Original article

Validity of four commercially available metabolic carts for assessing resting metabolic rate and respiratory exchange ratio in non-ventilated humans

J.M.A. Alcantara [a, *], J.E. Galgani [b, c], L. Jurado-Fasoli [a], M. Dote-Montero [a], E. Merchan-Ramirez [a], E. Ravussin [d], J.R. Ruiz [a, e, 1], G. Sanchez-Delgado [a, d, 1]

[a] PROFITH "PROmoting FITness and Health Through Physical Activity" Research Group, Sport and Health University Research Institute (iMUDS), Department of Physical and Sports Education, Faculty of Sport Sciences, University of Granada, 18011 Granada, Spain
[b] Department of Health Sciences, Nutrition and Dietetics Career, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile
[c] Department of Nutrition, Diabetes and Metabolism, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile
[d] Pennington Biomedical Research Center, Baton Rouge, LA 70808, USA
[e] Instituto de Investigación Biosanitaria, ibs.Granada, Granada, Spain

ARTICLE INFO

SUMMARY

Background & aims: The validity of most commercially available metabolic cart is mostly unknown. Thus, we aimed to determine the accuracy, precision, within-subject reproducibility, and concordance of RMR and RER measured by four commercially available metabolic carts [Cosmed Q-NRG, Vyaire Vyntus CPX, Maastricht Instruments Omnical, and Medgraphics Ultima CardiO2]. Further, we studied whether a previously proposed simulation-based post-calorimetric calibration of cart readouts [individual calibration control evaluation (ICcE)] modify the RMR and RER reproducibility and concordance.

Methods: Three experiments simulating different RMR and RER by controlled pure gas (N_2 and CO_2) infusions were conducted on 5 non-consecutive days. Moreover, 30-min methanol burns were performed on 3 non-consecutive days. Lastly, the RMR and RER of 29 young non-ventilated adults (11 women; 25 ± 4 years-old; BMI: 24.1 ± 3.2 kg/m^2) were assessed twice using each instrument, 24 hours apart, under standardized conditions.

Results: The Omnical presented the lowest measurement error for RER (Omnical = $1.7 \pm 0.9\%$; Vyntus = $4.5 \pm 2.0\%$; Q-NRG = $6.6 \pm 1.9\%$; Ultima = $6.8 \pm 6.5\%$) and EE (Omnical = $1.5 \pm 0.5\%$; Q-NRG = $2.5 \pm 1.3\%$; Ultima = $10.7 \pm 11.0\%$; Vyntus = $13.8 \pm 5.0\%$) in all in vitro experiments (controlled pure gas infusions and methanol burns). In humans, the 4 metabolic carts provided discordant RMR and RER estimations (all P < 0.001). No differences were detected in RMR within-subject reproducibility (P = 0.058; Q-NRG inter-day coefficient of variance = $3.6 \pm 2.5\%$; Omnical = $4.8 \pm 3.5\%$; Vyntus = $5.0 \pm 5.6\%$; Ultima = $5.7 \pm 4.6\%$), although the Ultima CardiO2 provided larger RER inter-day differences ($4.6 \pm 3.5\%$) than the others carts (P = 0.001; Omnical = $1.9 \pm 1.7\%$; Vyntus = $2.1 \pm 1.3\%$; Q-NRG = $2.4 \pm 2.1\%$). The ICcE procedure did not modify the RMR or RER concordance and did not reduce the inter-day differences in any of the carts.

Conclusions: The 4 metabolic carts provided discordant measurements of RMR and RER. Overall, the Omnical provides more accurate and precise estimations of RMR and RER than the Q-NRG, Vyntus and Ultima CardiO2, and might be considered the best for assessing RMR and RER in non-ventilated humans. Finally, our results do not support the use of an ICcE procedure.

* Corresponding author.
   E-mail address: alcantarajma@ugr.es (J.M.A. Alcantara).
[1] These authors contributed equally to this work.

List of abbreviations

ANOVA	Repeated-measures analyses of variance
BMI	Body mass index
CO_2	Carbon dioxide
CV	Coefficient of variation
CV_{D-to-D}	Day-to-day coefficient of variation
DTC	Deltatrac metabolic cart (Datex Instrumentarium Corp, Helsinki, Finland)
EE	Energy expenditure
ICcE	Individual calibration control evaluation
N_2	Nitrogen
O_2	Oxygen
Omnical	Omnical metabolic cart (Maastricht Instruments, Maastricht, The Netherlands)
Q-NRG	Q-NRG metabolic cart (Cosmed, Rome, Italy)
RER	Respiratory exchange ratio
RMR	Resting metabolic rate
STPD	Standard Temperature, Pressure, and Dry conditions
Ultima	Ultima CardiO2 metabolic cart (Medgraphics Corporation, St. Paul, MN, USA)
$\dot{V}CO_2$	Carbon dioxide production
$\dot{V}O_2$	Oxygen consumption
Vyntus	Vyntus CPX metabolic cart (Vyaire, Höchberg, Germany)

1. Introduction

The resting metabolic rate (RMR) is defined as the energy needed for maintaining a normal body function in an awake person resting in thermoneutrality [1]. Indirect calorimetry is the reference method for assessing human RMR [2–4] via measurement of oxygen consumption ($\dot{V}O_2$) and carbon dioxide production ($\dot{V}CO_2$). Indirect calorimetry also allows to determine the respiratory exchange ratio (RER), which gives information about the macronutrients being oxidized [5–7]. Metabolic carts are the most used indirect calorimeters for assessing RMR and RER in clinical and research settings [8,9]. The Deltatrac (DTC; Datex Instrumentarium Corp, Helsinki, Finland), the often-preferred metabolic cart, has been for long considered the reference metabolic cart for assessing RMR and RER in humans [3,10–15]. However, the DTC is no longer manufactured [2,10,13–16], and no other metabolic cart has been recognized yet as the new preferred instrument [15].

The validity of a metabolic cart can be determined [17] by (i) assessing its accuracy (i.e. the proximity of measurements to traceable standards [18]) and precision (i.e. the variability in repeated measures of the same magnitude [18]) by controlled pure gas infusions [nitrogen (N_2) and CO_2]; (ii) assessing its accuracy and precision by alcohol burning tests, and; (iii) assessing the within-subject reproducibility (i.e. the variability in repeated measures performed in individuals under the same conditions, thereinafter called *reproducibility*). For many of the commercially available metabolic carts, there is no published validity data for RMR and RER while some metabolic carts have provided unacceptable accuracy, precision, and/or reproducibility [13,19]. Of note, most studies examining the accuracy and/or precision of different metabolic carts have not compared them within the same settings and conditions [15] or have not used recently manufactured metabolic carts [19].

To improve the validity of RMR and RER measurements, Schadewaldt et al. [2] proposed a simulation-based post-calorimetric calibration of cart readouts procedure [individual calibration control evaluation (ICcE)]. In brief, this ICcE procedure consists of simulating the subject's $\dot{V}O_2$ and $\dot{V}CO_2$ by infusing pure gases [N_2, for diluting ambient O_2, and CO_2), using high-precision mass-flow controllers, immediately after the subject's indirect calorimetry testing [2]. The subject's $\dot{V}O_2$ and $\dot{V}CO_2$ can then be "corrected" by the measured metabolic cart "error" (i.e. the difference between the infused gases and the readouts of the metabolic cart) [2]. However, whether the application of the ICcE increases the reproducibility and concordance of RMR and RER using current commercially available metabolic carts remains to be determined.

The present study was designed to determine the accuracy, precision, reproducibility, and concordance of RMR and RER assessments provided by four different and commercially available metabolic carts: the Q-NRG (Cosmed, Rome, Italy), the Vyntus CPX (Vyaire, Höchberg, Germany; thereinafter called *Vyntus*), the Omnical (Maastricht Instruments, Maastricht, The Netherlands), and the Ultima CardiO2 (Medgraphics Corporation, St. Paul, MN, USA; thereinafter called *Ultima*). Further, we assessed whether the ICcE influences the reproducibility and concordance of the RMR and RER assessments.

2. Methods

2.1. Metabolic carts and procedures

The metabolic carts were calibrated (flow and gas analyzers) by the same researchers strictly following the manufacturers' instructions. We conducted four *in vitro* validation experiments (Fig. 1A), using controlled pure gas infusion and methanol burns, and one *in vivo* experiment in young non-ventilated adults (hereinafter *human* study; Fig. 1B). Supplementary Table 1 shows detailed information and characteristics of the metabolic carts. Briefly, all metabolic carts were calibrated every testing day except the Q-NRG which was monthly calibrated strictly following the manufacturer's instructions. Further, all metabolic carts were equipped with a ventilated plastic canopy to collect the gas exchange except the Ultima which used a face-tent system.

2.2. In vitro validation

Experiments 1, 2 and 3 were based on controlled pure gas infusion, whereas experiment 4 used both gas infusion and methanol burns (Fig. 1A). In *experiment 1*, we simulated energy expenditure (EE) of 800, 1200, 1600, 2000, and 2400 kcal/day while keeping RER constant (≈ 0.85). In *experiment 2* we simulated RER of 0.75, 0.80, 0.85, 0.90 and 0.95 while keeping EE constant (≈ 1500 kcal/day). In *experiment 3* we repeated the same simulated EE (≈ 1500 kcal/day) and RER (≈ 0.85) three times, 1-hour apart. In *experiment 4*, we performed 10-minute gas infusions simulating similar $\dot{V}O_2$ and $\dot{V}CO_2$ values than achieved during previous methanol burning tests routinely conducted in our laboratory, followed by 30-minutes methanol burns. *Experiments 1–3* were conducted in 5 non-consecutive days (within 10 days) with each cart, whereas *experiment 4* was performed in 3 non-consecutive days (within 30 days).

The controlled pure gas infusions were performed using two high-precision mass-flow controllers (358 Series, Analyt-MTC, Müllheim, Germany; 0–2 l/min). One controller was used for infusing pure N_2 (purity $\geq 99.9997\%$; Carburos Metálicos/Air
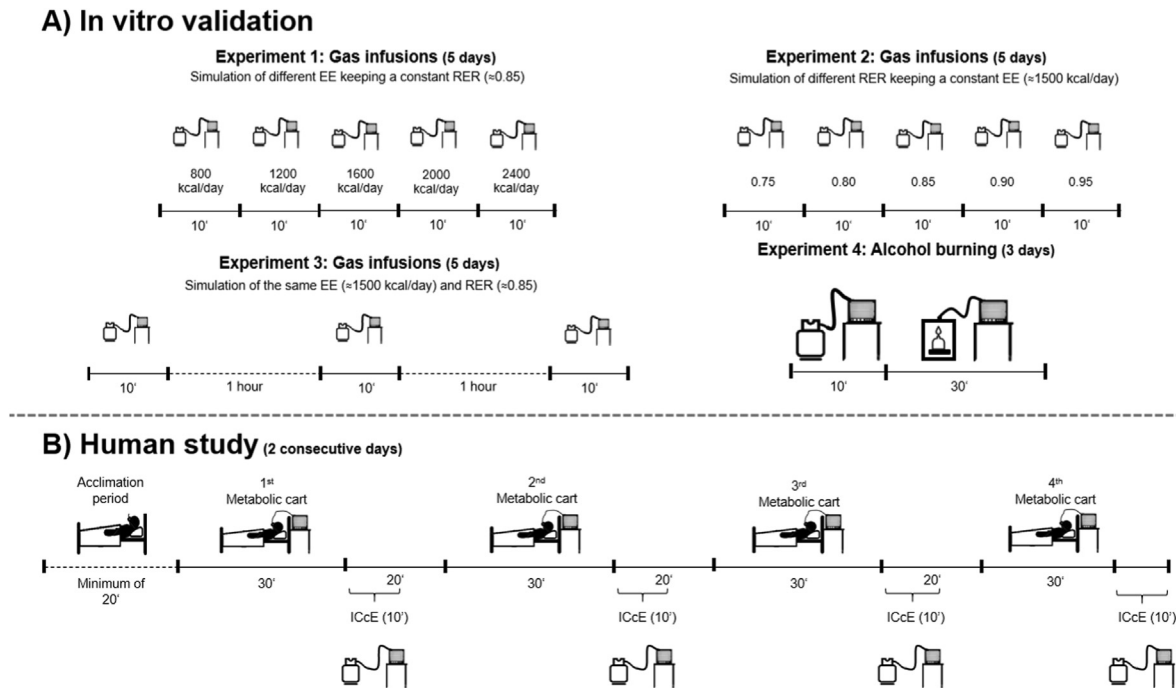
Fig. 1. Study and experiments design. Experiments 1, 2 and 3 were based on controlled pure gas infusion, whereas experiment 4 used both gas infusion and methanol burns (Panel A). The *in vivo* experiment consisted of gas exchange measurements in young non-ventilated adults (Panel B, *human* study). EE: energy expenditure (in kcal/day); RER: respiratory exchange ratio; ICcE: Individual calibration control evaluation procedure. Time is presented either as minutes or hours. The metabolic carts used were the Q-NRG (Cosmed, Rome, Italy), the Vyntus CPX (Vyaire, Höchberg, Germany), the Omnical (Maastricht Instruments, Maastricht, The Netherlands), and the Ultima CardiO2 (Medgraphics Corporation, St. Paul, MN, USA).

Products and Chemicals, Inc., Barcelona, Spain) and the other for infusing pure CO_2 (purity $\geq$99.995%; Carburos Metálicos/Air Products and Chemicals, Inc., Barcelona, Spain) directly into the hose tube of the metabolic cart [2]. Pure gas infusions lasted 10 minutes each. The first 5 minutes data of each gas infusion were discarded, and the remaining data were averaged and used for analysis. Although CO_2 is directly infused into the hose tube, the concomitant infusion of N_2 also dilutes CO_2. However, this dilution effect is negligible [2], and we assumed that the infused CO_2 equals $\dot{V}CO_2$.

N_2 infusion is used to dilute ambient O_2, and therefore, the simulated $\dot{V}O_2$ can be calculated by using the following equation [20]:

$$\dot{V}O_2\,(ml\,/\,min) = infused\,N_2\,(ml\,/\,min) \times 0.2646$$

The methanol [purity $\geq$99.9% and water $\leq$0.05% (EMSURE® ACS, ISO, Reag. Ph Eur, Merck, Darmstadt, Germany)] burning tests were performed by lighting the flame of the wick burning kit inside a methanol burning glass cage (Maastricht Instruments, Maastricht, The Netherlands) and letting the methanol burn for 30 minutes. The produced gases were continuously directed to the metabolic carts' hose tube, and the methanol weight was dynamically recorded using a calibrated scale (model MS 1602TS/00 precision scale, precision 0.01 g; Mettler Toledo, Giessen, Germany). The first 5 minutes data of the burn were discarded, and the remaining $\dot{V}O_2$ and $\dot{V}CO_2$ data were averaged for analysis. The methanol burn *expected value* considered for both $\dot{V}O_2$ and $\dot{V}CO_2$ recoveries were 100%, while the *expected value* for RER was 0.667 based on the following reaction [19,21]:

$$2CH_3OH + 3O_2 \rightarrow 2CO_2 + 4H_2O$$
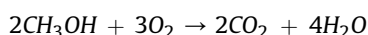
For the gas infusions and methanol burns, the measurement error was calculated by subtracting the *expected value* (i.e. infused/ simulated or produced by the burn) to the *measured value* (i.e. metabolic cart readouts). Later, we expressed this error as a percentage of the expected value (i.e. [(measured − expected)/ expected] × 100) to be used in the statistical analyses.

2.3. Human study

Twenty-nine non-ventilated young adults (see participant flowchart, Supplementary Fig. 1) participated in this observational study (Table 1). The inclusion criteria were: (i) older than 18 years; (ii) body mass index between 18.5 and 40 kg/m^2; (iii) stable body weight over the last 3 months (changes $\leq$3 kg) and not enrolled in a weight loss program; (iv) non-smokers; (v) no medication that could directly affect energy metabolism; (vi) free from chronic or acute illness; and (vii) not pregnant. All these criteria were verbally confirmed by the participants. Both, the study protocol and written informed consent followed the 2013 revised Declaration of Helsinki and were approved by the Human Research Ethics Committee of the University of Granada (n. 836).

On the first visit, height and weight were measured using a stadiometer and scale (Seca model 799, Electronic Column Scale, Hamburg, Germany) without shoes and with light clothing (replicated on both visits). Waist circumference was measured twice using a plastic tape while the subjects were in a standing position, and the average of both assessments was used. Lastly, body composition was assessed by whole-body dual-energy X-ray absorptiometry (Discovery Wi, Hologic, Inc., Bedford, MA, USA).

2.4. Indirect calorimetry assessment

Participants arrived at the research center by public transportation or motorized vehicle (avoiding any moderate or vigorous

Table 1
Subjects' characteristics.

	All (n = 29)					Men (n = 18)					Women (n = 11)				
	Mean	±	SD	Min	Max	Mean	±	SD	Min	Max	Mean	±	SD	Min	Max
Age (years)	25	±	4.3	18	36	24.9	±	4.2	18	34	25.8	±	5.0	20	36
Body weight (kg)	71.2	±	7.5	45.6	99.2	77.1	±	10.6	63.3	99.2	60.1	±	8.0	45.6	72.4
Height (cm)	171.0	±	12.9	154.6	184.5	174.9	±	5.5	160.5	184.5	164.7	±	5.6	154.6	174.3
BMI (kg/m^2)	24.1	±	3.2	19.1	31.9	25.2	±	3.4	21.6	31.9	22.1	±	2.3	19.1	25.9
Waist circumference (cm)	77.3	±	9.2	59.8	97.6	81.2	±	8.5	70.0	97.2	70.7	±	6.0	59.8	85.0
Lean mass (kg)	48.9	±	10.8	30.5	66.6	56.1	±	5.9	42.2	66.6	36.9	±	4.0	30.5	41.8
Fat mass (kg)	17.9	±	7.0	8.6	36.7	16.7	±	7.6	8.6	36.7	19.9	±	5.3	10.8	28.5
Fat mass (%)	26.0	±	8.6	13.5	41.1	21.5	±	7.2	13.5	40.4	33.4	±	5.6	24.2	41.1

SD: Standard deviation; Min: Minimum; Max: Maximum; BMI: body mass index.

physical activity since they woke up) and confirmed having consumed a standardized *ad-libitum* meal plan during the preceding 24 h, including a standardized dinner 12 h before the start of the indirect calorimetry assessment. Further, they abstained from both moderate (previous 24 h) and vigorous intensity (previous 48 h) physical activity. The RMR and RER were assessed with each metabolic cart on two consecutive days in the morning between 9 am and Noon. The assessment lasted 30 minutes on each cart, with a 20-minute period between measurements (Fig. 1B). The first and last 5 minutes data were discarded, and the remaining 20 minutes data were averaged for further analyses. The order of the 4 carts was randomly assigned and replicated on the second day.

Of note, the indirect calorimetry assessments were performed in agreement with current methodological recommendations [1]. Subjects stayed motionless on a reclined bed in the supine position, covered by a bed sheet, for a minimum of 20 minutes before the first indirect calorimetry assessment (Fig. 1B). Moreover, the subjects were asked to lay on the bed during the last 15 minutes of every period between measurements. Subjects were instructed not to sleep, talk, or fidget, and to breathe normally during the indirect calorimetry assessments.

Twelve-hour urine samples (i.e., starting the collection immediately after the dinner and continuing it during the fasting period) were collected before arriving to the research center. Total urine volume and urea concentration, by an enzymatic method (Spinreact, UREA-37_R1, Girona, Spain), were measured and urinary nitrogen estimated using a regression equation [22]. Estimated nitrogen urine concentration was multiplied by the urine volume and divided by the time of sample collection.

2.5. Individual calibration control evaluation procedure

Immediately after each indirect calorimetry assessment (without stopping the metabolic cart recording), pure N_2 and CO_2 gases were infused for 10 minutes into the metabolic cart hose tube in volumes mimicking the subject's $\dot{V}O_2$ and $\dot{V}CO_2$ (averaged from the 11th to the 20th minute of the indirect calorimetry assessment; *expected values*). Then, the $\dot{V}O_2$ and $\dot{V}CO_2$ readouts during the last 5 minutes of the infusion were averaged (*measured values*). Then, the $\dot{V}O_2$ and $\dot{V}CO_2$ *corrected values* were calculated.

2.6. RMR and RER calculations

All gas values are provided under standard temperature, pressure, and dry (STPD) conditions. The $\dot{V}O_2$ and the $\dot{V}CO_2$ data from the indirect calorimetry assessment were downloaded from all metabolic carts at their maximum data frequency (Supplementary Table 1). The RMR was calculated using the Weir abbreviated

equation [23], where N is urinary nitrogen excretion (N was considered to be 0 for the *in vitro* validation experiments):

$$RMR(kcal/day) = (3.941 \times \dot{V}O_2(l/min) + 1.106 \times \dot{V}CO_2\,(l/min) - 2.17 \times N*(g/min)) \times 1440$$

2.7. Statistical analysis

Results are presented as mean ± standard deviation, unless otherwise stated. Analyses were conducted using the Statistical Package for Social Sciences (SPSS, v. 22.0, IBM SPSS Statistics, IBM Corporation, Chicago, IL, USA) and the level of significance was set at $P < 0.050$. Figures were created using Graph Pad Prism (GraphPad Software, v. 8.4.1, CA, USA).

In vitro validation

Repeated-measures analyses of variance (ANOVA) with *post-hoc* LSD Tukey comparisons were used to compare the absolute value of the measurement error across metabolic carts, as determined by both gas infusions (pooling experiments 1–3 together, i.e. 50 infusions per metabolic cart) and methanol burns. The measurement error was calculated as [(measured − expected)/expected] × 100 (see Methods section). In addition, we compared the measurement error obtained by gas infusion and methanol burns using paired *t*-test analyses. Finally, we used repeated measures ANOVAs, with *post-hoc* LSD Tukey comparisons, to compare the measurement error across the simulated EE (experiment 1), RER (experiment 2), and the repeated infusions of the same magnitude (experiment 3) within each metabolic cart.

Human study

For every participant and cart, the day-to-day coefficient of variation (CV_{D-to-D}; expressed as %) were calculated for both *uncorrected* and *corrected* $\dot{V}O_2$, $\dot{V}CO_2$, RMR, and RER values. Concretely, every CV_{D-to-D} (%) was computed as: (e.g. [standard deviation *uncorrected* $\dot{V}O_2$/mean *uncorrected* $\dot{V}O_2$] × 100). Then, a two-factor repeated measures ANOVA [*metabolic cart* (i.e. Q-NRG; Vyntus; Omnical; Ultima) × *ICcE* (i.e. corrected; uncorrected)] with *post-hoc* LSD Tukey comparisons was used to test differences in $\dot{V}O_2$, $\dot{V}CO_2$, RMR, and RER reproducibility. We also conducted similar repeated measures ANOVAs to compare $\dot{V}O_2$, $\dot{V}CO_2$, RMR, and RER estimations among the metabolic carts (i.e. concordance). Bland−Altman analyses [24] of both *uncorrected* and *corrected* $\dot{V}O_2$, $\dot{V}CO_2$, RMR, and RER were also used to test reproducibility and concordance. Of note, whereas we used the measurement error for the *in vitro* experiments we used the CV_{D-to-D}, for analyzing the day-to-day reproducibility in the *in vivo* experiments.

3. Results

3.1. In vitro validation

The Omnical presented lower absolute value of the measurement error than the rest of carts on EE (Omnical = 1.4 ± 0.6%; Q-NRG = 1.6 ± 1.4%; Ultima = 7.3 ± 10.0%; and Vyntus = 11.9 ± 3.9%) and on RER (Omnical = 1.2 ± 0.8%; Vyntus = 3.1 ± 2.4%; Q-NRG = 6.6 ± 3.0%; and Ultima = 5.9 ± 7.7%), considering all the measurement errors retrieved by the 50 controlled pure gas infusion tests performed. The Omnical also presented the lowest standard deviation (SD = 0.6% and 0.8% for EE and RER respectively, considering all the measurement errors retrieved by the 50 controlled pure gas infusion tests performed), in the three gas infusion experiments indicating the highest precision among all carts. Similar results were observed in both $\dot{V}O_2$ and $\dot{V}CO_2$. Moreover, the Omnical presented an absolute measurement error lower than 2% in all variables except the RER assessed by methanol burns (absolute value of the measurement error = 2.2 ± 1.1%). The Q-NRG presented a measurement error lower than 2% in EE (and in $\dot{V}O_2$) when determined by gas infusion, but measurement errors above the 2% threshold for all other measurements. Finally, the measurement errors yield by the Vytnus and Ultima for EE and RER (and for $\dot{V}O_2$ and $\dot{V}CO_2$), either when determined by gas infusions or methanol burn, were higher than 2%. The measurement error determined with the gas infusions and methanol burns was similar in all metabolic carts (Supplementary Fig. 2).

When compared across the simulated EE (800–2400 kcal/d) the Q-NRG presented different measurement error (Fig. 2A), while the other carts did not (Fig. 2B-D). The measurement error on EE was found to be similar across the same day in all the metabolic carts (i.e. *in vitro* validation — *experiment 3*). The Q-NRG, the Omnical and the Ultima presented a trend toward significance on measurement errors when different RER were simulated (Fig. 2E, G and H), whereas all carts presented similar measurement errors on RER when identical infusions were repeated within a day (i.e. *in vitro* validation — *experiment 3*).

3.2. Human study

The RMR and RER uncorrected estimations were different across metabolic carts in day 1, with only the Vyntus and Ultima presenting comparable RMR estimations (Fig. 3A), and the Omnical and Ultima presenting comparable RER estimations (Fig. 3B). Moreover, the ICcE and the *metabolic cart × ICcE* interaction effects were significant (Fig. 3). *Corrected* RMR was different than *uncorrected* values in all the carts (Fig. 3A), whereas only the Q-NRG and the Vyntus presented differences between corrected and uncorrected RER (Fig. 3B). Although no statistically significant differences were observed, the *corrected* RER values for the Ultima presented a larger SD than these observed for the *uncorrected* RER values (Fig. 3B). These paired differences between *corrected* and *uncorrected* RMR and RER values are represented in Fig. 3 as *. Performing the ICcE procedure increased the RMR estimations yielded by the Q-NRG and the Omnical, while it reduced the RMR estimation yielded by the Vyntus and Ultima (Fig. 3A). These results observed for the day 1 were similar for the day 2. Differences in the $\dot{V}O_2$ and $\dot{V}CO_2$ were similar than these observed for RMR and RER respectively, and Bland and Altman plots comparing the *uncorrected* and *corrected* RMR and RER across metabolic carts are in Supplementary Figs. 3 and 4 respectively.

The RMR reproducibility was similar across metabolic carts and for *uncorrected* vs. *corrected* RMR (Fig. 4A). Moreover, there was no *metabolic cart × ICcE* interaction effect on RMR reproducibility

(Fig. 4A). In contrast, a significant metabolic cart effect was detected when comparing RER reproducibility, and post-hoc comparisons revealed that the Ultima yielded higher inter-day differences (i.e. lower reproducibility) than the other metabolic carts (Fig. 4B). The ICcE effect, as well as the *metabolic cart × ICcE* interaction effect, were also significant in the RER reproducibility analyses (Fig. 4B). Post-hoc analyses showed that the Ultima RER reproducibility was lower when using the corrected values (Fig. 4B). Finally, we detected a significant metabolic cart effect for $\dot{V}O_2$ and $\dot{V}CO_2$ reproducibility (P = 0.028 and P = 0.002 respectively). Post-hoc comparisons revealed that the $\dot{V}O_2$ and $\dot{V}CO_2$ reproducibility yielded by the Q-NRG was higher than the one yielded by the Ultima. Finally, Supplementary Fig. 5 shows the Bland–Altman plots for reproducibility of RMR and RER.

4. Discussion

This study analyzed the accuracy, precision, within-subject reproducibility and concordance of four different and commercially available metabolic carts (the Q-NRG, the Vyntus, the Omnical, and the Ultima) for assessing RMR and RER in non-ventilated young adults. The Omnical metabolic cart showed the most accurate and precise results, with a measurement error lower than 2% in all measured variables, cut-off threshold which has been suggested in literature as an accuracy criterion for methanol burning tests [19]. When assessing RMR and RER in non-ventilated young adults, all carts provided non comparable results, although these assessments were similarly reproducible within a 24-hour period. Finally, in contrast to our hypothesis, the application of the ICcE procedure did not modify the RMR and RER concordance, and did not decrease the within-subject reproducibility of the measurements obtained with any of the metabolic carts.

4.1. Validity of the four metabolic carts for assessing RMR and RER

Accuracy is defined as the proximity of measurements to traceable standards [18]. A previous study [19] stated a 2% as an acceptable measurement error for methanol burning tests, although this criterion might vary across laboratories. In our study, the Omnical was the only metabolic cart presenting an acceptable accuracy in all variables. The Omnical also presented the best precision as indicated by the pure gas infusions and methanol burning, as well as the most stable measurement error within a day. These results are in agreement with those of Kaviani et al. [19] and Schoffelen et al. [25]. However, it should be noted that Kaviani et al. [19] used 2 different Omnical units, and only one of them showed a measurement error lower than 2%. Noteworthy, in our study, the Q-NRG showed similar accuracy than the Omnical for the assessment of RMR, but worse accuracy for assessing $\dot{V}O_2$, $\dot{V}CO_2$ and RER. In a previous study [26], the accuracy of three Q-NRG units was tested by alcohol burning showing that the RER measurement error (−0.001, −0.012, and 0.008 for each unit) was lower than the mean RER measurement error observed in our study (−0.044). These differences could be partially explained because Delsoglio et al. [26] used ethanol (purity = 96%), while we used methanol (purity ≥99.9%). Finally, the measurement error showed by the Vyntus in our study is similar to that reported previously using butane burning tests [27].

Importantly, our study shows that different metabolic carts provide substantially different RMR and RER estimations when assessing energy metabolism in non-ventilated young adults. With differences exceeding 200 kcal/d for RMR and 0.05 for RER, our results prove that comparisons between studies using different metabolic carts should be avoided, and that using inaccurate
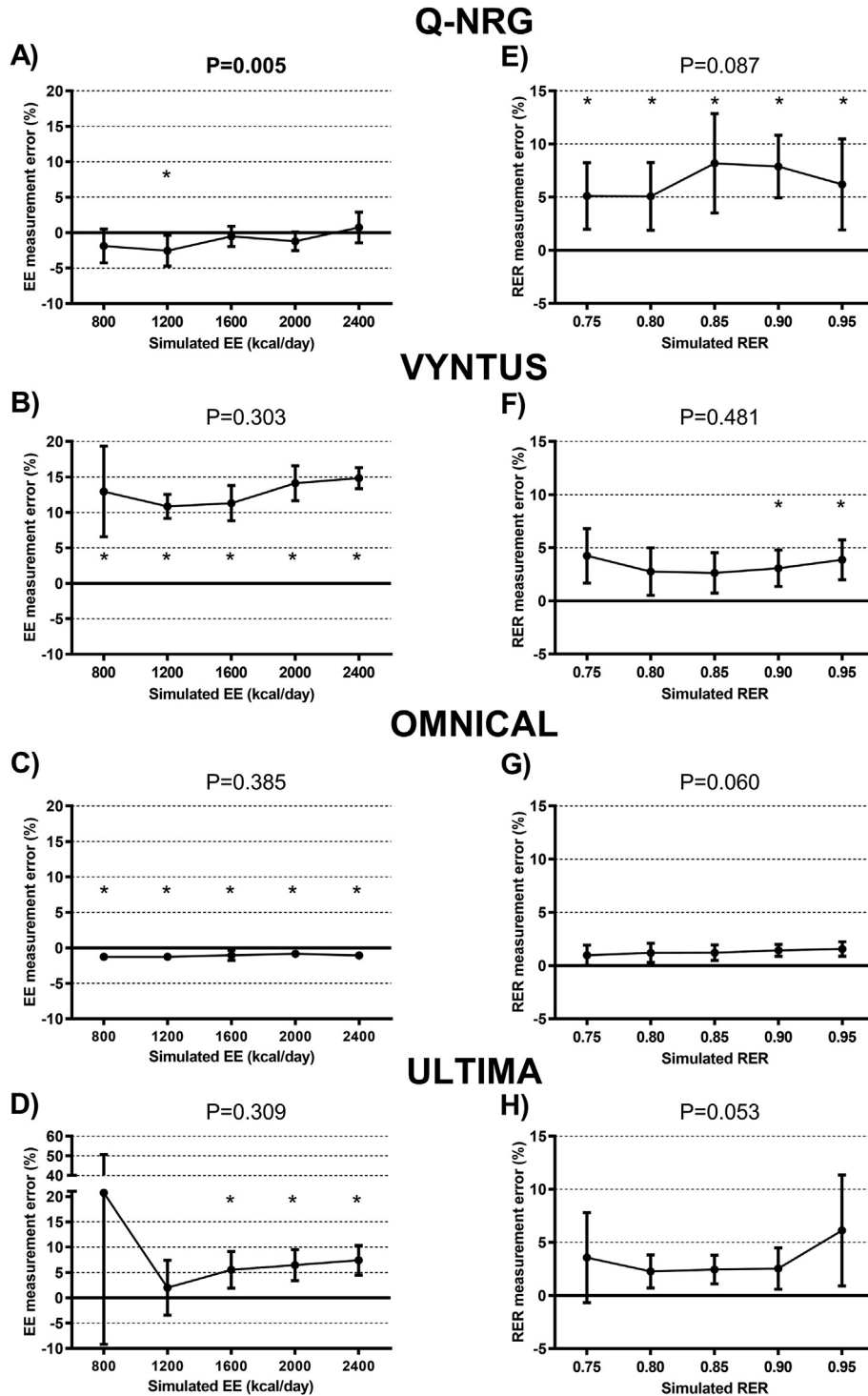
Fig. 2. Measurement error of the four metabolic carts for the determination of energy expenditure (EE) and respiratory exchange ratios (RER), determined by controlled pure gas infusions. Panels A−D show the measurement error detected when simulating different energy expenditure levels. Panels E−H represent the measurement error detected when simulating different RER levels. P values from repeated measures analysis of variance (ANOVA, n = 5). * represents significant differences vs. zero value (one-sample t-test). Results are presented as mean and standard deviation.

metabolic carts likely results in relevant measurement errors on the assessment of human energy metabolism. Surprisingly however, the large differences in accuracy and precision were not translated to differences in within-subject reproducibility. Previous studies have reported that the RMR reproducibility achieved by several metabolic carts is unacceptably high providing $CV_{D-to-D} \geq 10\%$

[3,10,14−16,28,29], whereas the DTC metabolic cart commonly provided CV_{D-to-D} below 4% [3,10−15]. Of note, in our study, all the metabolic carts showed similar CV_{D-to-D} (i.e. below 6%), and the Q-NRG achieved a RMR reproducibility similar to the one reported for the DTC ($CV_{D-to-D} = 3.6 \pm 2.5\%$). Moreover, we observed that the RMR estimates yield by the Q-NRG and the Omnical were better
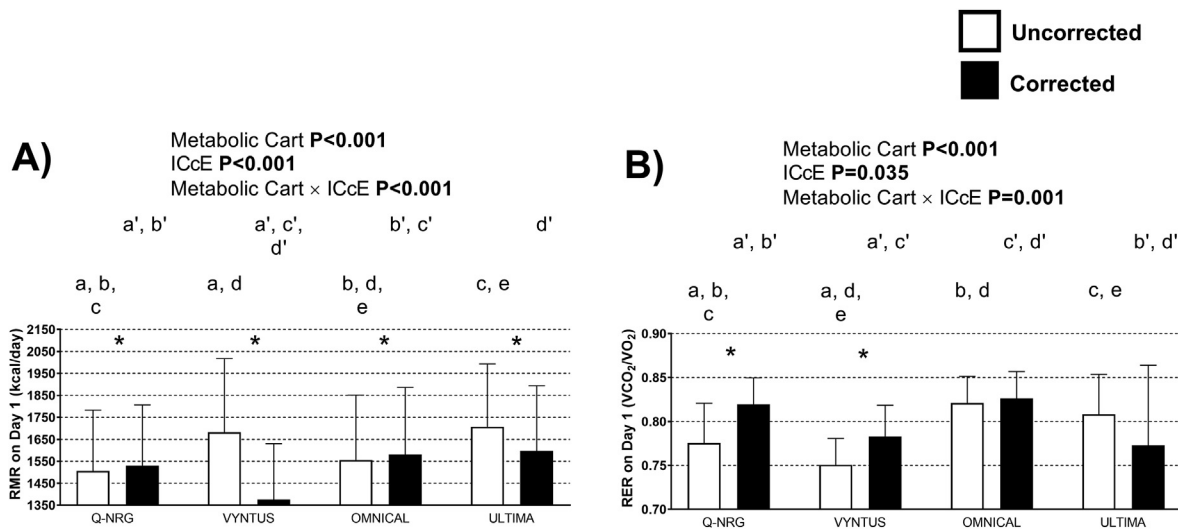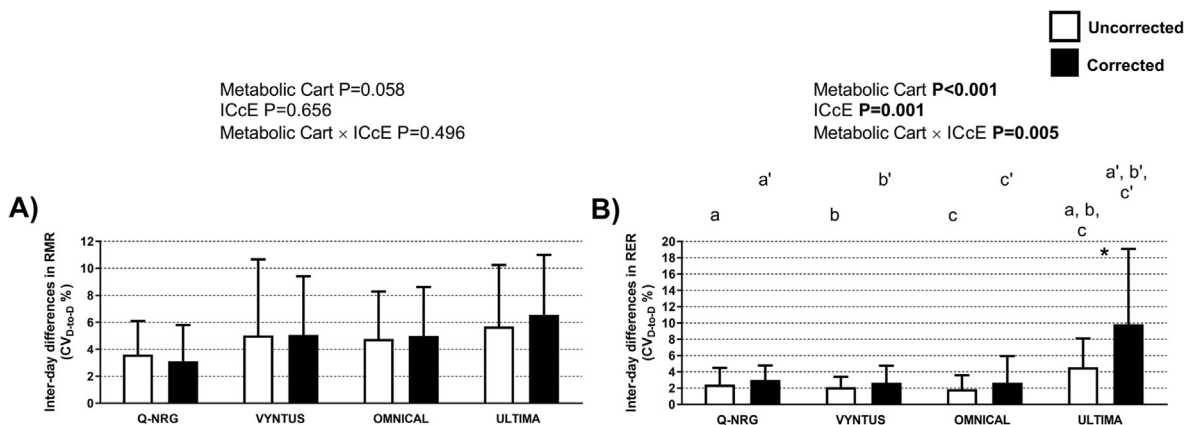
Fig. 3. Resting metabolic rate (RMR) and respiratory exchange ratio (RER) across metabolic carts on day 1 (Panels A and B respectively), with and without applying the individual calibration control evaluation procedure (ICcE). P values from two-factor (Metabolic Cart × ICcE) repeated measures analysis of variance (ANOVA, n = 29). Identical letters represent significant differences as determined by *post-hoc* LSD Tukey analysis for uncorrected values. Identical prime letters represent significant differences as determined by *post-hoc* LSD Tukey analysis for corrected values. * represents significant differences between the uncorrected vs. the corrected values. Results are presented as mean and standard deviation.

Fig. 4. Inter-day reproducibility of resting metabolic rate (RMR) and respiratory exchange ratio (RER) across metabolic carts (Panels A and B respectively), with and without applying the individual calibration control evaluation procedure (ICcE). P values from a two-factor (Metabolic Cart × ICcE) repeated measures analysis of variance (ANOVA, n = 29). Identical letters represent significant differences as determined by *post-hoc* LSD Tukey analysis for uncorrected values. Identical prime letters represent significant differences as determined by *post-hoc* LSD Tukey analysis for corrected values. * represents significant differences between the uncorrected and the corrected values. Results are presented as mean and standard deviation.

predicted by body weight, and by body composition and sex (i.e. its classical predictors [30]), than the RMR estimates yield by the Vyntus or the Ultima metabolic carts. This may suggest that the Q-NRG and the Omnical provided more valid RMR estimates than the Vyntus and the Ultima, which is in line with the higher accuracy observed in the *in vitro* validation experiments. This information, related to the associations of both *uncorrected* and *corrected* RMR with their classical predictors can be found elsewhere [31].

Unlike RMR, we detected differences in RER reproducibility among metabolic carts. The Ultima RER biological reproducibility was lower than the obtained by the three others metabolic carts, while similar RER reproducibility was observed for the Q-NRG, the Vyntus, and the Omnical.

Overall, our results suggest that the Omnical is the most valid metabolic cart for assessing RMR and RER and should be the option of choice for assessing RMR and RER in non-ventilated humans. Nevertheless, the Q-NRG metabolic cart performance was considerably worse than the performance by the Omnical for

RER measurements, but was similar for assessing RMR. Therefore, the Q-NRG can also be considered a valid option for RMR assessment.

4.2. Individual calibration control evaluation procedure

Schadewaldt et al. [2] proposed the ICcE procedure in an attempt to improve the validity of RMR and RER measurements, by correcting the RMR and RER estimations "using" the previously determined measurement error of the metabolic cart. Thus, we hypothesized that the application of this procedure would improve the concordance between metabolic carts and reduce the inter-day differences found for the RMR and RER assessments. However, the results clearly showed that the application of the ICcE procedure does not reduce neither the inter-device, nor the inter-day differences of the RMR and RER estimations. As noted earlier, we observed between-cart differences in the RMR and RER estimations. In contrast to our original hypothesis, the application of the

ICcE procedure did not eliminate those differences. Moreover, also in contrast with our initial expectations, the application of the ICcE procedure did not improve the inter-day reproducibility in any of the metabolic carts. Therefore, our results suggest that the application of the ICcE do not offer an advantage in terms of reproducibility (at least in the four metabolic carts used in the present study). In fact, the ICcE procedure is time consuming and requires specific instruments (e.g. high-precision mass-flow controllers), and in the end, the RMR and RER *uncorrected values* were similar to the *corrected values*.

Our results show that the measurement error determined by gas infusions was comparable to the measurement error determined by the alcohol burning test, which is commonly considered the reference method to validate indirect calorimeters *in vitro* [11,17,19,21]. Therefore, one may expect that at least the assessments made with those metabolic carts presenting higher measurement error (i.e. Vyntus and Ultima) would benefit from the application of the ICcE procedure, which is not supported by our data. It should be considered however that neither the gas infusion nor the alcohol burn fully simulate a person breath, e.g. none of them produces humidity in a similar extend to what is present in expired human gases [18], fact which may partially explain why our hypothesis was refuted.

4.3. Limitations

The performance of a single unit might not represent the performance of all manufactured units [19,32], and therefore, our study needs to be replicated before drawing firm conclusions. Secondly, our study did not include the DTC or, for example, other commercially available metabolic carts which have been reported to provide accurate results (e.g. ParvoMedics TrueOne 2400) [19]. Moreover, we used the methanol burning cage manufactured by Maastricht Instruments, and thus, it might have favored the Omnical, although the agreement between methanol burning and gas infusions suggest that this is not the case. Lastly, we did not control the menstrual cycle in female participants [33,34], although considering the within-subject design of the human study and that both assessments were performed within 24 h, its potential impact on RMR or RER reproducibility is likely negligible.

5. Conclusions

Our study shows that the Omnical provides more accurate and precise estimations of RMR and RER than the Q-NRG, the Vyntus CPX, and the Ultima CardiO2. Presenting a measurement error <2% in all variables during the *in vitro* experiments, our results suggest that the Omnical might be considered as a reference metabolic cart for assessing human energy metabolism. Our results also support that the Q-NRG can be considered a valid option for assessing RMR, but not RER, in non-ventilated humans. The four metabolic carts yielded different RMR and RER estimations, but similar RMR and RER reproducibility. Finally, our study also shows that using the ICcE procedure previously proposed by Schadewaldt et al. [2] does not improve the concordance between metabolic carts and does not increase the RMR and RER reproducibility obtained with any of the studied metabolic carts.

Conflict of interest

The authors report no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clnu.2022.01.031.

References

[1] Fullmer S, Benson-Davies S, Earthman CP, Frankenfield DC, Gradwell E, Lee PSP, et al. Evidence analysis library review of best practices for performing indirect calorimetry in healthy and non-critically ill individuals. J Acad Nutr Diet 2015;115:1417—46. https://doi.org/10.1016/j.jand.2015.04.003. e2.

[2] Schadewaldt P, Nowotny B, Strassburger K, Kotzka J, Roden M. Indirect calorimetry in humans: a postcalorimetric evaluation procedure for correction of metabolic monitor variability. Am J Clin Nutr 2013;97:763—73. https://doi.org/10.3945/ajcn.112.035014.

[3] Black C, Grocott MPW, Singer M. Metabolic monitoring in the intensive care unit: a comparison of the Medgraphics Ultima, Deltatrac II, and Douglas bag collection methods. Br J Anaesth 2015;114:261—8. https://doi.org/10.1093/bja/aeu365.

[4] Da Rocha EEM, Alves VGF, Da Fonseca RBV. Indirect calorimetry: methodology, instruments and clinical application. Curr Opin Clin Nutr Metab Care 2006;9:247—56. https://doi.org/10.1097/01.mco.0000222107.15548.f5.

[5] Ferrannini E. The theoretical bases of indirect calorimetry: a review. Metabolism 1988;37:287—301. https://doi.org/10.1016/0026-0495(88)90110-2.

[6] Livesey G, Elia M. Estimation of energy expenditure, net carbohydrate utilization, and net fat oxidation and synthesis by indirect calorimetry: evaluation of errors with special reference to the detailed composition of fuels. Am J Clin Nutr 1988;47:608—28. https://doi.org/10.1093/ajcn/47.4.608.

[7] Frayn KN. Calculation of substrate oxidation rates in vivo from gaseous exchange. J Appl Physiol Respir Environ Exerc Physiol 1983;55:628—34. https://doi.org/10.1152/jappl.1983.55.2.628.

[8] Oshima T, Berger MM, De Waele E, Guttormsen AB, Heidegger CP, Hiesmayr M, et al. Indirect calorimetry in nutritional therapy. A position paper by the ICALIC study group. Clin Nutr 2017;36:651—62. https://doi.org/10.1016/j.clnu.2016.06.010.

[9] Lam YY, Ravussin E. Indirect calorimetry: an indispensable tool to understand and predict obesity. Eur J Clin Nutr 2017;71:318—22. https://doi.org/10.1038/ejcn.2016.220.

[10] Kennedy S, Ryan L, Fraser A, Clegg ME. Comparison of the GEM and the ECAL indirect calorimeters against the Deltatrac for measures of RMR and diet-induced thermogenesis. J Nutr Sci 2014;3. https://doi.org/10.1017/jns.2014.58.

[11] Cooper BG, McLean JA, Taylor R. An evaluation of the Deltatrac indirect calorimeter by gravimetric injection and alcohol burning. Clin Phys Physiol Meas 1991;12:333—41. https://doi.org/10.1088/0143-0815/12/4/003.

[12] Tissot S, Delafosse B, Bertrand O, Bouffard Y, Viale JP, Annat G. Clinical validation of the Deltatrac monitoring system in mechanically ventilated patients. Intensive Care Med 1995;21:149—53. https://doi.org/10.1007/BF01726538.

[13] Schoeller DA. Making indirect calorimetry a gold standard for predicting energy requirements for institutionalized patients. J Am Diet Assoc 2007;107: 390—2. https://doi.org/10.1016/j.jada.2007.01.030.

[14] Graf S, Karsegard VL, Viatte V, Maisonneuve N, Pichard C, Genton L. Comparison of three indirect calorimetry devices and three methods of gas collection: a prospective observational study. Clin Nutr 2013;32. https://doi.org/10.1016/j.clnu.2013.08.012.

[15] Cooper JA, Watras AC, O'Brien MJ, Luke A, Dobratz JR, Earthman CP, et al. Assessing validity and reliability of resting metabolic rate in six gas analysis systems. J Am Diet Assoc 2009;109:128—32. https://doi.org/10.1016/j.jada.2008.10.004.

[16] Alcantara JMA, Sanchez-Delgado G, Martinez-Tellez B, Merchan-Ramirez E, Labayen I, Ruiz JR. Congruent validity and inter-day reliability of two breath by breath metabolic carts to measure resting metabolic rate in young adults. Nutr Metabol Cardiovasc Dis 2018;28:929—36. https://doi.org/10.1016/j.numecd.2018.03.010.

[17] Schoffelen PFM, Plasqui G. Classical experiments in whole-body metabolism: open-circuit respirometry—diluted flow chamber, hood, or facemask systems. Eur J Appl Physiol 2017. https://doi.org/10.1007/s00421-017-3735-5.

[18] Chen KY, Smith S, Ravussin E, Krakoff J, Plasqui G, Tanaka S, et al. Room indirect calorimetry operating and reporting standards (RICORS 1.0): a guide to conducting and reporting human whole-room calorimeter studies. Obesity 2020;28:1613—25. https://doi.org/10.1002/oby.22928.

[19] Kaviani S, Schoeller DA, Ravussin E, Melanson EL, Henes ST, Dugas LR, et al. Determining the accuracy and reliability of indirect calorimeters utilizing the methanol combustion technique. Nutr Clin Pract 2018;33:206—16. https://doi.org/10.1002/ncp.10070.

[20] 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—56. https://doi.org/10.1079/bjn19870104.

[21] Miodownik S, Melendez J, Carlon VA, Burda B, Melendez J, Arslan V, et al. Quantitative methanol-burning lung model for validating gas-exchange measurements over wide ranges of FIO2. J Appl Physiol 1998;84:2177—82. https://doi.org/10.1152/jappl.1998.84.6.2177.

[22] Sanchez-Delgado G, Alcantara JMA, Acosta FM, Martinez-Tellez B, Amaro-Gahete FJ, Merchan-Ramirez E, et al. Energy expenditure and macronutrient oxidation in response to an individualized nonshivering cooling protocol. Obesity 2020:1—9. https://doi.org/10.1002/oby.22972. 00.

[23] Weir JB de V. New methods for calculating metabolic rate with special reference to protein metabolism. J Physiol 1949;109:1—9. https://doi.org/10.1113/jphysiol.1949.sp004363.

[24] Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;327:307—10. https://doi.org/10.1016/S0140-6736(86)90837-8.

[25] Schoffelen PFM, den Hoed M, van Breda E, Plasqui G. Test-retest variability of VO 2 max using total-capture indirect calorimetry reveals linear relationship of VO 2 and Power. Scand J Med Sci Sports 2019;29:213—22. https://doi.org/10.1111/sms.13324.

[26] Delsoglio M, Dupertuis YM, Oshima T, van der Plas M, Pichard C. Evaluation of the accuracy and precision of a new generation indirect calorimeter in canopy dilution mode. Clin Nutr 2019:1—8. https://doi.org/10.1016/j.clnu.2019.08.017.

[27] Perez-Suarez I, Martin-Rincon M, Gonzalez-Henriquez JJ, Fezzardi C, Perez-Regalado S, Galvan-Alvarez V, et al. Accuracy and precision of the COSMED K5 portable Analyser. Front Physiol 2018;9:1—12. https://doi.org/10.3389/fphys.2018.01764.

[28] Poulsen MK, Thomsen LP, Kjærgaard S, Rees SE, Karbing DS. Reliability of, and agreement between, two breath-by-breath indirect calorimeters at varying levels of inspiratory oxygen. Nutr Clin Pract 2019:1—8. https://doi.org/10.1002/ncp.10250. 00.

[29] Sundström M, Tjäder I, Rooyackers O, Wernerman J. Indirect calorimetry in mechanically ventilated patients. A systematic comparison of three instruments. Clin Nutr 2013;32:118—21. https://doi.org/10.1016/j.clnu.2012.06.004.

[30] Johnstone AM, Murison SD, Duncan JS, Rance KA, Speakman JR. Factors influencing variation in basal metabolic rate include fat-free mass, fat mass, age, and circulating thyroxine but not sex, circulating leptin, or triiodothyronine. Am J Clin Nutr 2005;82:941—8. https://doi.org/10.1093/ajcn/82.5.941.

[31] Alcantara JMA. Assessment of resting energy expenditure and nutrient oxidation by indirect calorimetry: methodological implications. University of Granada; 2020.

[32] Roffey Darren M, Byrne Nuala M, Hills Andrew P. Day-to-Day variance in measurement of resting metabolic rate using ventilated-hood and mouthpiece & nose-clip indirect calorimetry systems. J Parenter Enteral Nutr 2006;30:426—32.

[33] Melanson KJ, Saltzman E, Russell R, Roberts SB. Postabsorptive and postprandial energy expenditure and substrate oxidation do not change during the menstrual cycle in young women. J Nutr 1996;126:2531—8. https://doi.org/10.1093/jn/126.10.2531.

[34] Ferraro R, Lillioja S, Fontvieille AM, Rising R, Bogardus C, Ravussin E. Lower sedentary metabolic rate in women compared with men. J Clin Invest 1992;90:780—4. https://doi.org/10.1172/JCI115951.