d) *v.* mean of energy intake and total energy expenditure. For details of subjects and procedures, see Table 1 and p. 602. Solid horizontal line (—) represents the bias, broken horizontal lines (- - -) the limits of agreement between the two methods.

agreement. Finally, mothers often reported quantities in household measures.

Accuracy of TEE measurements by doubly labelled water is excellent at the level of the group, but larger errors have been observed in individuals (Speakman, 1998). Error in the 'reference' method in the present study must therefore have contributed to the disagreement with EI estimates to some extent. However, we attempted to minimise error by following a standard, validated, protocol for the doubly-labelled water method (Schoeller *et al.* 1986). The study was based on the assumption that subjects were in energy balance. While our data on weight change over the course of the study supported this assumption, weight change is a crude index of energy balance and might have contributed to differences between methods. Imprecision would also contribute to differences between the methods, particularly imprecision in the dietary assessment. With 3 d of intake assessment and a mean within-subject CV of EI of 20 %, mean standard error of the estimate of EI in the present study was 11·5 % (Balogh *et al.* 1971). A longer period of dietary assessment would have improved precision but at the expense of subject compliance, obviating the practical advantages of the dietary assessment method chosen. The period of 3 d dietary assessment was chosen largely on practical grounds as in similar studies. An alternative approach to considering differences between the two methods is described by Davies *et al.* (1994). Applying this approach (with an imprecision of about 5 % for TEE and 12 % for EI) to the present study, suggested that 95 % of the individual differences should fall within 26 % of the mean value. In fact, for fourteen of the forty-one subjects (34 %) the difference was greater than 26 %.

For many applications, the degree of disagreement between the two methods in the present study would be more serious than the bias. The limits of agreement were sufficiently wide in the present study that we can conclude there was no agreement between the methods in individuals and that biologically or clinically significant errors in the assessment of EI by multiple pass 24 h recall are to be expected (Fig. 1). It is important to note that this conclusion is not specific to the present study, or to the method of dietary assessment used. Despite slightly better concordance between methods at the group level than observed in the present study, both Johnson *et al.* (1996) and Davies *et al.* (1994) reported similarly wide limits of agreement. Kaskoun *et al.* (1994) reported both a very large bias and wide limits of agreement. In older children, adolescents, and adults, the errors in dietary intake assessment at the individual level are also large and there is the additional problem of under-reporting.

At present it would appear that the methods available for assessing dietary EI in young children are, at best, able to produce unbiased estimates only for groups. When the amount of effort involved in obtaining such estimates is considered, particularly for weighed records, the techniques available offer a relatively poor return, inadequate for many applications. The imprecision of dietary assessment methods provides a further problem, adding to the uncertainty in the estimates obtained.

In summary, the inaccuracy of the available dietary assessment methods presents a serious barrier to measurement of individual EI. For clinical applications with individual patients all available methods may be inadequate. For applications involving group estimates the weighed intake might be adequate but is more demanding than simpler methods based on recalls and food-frequency questionnaires. For studies on causes and/or determinants of energy balance, the existing methods of assessing intake are limited. Many studies of energy balance should therefore consider concentration on more measurable components of energy balance such as TEE (Reilly *et al.* 1998, 1999). Alternatively, studying proxy measures (such as body weight or body composition) and/or the determinants of these might be a more useful means of making inferences in relation to causes of energy imbalance (Wells, 1998).

Acknowledgements

The SPARKLE Project is funded by Sport Aiding Medical Research for Kids. We are grateful to the families and nurseries involved for their enthusiastic co-operation.

References

Balogh M, Kahn HA & Medalie JH (1971) Random repeat 24 hour dietary recalls. *American Journal of Clinical Nutrition* **24**, 304–310.

Bandini LG, Schoeller DA, Cyr HN & Dietz WH (1990) Validity of reported energy intake in obese and non obese adolescents. *American Journal of Clinical Nutrition* **52**, 421–425.

Bland JM & Altman DG (1986) Statistical methods for assessing agreement between two measures of clinical measurement. *Lancet* **i**, 307–310.

Champagne CM, Delany JP, Harsha DW & Bray GA (1996) Under-reporting of energy intake in biracial children is verified by doubly-labelled water. *Journal of the American Dietetic Association* **96**, 707–709.

Cole TJ, Freeman JV & Preece MA (1995) Body mass index reference curves for the UK, 1990. *Archives of Disease in Childhood* **73**, 25–29.

Davies PSW, Coward WA, Gregory J, White A & Mills A (1994) Total energy expenditure and energy intake in the pre-school child: a comparison. *British Journal of Nutrition* **72**, 13–20.

General Register Office Scotland (1993) *1991 Census Scotland, Housing and Availability of Cars*, Edinburgh: H. M. Stationery Office.

Gregory JR, Collins DL, Davies PSW, Hughes JM & Clarke PC (1995) *National Diet and Nutrition Survey: Children Aged 1·5–4·5 Years*, vol. 1. London: H. M. Stationery Office.

Guenther PM, Cleveland LE, Ingwersen LA & Berline M (1998) Questionnaire design and data collection procedures. In *Design and Operation: The Continuing Survey of Food Intakes by Individuals and the Diet and Health Knowledge Survey 1994–1996*, chapter 4, pp. 42–63. *U.S. Department of Agriculture, Agriculture Research Service Nationwide Food Surveys Report* no. 96-1. Beltsville, MD: United States Department of Agriculture.

Jackson DM, Grant S, Paton JY & Reilly JJ (2000) Factors affecting physical activity in Scottish pre-school children. *International Journal of Obesity* **24**, S512.

Johnson RK, Driscoll P & Goran MI (1996) Comparison of multiple pass 24-hour recall estimates of energy intake with total energy expenditure estimated by the doubly-labelled water method in young children. *Journal of the American Dietetic Association* **96**, 1140–1144.

Kaskoun MC, Johnson RK & Goran MI (1994) Comparison of energy intake by semiquantitative food frequency questionnaire with total energy expenditure by the doubly-labelled water method in young children. *American Journal of Clinical Nutrition* **60**, 43–47.

Kroke A, Klipstein-Grobusch K, Voss S, Moseneder J, Thielecke F, Noack R & Boeing H (1999) Validation of a self administered food frequency questionnaire administered in the EPIC Study. *American Journal of Clinical Nutrition* **70**, 439–447.

Kuzawa CW (1998) Adipose tissue in human infancy and childhood: an evolutionary perspective. *Yearbook of Physical Anthropology* **41**, 177–209.

Livingstone MBE, Prentice AM, Coward WA, Strain JJ, Black AE, Davies PSW, Steward CM, McKenna PG & Whitehead RG (1992) Validation of estimates of energy intake by weighed dietary record and diet history in children and adolescents. *American Journal of Clinical Nutrition* **56**, 29–35.

Livingstone MBE, Prentice AM, Strain JJ, Coward WA, Black AE, Barker ME, McKenna PG & Whitehead RG (1990) Accuracy of weighed dietary records in studies of diet and health. *British Medical Journal* **300**, 708–712.

Livingstone MBE & Robson PJ (2000) Measurement of dietary intake in children. *Proceedings of the Nutrition Society* **59**, 279–293.

McMillan DC, Preston T & Taggart DP (1989) Analysis of oxygen-18 enrichment in biological fluids by continuous flow isotope ratio mass spectrometry. *Biomedical and Environmental Mass Spectrometry* **18**, 543–546.

Maffeis C, Schutz Y, Zaffanello M, Piccoli R & Pinelli L (1994) Elevated energy expenditure and reduced energy intake in obese prepubertal children: paradox or poor dietary reliability in obesity? *Journal of Pediatrics* **124**, 348–354.

Ministry of Agriculture Fisheries and Food (1993) *Food Portion Sizes*, 2nd ed. London: H. M. Stationery Office.

Nelson M, Atkinson M & Derbyshire S (1996) Food photography 2. Food photographs for estimating portion size and nutrient content of meals. *British Journal of Nutrition* **76**, 31–49.

Nelson M, Atkinson M & Meyer J (1997) *A Photographic Atlas of Food Portion Sizes*. London: MAFF Publications.

Prosser SJ & Scrimgeour CM (1995) High precision determination of deuterium and hydrogen in H_2 and H_2O by continuous flow mass spectrometry. *Analytical Chemistry* **67**, 1992–1997.

Reilly JJ, Ralston JM, Paton JY, Edwards CA, Weaver LT, Wilkinson J & Evans TJ (1999) Energy balance during acute respiratory exacerbations in children with cystic fibrosis. *European Respiratory Journal* **13**, 804–809.

Reilly JJ, Ventham JC, Ralston JM, Donaldson M & Gibson BES (1998) Reduced energy expenditure in pre-obese children treated for acute lymphoblastic leukemia. *Pediatric Research* **44**, 557–562.

Sawaya AL, Tucker K & Tsay R (1996) Evaluation of four methods of determining energy intake in young and older women. *American Journal of Clinical Nutrition* **63**, 491–499.

Schoeller DA, Ravussin E, Schutz Y, Acheson KJ, Baertschi P & Jequier E (1986) Energy expenditure by doubly-labelled water: validation in humans and proposed calculation. *American Journal of Physiology* **250**, R823–R830.

Speakman JR (1998) The history and theory of the doubly labeled water technique. *American Journal of Clinical Nutrition* **68**, Suppl., 932S–938S.

Scrimgeour CM, Rollo MM, Mudambo S, Handley L & Prosser SJ (1993) A simplified method for deuterium:hydrogen isotope ratio measurements in water samples of biological origin. *Biological Mass Spectrometry* **32**, 383–387.

Wells JCK (1998) Is obesity really due to high energy intake or low energy expenditure? *International Journal of Obesity* **22**, 1139–1140.