

## ORIGINAL COMMUNICATION

# Precision of the doubly labeled water method in a large-scale application: evaluation of a streamlined-dosing protocol in the Observing Protein and Energy Nutrition (OPEN) study

J Trabulsi<sup>1</sup>, RP Troiano<sup>2</sup>, AF Subar<sup>2</sup>, C Sharbaugh<sup>3</sup>, V Kipnis<sup>4</sup>, A Schatzkin<sup>5</sup> and DA Schoeller<sup>1\*</sup>

<sup>1</sup>Department of Nutritional Sciences, University of Wisconsin, Madison, WI, USA; <sup>2</sup>National Cancer Institute, Division of Cancer Control and Population Sciences, Applied Research Program, Bethesda, MD, USA; <sup>3</sup>Westat, Rockville, MD, USA; <sup>4</sup>National Cancer Institute, Division of Cancer Prevention, Biometry Research Group, Bethesda, MD, USA; and <sup>5</sup>National Cancer Institute, Division of Cancer Epidemiology and Genetics, Nutritional Epidemiology Branch, Bethesda, MD, USA

**Objective:** To evaluate whether the doubly labeled water (DLW) method is precise under conditions required for a large-scale evaluation of dietary intake instruments.

**Design:** Energy expenditure was measured in 484 subjects (main study). Subjects received one of five different weight DLW dose bottles prepared in advance of the study. A repeat energy expenditure measure was obtained in a subset of 24 subjects (substudy). DLW measures of energy expenditure were performed over a 2-week interval with urine collection at the beginning and end.

**Setting:** Free-living environment with three clinic visits in the Maryland suburban area of Washington, DC.

**Subjects:** A total of 484 subjects (261 men and 223 women) aged 40–69 y, 24 of whom (13 men and 11 women) participated in a substudy in which DLW was administered a second time.

**Results:** The coefficient of variation of the DLW energy expenditure measurement was 5.1%. This included a 2.9% analytical and a 4.2% physiologic variation. Based on observed initial isotopic enrichment, the preweighed dosages were optimal in 70% of the main study subjects, and 9% received a dose that was less than optimal. Only six subjects (1%) were excluded because the final isotopic enrichment was too low to conduct precise measurement.

**Conclusions:** Use of preweighed DLW dosages did not compromise the precision of the DLW method. The DLW method is a reliable measure of energy expenditure for large-scale evaluations of dietary intake instruments.

**Sponsorship:** The National Cancer Institute.

*European Journal of Clinical Nutrition* (2003) **57**, 1370–1377. doi:10.1038/sj.ejcn.1601698

**Keywords:** doubly labeled water; total energy expenditure; stable isotopes; reproducibility

### Introduction

The doubly labeled water (DLW) method is a technique used to measure the average daily energy expenditure of free-

living humans. The method involves the administration of dose water containing enriched quantities of the stable isotopes deuterium and <sup>18</sup>O, which equilibrate with body water. Over time, the deuterium is eliminated from the body in the form of water and the <sup>18</sup>O is eliminated as both water and carbon dioxide. The difference in the elimination rates of the two isotopes, after adjusting for isotopic fractionation, is a measure of CO<sub>2</sub> production rate, which in turn is used to calculate total energy expenditure (TEE). The DLW method has been validated against near continuous indirect calorimetry and it has been found to be accurate to 1–2% with a coefficient of variation of 2–12% (Schoeller, 1988).

\*Correspondence: DA Schoeller, Department of Nutritional Sciences, University of Wisconsin-Madison, 1415 Linden Drive, Madison, WI 53715-1571, USA.

E-mail: dschoell@nutrisci.wisc.edu

Guarantor: DA Schoeller.

Contributors: RPT, AFS, CS, VK, AS, and DAS designed the study. JT and CS collected the data. JT and DAS conducted data analysis. JT, RPT, AFS, and DAS wrote the manuscript. CS, VK, and AS critically revised the manuscript for intellectual content.

More recently, the DLW method has been used as a biomarker of usual energy intake. Based on the First Law of Thermodynamics, metabolizable energy intake may be equated with energy expenditure in weight stable adults. As such, the DLW method may be used as a standard with which to validate reported energy intake. To be useful as a biomarker of usual energy intake, the DLW method must be both accurate and precise in measuring TEE.

There are a number of factors that may affect the accuracy and precision of the DLW method. Variation in the measurement of TEE is comprised of both analytic and physiologic error. Analytical error is determined by factors the investigator may manipulate such as the dose of DLW provided, the duration of the metabolic period (Schoeller, 1983; Cole & Coward, 1992), sample processing (Ritz *et al*, 1994), and measurement error during mass spectrometry (Schoeller *et al*, 1995; Roberts *et al*, 1995). This type of error may be assessed by repeating the isotopic analyses and comparing the results of the two analyses. Physiologic error is the intraindividual variation in TEE. Factors that may affect this error include changes in physical activity, health, and menstrual cycle status. In most study paradigms, physiologic error also includes error because of variation in isotopic fractionation (Schoeller *et al*, 1986) and isotopic background (Horvitz & Schoeller, 2001). Physiologic error may be quantified by repeated measurement of TEE in an individual. Each repeat measure of TEE, however, includes both analytical error and physiologic error, and the sum of these errors (analytical + physiologic) comprises total error (Schoeller & Hnilicka, 1996). Precision of the DLW method has been investigated, yet to date, only one small study ( $n=6$  females) of physiologic error also included a measure of analytical error (Schoeller & Hnilicka, 1996). The present study further investigates the precision of the DLW method in a substantially larger ( $n=24$ ) and more heterogeneous (males and females) sample of adults. The measurement of analytical error in this study allows for the determination of true physiologic error in the DLW method. An intended application of this error quantification is for mathematical modeling of person-specific biases in dietary assessment methodologies.

The amount of dose water administered to each subject is one factor that can alter the precision of the DLW method. The administration of too little isotope may result in a low enrichment of body fluid at the conclusion of the study, which leads to increased measurement error (Schoeller, 1983). To ensure the provision of a sufficient dose of isotope, the DLW dose is typically calculated for each subject individually based on body weight. This is a time intensive process as each dose must be filtered and weighed to the nearest 0.3% or better. We have recently administered DLW on a large-scale in the Observing Protein and Energy Nutrition (OPEN) study, which involved 484 subjects. To accommodate the time constraints of this large DLW study, in which several subjects were dosed each hour, we developed a streamlined-dosing protocol that involved the

preparation of five different dose sizes of DLW in advance of the study. Such a method leaves open the possibility that an insufficient quantity of isotope could be administered to some of the participants, leading to a reduction in the precision of the energy expenditure measurement.

The aim of this analysis was to determine if the use of a streamlined-dosing procedure in a large-scale study affected the precision of the DLW method. To do so, we evaluated both dose adequacy and the reproducibility of energy expenditure measurement. For dose adequacy, we compared the urinary  $^{18}\text{O}$  enrichment at 4 h postdose to an optimal value of 98 permil  $\pm 10\%$ , and the end point urinary  $^{18}\text{O}$  enrichment to a quality control cutoff of 8 permil<sup>1</sup> (Schoeller *et al*, 1983). For reproducibility, a repeat measure of energy expenditure was obtained in a subset of study subjects.

## Subjects and methods

### Subjects

**Study subjects.** Subjects ( $n=484$ ) were participants of the National Cancer Institute's OPEN study (Subar *et al*, submitted). The OPEN study was approved by the National Cancer Institute Special Studies Institutional Review Board for Human Subjects Research. Subject recruitment and data collection for the OPEN study began in September 2000 and concluded in March 2001. Men and women of ages 40–69 y were recruited from the Maryland suburban area of Washington, DC. Households to be contacted were selected at random from a purchased list of households with telephone numbers listed in the white pages. To be eligible for the study participants had to be a resident in the target geographical area, able to read English, not pregnant or planning to become pregnant in the next year, not on a weight loss diet, not having formal training in nutrition. Participants were excluded if they had the following medical conditions: insulin-dependent diabetes mellitus, congestive heart failure, kidney failure requiring dialysis, fluid retention, malabsorption, hemophilia, or any condition requiring supplemental oxygen. Subjects who did not complete the DLW protocol ( $n=1$ ) or whose DLW samples were not complete ( $n=2$ ) were excluded from this analysis, resulting in a sample size of 481 subjects.

**Substudy subjects.** At the conclusion of the first visit, a convenience sample of 25 OPEN study participants (roughly half men and half women) were offered the opportunity to enroll in the substudy. The purpose of the substudy was to determine the precision of energy expenditure measurement

<sup>1</sup>Permil is a one part per thousand (‰) change in the ratio of the heavy to light isotope—that is  $(R_{\text{sample}}/R_{\text{standard}}-1) \times 1000$ , where  $R$  is the molar ratio of the heavy ( $^2\text{H}$ ) to light ( $^1\text{H}$ ) isotope abundance. A 1 permil change in  $^{18}\text{O}$  or  $^2\text{H}$  enrichment corresponds to the addition of 2 mol of  $^{18}\text{O}$ -hydride or 0.16 mol  $^2\text{H}$  oxide to  $10^6$  mol of water, respectively.

in free-living subjects using a test, re-test design with DLW dosing 2 weeks apart. The sample size was calculated so that the estimated CV for the DLW method did not differ from the true average CV by more than 3% with 95% confidence. Assuming a true average CV of 9% for the DLW method, 25 subjects were therefore required. One subject was excluded from the final substudy data set because the urine sample isotopic analyses demonstrated failure to achieve equilibration at 4 h postdose. This reduced the sample size of the substudy to 24 subjects.

## Methods

**Study protocol.** Study protocols were carried out at the Westat research facility in Rockville, Maryland. At the first visit: a predose spot urine specimen was collected, height and weight were measured, and a dose of DLW was given. Three postdose spot urines were collected at approximately 2, 3, and 4 h after the dose. In subjects  $\geq 60$  y of age, a blood sample was taken at 3 h postdose (Blanc *et al*, 2002). Between 1 and 3 h after the DLW dose, subjects were allowed to consume up to 600 ml of a liquid meal replacement (Boost, Mead Johnson Nutritionals, Evansville, IN, USA), coffee, tea, juice, or water. The type, quantity, and time of beverage ingestion were recorded to correct future total body water (TBW) calculations (Schoeller, 1996). Liquid consumption was not allowed between 3 and 4 h postdose. Approximately 14 days ( $\bar{x} = 13.8 \pm 0.7$  days) after the first visit, subjects returned for their second visit. At this visit, weight was measured and two end point spot urines were collected (1 h apart).

**Substudy protocol.** At the conclusion of their second visit, substudy subjects were given another dose of DLW (dose 2) and three postdose spot urines were collected. Postdose and end point spot urines were collected as outlined above, with the dose 2 end point spot urines collected approximately 14 days after dose 2. The final end point specimen from dose 1 was used as the baseline specimen for the second TBW measurement (dose 2). The initial baseline specimen from dose 1 was used for elimination rate calculations.

**DLW protocol.** Owing to the large number of subjects, five different dose weights of DLW were prepared prior to the start of the study (Table 1) to eliminate the delays involved in weighing the dose water during clinic time. The doses were calculated to provide approximately 2.0 g 10%  $\text{H}_2^{18}\text{O}$  (Cambridge Isotope Laboratories, Andover, MA, USA) and 0.14 g 99.9%  $^2\text{H}_2\text{O}$  (Cambridge Isotope Laboratories, Andover, MA, USA) per kilogram of estimated TBW. The dose weights were based on TBW for the median of gender-dependent range of body weight and TBW calculated from adult Bioelectrical Impedance Analysis (BIA) data from the third National Health and Nutrition Examination Survey (unpublished data). The doses were calculated to achieve an average urinary enrichment of 98 and 600 permil, for  $^{18}\text{O}$  and deuterium, respectively.

**Table 1** Gender- and weight-specific DLW dose categories

	DLW dose				
	A	B	C	D	E
<i>Male</i>					
Body weight (kg)	NA	<60.0	60.1–70.0	70.1–95.0	>95.1
<i>Female</i>					
Body weight (kg)	<55.0	55.1–75.0	75.1–110	>110.1	NA
DLW (g)/bottle	58	67	79	95	114

DLW: doubly labeled water.

NA: not applicable.

**Isotopic analysis.** Urine samples were mixed with 200 mg dry carbon black (Fischer Scientific Chemical Co., Itaska, IL, USA) and filtered (0.45  $\mu\text{m}$ , Cameo 25Gas, Osmonics, Inc., Minnetonka, MN, USA). Plasma samples were untreated for  $^{18}\text{O}$  analysis and centrifuged (4°C, 1-h at 10 000 rpm) on regenerated cellulose filters (YM-50, Centricon, Bedford, MA, USA) to extract water for deuterium analysis. A 1.4% exchange of hydrogen isotopes was observed on these filters, independent of the volume filtered, and all plasma analyses were subsequently corrected for this exchange.

For  $^{18}\text{O}$  analysis, 1 ml of the cleaned urine sample was transferred to a Vacutainer (3 ml, Becton Dickinson and Co., Franklin Lakes, NJ, USA) containing 1 ml  $\text{CO}_2$  at standard temperature and pressure and equilibrated for at least 24 h at 30°C. Approximately, 15  $\mu\text{l}$  of equilibrated  $\text{CO}_2$  was chromatographed to separate it from air and introduced into a continuous flow inlet system (Luke & Schoeller, 1997) using helium as a carrier gas and analyzed on a Delta-S isotope ratio mass spectrometer (Finnigan MAT, San Jose, CA, USA). Analyses were performed in duplicate. The standard deviations for duplicate oxygen-18 analyses were 0.17 and 0.4 permil at low and high abundances, respectively. For deuterium analysis, 1 ml of the cleaned urine sample was transferred into a vial (Target I-D™ Vials, National Scientific Company, Lawrenceville, GA, USA) and sealed. An auto-sampler injected 0.8  $\mu\text{l}$  of sample into a quartz tube containing chromium metal powder (Mesh size 100 and finer, Fisher Scientific Chemical Co., Itaska, IL, USA) to reduce water to hydrogen gas (Schoeller *et al*, 2000). Samples were injected in duplicate. The analyses were corrected for  $\text{H}_3^+$  as well as memory from the reduction system. The standard deviations for duplicate deuterium analyses were 0.6 and 1.0 permil at low and high abundances, respectively.

**TBW calculation.** Isotope dilution spaces were determined using the urinary enrichment above baseline of each isotope at 3- and 4-h postdose. In subjects greater than or equal to 60 y of age, blood samples were taken 3 h postdose and analyzed for isotopic enrichment (Blanc *et al*, 2002). If the enrichment of the urine and blood samples agreed within 2% (approximately 3 s.d.), the urinary value was used in

dilution space calculations. If the 3- and 4-h postdose enrichment differed by more than 2% (main study  $n = 32$ ; substudy  $n = 1$ ), the plasma enrichment was used for dilution calculation. Isotope dilution spaces ( $N$ , kg) were calculated according to Cole and Coward (1992),

$$N = (WA/1000a)(d_a - d_t)/(d_s - d_p) - w,$$

where  $W$  is the grams of water used in the dilution of the dose water,  $A$  is the grams of dose water administered to the subject,  $a$  is the grams of dose water used in the dilution,  $d$  is the permil isotopic abundance of the diluted dose water ( $d_a$ ), the tap water used in the dilution ( $d_t$ ), the postdose urine/blood specimen ( $d_s$ ), the predose urine specimen ( $d_p$ ), and  $w$  is the amount of water (kg) consumed between the DLW dose and the 3 h postdose urine specimen. TBW was calculated from the average of the deuterium and oxygen dilution space divided by 1.041 and 1.007, respectively, to correct for *in vivo* isotopic exchange (Racette et al, 1994).

**Total energy expenditure.** The  $^{18}\text{O}$  and deuterium elimination rates ( $k_o$  and  $k_d$ ) were calculated from the change in the natural logarithm of isotope enrichment as a function of time elapsed after dose administration. The mean enrichment of two postdose urine samples (collected 3- and 4-h after the dose) and two end point urine samples (collected approximately 14 days later, 1 h apart), and the actual number of days elapsed, were used in the elimination rate calculations, except in participants (main study  $n = 32$ ; substudy  $n = 1$ ) whose plasma enrichment was used, as detailed previously, for the postdose enrichment. Carbon dioxide production was calculated according to the equation of Racette et al (1994),

$$r\text{CO}_2 = 0.455\text{TBW}(1.007k_o - 1.041k_d)$$

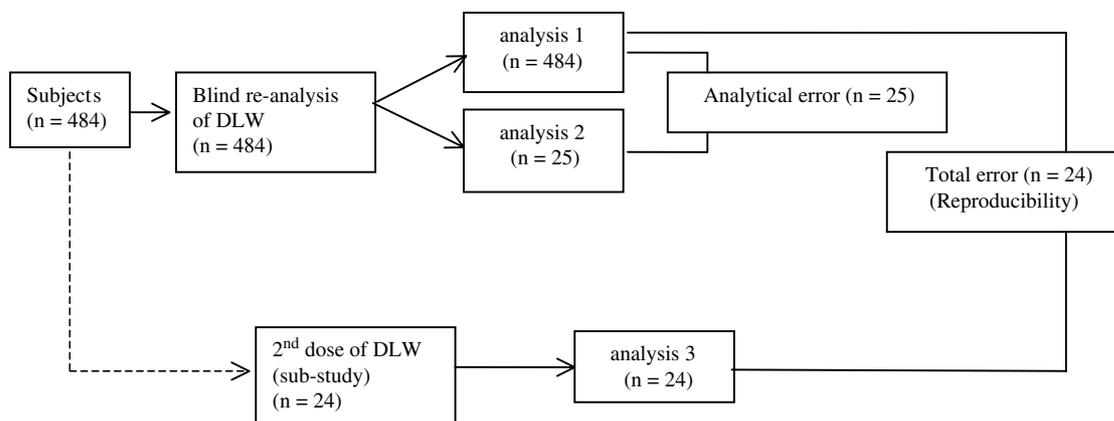
where TBW is the total body water in moles,  $k_o$  and  $k_d$  are the oxygen and deuterium elimination rate in pools/day, respectively. From  $r\text{CO}_2$ , total energy expenditure was calculated using the modified Weir equation (Weir, 1949)

using a respiratory quotient of 0.86 (based on dietary RQ of accurate reporters of energy intake in the OPEN study).

**Error calculations.** Analytical error (Figure 1) was determined by repeating the isotope analyses (analysis 2) in a subset of participants ( $n = 25$ ) from the main study. Analysis 2, conducted in a blinded fashion, was completed a minimum of 2 weeks after analysis 1.

Total error or reproducibility (Figure 1) was determined by re-dosing a subset of participants ( $n = 24$ ) to repeat the DLW method (re-test), as described above.

**Data organization and statistical analyses.** The dose adequacy for individuals was evaluated by comparing the urinary  $^{18}\text{O}$  enrichment at 4-h postdose to an optimal value of 98 permil  $\pm 10\%$  (Schoeller et al, 1995; Horvitz & Schoeller, 2001). Those subjects with enrichment below this cutoff ( $< 88.2$  permil) were considered suboptimally dosed, and those with enrichment above this cutoff ( $> 107.8$  permil) were considered supraoptimally dosed. Subjects with an end point urinary  $^{18}\text{O}$  enrichment of  $< 8$  permil above baseline were considered to have failed quality control guidelines and were removed from further analyses. Statistical analyses were performed using StatView<sup>®</sup> 5.01 (SAS Institute Inc., Cary, NC, USA) and SAS<sup>®</sup> (SAS Institute Inc., Cary, NC, USA). Values are presented as the mean and standard deviation unless otherwise stated. To determine if the physical characteristics differed among those receiving a suboptimal, optimal, or supraoptimal dose of DLW, ANOVA analyses and  $\chi^2$  testing were used for continuous and categorical variables, respectively. To identify predictors of a low end point enrichment, forward step-wise linear regression was performed. A  $P$ -value of 0.05 was considered significant. To quantify the analytical, physiologic and total error, the within-subject coefficient of variation ( $\text{CV}_w$ ) was computed for each subject and the geometric mean of the individual  $\text{CV}_w$  was determined.



**Figure 1** Analytical error was determined from blinded repeat analysis of the same urines from 25 subjects. Total error was determined from separate DLW dosing periods in 24 subjects.

**Results**

**Dose adequacy**

Overall, 70% ( $n = 338$ ) of participants were optimally dosed (Figure 2a). For the 9% receiving a suboptimal dose ( $n = 44$ ), the 4-h postdose urinary  $^{18}\text{O}$  enrichment was only slightly suboptimal ( $84 \pm 5$  permil). For the 21% receiving a supraoptimal dose ( $n = 99$ ), the 4-h postdose enrichment was 114 ( $\pm 7$ ) permil (Figure 2a). Only four of those receiving a suboptimal DLW dose, and two of those receiving an optimal DLW dose, had an end point isotope enrichment that was considered unacceptable ( $< 8$  permil) according to quality control guidelines (Figure 2b). Dosing adequacy differed significantly by age and fat-free mass (FFM), but was not affected by body mass index (BMI) or gender (Table 2). Deuterium turnover ( $k_d$ ) explained much of the variance in

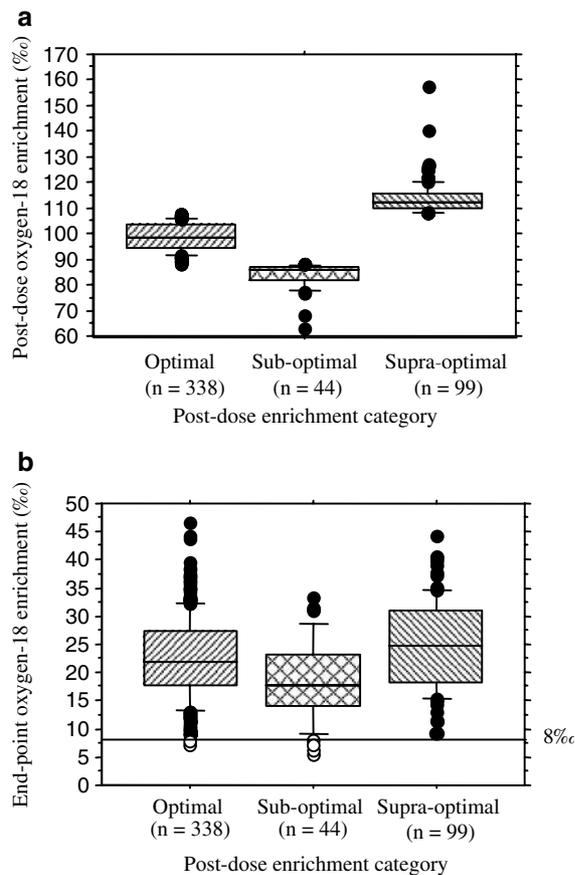
end point enrichment (Table 3); TBW/weight, gender, and weight were small, but significant predictors (Table 3).

**Error in measurements of energy expenditure**

The analytical error in the components of energy expenditure was small. A CV near 1% was found for most variables with the exception of TEE, which had a CV of 2.9% (Table 4). The total error in the components of energy expenditure was slightly larger, ranging from a 1.2% CV for the dilution space ratio (DS ratio) to a 5.1% CV (95%CI 4.2–6.8) for TEE (Table 5).

**Discussion**

The aim of this study was to determine if the DLW method maintained its precision in a large-scale study where a streamlined-dosing protocol was employed. To do so, we evaluated dose adequacy and reproducibility. For dose adequacy, we compared the urinary  $^{18}\text{O}$  enrichment at 4-h postdose to an optimal value of 98 permil  $\pm 10\%$  (optimally dosed), and the end point urinary  $^{18}\text{O}$  enrichment to a quality control cutoff of 8 permil in all main study



**Figure 2** (a) Postdose oxygen-18 enrichment level (permil) in subjects based on postdose enrichment category. Box plot: displays the 10th, 25th, 50th, 75th, and 90th percentiles. Solid circles (●) represent outliers below the 10th and above the 90th percentiles plotted separately. (b) End point oxygen-18 enrichment level (per thousand) based on postdose enrichment category. Open circles (○) represent outliers ( $n = 6$ ) with end point enrichments below 8‰; these subjects failed laboratory quality control standards and were excluded from further analyses.

**Table 2** Physical characteristics of subjects receiving suboptimal, optimal, and supraoptimal doses of DLW

	Suboptimal (n=44)	Optimal (n=338)	Supraoptimal (n=99)
BMI ( $\text{kg}/\text{m}^2$ )*	28.8 (6.5)	27.7 (5.2)	27.7 (4.5)
Age (years)†	48.8 (7.1)	53.3 (8.3)	55.8 (8.0)
FFM ( $\text{kg}$ )†	59.6 (11.9)	51.8 (10.6)	45.9 (10.5)
%male**	54.5%	54.1%	55.6%

BMI: body mass index.

FFM: fat-free mass.

\*No significant difference in mean BMI among groups by ANOVA ( $P = 0.3937$ ).

†Significant difference in mean age and FFM among groups by ANOVA ( $P < 0.0001$ ).

\*\*No significant difference in %males among groups by  $\chi^2$  ( $P = 0.9696$ ).

Data for continuous variables are expressed as mean (s.d.).

**Table 3** Predictors of low end point enrichment in OPEN study participants

Variable	Partial $R^2$	P-value
$K_d$ ( $\text{day}^{-1}$ )	0.8472	$< 0.0001$
TBW/weight	0.0019	$< 0.0001$
Gender	0.0085	$< 0.0001$
Weight	0.0258	$< 0.0001$
Height	0.0003	0.2652
Age	0.0001	0.4976

Forward step-wise regression of the variables:  $K_d$ , TBW/weight, gender, weight, height, and age.

$K_d$ : deuterium turnover.

TBW: total body water.

**Table 4** Analytical error in the measurement of elimination rate, DS, TBW, and TEE in a subset of participants ( $n=25$ )

	Dose 1 (analysis 1)	Dose 1 (analysis 2)	CV <sub>w</sub> (%)
$k_o$ (d <sup>-1</sup> )	-0.1157	-0.1154	0.7
$k_d$ (d <sup>-1</sup> )	-0.0895	-0.0895	0.6
$N_o$ (kg)	36.81	37.10	1.0
$N_d$ (kg)	38.33	38.03	1.1
DS ratio	1.0409	1.0243	1.5
TBW (kg)	36.19	36.19	0.8
TEE (MJ/d)	11.11	11.00	2.9

$k$ : elimination rate; <sup>18</sup>O ( $k_o$ ), deuterium ( $k_d$ ).  
 $N$ : dilution space; <sup>18</sup>O ( $N_o$ ), deuterium ( $N_d$ ).  
 DS ratio: dilution space ratio.  
 TBW: total body water.  
 TEE: total energy expenditure.  
 CV<sub>w</sub>: the geometric mean of the with-in subject coefficients of variation.  
 Values expressed are the group means (analyses 1 and 2) and geometric mean of the within subject coefficients of variation (CV<sub>w</sub>).

**Table 5** Total error (reproducibility) in elimination rate, DS, TBW, and TEE in substudy participants ( $n=24$ )

	Dose 1	Dose 2	CV <sub>w</sub> (%)
$k_o$ (d <sup>-1</sup> )	-0.1189	-0.1139	6.6
$k_d$ (d <sup>-1</sup> )	-0.0940	-0.0889	6.6
$N_o$ (kg)	37.97	38.70	1.6
$N_d$ (kg)	39.23	40.23	2.0
DS ratio	1.0346	1.0404	1.2
TBW (kg)	37.23	37.98	1.8
TEE (MJ/d)	10.79	11.12	5.1

$k$ : elimination rate; <sup>18</sup>O ( $k_o$ ), deuterium ( $k_d$ ).  
 $N$ : dilution space; <sup>18</sup>O ( $N_o$ ), deuterium ( $N_d$ ).  
 DS ratio: dilution space ratio.  $N_d/N_o$   
 TBW: total body water.  
 TEE: total energy expenditure.  
 CV<sub>w</sub>: the geometric mean of the within subject coefficients of variation.  
 Values expressed are the group means (doses 1 and 2) and geometric mean of the within subject coefficients of variation (CV<sub>w</sub>).

participants ( $n=484$ ). For reproducibility, a repeat measure of energy expenditure was obtained in a subset of study subjects ( $n=24$ ).

**Dose adequacy**

The DLW dosing methodology created for the OPEN study had to satisfy two criteria. First, to allow several subjects to be dosed each hour by one research technician, dose bottles had to be prepared in advance of the study. The prepared DLW dose bottles allowed up to six subjects to be dosed per hour. If these doses had not been prepared in advance, and instead, filtered and weighed during the clinic day, approximately

half the number of subjects could have been dosed per hour without additional personnel.

Second, the doses had to provide sufficient quantities of isotopes to maintain the precision of the DLW method, without providing excessive quantities of isotopes, which would needlessly increase the cost of the study. The adequacy of the new dosing methodology was determined by categorizing all subjects (main study) into those receiving a sub-, supra-, or optimal dose of DLW. Over 90% of subjects received either an optimal or supraoptimal dose, a sufficient quantity to maintain the precision of the DLW methodology. In the 21% subjects who received a supraoptimal dose, a slight increase in the precision of TEE measurement was gained at the expense of an increased cost of the study (3% when averaged across all subjects). Only 9% ( $n=44$ ) of subjects received a suboptimal DLW dose. More importantly, only two of these subjects were dropped from further analysis because of a low end point enrichment, which would have led to a significant loss of precision. Although not documented in the literature, this is similar to the underdosing failure rate we have observed in smaller studies in which we weighed each dose individually. It should be noted that in small-scale DLW studies, a mid-point urine sample may be collected to prevent loss of subjects because of low end point enrichment. This is especially important in DLW studies conducted in tropical climates where the water turnover rate is higher. One of the physical characteristics associated with under-dosing was younger age. This is consistent with the observation that younger adults average a higher water turnover than older adults (unpublished data). Additionally, there was a significant difference in FFM among the groups. Subjects who received a suboptimal DLW dose had the highest mean FFM and those receiving a supraoptimal dose had a lower mean FFM. This finding is consistent with the fact that those with a higher FFM, also have a higher TBW, and therefore would require a larger DLW dose to achieve an optimal postdose enrichment.

**Variation in the measurement of energy expenditure**

The new dosing methodology did not adversely affect the reliability of the DLW method. In free-living subjects, investigators have reported the within-subject CV in TEE in a test, retest protocol to be as low as 4.6% in 28 to 32-y-old women and men ( $n=4$ ) (Calazel *et al*, 1993) and as high as 11.9% in 18 to 30-y-old men ( $n=19$ ) (Goran *et al*, 1993). These studies did not include an independent measure of analytical error and as recognized by Schoeller and Hnilicka (1996), each measurement of TEE conducted to quantify physiologic error also contained analytical error. Therefore test, retest protocols designed to determine physiologic variation in energy expenditure are actually a measure of total variance (ie, total variance=analytical variance+physiologic variance). The analytical and total CV in TEE in our study were 2.9 and 5.1%, respectively. Given an

analytical error of 2.9%, the physiologic error (physiologic  $CV^2 = \text{total } CV^2 - \text{analytical } CV^2$ ) in our study was 4.2%.

It is possible that our physiologic reliability study (substudy) was enhanced by the consecutive manner in which subjects were dosed, with the dose 2 given 2 weeks after dose 1. It may be questioned whether this accurately represents all possible physiologic variations in energy expenditure, as seasonal variations were not accounted for. A seasonal effect on TEE, as measured by DLW, has been reported in children with a 6% increase in expenditure during the spring vs fall months (Goran *et al*, 1998). The Goran study was limited, however, in that the design was cross-sectional, rather than longitudinal. In adults of industrialized societies, seasonal disparity of energy expenditure has not been found. Schoeller and Hnilicka (1996) measured energy expenditure via the DLW method in six women residing in an urban area during the winter (January or February) and summer (July, August, or September) months. Seasonal differences in total daily energy expenditure (TDEE) were not significant (winter TDEE = 9.08 MJ/d, summer TDEE = 9.01 MJ/d). Even though no seasonal variation was detected, the physiologic CV was 6.4%, suggesting that the use of consecutive measures of TEE may have reduced the physiologic error by a few percent. It should be noted that seasonal variation can be quite large in agrarian societies, with an increase in TEE of 36% found during the harvest season (Brun *et al*, 1981).

The physiologic CV for TEE, however, is considerably smaller than the error observed in repeat measures of self-reported energy intake. Repeat administration of 24-h recalls have demonstrated an intraindividual variation in daily energy intake of 26 and 31% in males and females, respectively (Beaton *et al*, 1979). In another study, administration of a 3-day diet record twice over 2 years found an intraindividual variation in reported energy intake of 47 and 45.0% in males and females, respectively (Hunt *et al*, 1983). These higher intraindividual coefficients of variation in reported energy intake demonstrate that the DLW method is indeed a far more reliable marker of habitual energy intake.

Finally, the importance of determining both the analytical and physiologic error in studies such as OPEN, where DLW is used to validate and/or calibrate dietary intake data should be recognized. Although the accuracy and precision of the DLW method has been established (Schoeller, 1988), an interlaboratory evaluation demonstrated that this precision is not universal among all laboratories (Roberts *et al*, 1995). The analytical error evaluation conducted for this study demonstrates that energy expenditure data derived from our laboratory are valid. Additionally, the results of the substudy proved that energy expenditure as determined by DLW is a precise biomarker of habitual energy intake. In future studies where DLW is used to validate dietary intake data, analysis of both the analytical error of the laboratory and the inclusion of a substudy to determine physiologic error in the measurement of energy expenditure may be warranted.

In summary, the large-scale application including the streamlined-dosing approach created for the OPEN study, achieved its goal of providing a sufficient quantity of tracer, to a large number of people, in a limited amount of time. Greater than 90% of subjects received a sufficient dose of DLW. Only 1% of subjects were excluded from the final data set because of insufficient end point enrichment. Additionally, the new dosing approach did not adversely affect the precision of the method as demonstrated by an analytical and total error of 2.9 and 5.1%, respectively. These data suggest that the streamlined-dosing procedure may be used in large epidemiologic studies.

## References

- Beaton G, Milner J, Corey P, McGuire V, Cousins M, Stewart E, deRamos M, Hewitt D, Grambsch P, Kassim N & Little J (1979): Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. *Am. J. Clin. Nutr.* **32**, 2456–2559.
- Blanc S, Colligan A, Trabulsi J, Harris T, Everhart J, Bauer D & Schoeller DA (2002): Influence of delayed isotopic equilibration in urine on the accuracy of the  $^2\text{H}_2^{18}\text{O}$  method in the elderly. *J. Appl. Physiol.* **92**, 1036–1044.
- Brun T, Bleiberg F & Goihman S (1981): Energy expenditure of male farmers in dry and rainy seasons in Upper-Volta. *Br. J. Nutr.* **45**, 67–75.
- Calazel C, Young V, Evans W & Roberts S (1993): Effect of fasting and feeding on carbon dioxide production using doubly labeled water. *J. Appl. Physiol.* **74**, 1824–1829.
- Cole T & Coward W (1992): Precision and accuracy of doubly labeled water energy expenditure by multipoint and two-point methods. *Am. J. Physiol.* **263**, E965–E973.
- Goran M, Beer W, Wolfe R, Poehlman E & Young V (1993): Variation in total energy expenditure in young healthy free-living men. *Metabolism* **42**, 487–496.
- Goran M, Nagy T, Gower B, Mazariegos M, Solomons N, Hood V & Johnson R (1998): Influence of sex, seasonality, ethnicity, and geographic location on the components of total energy expenditure in young children: implications for energy requirements. *Am. J. Clin. Nutr.* **68**, 675–682.
- Horvitz M & Schoeller DA (2001): Natural abundance deuterium and  $^{18}\text{O}$ -oxygen effects on the precision of the doubly labeled water method. *Am. J. Physiol.* **280**, E965–E972.
- Hunt W, Leonard A, Garry P & Goodwin J (1983): Components of variance in dietary data for an elderly population. *Nutr. Res.* **3**, 433–444.
- Luke A & Schoeller DA (1997): Rapid  $^{18}\text{O}$  analysis of  $\text{CO}_2$  samples by continuous-flow isotope ratio mass spectrometry. *J. Mass Spectrom.* **32**, 1332–1336.
- Racette S, Schoeller DA, Luke A, Shay K, Hnilicka J & Kushner R (1994): Relative dilution spaces of  $^2\text{H}$ - and  $^{18}\text{O}$ -labeled water in humans. *Am. J. Physiol.* **267**, E585–E590.
- Ritz P, Johnson P & Coward W (1994): Measurements of  $^2\text{H}$  and  $^{18}\text{O}$  in body water: analytical considerations and physiological implications. *Br. J. Nutr.* **72**, 3–12.
- Roberts S, Dietz W, Sharp T, Dallal G & Hill J (1995): Multiple laboratory comparison of the doubly labeled water technique. *Obes. Res.* **3**, 3–13.
- Schoeller DA (1983): Energy expenditure from doubly labeled water: some fundamental considerations in humans. *Am. J. Clin. Nutr.* **38**, 999–1005.
- Schoeller DA (1988): Measurement of energy expenditure in free-living human beings by using doubly labeled water. *J. Nutr.* **118**, 1278–1289.

- Schoeller DA (1996): Hydrometry. In *Human Body Composition*. eds Roche AF, Heymsfield SB & Lohman TG, pp 25–43. Champaign: Human Kinetics.
- Schoeller DA, Colligan AS, Shriver T, Hargish H & Bartok-Olson C (2000): Use of an automated reduction system for hydrogen isotope analysis of physiologic fluids applied to doubly labeled water. *J. Mass Spectrom.* **35**, 1128–1132.
- Schoeller DA & Hnilicka J (1996): Reliability of the doubly labeled water method for the measurement of total daily energy expenditure in free-living subjects. *J. Nutr.* **126**, 348S–354S.
- Schoeller DA, Leitch C & Brown C (1986): Doubly labeled water method: *in vivo* oxygen and hydrogen isotope fractionation. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **251**, R1137–R1143.
- Schoeller DA, Taylor P & Shay K (1995): Analytical requirements for the doubly labeled water method. *Obes. Res.* **3**, 15–20.
- Subar AF, Kipnis V, Troiano RP, Midthune D, Schoeller DA, Bingham S, Sharbaugh C, Trabulsi J, Runswick S, Ballard-Barbash R, Sunshine J & Schatzkin A 2003, in press: Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: The Observing Protein and Energy Nutrition (OPEN) study.
- Weir J (1949): New methods for calculating metabolic rate with special reference to protein metabolism. *J. Physiol.* **109**, 1–9.