


ORIGINAL ARTICLE

Clinical Trials and Investigations

Energy expenditure measurements are reproducible in different whole-room indirect calorimeters in humans

Emma J. Stinson¹ | Theresa Rodzevik¹ | Jonathan Krakoff¹ | Paolo Piaggi^{1,2}  | Douglas C. Chang¹

¹Obesity and Diabetes Clinical Research Section, Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, Phoenix, Arizona, USA

²Department of Information Engineering, University of Pisa, Pisa, Italy

Correspondence

Paolo Piaggi, Obesity and Diabetes Clinical Research Section, National Institute of Diabetes and Digestive and Kidney Diseases, 4212 N. 16th St., Phoenix, AZ 85016, USA. Email: paolo.piaggi@nih.gov, paolo.piaggi@unipi.it

Funding information

Ministero dell'Istruzione, dell'Università e della Ricerca; National Institute of Diabetes and Digestive and Kidney Diseases, Grant/Award Number: DK069015-36

Abstract

Objective: This study aimed to evaluate the agreement of commonly reported energy metabolism measurements obtained from two different whole-room indirect calorimeters (WRICs).

Methods: Nine healthy adult volunteers were evaluated over four separate 24-hour periods in a crossover design, twice in two different WRICs of different sizes, each operated according to the Room Indirect Calorimetry Operating and Reporting Standards published in 2020. The reproducibility of repeated measurements was quantified by the coefficient of variation (CV) and intraclass correlation coefficient (ICC).

Results: The CVs between and within each WRIC for average 24-hour carbon dioxide production rate (VCO_2) and oxygen consumption rate (VO_2), 24-hour energy expenditure (EE), and respiratory exchange ratio ranged from 1.5% to 3.6%, whereas sleep EE ranged from 3.1% to 5.5%. CVs for macronutrient oxidation rates and spontaneous physical activity were higher, ranging from 9.2% to 38.1%. ICCs of VCO_2 , VO_2 , 24-hour EE, and energy expenditure at zero activity were >0.95 , indicating excellent reproducibility, whereas ICCs for lipid oxidation, awake and fed thermogenesis, and sleep EE ranged from 0.55 to 0.92, indicating moderate to high reproducibility. ICCs for respiratory exchange ratio and carbohydrate and protein oxidation rates were lower (<0.70). Spontaneous physical activity showed high reproducibility within chambers (ICC = 0.88) but differed substantially between chambers (ICC = 0.23).

Conclusions: Cross-chamber reproducibility is high for common outcome measures assessed in the respiratory chamber. The results support efforts to promote standardization across WRICs to allow multicenter studies.

INTRODUCTION

A whole-room indirect calorimeter (WRIC), also called a respiratory chamber, is a unique tool to study energy metabolism. Multicenter

studies involving different calorimeters would be beneficial to increase sample size and increase generalizability of research findings. However, to our knowledge, only one multicenter study has been published [1]. Although the limited number of operational calorimeters worldwide is one factor, the lack of consistency in technical operation and validation is another important reason. To improve consistency across calorimeters, recommendations to

Emma J. Stinson and Theresa Rodzevik contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Published 2022. This article is a U.S. Government work and is in the public domain in the USA. *Obesity* published by Wiley Periodicals LLC on behalf of The Obesity Society (TOS).

standardize calorimetry operations and reporting were recently published by an expert panel [2].

Since measurement of biological variability is a goal of human calorimetry studies, the variability in repeated measurements of the same individual in a single WRIC has been used to estimate the number of participants required for studies. The coefficient of variation (CV) for 24-hour (24-h) energy expenditure (EE), which is often the main study outcome, has been reported to be 1% to 5% [3–8], whereas the CV of macronutrient substrate oxidation has been higher, between 5% and 25% [7, 8]. Prior studies have not examined reproducibility across different WRICs. If measurements can be demonstrated to be reproducible across different calorimeters by applying the current quality standards, confidence in replicability of research results is increased and larger multicenter studies can be planned.

In the current study, we sought to evaluate whether measurements of human energy metabolism can be reproduced in two whole-room calorimeters located on the same inpatient unit. These calorimeters differ in terms of room size and wall construction but otherwise they used similar analytical equipment. Each WRIC was separately validated with the same methods (i.e., dry gas infusion and propane combustion), and operated in the same manner. Participants were studied over four 24-h periods in two calorimeters (twice in each calorimeter) in a crossover design and in conditions of energy balance and weight stability.

METHODS

Study population and design

Adult male and female volunteers were recruited from the Phoenix, Arizona, metropolitan area in 2019 and admitted to the clinical research unit (CRU) as part of a larger observational inpatient study investigating risk factors for obesity ([ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT00340132) identifier NCT00340132). The Institutional Review Board of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) approved the study. Written informed consent was obtained prior to participation. All volunteers were healthy and not currently using drugs, prescription medication, or nicotine products by history, physical examination, and routine laboratory tests, including urine drug and nicotine tests. Female participants were verified to be nonpregnant with a urine pregnancy test.

An overview of the inpatient study is shown in Supporting Information Figure S1. During the inpatient study, a standard weight-maintaining diet (50% carbohydrate, 30% fat, and 20% protein), based on a CRU-specific equation using weight and sex, was provided to participants from the day of admission [9]. On the second day, body composition was assessed using dual-energy x-ray absorptiometry (iDXA, GE Lunar Healthcare). After at least 3 days on the weight-maintaining diet, glucose tolerance was assessed by a 75-g oral glucose tolerance test. Participants with diabetes were excluded according to 2003 American Diabetes Association criteria [10]. Plasma glucose concentrations were determined by glucose oxidase method

Study Importance

What is already known?

- Whole-room indirect calorimeters (WRICs), also called respiratory or calorimetric chambers, are important tools to study energy metabolism.
- The Room Indirect Calorimetry Operating and Reporting Standards were recently published, promoting standardization across WRICs around the world.

What does this study add?

- Measurements were obtained from two different WRICs, operated according to recently published guidelines.
- Twenty-four-hour energy expenditure, oxygen consumption, and carbon dioxide production were highly reproducible across different chambers, whereas substrate oxidation rates were less reproducible between chambers.
- Spontaneous physical activity differed between two chambers of different sizes.

How might these results change the direction of research or the focus of clinical practice?

- Standardizing operations and establishing reproducibility of WRICs may increase confidence in developing multicenter studies which include these measurements.
- Findings support efforts to promote standardization in WRIC operations and reporting.

(Analog GM9, Analog Instruments). Volunteers had four WRIC studies of approximately 23.5 hours each, with two being in one WRIC and two in a second WRIC (WRIC A and B). The four WRIC studies were separated by at least one inpatient washout day (range 1–3 days). Volunteers were randomly assigned to one of six chamber sequences (i.e., AABB, ABAB, ABBA, BBAA, BABA, or BAAB).

Chamber diet and instructions

On the day of the entry into the chamber, participants ate a standardized breakfast at 7:00 am and entered the chamber at 7:30 am. Prescribed energy intake in the chamber was based on previously reported equations, developed specifically for the CRU [11], and was approximately 80% of the weight-maintaining diet provided outside the chamber to account for restricted physical activity within the chamber. For accurate calculation of energy intake, all unconsumed food was returned to the metabolic kitchen for weighing. Macronutrient composition was 50% carbohydrate, 30% fat, and 20% protein

with a food quotient of 0.87; specific details of the meals are provided in online Supporting Information Methods. Meals were prepared by the CRU metabolic kitchen in accordance with prescribed energy intake and macronutrient composition. Participants drink ad libitum from water bottles provided in the chamber.

Participants were instructed not to exercise but they were able to move freely about the chamber. Volunteers could bring in reading materials, but cell phones and computers were not permitted inside the chamber. Volunteers were permitted to call the nurses station or externally (e.g., to family) from the phone provided in the chamber. Bedtime and “lights out” periods were determined by the volunteer.

WRIC description

WRIC A was first constructed in 1983, with walls consisting of two layers of 1-mm-thick aluminum and 10 cm of urethane between them as previously described [3]. The dimensions of WRIC A are 3.33 m in length, 2.45 m in width, and 2.39 m in height. WRIC B was initially built in 2009 from standard construction materials (studs and drywall), with an insulated door and walls that are 12.7 cm thick. The measured dimensions are 3.30 m in length, 2.72 m in width, and 2.46 m in height. For both calorimeters, “washout” tests in which fresh air flowed into the chamber with elevated CO₂ concentrations were performed to determine the effective chamber volume (V_e) including furniture [12]. Volumes were determined to be 16,940 and 19,000 L for WRIC A and B, respectively. Each WRIC is furnished with a sink, toilet, bed, small refrigerator, microwave, television, phone, desk, and chair and a two-door air lock through which meals may be passed. The doors for human passage are not locked and may be opened from either side.

Both WRIC A and B were initially configured as a pull calorimeter (i.e., air pulled from the room) but changed to a push configuration (i.e., air pushed into the room) prior to the start of the current study [13]. Inflow air is drawn by separate voltage-controlled blowers (Solberg) from the same closed reference room (volume approximately 20 m³) that is supplied by the hospital HVAC (heating, ventilating, and air-conditioning) system, which in turn sources air from within the hospital. The inflow rate for each WRIC is controlled by separate mass flow controllers (MFCs; Alicat, MCRW-250SLPM-D-I) at a fixed rate of 60 L/min. Samples of air from the reference room and from within the chamber are dried by gas sample dryers (Perma Pure LLC, PD-50 T-48MSS) driven by counterflowing dry air generated by an air compressor (JUN-AIR). Air samples are analyzed by differential CO₂ (ABB Automation, Uras 26) and O₂ (Siemens Oxymat 6E) analyzers at a flow rate of 0.7 L/min controlled by MFC (Alicat, Basis). Separate HVAC systems, set to 23.5 °C for human studies for the comfort of our participants, maintain temperature within each chamber. Temperature and humidity are continuously monitored (Vaisala, HMP 60C12AOA3B0) inside each WRIC.

As the calorimeters perform optimally when pressure is higher in the chamber room relative to the adjacent air space, the barometric pressure difference between the WRIC and the adjacent hallway

(Omega, PX653-2.5BD5V) is continuously monitored during human studies.

Spontaneous physical activity (SPA) was detected by two microwave radar devices (BB-150-NH Motion Sensor, Museum Technology Source) within each chamber and expressed as the percentage of time spent in motion, averaged over the entire 24-h stay in the chamber [3]. For example, a 5% average activity over 24 hours would represent 72 minutes of continuous motion. Additional details about the radar device configuration and sensitivity are provided in online Supporting Information Methods.

The variables measured by the WRIC system were recorded by customized software developed in LabVIEW (CalRQ, MEI Research). The equations for calculating VCO₂ and VO₂ are described in online Supporting Information Methods. Respiratory exchange ratio (RER) was calculated as the ratio of VCO₂ divided by VO₂, whereas the metabolic rate was calculated using the equations of Lusk [14]. Sleeping EE was calculated based on the average EE spent between 01:00 and 05:00 when SPA <1.5%. The average SPA for each chamber is shown over the time course during the chamber in Supporting Information Figure S2. Awake and fed thermogenesis (AFT), as a measure of a subject's EE taking into account the cost of being awake and in the fed state (thermic effect of food) during daytime hours, was calculated by taking the difference between EE in the inactive state and sleeping EE, as previously described [15]. EE measurements were extrapolated to 24 hours. From the 24-h RER, 24-h carbohydrate and lipid oxidation was calculated accounting for 24-h protein oxidation, which was obtained by measuring 24-h urinary nitrogen excretion as previously described [16].

The accuracy of each calorimeter system was regularly verified prior to and during the study period using two methods: combustion of propane and infusion of a mixture of dry gases (nitrogen and CO₂) of duration of at least 19 hours. Shorter infusions of 13 hours were performed as an additional check in between the longer validation studies. For infusions, a gas blender (MEI Research) that incorporated calibrated MFCs (Alicat Scientific) was used to simulate oxygen consumption rate (VO₂) and carbon dioxide production rate (VCO₂) that are typical of past similar human WRIC studies on the CRU. The approximate height of the propane flame targeted VCO₂ similar to past participants on the CRU.

Statistical analysis

Analyses were performed in SAS 9.4 (SAS Institute Inc.). Normally distributed data are presented as mean ± standard deviation (SD), whereas skewed distributions are shown as median with interquartile range. Agreement between first chamber A and chamber B (A1 and B1) as well as second chamber A and chamber B (A2 and B2) was evaluated by Pearson correlation and Bland-Altman plots.

The reproducibility of outcome measurements in duplicate chambers was quantified by the CV and intraclass correlation coefficient (ICC). The intraindividual CVs were calculated as the standard deviation divided by the mean for each variable. ICCs were estimated to

TABLE 1 Participant characteristics

Variable	Value
Sex, n (%)	
Female	5 (56)
Male	4 (44)
Age (y)	35.5 ± 9.8
Race and ethnicity, n (%)	
Indigenous American	7 (78)
White	1 (11)
Hispanic	1 (11)
Body weight (kg)	98.8 ± 29.6
Height (cm)	166.2 ± 6.0
BMI (kg/m ²)	36.0 ± 11.2
Fat-free mass (kg)	55.0 ± 6.9
Fat mass (kg)	43.8 ± 25.3
Body fat (%)	41.0 ± 14.7
Plasma glucose, fasting (mg/dL)	92.8 ± 6.8
Plasma glucose, 2-hour (mg/dL)	122.6 ± 22.8
Energy intake (kcal/d)	2,311 ± 321

Note: Data are presented as mean ± SD or n (%).

assess the reproducibility of the chamber measurements using linear mixed effects models. The modeling approach consisted of three-level random effects models to account for clustering of the data: observations nested within chambers (A/B) nested within participants. We calculated ICCs for each level which reflected the proportion of total variance in each outcome accounted for by the variation between participants (ICC_{participant}, level 3) and between chambers within participant (ICC_{chamber}, level 2).

RESULTS

Results of accuracy studies using dry gas infusions and combustion are provided in Supporting Information Table S1. Both infusions and combustion provided comparable results. Participant characteristics are presented in Table 1. Participants ate all food prescribed (i.e., actual intake was equal to prescribed intake). Respiratory chamber measurements are presented in Supporting Information Table S2. The average percentage deviation from energy balance in chambers A and B ranged from 5.7% to 8.2%. Weights at chamber entry and exit were highly similar during the study (average CV of 0.36% and 0.47%, respectively), indicating that participants were close to energy balance during the study period.

Within-chamber agreement

The Bland-Altman plots for agreement between the first and second chamber A as well as between the first and second chamber B for all

chamber measurements are displayed in Supporting Information Figures S3 to S5. The CVs for chamber variables are shown in Table 2. The CVs of VCO₂, VO₂, 24-h EE, RER, energy expenditure at zero activity (EE_{0-activity}), and sleep EE showed high reproducibility with CVs less than 5%. The CVs for SPA, AFT, carbohydrate oxidation, lipid oxidation, and protein oxidation were moderate, ranging between 10% and 25%.

Between-chamber agreement

Figure 1A-C shows Bland-Altman plots between chambers A1 and B1 and Figure 1D-F shows similar plots between chambers A2 and B2 for VCO₂, VO₂, and 24-h EE measurements. For RER, carbohydrate oxidation, lipid oxidation, and protein oxidation measurements, Bland-Altman plots between chambers A1 and B1 and between chambers A2 and B2 are shown in Figures 2A-D and Figures 2E-H, respectively. Between-chamber Bland-Altman plots for SPA, AFT, EE_{0-activity}, and sleep EE are displayed Figure 3A-H. Visual assessment of the Bland-Altman plots for SPA (Figure 3A,E) indicated a bias with higher measured SPA in chamber A compared with chamber B. The correlation between the differences and averages of SPA was significant in A1 versus B1 ($r = -0.67$, $p = 0.05$) although this was likely due to an outlier because there was no association for A2 versus B2 (Supporting Information Table S3). The correlation between differences and averages of sleep EE was also significant in A2 and B2 ($r = 0.75$, $p = 0.02$) although this was not observed in the comparison between A1 and B1 (Supporting Information Table S3). No proportional biases were observed between chambers for other measured variables, as indicated by the absence of correlation between differences and averages of these variables (Supporting Information Table S3).

The CVs of VCO₂, VO₂, 24-h EE, RER, EE_{0-activity}, and sleep EE between the first chamber A and B measurements (CVs: 1.83%, 2.45%, 2.10%, 2.51%, 3.22%, and 3.07%) and between the second chamber A and B measurements (CVs: 3.46%, 3.57%, 3.45%, 1.90%, 3.91%, 5.46%) were all below 5.5%, indicating good precision across chambers for these measurements (Table 2). The CVs between chambers for carbohydrate oxidation, lipid oxidation, protein oxidation, SPA, and AFT were higher and ranged from 9.24% to 38.07% (Table 2).

The ICC values for the chamber variables are also shown in Table 2. The ICC_{participant} and ICC_{chamber} estimates reflect the proportion of total variance in repeated measurements that occurs between participants (level 3) and between chambers (A/B) within participants (level 2), respectively. ICC_{participant} is the correlation among repeated measurements within the participant; ICC_{chamber} is the correlation among repeated measurements from the same chamber and within the same participant. The ICC_{participant} estimates for VCO₂, VO₂, 24-h EE, and EE_{0-activity} were all estimated to be over 0.95, indicating that almost all of the variance can be attributed to variation between participants and that there is excellent intraindividual consistency across repeated measurements. Because almost all of the variation in VCO₂, VO₂, 24-h EE, and EE_{0-activity} was accounted for by variation between

TABLE 2 Coefficient of variation and intraclass correlation coefficient for measurements from both respiratory chambers

Chamber measure	A1 and A2 CV (%)	B1 and B2 CV (%)	A1 and B1 CV (%)	A2 and B2 CV (%)	ICC _{participant}	ICC _{chamber}
VCO ₂ (L/d)	2.82	2.28	1.83	3.46	0.97	0.98
VO ₂ (L/d)	1.94	2.03	2.45	3.57	0.97	0.99
24-ho EE (kcal/d)	2.02	1.89	2.10	3.45	0.98	0.99
RER (ratio)	1.50	2.09	2.51	1.90	0.35	0.53
Carbohydrate oxidation (kcal/d)	10.17	11.48	13.64	10.84	0.24	0.41
Lipid oxidation (kcal/d)	12.48	24.39	24.17	20.31	0.84	0.88
Protein oxidation (kcal/d)	11.98	14.30	9.24	19.44	0.50	0.69
SPA (%)	14.43	14.33	38.07	37.69	0.23	0.88
AFT (kcal/15 h)	21.35	17.10	21.09	25.08	0.55	0.63
EE _{0-activity} (kcal/15 h)	3.44	2.10	3.22	3.91	0.96	0.97
Sleep EE (kcal/d)	3.47	4.94	3.07	5.46	0.91	0.92

Abbreviations: A1, first chamber A; A2, second chamber A; AFT, awake and fed thermogenesis; B1, first chamber B; B2, second chamber B; CV, coefficient of variation; EE, energy expenditure; EE_{0-activity}, energy expenditure at zero activity; ICC, intraclass correlation coefficient; RER, respiratory exchange ratio; SPA, spontaneous physical activity; VCO₂, carbon dioxide production rate; VO₂, oxygen consumption rate.

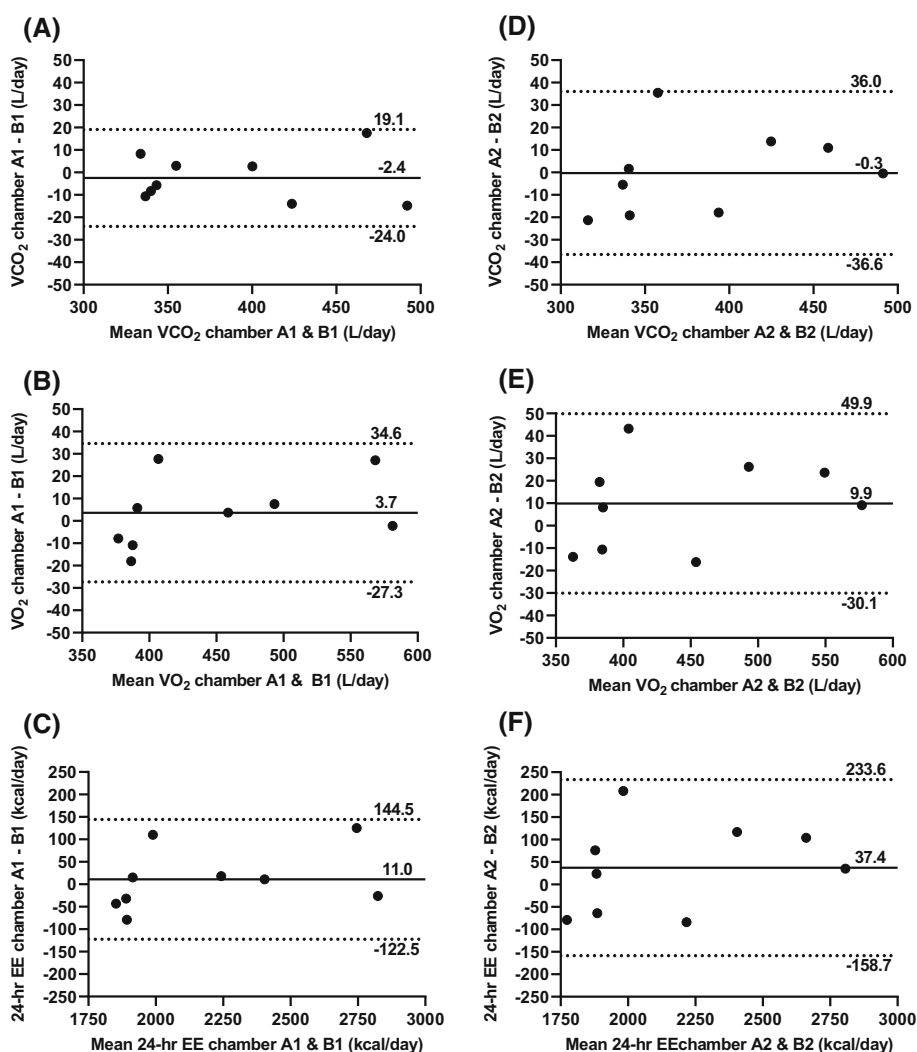


FIGURE 1 Bland-Altman plots showing agreement between chambers A and B: (A) VCO₂ for A1 and B1, (B) VO₂ for A1 and B1, (C) 24-h energy expenditure (EE) for A1 and B1, (D) VCO₂ for A2 and B2, (E) VO₂ for A2 and B2, and (F) 24-h EE for A2 and B2. Solid line represents average difference between the two measurements, whereas dotted lines represent upper and lower limits of agreement (mean and 95% confidence interval). VCO₂, carbon dioxide production rate; VO₂, oxygen consumption rate

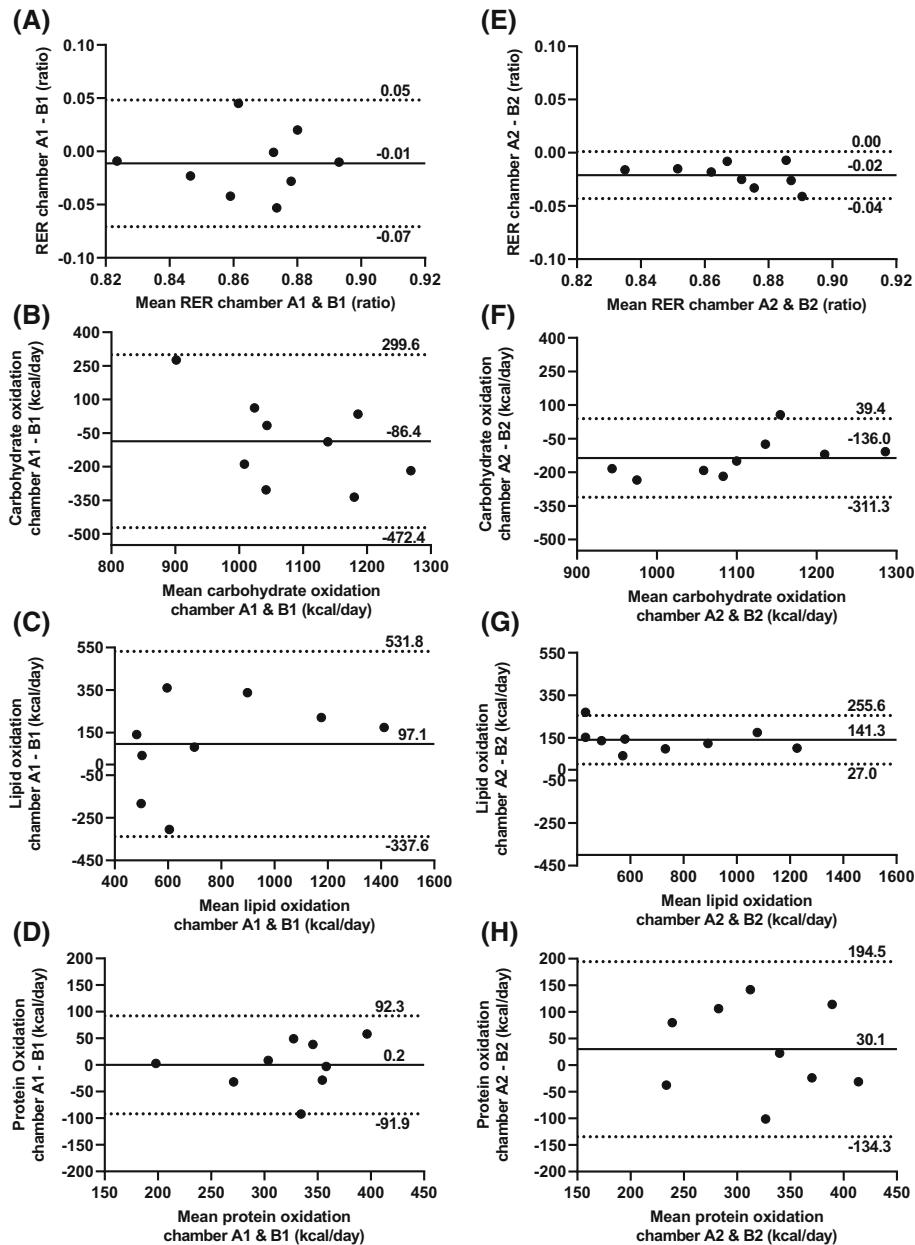


FIGURE 2 Bland-Altman plots showing agreement between chambers A and B: (A) respiratory exchange ratio (RER) for A1 and B1, (B) carbohydrate oxidation for A1 and B1, (C) lipid oxidation for A1 and B1, (D) protein oxidation for A1 and B1, (E) RER for A2 and B2, (F) carbohydrate oxidation for A2 and B2, (G) lipid oxidation for A2 and B2, and (H) protein oxidation for A2 and B2. Solid line represents average difference between the two measurements, whereas dotted lines represent upper and lower limits of agreement (mean and 95% confidence interval).

participants, the ICC_{chamber} estimates were nearly identical, as there is little to no additional variance to account for, further indicating that repeated measurements within the same participant, regardless of chamber, are highly consistent. The ICC estimates for lipid oxidation ($ICC_{\text{participant}} = 0.84$, $ICC_{\text{chamber}} = 0.88$), AFT ($ICC_{\text{participant}} = 0.55$, $ICC_{\text{chamber}} = 0.63$), and sleep EE ($ICC_{\text{participant}} = 0.91$, $ICC_{\text{chamber}} = 0.92$) were all relatively similar and showed moderate to high reproducibility, with differences between participants accounting for the majority of variance in measurements. In contrast, the ICC estimates were lower for RER ($ICC_{\text{participant}} = 0.35$, $ICC_{\text{chamber}} = 0.53$),

carbohydrate oxidation ($ICC_{\text{participant}} = 0.24$, $ICC_{\text{chamber}} = 0.41$), protein oxidation ($ICC_{\text{participant}} = 0.50$, $ICC_{\text{chamber}} = 0.69$), and SPA ($ICC_{\text{participant}} = 0.23$, $ICC_{\text{chamber}} = 0.88$) compared with other chamber measurements, and a proportion of the variance in measurements was attributable to variability between chambers within participants, reflective of poorer reproducibility among chambers.

As shown in Table 2, chamber measures with excellent CVs generally had high $ICC_{\text{participant}}$, except for RER, which had excellent CV but low $ICC_{\text{participant}}$. Compared with total variation in other chamber measures (between chambers, Figures 4-6; within

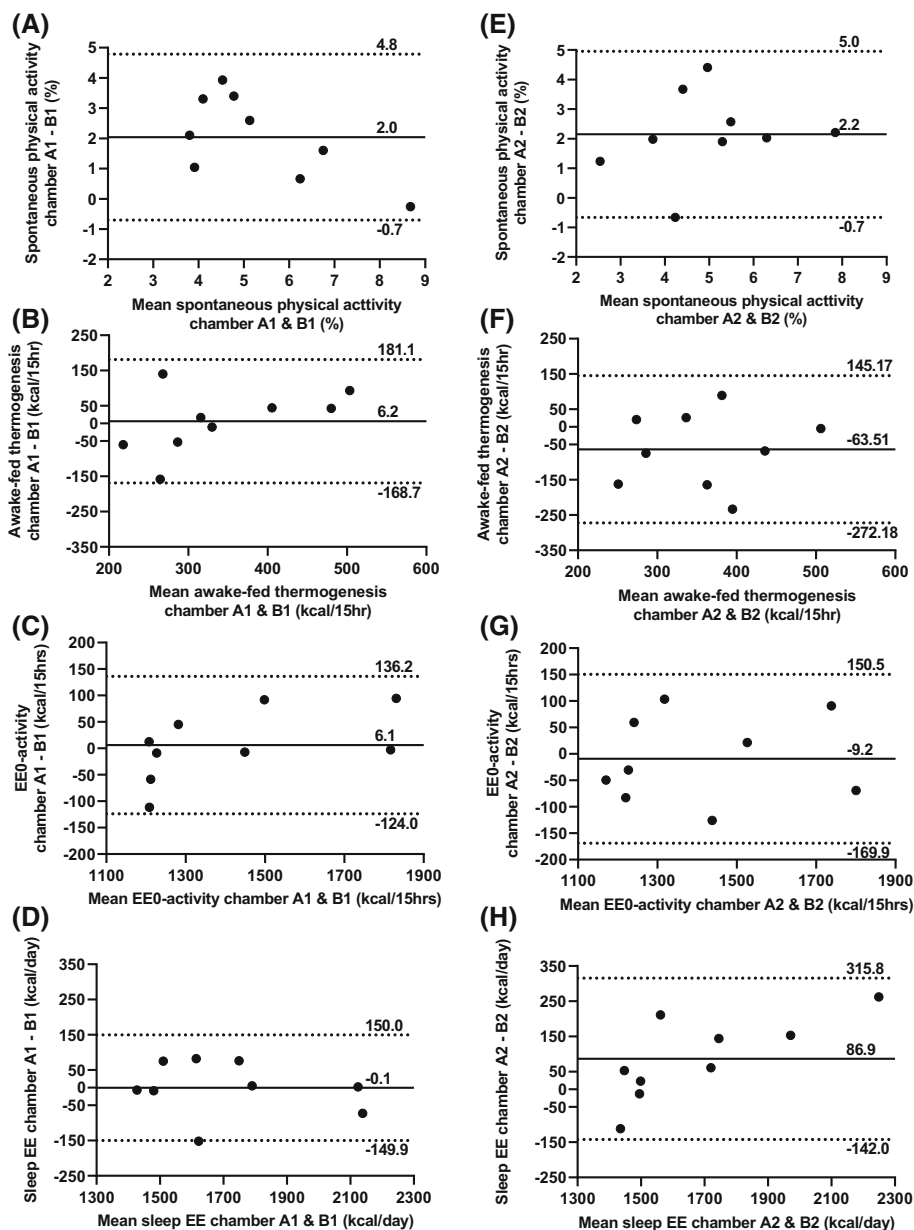


FIGURE 3 Bland-Altman plots showing agreement between chambers A and B: (A) spontaneous physical activity for A1 and B1, (B) awake-fed thermogenesis for A1 and B1, (C) energy expenditure at zero activity (EE0-activity) for A1 and B1, (D) sleep energy expenditure (EE) for A1 and B1, (E) spontaneous physical activity for A2 and B2, (F) awake-fed thermogenesis for A2 and B2, (G) EE0-activity for A2 and B2, and (H) sleep EE for A2 and B2. Solid line represents average difference between the two measurements, whereas dotted lines represent upper and lower limits of agreement (mean and 95% confidence interval).

chambers, Supporting Information Figures S6-S8), total variation for RER (Figure 5A,E; Supporting Information Figure S7A,E) was small, indicating that the low $ICC_{\text{participant}}$ is due to this limited variation.

DISCUSSION

This study compared the accuracy of repeated measurements within and between two WRICs. As in prior studies, 24-h EE showed high precision

in each chamber. The current investigation further demonstrated that 24-h EE, along with VO_2 and VCO_2 , can be accurately reproduced in different WRICs with different chamber volumes, a crucial variable used in the calculations. By studying the same participants in both WRICs, the results indicate that it is feasible to obtain comparable results between different WRICs with respect to a common outcome measure.

We expected that VO_2 and VCO_2 would be highly reproducible between chambers since system accuracy was verified based on expected VO_2 and VCO_2 from two methods (i.e., dry gas infusions and propane combustion). The high reproducibility of 24-h EE is

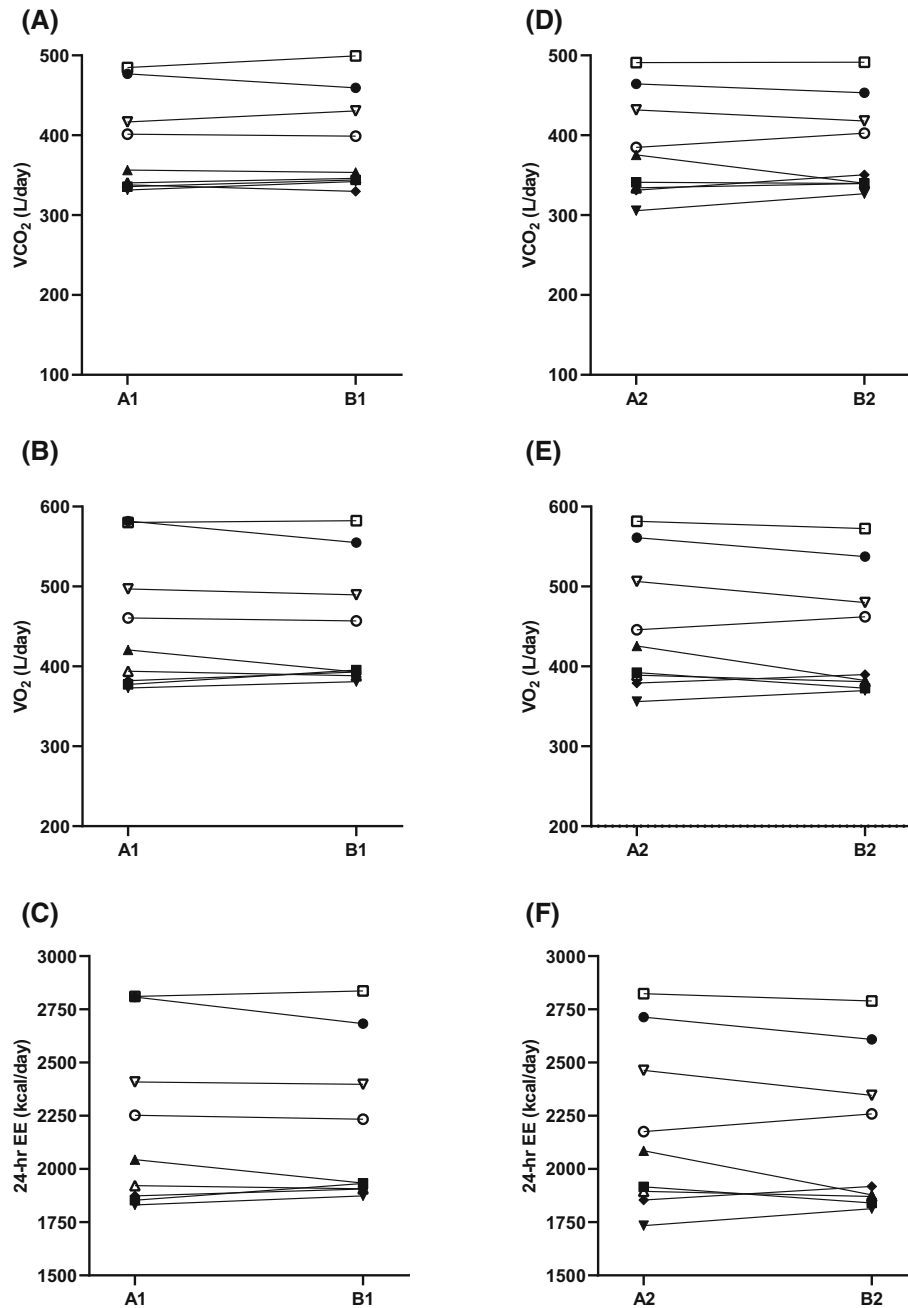


FIGURE 4 Repeated measures between chambers A and B: (A) VCO₂ for A1 and B1, (B) VO₂ for A1 and B1, (C) 24-h energy expenditure (EE) for A1 and B1, (D) VCO₂ for A2 and B2, (E) VO₂ for A2 and B2, and (F) 24-h EE for A2 and B2. VCO₂, carbon dioxide production rate; VO₂, oxygen consumption rate (VO₂). Symbols denote different subjects.

consistent with the use of these methods to verify system accuracy. Since the verification methods followed recently published guidance [2], the results demonstrate that cross-chamber reproducibility can be achieved by following these operational practices. These recent guidelines also recommended that combustion results be the minimum standard establishing system accuracy and that infusions be considered an additional measurement. In the current study, dry gas infusions were consistent with combustion (Supporting Information Table S1), indicating that the recommendation to require combustion may be overly conservative. Requiring only one method would save

resources. Unlike combustion, dry gas infusions also allowed us to target VO₂ and VCO₂ rates similar to humans and avoid the need for safety monitoring of an open flame during combustion; however, infusions require a relatively expensive blender with calibrated MFCs, whereas combustion requires only a high-precision scale. Moreover, because the flame for combustion of pure gases is typically set by eye, there is variation in expected VO₂ and VCO₂ rates, whereas the rates for infusions can be precisely targeted. Thus, the recommendation for dry gas infusion alone should be considered as an alternative minimum standard besides combustion.

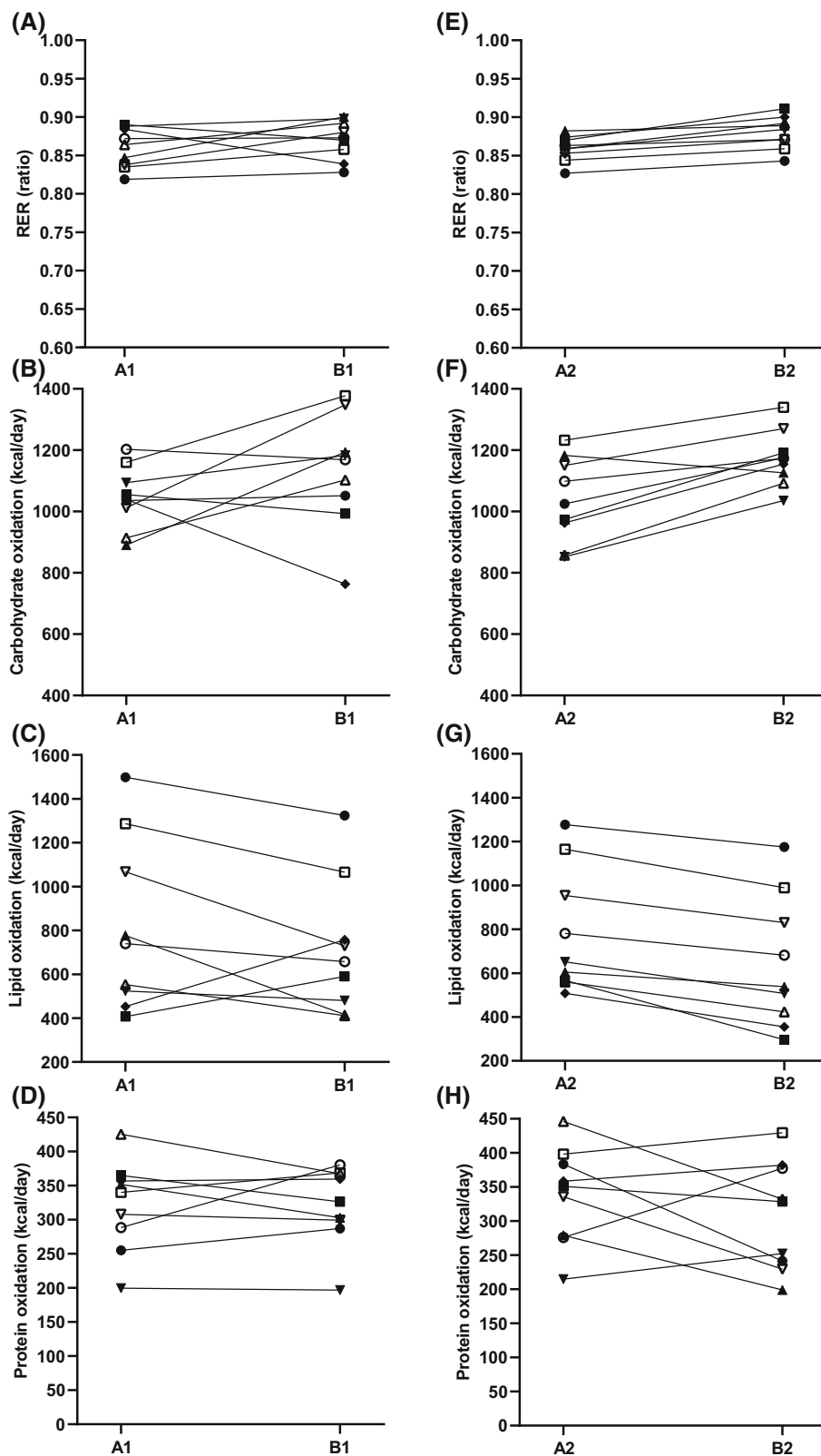


FIGURE 5 Repeated measures between chambers A and B: (A) respiratory exchange ratio (RER) for A1 and B1, (B) carbohydrate oxidation for A1 and B1, (C) lipid oxidation for A1 and B1, (D) protein oxidation for A1 and B1, (E) RER for A2 and B2, (F) carbohydrate oxidation for A2 and B2, (G) lipid oxidation for A2 and B2, and (H) protein oxidation for A2 and B2. Symbols denote different subjects.

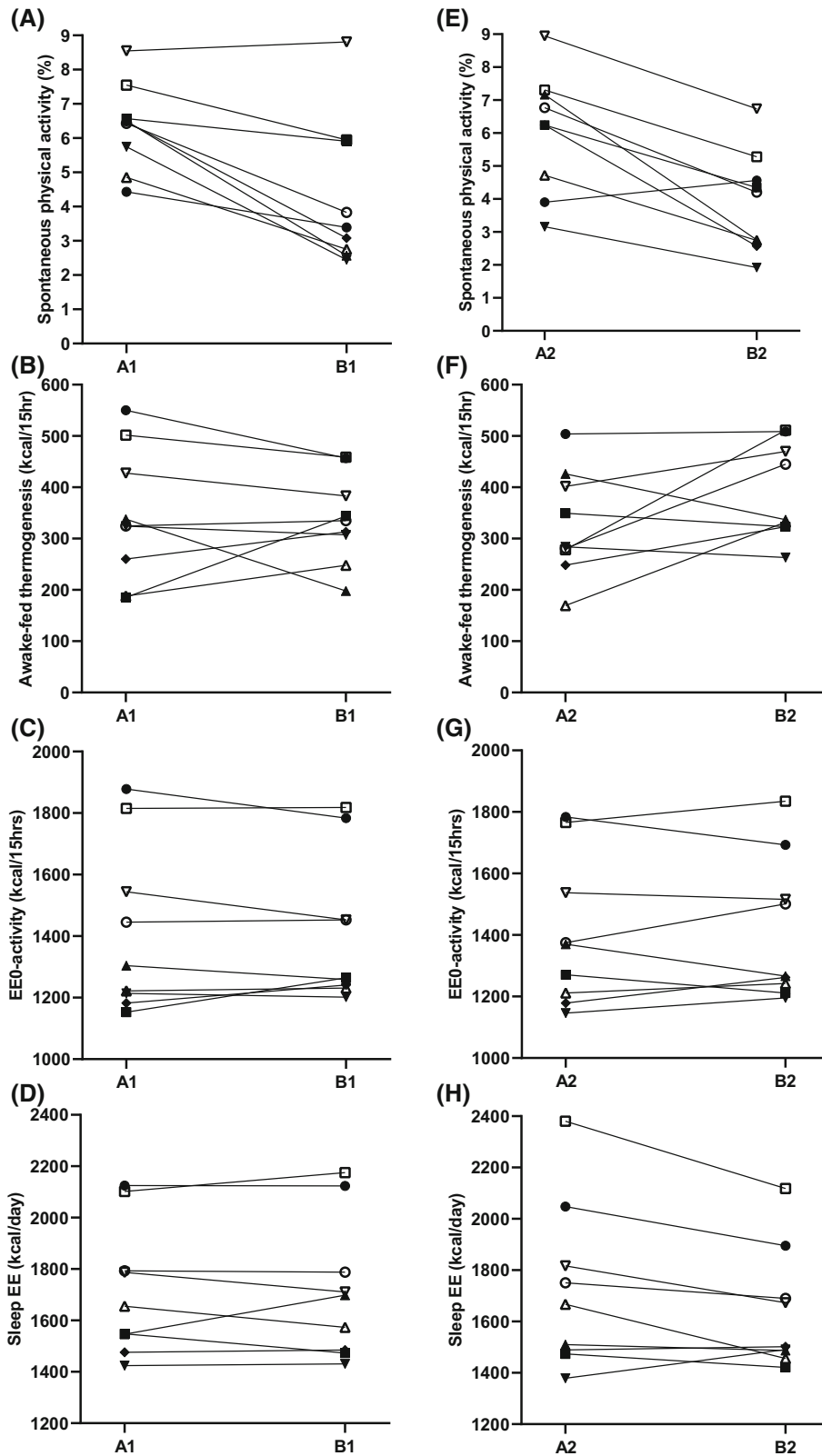


FIGURE 6 Repeated measures between chambers A and B: (A) spontaneous physical activity for A1 and B1, (B) awake-fed thermogenesis for A1 and B1, (C) energy expenditure at zero activity (EEO-activity) for A1 and B1, (D) sleep energy expenditure (EE) for A1 and B1, (E) spontaneous physical activity for A2 and B2, (F) awake-fed thermogenesis for A2 and B2, (G) EEO-activity for A2 and B2, and (H) sleep EE for A2 and B2. Symbols denote different subjects.

RER had excellent CVs but low ICC_{participant} values. The CVs are in line with what we previously found [17]. RER values were consistent within an individual in absolute terms, as reflected by the excellent CVs, but the repeated values are not strongly correlated with each other. This is due to the small total variance in RER, which is reflected in the lower ICCs. The interindividual variance in RER is small because RER depends primarily on expected food quotient, which was constant for this study. Moreover, the mean RER of the volunteers' first chamber session (0.87) was close to the expected food q(0.86), thus confirming the validity of our metabolic measurements. Thus, CVs for RER are a better reflection of reproducibility in this situation.

Macronutrient substrate oxidation rates were relatively less reproducible than other VO₂, VCO₂, and 24-ho EE. The lower reproducibility is likely due the additional calculations required for these measurements. The additional deviations in urine collection or measurements of urea plus the higher complexity of the estimate in the cases of lipid and carbohydrate oxidation add to variation in these measurements. Assessment of protein oxidation relied on urine collection and measurement of urine urea concentration whereas lipid and carbohydrate oxidation rates rely on indirect calorimetry in addition to those urinary measurements. Since markers for completeness of 24-h urine collection were not available, it is unclear whether missed voids were influencing these estimates. However, because the equipment to collect urine is readily available for the volunteer within the chamber, incomplete urine collection is relatively minimized compared with studies requiring outpatient collection of urine. Since sleeping and lights-out periods were established by the participant, subtle variation between the chambers possibly affecting sleep may contribute to increased variation in macronutrient substrate oxidation rates.

Physical activity measured by radar showed moderate to high reproducibility within each chamber but differed substantially between the chambers. The reasons are unclear although it may be due to differences in sensitivity of the motion sensors in chambers of different size. Despite this, EE_{0-activity}, which depends on measurements of EE and activity by radar, was highly reproducible within and between chambers. This suggests that the activity detection systems in the chambers agree with each other at zero or very low activity levels. Since accelerometry using wearable devices was not assessed in these participants, it is uncertain whether reproducibility could be improved using alternate methods of measuring physical activity. Moreover, if there are differences in actual SPA between chambers A and B (i.e., not due to differences in radar sensitivity), these differences may be contributing to variability in macronutrient oxidation rates observed between chambers. Humidity differences between chamber A and B were also observed (Supporting Information Table S2). The differences in humidity may be due to differences in HVAC design and construction between the chambers. Although it is unlikely to influence substantial measurement error because the sample gas dryers achieved desirable levels of dryness for each chamber, it is possible that humidity differences influence human behavior in the chamber.

The current study has several limitations. Because the two calorimeters are from the same laboratory, the reproducibility estimates may represent an ideal situation as the chambers were operated by

the same research staff, using the same device manufacturers and sharing the same inflow air source. It is unclear whether results would differ if the study was conducted at different geographical sites and markedly different operating characteristics (e.g., push vs. pull configuration; different data acquisition systems) compared with the current study. In addition, menstrual cycle information was not available for the women in this study. Since the chambers were done over a maximum of an 18-day span, it is possible that women were in different menstrual phases when they were in the successive chambers. Menstrual phase has been reported to be associated with differences in EE [18]. We attempted to minimize this potential effect by varying the sequence of the chambers. It is possible that studying the women in the same menstrual phase would have produced higher reproducibility estimates.

In summary, we found that cross-chamber reproducibility is high for many measurements often studied in the respiratory chamber. The chambers, though not identical, were operated and validated in a similar manner. These results support recent efforts to promote standardization across chambers.○

AUTHOR CONTRIBUTIONS

Emma J. Stinson, Theresa Rodzevik, Douglas C. Chang: conceptualization, methodology, formal analysis, investigation, writing (original draft), and visualization. Jonathan Krakoff, Paolo Piaggi: conceptualization, methodology, formal analysis, investigation, and writing (review and editing). Douglas C. Chang is the guarantor of the work and, as such, had full access to all the study data and takes responsibility for the integrity of the data and the accuracy of the data analysis.

ACKNOWLEDGMENTS

The authors thank the volunteers who participated in the study. They also thank the nursing, clinical, dietary, and laboratory staff of the Phoenix Epidemiology and Clinical Research Branch for assistance conducting the study visits and for care of the volunteers.

CONFLICT OF INTEREST

The authors declared no conflict of interest.

ORCID

Paolo Piaggi  <https://orcid.org/0000-0003-2774-9161>

REFERENCES

1. Hall KD, Chen KY, Guo J, et al. Energy expenditure and body composition changes after an isocaloric ketogenic diet in overweight and obese men. *Am J Clin Nutr*. 2016;104:324-333.
2. Chen KY, Smith S, Ravussin E, et al. Room Indirect Calorimetry Operating and Reporting Standards (RICORS 1.0): a guide to conducting and reporting human whole-room calorimeter studies. *Obesity (Silver Spring)*. 2020;28:1613-1625.
3. Ravussin E, Lillioja S, Anderson TE, Christin L, Bogardus C. Determinants of 24-hour energy expenditure in man. Methods and results using a respiratory chamber. *J Clin Invest*. 1986;78:1568-1578.
4. Murgatroyd PR, Davies HL, Prentice AM. Intra-individual variability and measurement noise in estimates of energy expenditure by whole body indirect calorimetry. *Br J Nutr*. 1987;58:347-356.

5. Astrup A, Thorbek G, Lind J, Isaksson B. Prediction of 24-h energy expenditure and its components from physical characteristics and body composition in normal-weight humans. *Am J Clin Nutr.* 1990; 52:777-783.
6. Rumpler WV, Seale JL, Conway JM, Moe PW. Repeatability of 24-h energy expenditure measurements in humans by indirect calorimetry. *Am J Clin Nutr.* 1990;51:147-152.
7. White MD, Bouchard G, Buemann B, et al. Reproducibility of 24-h energy expenditure and macronutrient oxidation rates in an indirect calorimeter. *J Appl Physiol (1985).* 1996;80:133-139.
8. Toubro S, Christensen NJ, Astrup A. Reproducibility of 24-h energy expenditure, substrate utilization and spontaneous physical activity in obesity measured in a respiration chamber. *Int J Obes Relat Metab Disord.* 1995;19:544-549.
9. Ferraro R, Boyce VL, Swinburn B, De Gregorio M, Ravussin E. Energy cost of physical activity on a metabolic ward in relationship to obesity. *Am J Clin Nutr.* 1991;53:1368-1371.
10. Genuth S, Alberti KG, Bennett P, et al. Follow-up report on the diagnosis of diabetes mellitus. *Diabetes Care.* 2003;26:3160-3167.
11. Abbott WG, Howard BV, Christin L, et al. Short-term energy balance: relationship with protein, carbohydrate, and fat balances. *Am J Physiol.* 1988;255:E332-E337.
12. Chen S, Wohlers E, Ruud E, Moon J, Ni B, Celi FS. Improving temporal accuracy of human metabolic chambers for dynamic metabolic studies. *PLoS One.* 2018;13:e0193467.
13. Schoffelen PFM, Plasqui G. Classical experiments in whole-body metabolism: open-circuit respirometry-diluted flow chamber, hood, or facemask systems. *Eur J Appl Physiol.* 2018;118:33-49.
14. Lusk G. Animal calorimetry twenty-fourth paper. Analysis of the oxidation of mixtures of carbohydrate and fat. *J Biol Chem.* 1924;59: 41-42.
15. Piaggi P, Krakoff J, Bogardus C, Thearle MS. Lower "awake and fed thermogenesis" predicts future weight gain in subjects with abdominal adiposity. *Diabetes.* 2013;62:4043-4051.
16. Jequier E, Acheson K, Schutz Y. Assessment of energy expenditure and fuel utilization in man. *Annu Rev Nutr.* 1987;7:187-208.
17. Thearle MS, Pannacciulli N, Bonfiglio S, Pacak K, Krakoff J. Extent and determinants of thermogenic responses to 24 hours of fasting, energy balance, and five different overfeeding diets in humans. *J Clin Endocrinol Metab.* 2013;98:2791-2799.
18. Bisdee JT, James WP, Shaw MA. Changes in energy expenditure during the menstrual cycle. *Br J Nutr.* 1989;61:187-199.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Stinson EJ, Rodzevik T, Krakoff J, Piaggi P, Chang DC. Energy expenditure measurements are reproducible in different whole-room indirect calorimeters in humans. *Obesity (Silver Spring).* 2022;30(9):1766-1777. doi:[10.1002/oby.23476](https://doi.org/10.1002/oby.23476)