Articles in PresS. Am J Physiol Endocrinol Metab (October 3, 2017). doi:10.1152/ajpendo.00241.2017

Validation of the doubly labeled water method using off-axis integrated cavity output spectroscopy and isotope ratio mass spectrometry.

Edward L. Melanson, Tracy Swibas, Wendy M. Kohrt, Vicki A. Catenacci, Seth A. Creasy, Guy Plasqui, Loek Wouters, John R. Speakman, and Elena S.F. Berman

Author Affiliations:

- Division of Endocrinology, Metabolism, and Diabetes (EM, VC, SC) and Division of Geriatric Medicine (EM, TS, WK), Department of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO;
- Geriatric Research, Education, and Clinical Center, VA Eastern Colorado Health Care System, Denver, CO (EM, WK)
- NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University, Maastricht, Netherlands (GP, LW)
- Institute of Biological and Environmental Sciences, Aberdeen University, Aberdeen, United Kingdom and State Key Laboratory of Molecular Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China (JRS),
- Los Gatos Research/ABB, San Jose, CA (EB)

Author's last names: Melanson, Swibas, Kohrt, Catenacci, Creasy, Plasqui, Wouters, Speakman, Berman

Corresponding Author: Edward L. Melanson, Ph.D. (Ed.melanson@ucdenver.edu)
MS 8106, 12801 East 17th Ave, RC1 South RM 7103
University of Colorado Anschutz Medical Campus, Aurora, CO 80045
Phone: (303) 724-0935 FAX: (303) 724-3920

Sources of Support: This work was supported by an NIH Small Business Innovation (SBIR) research Grant (R44 DK093362), as well as support from the Colorado Nutrition and Obesity Research Center (P30 DK048520) and the Colorado Clinical and Translational Science Institute (UL1 RR025780). Dr. Melanson is also supported by resources from the Geriatric Research, Education, and Clinical Center at the Denver VA Medical Center

Clinical Trial Registry: The study was registered on ClinicalTrials.gov (NCT01938794)

Running Head: Doubly-labeled water measurements by OA-ICOS and IRMS

ABSTRACT

When the doubly-labeled water (DLW) method is used to measure total daily energy expenditure (TDEE), isotope measurements are typically performed using isotope ratio mass spectrometry (IRMS). New technologies, such as off-axis integrated cavity output spectroscopy (OA-ICOS) provide comparable isotopic measurements of standard waters and human urine samples, but the accuracy of carbon dioxide production (VCO_2) determined with OA-ICOS has not been demonstrated. We compared simultaneous measurement VCO_2 obtained using whole-room indirect calorimetry (IC) with DLW-based measurements from IRMS and OA-ICOS. 17 subjects (10 female; 22 to 63 yrs.) were studied for 7 consecutive days in the IC. Subjects consumed a dose of 0.25 g $H_2{}^{18}O$ (98% APE) and 0.14 g ^2H_2O (99.8% APE) per kg of total body water, and urine samples were obtained on days 1 and 8 to measure average daily CO_2 production (VCO_2) using OA-ICOS and IRMS. VCO_2 was calculated using both the plateau and intercept methods. There were no differences in VCO_2 measured by OA-ICOS or IRMS compared with IC when the plateau method was used. When the intercept method was used, VCO_2 using OA-ICOS did not differ from IC, but VCO_2 measured using IRMS was significantly lower than IC. Accuracy (~1-5%), precision (~8%), intraclass correlation coefficients (R=0.87-90), and root mean squared error (30-40 L/day) of VCO_2 measured by OA-ICOS and IRMS were similar. Both OA-ICOS and IRMS produced measurements of VCO_2 with comparable accuracy and precision when compared to IC.

Key Words: Adult, Humans, Oxygen Isotope, Deuterium, Respiratory Gas Exchange

INTRODUCTION

The gold-standard for measuring total daily energy expenditure (TDEE) in free-living individuals is the doubly-labeled water (DLW) method, which is based on the principle that differential elimination rates of isotopic labels of hydrogen and oxygen provides a measure of carbon dioxide (CO_2) production, subject to certain limiting assumptions (10, 19). TDEE measured using the DLW method has been shown to have an accuracy in humans of ±1-5% against whole room indirect calorimetry (IC) (5, 8, 15, 17-19, 23). Although the number of DLW studies in humans has increased over time (approximately 100 per year), widespread adoption of the DLW method in humans has been limited by the costs of the isotopic labels, and challenges related to sample collection, preparation, and analysis using isotope ratio mass spectrometry (IRMS).

An alternative approach to IRMS for water isotope analysis is laser absorption spectroscopy. These instruments are less expensive than IRMS (~$100,000 vs. $250,000), do not require highly trained technicians for their operation (1), and provide simultaneous measurement of multiple isotopes with less tedious sample preparation (20). There are two commercially-available forms of laser absorption spectroscopy for water isotope analysis, cavity ring-down spectroscopy (CRDS) and Off-Axis Integrated Cavity Output Spectroscopy (OA-ICOS). With CRDS, a laser pulse is trapped in a highly reflective optical cavity. The exponential decay of the light intensity is measured ("ringdown" time) and used to calculate the concentration of the absorbing substance in the gas mixture in the cavity. Although CRDS water isotope analyzers provide accurate and precise measurements of total body water (0.5 ± 1%) and TDEE (0.5 ± 6%)

compared with IRMS, commercial CRDS analyzers have substantial instrumental memory effects, necessitating both careful considerations for reducing isotopic disparity between measured samples and mathematical correction (21). Furthermore, in the above referenced study CRDS was validated against IRMS, but not against the criterion measure of near continuous respiratory gas exchange.

The other commercially-available form of laser absorption spectroscopy for water isotopes, OA-ICOS, uses a laser light source that is coupled to an optical cavity in an off-axis fashion. The laser light wavelength is scanned over absorption features of interest, providing a direct measurement of the absorbing substances in the gas mixture (1). As with IRMS and CRDS, OA-ICOS also suffers from memory issues between adjacent samples. However, because the time to measure each sample (100 seconds) with OA-ICOS is relatively short and requires only a small volume of sample per injection (~1000 nL), memory issues can be circumvented using a higher number of injections per sample, negating the need to perform mathematical corrections. We have previously shown this approach to be accurate and precise when compared to IRMS for both measuring isotopic measurements of pure water and of human urine samples at both enriched and natural abundances (1-3). However, the accuracy and precision of measuring daily carbon dioxide production (VCO_2) using the DLW method with samples measured using OA-ICOS by comparison to whole room indirect calorimetry has not yet been determined. Thus, the purpose of this study was to compare measurement of daily carbon dioxide production (VCO_2) L/day in a whole-room indirect calorimeter with VCO_2 measured simultaneously using the doubly-labeled water (DLW) method with the resultant body water samples (urine) analyzed using OA-ICOS. We also compared the accuracy and precision of OA-ICOS to that of IRMS.

METHODS

Institutional Approval and Ethics - Procedures followed were in accordance with the ethical standards of the Helsinki Declaration of 1975 as revised in 1983. The study was approved by the Colorado Multiple Institutional Review Board on May 2, 2013. The study was registered on ClinicalTrials.gov (NCT01938794) on September 5, 2013. Subject recruitment and enrollment commenced in September, 2013, and the last study visit occurred in February, 2017.

Subjects and screening procedures – Adult volunteers (≥ 18 years) were recruited from the University of Colorado Anschutz Medical Campus (CU-AMC) and local communities. After providing informed, written consent, a Health History and Physical Examination was performed to confirm that volunteers were in a good state of health and that they met criteria for inclusion or exclusion. Primary study exclusion criteria were self-reported smoking or use of smokeless tobacco products, self-reported chronic disease (e.g. heart disease, diabetes, or thyroid disease), or currently pregnant. Body composition was then assessed using whole-body dual-energy x-ray absorptiometry (DXA, Hologic Delphi-W, Hologic, Inc., Bedford, MA). Because of weight limitations of the DXA, volunteers with a body weight greater than 135 kg were also excluded.

Experimental design and study procedures – Subjects were studied for 1 week in the whole-room indirect calorimeter located at the University of Colorado Anschutz Medical Campus. Upon subject arrival on day 1, body weight was measured to ± 0.1 kg and a baseline urine sample was obtained for determination of background abundances of δ^{2}H and δ^{18}O. Subjects were then

given an oral dose of 0.25 g of 98 atom percent (98% APE) ^{18}O labeled water and 0.14 g 99.8% APE ^{2}H labeled water (Sigma Aldrich) per kg of total body water (estimated as 73% of FFM derived from DXA). The dosing cup was twice rinsed with 30 mL of tap water and consumed to ensure complete dosing. After the dose was provided, subjects entered the room calorimeter to begin the 7 day study. Subjects were instructed to completely void ~1 hour after the dose was delivered. Post-dosing urine samples were obtained 4 (PD4) and 5 hours (PD5) after the DLW dosing. On days 2-7, subjects exited the calorimeter for 1 h each day (0700-0800), during which time body weight was measured and then subjects were permitted to shower. For the entire 7 day study, *ad libitum* meals were provided each day at 9 AM, 1 PM, and 6 PM. Subjects were instructed to perform exercise (30 min of treadmill walking at a brisk walking pace) each day to increase TDEE above sedentary levels. On Day 8, subjects exited the calorimeter and end-dose urine and blood samples were obtained at the same time of day as on Day 1 (ED4 and ED5). Approximately 20 ml of each urine sample was immediately pipetted into airtight cryotube and stored at ~-10°C until transferred to a -80°C freezer. Duplicate samples remained frozen at -80°C until analysis.

Whole room-indirect calorimetry – Average daily VCO_2 and 24 h EE over the 7 day period were measured using the whole-room indirect calorimeter located at CU-AMC using a previously described indirect calorimetry system (Sable Systems, International, Las Vegas, NV) (13). O_2 consumption (VO_2) and VCO_2 were calculated in 1-minute intervals using the flow rate and the differences in CO_2 and O_2 concentrations between entering and exiting air, and minute by minute energy expenditure (EE) was calculated using the equations of Jequier et al. (7). Daily 24 h VCO_2 and EE were obtained by summing minute values over the 23 hour measurement period

and extrapolating to 24 h values. The accuracy and precision of the system was tested monthly using propane combustion tests. The average O_2 and CO_2 recoveries during the study were ≥97.0%. While this study was being performed, we also performed several tests using infusions of nitrogen and CO_2 using high precision mass flow controllers, and those tests yielded an accuracy of the IC within 1% of the expected values (unpublished).

OA-ICOS analysis of urine samples - Previously frozen urine samples were prepared by centrifugation, as previously described (3); no distillation or decolorizing steps were undertaken. The OA-ICOS instrument was calibrated using deionized working standards that had been previously calibrated by OA-ICOS against the VSMOW2 and SLAP2 international standards, as previously described (1, 3). Briefly, centrifuged urine samples were injected into heated (~85 °C) stainless steel injection block to produce water vapor, which was then introduced into the OA-ICOS optical cavity. Simultaneous measurements of δ^{2}H and δ^{18}O were performed on each individual injection. Isotope range within each run was minimized by grouping samples expected to have similar enrichments (e.g. PD4/PD5, ED4/ED5) and by using working standards that closely bracketed the expected isotope ratios. Samples, working standards, and internal controls were interleaved throughout each analysis to ensure high accuracy by frequent intra-run calibration. For every individual measurement within a run, samples, working standards, and internal controls were injected 8-12 times, depending on the total isotope range of the run (*e.g.* runs with high enriched samples were injected 12 times,). We have previously shown this approach to produce accurate and precise measurements without memory correction when compared to IRMS (1, 3). 3-5 urine samples were typically included in an individual run which took ~5 to 7 hours to complete. At the conclusion of each OA-ICOS run, the syringe, injector

block, tubes, and filters were cleaned as previously described (1). Each sample was analyzed in a duplicate run on a subsequent day (typically within the same week). If the difference between duplicate runs exceeded 2 δ‰ for $^{2}H:^{1}H$ or 1 δ‰ for $^{18}O:^{16}O$ for a given sample, then that sample was run again and only duplicate values that fell within this range were used.

Isotopic data from the OA-ICOS analyzer were processed using commercially-available Post Analysis Software (LGR, version 3.1.0.9) as previously described (1, 2). Within each run, working standard measurements were utilized with a cubic spline standardization to calibrate urine sample measurements. Specifically, a cubic spline was fit to all measurements of a single standard throughout the run. For each sample injection, an individual calibration curve was constructed from the splined values of each of the working standards. This approach maximally corrects for any instrument drift over the course of the run. To mitigate the effects of sample to sample memory on the OA-ICOS measurements, several procedures were employed (1, 3). First, to account for memory effects between successive samples, the last 4 injections for each sample were averaged, ignoring the first 4-8 injections. Second, to monitor instrument performance, including memory effects between successive samples, an internal control water of known isotopic composition within the range of the isotope ratios of the working standards was measured periodically within each run. Internal controls were checked against the known values. Runs where the internal controls differed from known values by more than ±1.0-2.0 δ per mil (‰) (for low and high-enriched samples) for $\delta^{2}H$, or ±0.3‰ for $\delta^{18}O$ from the known value were repeated. Precision of the urine samples was assessed using these same parameters. Finally, an injection volume (linearity) correction was employed to reduce the effects of different water concentrations (due to syringe volume fluctuations) on the measured isotope ratios. The post-

analysis software also identified any individual injections that were outliers (isotope ratio ±3.0 SD within an injection set) and for the presence of any organic contamination using the integrated Spectral Contamination Identifier feature (9). The presence of any outliers also identified samples where memory effects had not been eliminated.

IRMS analysis of urine samples – Frozen urine samples were shipped from University of Colorado Anschutz Medical Campus to Maastricht University using airtight sealed glass vials and kept frozen using dry ice. Samples were transferred to a -80° freezer and remained frozen until analyzed. For the analysis of $^{2}H:^{1}H$, a 2 ml glass vial containing 300µl of urine was filled with hydrogen gas and equilibration occurred for 1 day at room temperature with a catalyst (5% platinum-on alumina, 325 mesh; Aldrich Chemical Company Ltd) placed in an insert in the vial. For the analysis of ^{18}O, 300µl of urine was put in a glass vial, which was then filled with CO_2 gas. Equilibration then took place for 4 hours at 40° C. The relative amounts of $^{2}H:^{1}H$ in hydrogen gas and $^{18}O:^{16}O$ in CO_2 were then determined using isotope ratio mass spectrometry (Micromass Optima Dual Inlet mass spectrometer with a Multiprep; Manchester,UK, 1998). Each run contained a total of 60 samples of which 12 were working standards with isotope concentrations that bracketed the expected isotope ratios of the urine samples. Each sample was analyzed in a duplicate run on a subsequent day (typically within the same week).

Calculation of CO_2 production (VCO_2) and TDEE – For both OA-ICOS and IRMS, total body water (TBW) was calculated as the average of the dilution spaces of ^{2}H and ^{18}O after correction for isotopic exchange with other body pools (14). Deuterium (K_D) and oxygen (K_O) turnover rates were calculated by linear regression of the natural logarithm of isotope enrichment as a

function of time. All 4 time points were used in the calculation of K_D and K_O. TBW and VCO_2 were calculated using the plateau and intercept methods (using the average of the PD4 and PD5 enrichments) and the equation A6 of Schoeller et al. (15):

$$rCO_2(mol/d) = (N/2.078) * (1.01k_O - 1.041k_D) - 0.0246 * rGF$$

where 1.01 and 1.04 represent the dilution spaces for deuterium and ^{18}O, respectively, N is the body water dilution space, and rGF is the rate of gas fractionation estimated as $1.05N(k_O - k_D)$ (5). TDEE from OA-ICOS and IRMS was calculated using the calculated VCO_2 and the equation of Weir [TDEE = 3.94 x VO_2 + 1.1 VCO_2, where $VO_2 = VCO_2$/ RQ] (22), assuming a respiratory quotient of 0.86, and averaged over 7 days.

Sample Size Justification - Samples size estimates were based on repeated measures on 15 individuals studied in the room calorimeter located at the University of Colorado Anschutz Medical Campus (unpublished data). The difference between the two 24 h VCO_2 measurements was ~12.7 ± 7.5 L/day (~3% of mean values). A total sample of 16 paired measurements was estimated to achieve ~80% power to detect equivalence in 24 h VCO_2 between IC and either IRMS or OA-ICOS when the margin of equivalence is ±7.7 L/day with a 0.05 significance level.

Statistics - Prior to analysis, all data were tested for normality. Differences between IC, OA-ICOS, and IRMS were determined using a repeated measures ANOVA. Post-hoc comparisons were performed using Tukey's multiple comparison test. Because our primary objective was to compare each instrument type to the criterion measure IC, we report only the comparison between IC and OA-ICOS and IC and IRMS. Level of agreement was evaluated using the

difference between the criterion and observed values (percent error, a measure of accuracy), the variance around the accuracy (a measure of precision), intraclass correlation coefficient (a measure of level of agreement), root mean squared error (rMSE, a measure of the magnitude of errors resulting from both bias and variability), and Bland-Altman plots (which provides a measure of bias and limits of agreement, as well as determining whether the error is associated with the magnitude of the criterion measure). The Bland-Altman analyses were performed using the IC as the criterion measure. Associations between subject characteristics and measurement error were determined using the Pearson's correlation coefficient. Significance for all tests was set at P=0.05. Analyses were performed using GraphPad Prism (5.03, La Jolla, CA). Data are reported as mean ± SD.

RESULTS

19 subjects participated in the study. One subject withdrew after one day in the calorimeter. Due to technical issues, two days of data were lost on another subject, and this subject was excluded from the analysis. Thus, the final study sample consisted of 17 participants (Table 1).

Average daily turnover rates of deuterium (k_D/day) and oxygen (k_O/day) determined using OA-ICOS (0.118 ± 0.031/day and 0.142 ± 0.034/day, respectively) were nearly identical to those determined using IRMS (0.118 ± 0.032/day, 0.141 ± 0.033/day). The individual k_O, k_D, N_O, and N_D data used to perform these calculations is contained in the supplementary data file.

Results using the plateau method

TBW, fat free mass (FFM), fat mass (FM), and body fat percentage (%fat) measured by DXA, OA-ICOS, and IRMS are shown in Table 2. There were no differences in TBW, FFM, FM, or %fat measured by OA-ICOS or IRMS when compared with DXA. Regardless of approach N_D and N_O were similar (Table 3), and the average dilution space ratios were close to the empirically derived value in adult humans of 1.031 (15).

There were no significant differences in average VCO_2 measured by OA-ICOS (433.0 ± 72.7 L/day) or IRMS (418.3 ± 73.0 L/day) when compared with IC (411.2 ± 62.1 L/day) (Figure 1, Table 1). To demonstrate the effect on calculated TDEE, 24 h EE from IC (calculated using the measured RQ) was compared to TDEE calculated from OA-ICOS and IRMS using the assumed RQ of 0.86, as would be done in a standard DLW study. Mean TDEE measured by OA-ICOS (10.16 ± 1.70 MJ/day) and IRMS (9.91 ± 1.70 MJ/day) did not significantly differ from IC (9.88 ± 1.56 MJ/day).

The accuracy of VCO_2 measured by OA-ICOS (mean % error) and IRMS was 5.4 and 1.7%, respectively (Table 4). The accuracy of OA-ICOS was significantly different from zero (95% CI does not cross zero). However, the size of the 95% CIs around the percent error were similar for OA-ICOS (+1.1 to +9.6 L/day) and IRMS (-2.5 to +5.8 L/day), indicating a similar level of precision. The ICC between OA-ICOS and IC [0.87 (95% CI = 0.67 – 0.95)] was similar to the ICC between IRMS and IC [0.89 (0.72 – 0.96)]. The RMSE was 40.2 L/day for OA-ICOS and 31.5 L/day for IRMS. Results of the Bland-Altman analysis are presented in Figure 2. There was a significant bias for OA-ICOS (+21.8 L/day, 95% CI = +3.9 to +39.8 L/day) compared to

IC, but not for IRMS (+7.1 L/day, 95% CI = -9.1 to +23.4 L/day). The reduced accuracy and significant bias for OA-ICOS was driven by a single outlier. The Bland-Altman correlations for OA-ICOS and IRMS were not significant indicting no bias with absolute level of VCO_2. VCO_2 for each individual measured by IC, OA-ICOS, and IRMS are shown in Table 1. For most individuals, all three methods produced similar results.

Results using the intercept method

When the intercept method was used, TBW and FFM estimated using IRMS were significantly lower, and FM and %fat significantly higher compared to DXA ($P<0.001$) (Table 2). There were no differences in TBW, FFM, FM, and %fat measured by DXA compared with OA-ICOS. N_D and N_O were similar, and the average dilution space ratios were close to the theoretical value in adult humans of 1.031 (15) (Table 3). There was no difference in average VCO_2 measured by OA-ICOS (422.9 ± 70.7 L/day) when compared with IC (411.2 ± 62.1 L/day), but VCO_2 measured by IRMS (381.9 ± 69.2 L/day) was significantly different compared with IC (Figure 3). Similarly, mean TDEE measured by OA-ICOS (10.40 ± 1.70 MJ/day) was not different than 24 h EE. However, mean TDEE measured by IRMS using the intercept method (9.05 ± 1.62 MJ/day) was significantly lower than 24 h EE. Individual subject VCO_2 results calculated using the intercept method are presented in the Supplemental Table 1.

As with the plateau method, there was a similar level of agreement when VCO_2 measured using OA-ICOS and IRMS were compared with IC (Table 4). Interestingly, accuracy between OA-ICOS and IC tended to be better using the intercept method, whereas accuracy between IRMS

and IC tended to be better using the plateau method. Precision, ICC, and RMSE were similar for OA-ICOS and IRMS using the intercept method. Results of the Bland-Altman analysis are presented in Figures 4. There was a significant bias for IRMS (-29.2 L/day, 95% CI = -44.6 to -13.9 L/day) compared to IC, but not for OA-ICOS (+11.7 L/day, 95% CI = -5.1 to +28.5 L/day). The Bland-Altman correlations between average VCO_2 from IC and both IRMS and OA-ICOS were not significant indicting no bias with absolute level of EE.

Additional Analyses

To determine if %fat, BMI, or age were contributing factors to differences between IC and IRMS or OA-ICOS, correlations between these variables and the differences in VCO_2 between IC and OA-ICOS and IC and IRMS were determined (using the plateau data). The differences in VCO_2 between IC and OA-ICOS were not significantly correlated with %fat (r=0.41) or BMI (r=0.42), but were positively and significantly ($P<0.05$) associated with age (r=0.59). However, this significant correlation was driven solely by one subject (S12, 60 yr. old female) where OA-ICOS substantially overestimated IC (+54 L/day). The differences between IC and IRMS were not significantly correlated with %fat (r=-0.07), BMI (r=-0.20), or age (r=0.04). We also examined the association between the differences in VCO_2 (IC – OA-ICOS, IC – IRMS) with measured RQ. The differences (TDEE – 24 h EE) between IC and OA-ICOS (r=0.19) and IRMS (r=0.46) were positively but weakly ($P>0.05$) correlated with average daily 24 hr RQ. We performed these same analyses using the intercept data, and results were similar (data not shown).

DISCUSSION

Because of the high costs of operation and technical expertise required for operation of IRMS, only a few specialized labs are equipped to perform DLW measurements of TDEE. Although new approaches such as OA-ICOS are available, they have not yet been validated against room calorimetry. We compared VCO_2 calculated using isotopic measurements obtained using OA-ICOS against 24 h VCO_2 measured using whole-room indirect calorimetry as the criterion measure. We also compared VCO_2 calculated using isotopic measurements obtained using IRMS on the same samples to then evaluate if the techniques provide comparable accuracy and precision compared to IC. Mean VCO_2 measured using OA-ICOS did not differ significantly from IC, whether using plateau or intercept calculation approach. Mean VCO_2 measured using IRMS did not differ from IC when the plateau method was used, but was significantly lower than IC when the intercept method was used. Nonetheless, measurements of accuracy (% error), precision (SD of mean % error), ICC, RMSE, and Bland-Altman analyses suggested that level of agreement with IC was similar for both IRMS and OA-ICOS. Thus, results of this study demonstrate that off-axis integrated cavity output spectroscopy provides estimates of VCO_2 from DLW studies in humans that are as accurate and precise as estimates derived from IRMS.

Initial validation work of the DLW method performed in the 1950's in several small animal species showed that VCO_2 was within ~3% of that measured simultaneously by indirect calorimetry (11, 12). Schoeller and van Santen (16) performed the first validation studies in humans in 1982, and reported that TDEE from the DLW method differed from measured energy intake (adjusted for changes in body composition) by an average of 2%. Subsequent validation studies against near continuous respiratory gas exchange measured over 4-7 days reported

precisions of ~1-8% for measuring VCO_2 and TDEE (5, 8, 15, 17, 18, 23). The range of accuracies for both OA-ICOS and IRMS in the current study (Table 4), using both the plateau and intercept method, were similar to these previous studies. Surprisingly, when using the intercept method, we observed a significant difference between mean VCO_2 measured by IC and IRMS, which is not consistent with previous validation studies.

To more thoroughly compare the IC to OA-ICOS (and IC to IRMS), we performed several statistical tests to assess the levels of agreement between instruments, some of which are more reflective of individual errors. Specifically, both the ICC and RMSE describe how concentrated the data are around the line of best fit (in this case, the line of identity) , whereas the Bland-Altman allows identification of systematic differences between two measurements (4). Both the RMSE and Bland-Altman can be also used to identify where measurement errors are driven by the presence of outliers. Because DLW studies are performed on groups of individuals (e.g., to compare differences between groups to determine the effect of some intervention), more weight should be given to tests that are based on mean differences. For example, even though the Bland-Altman test indicated a significant, positive bias in measuring VCO_2 using the plateau method with OA_ICOS (+21.8 L/day), there was no difference in mean VCO_2 measured by OA-ICOS and IC. Based on the current analyses, we conclude that OA-ICOS provides a measure of average daily VCO_2 that is accurate (1% to 5%) and precise (8%) without systematic bias. We also conclude that accuracy, precision, and bias are similar to those observed with IRMS.

It has been suggested that adiposity and nutritional status affect the dilution space ratio (N_d/N_o) between ^{2}H and ^{18}O, causing potential errors in VCO_2 when the DLW method is used (6). In that

study, it was reported that there was an overestimation of VCO_2 by the DLW method in high fat (HF) diet fed mice compared with measured VCO_2 using continuous measurements with IC. This overestimation occurred in both a diet-induced obesity-prone (DIO) and diet-induced obesity-resistant (DR) groups, suggesting that the overestimation is independent of body fat gain during a HF diet. In the current study, we found no association between either %fat or BMI and the difference in VCO_2 measured with IC and DLW. We also explored the association between measured RQ and the difference in VCO_2 measured with IC and DLW. These associations were also non-significant with both OA-ICOS and IRMS. Although we did not measure energy intake (subjects consumed an *ad libitum* diet), our subjects were weight stable throughout the 7 day study (-0.5 ± 0.8 kg, mean ± SD), suggesting that individual differences in average 24 hr RQ reflected differences in habitual energy macronutrient intake rather than energy balance. Under this assumption, if VCO_2 is overestimated during consumption of a high fat diet, a negative correlation would be expected when the differences between the DLW and IC VCO_2 are plotted against RQ (with a lower RQ indicative of a higher fat intake). Thus, results of the current study do not support the conclusion that VCO_2 from the DLW method is overestimated during a high-fat diet, but we concede that this can only be determined during studies in which energy and macronutrient intake is highly controlled.

Strengths and limitations: A strength of the current study is the sample size, which is larger (N=17) than previous validation studies performed using near continuous measurements of respiratory gas exchange (N<10) (5, 8, 15, 17, 18, 24). A limitation of the current study, as in all validation studies, is the validity of the criterion measure (IC). However, as described in the Methods section, the room calorimeter system at AMC consistently measures within 1-3% of

expected values using gas infusion and propane combustion tests. In addition to costs, OA-ICOS offers several advantages over IRMS including easier sample preparation and reducing the need for highly trained technicians. However, it should be noted that the sample measurement configuration used in the current study (e.g. 8-12 injections per sample, with multiple interleaved measurements of working standards and internal controls) does not increase the throughput compared to IRMS and CRDS. The advantage of this approach is that it negates the need for mathematical correction due to memory effects. Throughput could be increased by reducing the number of injections per sample, but the tradeoff would then be the need to apply mathematical correction for memory effects.

In conclusion, , mean VCO_2 measured using OA-ICOS did not differ significantly from concurrently measured 24 h VCO_2 using whole room indirect calorimetry, whether using plateau or intercept calculation approach. Furthermore, both OA-ICOS and IRMS produced measurements of VCO_2 with comparable accuracy and precision when compared to whole room indirect calorimetry. Based on these results, we conclude that off-axis integrated cavity output spectroscopy provides a valid and viable alternative to IRMS for measuring TDEE using DLW in humans.

Acknowledgements: This work was supported with resources and use of facilities from the Geriatric Research, Education, and Clinical Center at the Denver VA Medical Center

Conflict of Interest (COI) Statement: Elena Berman is employed by ABB/Los Gatos Research, the company that manufactures the OA-ICOS analyzer.

Disclaimer: The contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government.

396 Table 1. Subject characteristics and individual average total daily carbon dioxide production (VCO_2) measured by measured by IC
397 and by OA-ICOS and IRMS using the plateau method.

					VCO_2 (L/day)		
Subject #	Sex	Age (yrs.)	Weight (kg)	BMI (kg/m^2)	IC	OA-ICOS	IRMS
1	F	46	63.0	24.0	310.6	307.1	267.3
2	M	32	82.8	23.9	457.4	456.2	440.2
3	M	43	74.8	25.1	455.8	484.1	487.8
4	M	28	61.0	22.4	346.8	374.9	367.8
5	F	24	93.8	31.9	471.3	474.7	476.9
6	F	60	48.9	19.4	293.4	339.0	334.7
7	F	62	53.3	21.8	349.9	372.8	351.3
8	M	34	91.5	32.2	444.2	458.3	436.9
9	M	40	71.6	23.0	442.8	448.3	390.5
10	F	27	111.6	46.4	514.4	568.2	529.5
11	F	60	95.0	34.8	437.1	560.5	421.6
12	F	63	115.0	42.8	423.1	474.7	453.7
13	F	34	101.4	36.1	433.7	449.7	453.9
14	M	24	73.9	23.0	473.4	450.7	545.9
15	F	30	72.0	28.5	394.0	367.3	391.7
16	F	22	61.7	24.5	353.8	358.4	336.6
17	M	43	69.6	20.8	387.9	415.9	424.4
Mean (SD)		**39 (14)**	**78.8 (19.7)**	**28.3 (7.9)**	**411.2 (62.1)**	**433.0 (72.7)**	**418.3 (73.0)**

Table 2. Total body water (TBW), fat free mass (FFM), fat mass (FM), and percent body fat (%Fat) measured by DXA, OA-ICOS, and IRMS. OA-ICOS and IRMS results are presented for both plateau and intercept methods. Mean (SD).

		Intercept Method		Plateau Method	
	DXA	OA-ICOS	IRMS	OA-ICOS	IRMS
TBW (kg)	38.3 (7.3)	38.3 (6.7)	35.6 (6.5)[a]	38.3 (6.7)	39.0 (6.7)
FFM (kg)	52.5 (10.0)	52.2 (9.4)	48.8 (8.9)[a]	52.5 (10.0)	53.4 (9.2)
FM (kg)	25.9 (15.8)	26.6 (15.8)	29.9 (16.0)[a]	26.3 (16.0)	25.3 (15.8)
%Fat	31.0 (12.5)	31.9 (11.9)	36.8 (11.3)[a]	31.5 (12.6)	30.2 (12.0)

[a] significantly different from DXA

403 Table 3. Deuterium (N_D) and oxygen (N_O) dilution spaces and dilution space ratio (N_D:N_o) measured by OA-

404 ICOS and IRMS. Results are presented for both plateau and intercept methods. Mean (SD).

	Intercept Method		Plateau Method	
	OA-ICOS	IRMS	OA-ICOS	IRMS
N_D (kg)	38.0 (6.7)	37.9 (6.9)	38.9 (6.8)	40.4 (6.7)
N_O (kg)	36.8 (6.6)	36.8 (6.7)	37.8 (6.6)	39.0 (6.8)
N_D:N_o	1.033 (0.005)	1.030 (0.006)	1.029 (0.0068)	1.037 (0.013)

405

406

Table 4. Limits of agreement for CO_2 production (VCO_2) measured by OA-ICOS and IRMS. Results are presented for both plateau and intercept methods.

	Error (%) Mean (95% CI)	ICC (95% CI)	RMSE (L/day)
OA-ICOS - Plateau	5.4 (+1.1, +9.6)	0.87 (0.67, 0.95)	40.2
IRMS - Plateau	1.7 (-2.5, +5.8)	0.89 (0.72, 0.96)	31.5
OA-ICOS – Intercept	2.9 (-1.1, +6.9)	0.88 (0.70, 0.90)	33.8
IRMS – Intercept	-7.2 (-11.2, -3.3)	0.90 (0.74, 0.96)	35.9

ICC – interclass correlation; RMSE – Root mean Square Error

FIGURE LEGENDS

Figure 1. VCO_2 (Mean ± SEM) measured by IC and by OA-ICOS and IRMS using the plateau method.

Figure 2. Bland-Altman plots of OA-ICOS (A) and IRMS (B) using the plateau method vs. the criterion measure IC.

Figure 3. VCO_2 (Mean ± SEM) measured by IC and by OA-ICOS and IRMS using the intercept method. * Significantly different than IC.

Figure 4. Bland-Altman plots of OA-ICOS (A) and IRMS (B) using the intercept method vs. the criterion measure IC.

REFERENCES

1. Berman ES, Fortson SL, Snaith SP, Gupta M, Baer DS, Chery I, Blanc S, Melanson EL, Thomson PJ, and Speakman JR. Direct analysis of delta2H and delta18O in natural and enriched human urine using laser-based, off-axis integrated cavity output spectroscopy. *Analytical chemistry* 84: 9768-9773, 2012.
2. Berman ES, Levin NE, Landais A, Li S, and Owano T. Measurement of delta18O, delta17O, and 17O-excess in Water by Off-Axis Integrated Cavity Output Spectroscopy and Isotope Ratio Mass Spectrometry. *Analytical chemistry* 2013.
3. Berman ES, Melanson EL, Swibas T, Snaith SP, and Speakman JR. Inter- and intraindividual correlations of background abundances of (2)H, (18)O and (17)O in human urine and implications for DLW measurements. *European journal of clinical nutrition* 69: 1091-1098, 2015.
4. Bland JM, and Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1: 307-310, 1986.
5. Coward WA, and Prentice AM. Isotope method for the measurement of carbon dioxide production rate in man. *The American journal of clinical nutrition* 41: 659-663, 1985.
6. Guidotti S, Meijer HA, and van Dijk G. Validity of the doubly labeled water method for estimating CO2 production in mice under different nutritional conditions. *American journal of physiology Endocrinology and metabolism* 305: E317-324, 2013.
7. Jequier E, Acheson K, and Schutz Y. Assessment of energy expenditure and fuel utilization in man. *Annual review of nutrition* 7: 187-208, 1987.
8. Klein PD, James WP, Wong WW, Irving CS, Murgatroyd PR, Cabrera M, Dallosso HM, Klein ER, and Nichols BL. Calorimetric validation of the doubly-labelled water method

for determination of energy expenditure in man. *Human nutrition Clinical nutrition* 38: 95-106, 1984.

9. Leen JB, Berman ESF, Liebson L, and Gupta M. Spectral contaminant identifier for off-axis integrated cavity output spectroscopy measurements of liquid water isotopes. *Rev Sci Instrum* 83: 2012.

10. Lifson N. Theory of use of the turnover rates of body water for measuring energy and material balance. *J Theor Biol* 12: 46-74, 1966.

11. Lifson N, Gordon GB, and Mc CR. Measurement of total carbon dioxide production by means of D2O18. *J Appl Physiol* 7: 704-710, 1955.

12. McClintock R, and Lifson N. Measurement of basal and total metabolism in hereditarily obese-hyperglycemic mice. *The American journal of physiology* 193: 495-498, 1958.

13. Melanson EL, Ingebrigtsen JP, Bergouignan A, Ohkawara K, Kohrt WM, and Lighton JR. A new approach for flow-through respirometry measurements in humans. *American journal of physiology Regulatory, integrative and comparative physiology* 298: R1571-1579, 2010.

14. Racette SB, Schoeller DA, Luke AH, Shay K, Hnilicka J, and Kushner RF. Relative dilution spaces of 2H- and 18O-labeled water in humans. *The American journal of physiology* 267: E585-590, 1994.

15. Schoeller DA, Ravussin E, Schutz Y, Acheson KJ, Baertschi P, and Jequier E. Energy expenditure by doubly labeled water: validation in humans and proposed calculation. *The American journal of physiology* 250: R823-830, 1986.

16. Schoeller DA, and van Santen E. Measurement of energy expenditure in humans by doubly labeled water method. *Journal of applied physiology: respiratory, environmental and exercise physiology* 53: 955-959, 1982.

17. Schoeller DA, and Webb P. Five-day comparison of the doubly labeled water method with respiratory gas exchange. *The American journal of clinical nutrition* 40: 153-158, 1984.

18. Seale JL, Conway JM, and Canary JJ. Seven-day validation of doubly labeled water method using indirect room calorimetry. *Journal of applied physiology* 74: 402-409, 1993.

19. Speakman JR. *Doubly labelled water: Theory and Practice*. London: Chapman Press, 1997.

20. Steig EJ, Gkinis V, Schauer AJ, Schoenemann SW, Samek K, Hoffnagle J, Dennis KJ, and Tan SM. Calibrated high-precision O-17-excess measurements using cavity ring-down spectroscopy with laser-current-tuned cavity resonance. *Atmos Meas Tech* 7: 2421-2435, 2014.

21. Thorsen T, Shriver T, Racine N, Richman BA, and Schoeller DA. Doubly labeled water analysis using cavity ring-down spectroscopy. *Rapid communications in mass spectrometry : RCM* 25: 3-8, 2011.

22. Weir JB. New methods for calculating metabolic rate with special reference to protein metabolism. 1949. *Nutrition* 6: 213-221, 1990.

23. Westerterp KR, Brouns F, Saris WH, and ten Hoor F. Comparison of doubly labeled water with respirometry at low- and high-activity levels. *Journal of applied physiology* 65: 53-56, 1988.

24. Westerterp KR, Lafeber HN, Sulkers EJ, and Sauer PJ. Comparison of short term indirect calorimetry and doubly labeled water method for the assessment of energy expenditure in preterm infants. *Biology of the neonate* 60: 75-82, 1991.

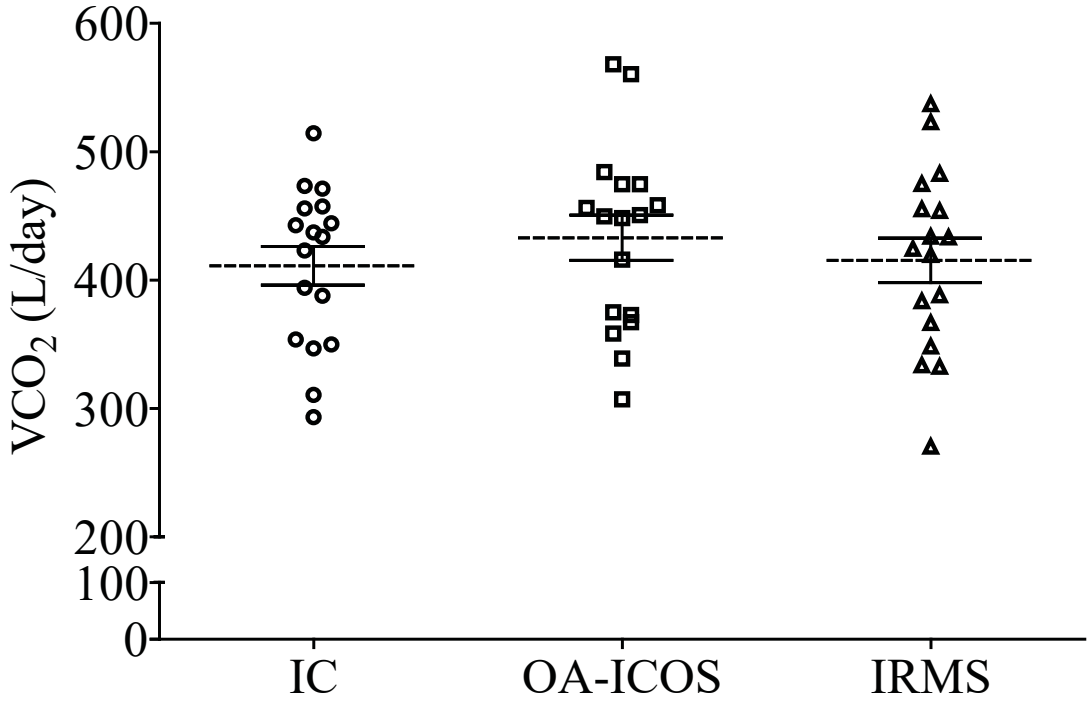

Figure 1

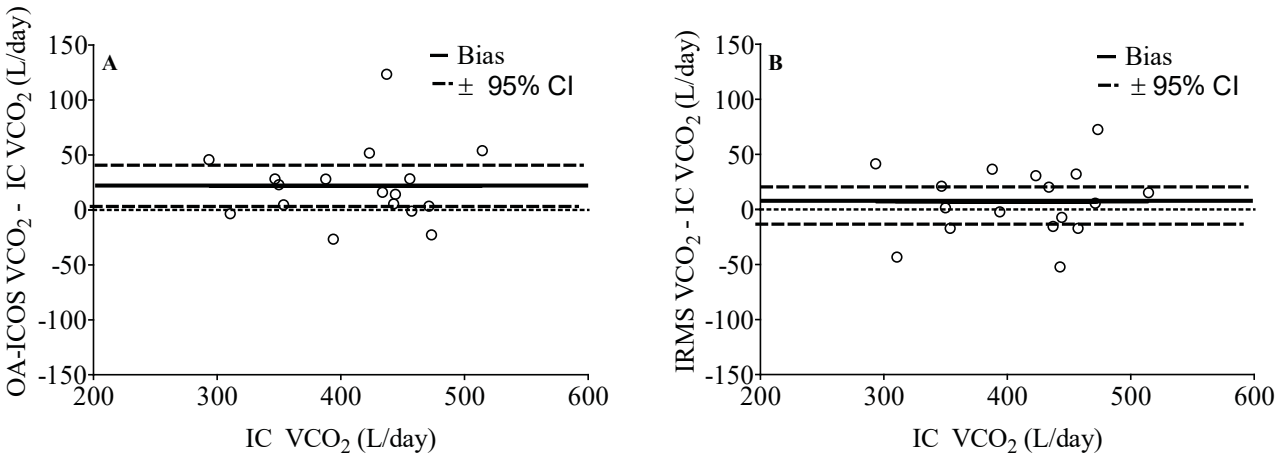
A
Bias
± 95% CI
OA-ICOS VCO2 - IC VCO2 (L/day)
IC VCO2 (L/day)
150
100
50
0
-50
-100
-150
200
300
400
500
600
B
Bias
± 95% CI
IRMS VCO2 - IC VCO2 (L/day)
IC VCO2 (L/day)

Figure 2

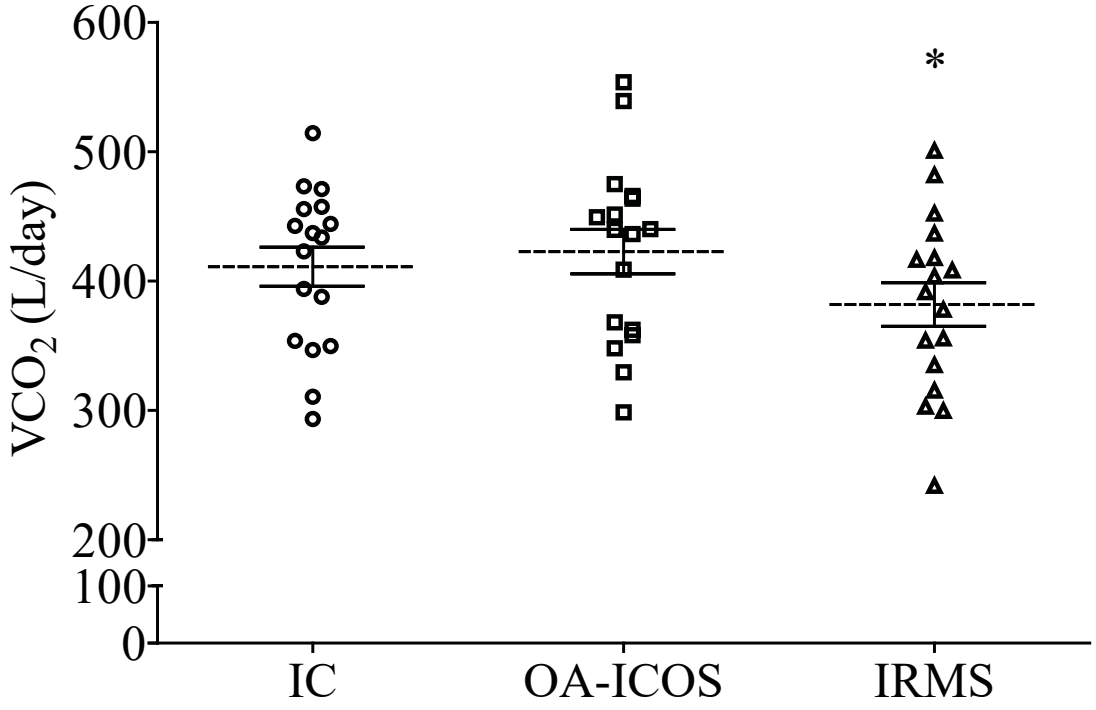

Figure 3

Figure 3

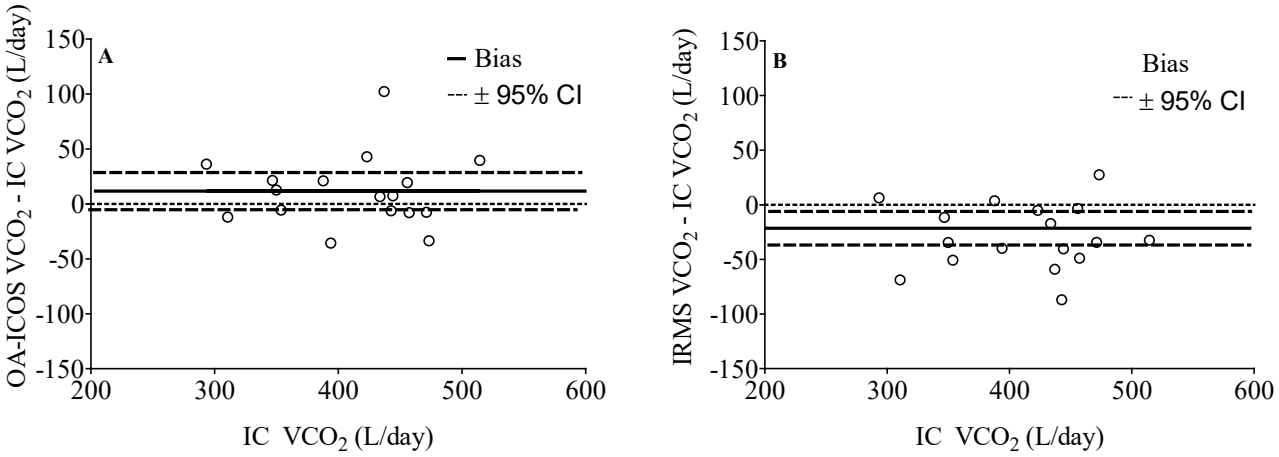
A
Bias
± 95% CI
OA-ICOS VCO2 - IC VCO2 (L/day)
IC VCO2 (L/day)
150
100
50
0
-50
-100
-150
200
300
400
500
600
B
Bias
± 95% CI
IRMS VCO2 - IC VCO2 (L/day)
IC VCO2 (L/day)

Figure 4