Recent Advances from Application of Doubly Labeled Water to Measurement of Human Energy Expenditure$^1,2$

Dale A. Schoeller

Department of Nutritional Sciences, University of Wisconsin, Madison, WI 53706

ABSTRACT  The doubly labeled water (DLW) method was developed 50 years ago, but nearly 40 years passed before it became a major tool for human nutrition research. Its use has grown dramatically, however, since the first human applications. The DLW method is now the preferred method for determining energy requirements of healthy and clinical populations. The method has been applied extensively to the study of the growing problem of obesity in order to determine the role of energy expenditure and physical activity in weight control. Finally, DLW has provided a new means of validating methods for assessing dietary intake. The accuracy and noninvasive nature of the DLW method makes it ideal for the study of human energy metabolism. J. Nutr. 129: 1765–1768, 1999.

KEY WORDS:  • energy • metabolism • calorimetry • requirements • isotope

In 1985, the FAO/WHO/UNU Expert Committee on Energy Requirements (1985) recommended that energy requirements be based on assessments of energy expenditure rather than energy intake. The assessment of energy expenditure, however, was a daunting task because of the difficulty associated with the measurement of physical activity (PA), the most variable component of total daily energy expenditure (TEE). Earlier, we documented the first application of a stable isotope technique that had the potential for accurate, objective measurement of TEE in free-living individuals (Schoeller and van Santen 1982). This technique is the doubly labeled water method (DLW), which was developed by Nathan Lifson in the early 1950s (Lifson and McClintock 1966). The doubly labeled water method, in essence, turns the body into its own metabolic recorder. After a loading of water labeled with deuterium and $^{18}$O, the deuterium washes out of the body as carbon dioxide, but simply collect periodic urine specimens to determine the amount of stable isotope left in the body. Subjects are, therefore, free to live at home and engage in normal activities of daily life.

During the 10 years after its initial use in humans, the DLW method was subjected to extensive validation (Schoeller 1988) as well as a lively debate over the number of specimens and exact assumptions to be used in the model (Coward and Prentice 1985, Schoeller 1984). These discussions have been largely resolved (Coward et al. 1994, Roberts et al. 1995). The method is now considered the gold standard for the measurement of TEE and it has become a common tool. The method is also used widely in comparative biology, but that is beyond the scope of this review.

In recent years, DLW research has grown dramatically. Some of this growth results from advances in isotope ratio instrumentation (Coplen and Harper 1994, Schoeller and Luke 1997). These improvements have made it possible to apply the DLW to studies involving hundreds of participants rather than the 5- to 20-participant studies that characterized earlier applications. The DLW method, however, is still very sensitive to the precision of the isotopic analyses (Speakman 1995). The number of participants in a given study is not a guarantee that the method is or is not sufficiently powered for correlation analysis because precision of the method varies among laboratories (Roberts et al. 1995, Schoeller et al. 1995).

The application of the DLW method has expanded human nutrition knowledge. Most of the recent advances can be placed in one of three categories as follows: the assessment of energy requirements in health and disease, the study of the etiology of obesity, and the validation of tools for the assessment of dietary intake, an application that takes the FAO/WHO/UNU recommendation to its extreme.

Assessment of Energy Requirements. Black et al. (1996) compiled results from 574 DLW studies that had been published over the previous 12 years. These represented only a third of the DLW measurements that had been published at the time because selection criteria excluded data that did not include a resting metabolic rate under near basal conditions, data from developing countries and data that did not represent "normal" free-living conditions. These data indicated that the recommended energy intake for the United States (NRA 1989) for infants and children were high, whereas those for adolescents and adults were usually low. In the three years since the compilation was published, the data available in the literature have tripled. Care should be exercised in reviewing these data, however, because many of the DLW results are included in more than one data analysis. Among the recent publications are four that are notable for their size. Davis (1998) recently reviewed the growing body of data on energy requirements in infants ($n > 400$) and reported that current recommendations for energy intake in 1 of life remain $\sim 10\%$ higher than indicated by DLW measured expenditure plus estimated increase in body energy stores. A study of 8 to 12-year-old girls ($n = 109$) by Bandini et al. (1997) suggests that TEE is $100-200$ kcal/d greater than the earlier estimate of Black et al. (1996), whereas a study of 46 adolescents (Bratteby et al. 1998) supports the values accumulated by Black et al. (1996), but indicates that the gender difference may be smaller than...
are lost from the body each day. The product of total body water (TBW) due to the training sessions themselves. Only Treuth et al. (1998) reported that a weight training program in 7- to 10-y-old girls resulted in no increase in TEE despite a 19% increase in strength. Thus most studies continue to find that programmed changes in physical activity do alter TEE in the expected direction (Westerterp 1998).

An area of study that is still significantly limited is that of energy requirements in developing countries (Coward 1998). Only two recent papers have appeared. Jiang et al. (1998) studied 41 infants at 4 and 6 mo of age. They reported TEE values that are not different from those in developed countries, but do find evidence of a decline in growth rate at 6 mo, suggesting that energy intake at the time of weaning may be less than optimal. Wren et al. (1997) compared the expenditures of small-for-age children in Guatemala with those who were normal weight-for-age and a second control group of normal weight-for-age children in the United States. Although TEE was decreased in the small weight-for-age children, it was not reduced after adjusting for fat-free mass, indicating that factors leading to the growth stunting did not alter cellular determinant of energy expenditure.

Assessment of Energy Requirements of Clinical Populations. The DLW method provides a unique tool for assessing energy requirements for clinical populations, many of which suffer weight loss or failure to thrive. Infants with cyanotic congenital heart disease or broncopulmonary disease fail to grow at reference rates. Leitch et al. (1998) and de Meeg (1997), used the DLW method to document increases in TEE of 30 and 15%, respectively, in 1- to 2-mo-old infants. These increase the risk of undernutrition and require that special attention be paid to maintaining energy intake.

Among adults, however, a different relationship is being noted. Chronic diseases such as chronic obstructive pulmonary disease or heart failure have often been classified as hypermetabolic diseases because of increases in resting metabolic rate (RMR). DLW measurements, however, have demonstrated that TEE either is not different from controls or is less due to decreases in physical activity (Baarends et al. 1997, McCallan et al. 1995). Thus, losses of weight or muscle mass in these individuals is not usually caused by increased energy needs, and greater attention should be focused on maintaining normal levels of dietary intake. Investigators interested in determining energy requirements of clinical populations are finding that individual levels of TEE are quite variable; they are focusing on identifying factors other than simple anthropometrics that may help predict an individual’s energy requirements (Johnson et al. 1997, Toth et al. 1997).

Obesity Research. Before the DLW method became available for human research, many believed that obese individuals had lower absolute energy requirements than their lean controls. It is now clear that TEE tends to increase with weight and that the obese as a group expend more energy than lean controls (Prentice et al. 1996). The DLW method has also made it possible to quantitate the energy costs of physical activity in free-living individuals, thus helping to focus research attention on the role of physical activity in the development, maintenance and treatment of obesity (Schoeller 1998). The DLW method has also been used to demonstrate that the interindividual differences in energy expenditure associated with physical activity are larger than those for RMR or the thermic effect of meals (TEM) and thus more likely to lead to major deficits or surfeits in energy balance. Care is required in the interpretation of such findings, however, because the energy expended in physical activity is calculated by taking the difference between TEE and RMR (with or without

\[ r_{\text{H}_2\text{O}} = \text{TBW} \cdot k_H \]
\[ r_{\text{H}_2\text{O}} + 2r_{\text{CO}_2} = \text{TBW} \cdot k_O \]

\[ r_{\text{CO}_2} = \text{TBW}(k_O k_H)/2 \]

Each day, 3-9 mo infants lose about 3% of their TBW, whereas 1- to 2-y-old children and 16- to 18-y-old children lose about 2% of their TBW due to the isotope fractionation. Corrections are required for differences in isotope dilution spaces and isotope fractionation.

FIGURE 1 After a loading dose of $^{2}{\text{H}}_2$, $^{18}{\text{O}}$, 5–20% of the tracers are lost from the body each day. The product of total body water (TBW) and the deuterium elimination rate $k_D$ is a measure of the rate of water output $v_{\text{H}_2\text{O}}$. The product of TBW and the oxygen elimination rate $k_O$ is a measure of the sum of water output and two times carbon dioxide output $v_{\text{CO}_2}$. Taking the difference between these two equations and rearranging terms yields the equations for calculation of $v_{\text{CO}_2}$. Small corrections are required for differences in isotope dilution spaces and isotope fractionation.
correction for TEM). TEE has an uncertainty of 5–10% as a result of random measurement errors; when the difference from RMR is taken, this uncertainty is transferred to the physical activity term. This inflates the apparent between-individual variability in the energy expended in physical activity (Schoeller and Hnilicka 1996). Other difficulties in the interpretation of TEE data also occur because there is controversy regarding the means of comparing TEE among individuals of different sizes (Dietz 1998). In planning future studies, there is considerable value in the inclusion of a measurement of RMR so that expenditure can be partitioned between resting and physical activity. Inclusion of TEM is also of some value, but in many instances it is necessary to measure TEM only in a subset of the study population or use literature values because the individual measurement error for TEM is comparable to the interindividual variation.

Considerable research efforts have been directed recently at the study of ethnic differences in energy metabolism to determine whether these may explain the differences in the prevalence of obesity among ethnic groups. During the last few years alone, studies have been performed in three ethnic groups with high prevalences of obesity. Rush et al. (1999) compared adult New Zealanders of European and Polynesian origins. TEE was noted to correlate with body weight in all but the obese subjects of Polynesian origin, suggesting a difference in the factors that control TEE in this group. They also noted that the nonobese Polynesians expended 50% more energy in physical activity than the nonobese Europeans. Sable et al. (1997) compared TEE and RMR between 5-y-old Pima Indians and Caucasians. They could not find differences in energy expenditure after adjusting for body size that would explain the already greater weight and fatness in the Pima Indian children. Three studies comparing African-American and Caucasian children have also been recently performed. Sun et al. (1998) studied children between 5 and 10 y of age and found that self-reported dietary intake was unreliable for the study of the role of energy intake in the etiology of obesity among 10-y-old children. Underreporting ranged from 17 to 30% for groups of obese or nonobese children of different ethnicity. Ambler et al. (1998) reported that self-report was unreliable in determining the effect of physical training on energy intake in adolescents by comparing self-reported energy intake against TEE in a group during training and a control group. Reported energy intakes were comparable in both groups even through expenditure was greater in the group receiving physical training and weight did not change. Finally, Martin et al. (1996) compared self-reported energy intake with TEE in middle-aged women during y 2 of a dietary intervention study. These nonobese women underreported energy intake by 20%, raising further concern about the general accuracy of the intake questionnaires.

Johnson et al. (1998) attempted to validate a multipass 24-h recall method for determining energy intake. They found that underreporting correlated with women being overweight or having poor reading and spelling skills. Finally, Black et al. (1997) compared the utility of three methods of defining underreporting of dietary intake. The found that urinary nitrogen could identify the worst underreporters, but that it was less sensitive than TEE from DLW. They also found that the use of the cut-offs based on calculated RMR (Goldberg et al. 1991) helped to identify major underreporting, but that these too had limitations.

The application of the DLW method has come full circle. Early studies used the ability to measure expenditure to replace energy intake as a proxy measure of energy requirements. Now, expenditure from DLW has become a benchmark for validating methods to assess habitual energy intake. Other applications have also proliferated to such a degree that DLW has become a victim of its own success. The demand for $^{18}$O has exceeded world-wide production for the past two years, creating a world-wide shortage of the tracer. This is currently curtailting further research, and increased production is required to meet this demand.

**LITERATURE CITED**


Validating Measurement of Energy Intake. Before the development of the DLW method, self-reported habitual energy intake was often used as a proxy measure of TEE. Now it has become clear that TEE obtained from DLW is accurate, and TEE is being used as a biomarker of energy intake. In so doing, significant underreporting of dietary energy intake has been documented (Schoeller 1995). Because of the importance of self-reported dietary intake in the study of the relationships between diet and health, considerable attention has been directed toward characterizing and understanding the roots of the underreporting.

Three studies characterized the degree of underreporting of dietary energy as part of studies in which the use of self-reported dietary intake is common. Champagne et al. (1998) found that self-reported dietary energy intake was unreliable for the study of the role of energy intake in the etiology of obesity among 10-y-old children. Underreporting ranged from 17 to 30% for groups of obese or nonobese children of different ethnicity. Ambler et al. (1998) reported that self-report was unreliable in determining the effect of physical training on energy intake in adolescents by comparing self-reported energy intake against TEE in a group during training and a control group. Reported energy intakes were comparable in both groups even through expenditure was greater in the group receiving physical training and weight did not change. Finally, Martin et al. (1996) compared self-reported energy intake with TEE in middle-aged women during y 2 of a dietary intervention study. These nonobese women underreported energy intake by 20%, raising further concern about the general accuracy of the intake questionnaires.

Johnson et al. (1998) attempted to validate a multipass 24-h recall method for determining energy intake. They found that underreporting correlated with women being overweight or having poor reading and spelling skills. Finally, Black et al. (1997) compared the utility of three methods of defining underreporting of dietary intake. The found that urinary nitrogen could identify the worst underreporters, but that it was less sensitive than TEE from DLW. They also found that the use of the cut-offs based on calculated RMR (Goldberg et al. 1991) helped to identify major underreporting, but that these too had limitations.

The application of the DLW method has come full circle. Early studies used the ability to measure expenditure to replace energy intake as a proxy measure of energy requirements. Now, expenditure from DLW has become a benchmark for validating methods to assess habitual energy intake. Other applications have also proliferated to such a degree that DLW has become a victim of its own success. The demand for $^{18}$O has exceeded world-wide production for the past two years, creating a world-wide shortage of the tracer. This is currently curtailing further research, and increased production is required to meet this demand.


