

## ORIGINAL ARTICLE

# Test-retest variability of $\text{VO}_{2\text{max}}$ using total-capture indirect calorimetry reveals linear relationship of $\text{VO}_2$ and Power

Paul F. M. Schoffelen<sup>1,2</sup>  | Marcel den Hoed<sup>3</sup> | Eric van Breda<sup>4</sup> | Guy Plasqui<sup>2</sup>

<sup>1</sup>Department of Human Biology, NUTRIM School for Nutrition, Toxicology and Metabolism, Maastricht University Medical Centre, Maastricht, The Netherlands

<sup>2</sup>Department of Nutrition and Movement Sciences, NUTRIM School for Nutrition, Toxicology and Metabolism, Maastricht University Medical Centre, Maastricht, The Netherlands

<sup>3</sup>Department of Immunology, Genetics and Pathology and SciLifeLab, Uppsala University, Uppsala, Sweden

<sup>4</sup>Faculty of Medicine and Health Sciences, Department of Rehabilitation Sciences and Physiotherapy/MOVANT, University of Antwerp, Antwerp, Belgium

## Correspondence

Paul F. M. Schoffelen, Department of Human Biology, Nutrition and Movement Sciences, Maastricht University, Maastricht, The Netherlands.  
Email: p.schoffelen@maastrichtuniversity.nl

This study aimed to analyze the intra-individual variation in  $\text{VO}_{2\text{max}}$  of human subjects using total-capture and free-flow indirect calorimetry. Twenty-seven men ( $27 \pm 5$  year;  $\text{VO}_{2\text{max}}$  49-79  $\text{mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ) performed two maximal exertion tests (CPETs) on a cycle ergometer, separated by a  $7 \pm 2$  day interval.  $\text{VO}_2$  and  $\text{VCO}_2$  were assessed using an indirect calorimeter (Omnical) with total capture of exhalation in a free-flow airstream. Thirteen subjects performed a third maximal exertion test using a breath-by-breath calorimeter (Oxycon Pro). On-site validation was deemed a requirement. For the Omnical, the mean within-subject CV for  $\text{VO}_{2\text{max}}$  was  $1.2 \pm 0.9\%$  (0.0%-4.4%) and for ergometer workload  $P_{\text{max}}$   $1.3 \pm 1.3\%$  (0%-4.6%).  $\text{VO}_{2\text{max}}$  values with the Oxycon Pro were significantly lower in comparison with Omnical ( $P < 0.001$ ; *t* test) with mean 3570 vs 4061 and difference SD 361  $\text{mL}\cdot\text{min}^{-1}$ . Validation results for the Omnical with methanol combustion were  $-0.05 \pm 0.70\%$  (mean  $\pm$  SD;  $n = 31$ ) at the 225  $\text{mL}\cdot\text{min}^{-1}$   $\text{VO}_2$  level and  $-0.23 \pm 0.80\%$  ( $n = 31$ ) at the 150  $\text{mL}\cdot\text{min}^{-1}$   $\text{VCO}_2$  level. Results using gas infusion were  $0.04 \pm 0.75\%$  ( $n = 34$ ) and  $-0.99 \pm 1.05\%$  ( $n = 24$ ) over the respective 500-6000  $\text{mL}\cdot\text{min}^{-1}$   $\text{VO}_2$  and  $\text{VCO}_2$  ranges. Validation results for the Oxycon Pro in breath-by-breath mode were  $-2.2 \pm 1.6\%$  ( $n = 12$ ) for  $\text{VO}_2$  and  $5.7 \pm 3.3\%$  ( $n = 12$ ) for  $\text{VCO}_2$  over the 1000-4000  $\text{mL}\cdot\text{min}^{-1}$  range. On a Visual analog scale, participants reported improved breathing using the free-flow indirect calorimetry (score  $7.6 \pm 1.2$  vs  $5.1 \pm 2.7$ ,  $P = 0.008$ ). We conclude that total capturing free-flow indirect calorimetry is suitable for measuring  $\text{VO}_2$  even with the highest range.  $\text{VO}_{2\text{max}}$  was linear with the incline in  $P_{\text{max}}$  over the full range.

## KEYWORDS

breath-by-breath, exercise testing, gas exchange, intra-individual variability, validation

## 1 | INTRODUCTION

Maximal oxygen uptake ( $\text{VO}_{2\text{max}}$ ), that is, the maximal capacity of the cardiovascular system to provide  $\text{O}_2$  to the working muscle and the capacity of the muscle to use  $\text{O}_2$  during sustained exercise, is considered the physiological gold standard of cardiovascular fitness.<sup>1</sup> It is an important health parameter as well,

as it has been shown that a high  $\text{VO}_{2\text{max}}$  is inversely related with the risk of developing cardiovascular diseases and mortality.<sup>2,3</sup> Improvements in  $\text{VO}_{2\text{max}}$  have been shown to have a significant beneficial effect on the reduction in cardiometabolic risk factors, that is, a decrease in blood pressure and an increased sensitivity to insulin.<sup>4</sup> As such, highly accurate gas exchange analysis is mandatory for assessing reliable information.

Cardiopulmonary exercise tests (CPETs) are usually performed under laboratory conditions using an incremental exercise protocol and technological advanced gas analysis devices.<sup>5</sup> Breath-by-breath analysis has been recommended as a reliable and accurate method for O<sub>2</sub> and CO<sub>2</sub> analysis during exercise testing.<sup>1</sup> In order to accurately analyze gas exchange in an airflow up to and beyond athletic levels of 6000 mL•min<sup>-1</sup>, it is mandatory to assess concentrations of O<sub>2</sub> and CO<sub>2</sub> in both the inhaled and exhaled air, synchronously.<sup>6</sup> However, breath-by-breath analysis during maximum exercise testing or even below that<sup>7,8</sup> requires fast response analyzers, and the response time of the most rapid O<sub>2</sub> analyzers is much lower than the physiological maximum of the athlete.<sup>9</sup> In such devices, airflow is typically measured using a bi-directional flow sensor at the mouth, calibrated with a 2-3 L volume syringe prior to testing. A disadvantage of this method is that abnormal or maximum exertion breathing patterns may differ from that of simulated breathing with a calibration syringe or other validation devices.<sup>10</sup> A more physiological concern is that gas exchange in the lungs is a continuous process even during exhalation and hence, alveolar gas concentrations continually change.<sup>9</sup> Since gas concentrations at the start of exhalation overlap with the exhalation peak-flow, it is necessary in breath-by-breath calorimetry to measure airflow and gas concentrations (O<sub>2</sub> and CO<sub>2</sub>) in parallel and with parallel timing. Subsequently, the increased breathing frequency that is related to VO<sub>2</sub> during a incremental CPET will cause breath-by-breath systems to become inaccurate at higher breathing frequencies.<sup>11</sup> Furthermore, tidal volumes during breathing are not always constant and therefore require multiple breaths to accurately determine the average VO<sub>2</sub> level.<sup>12-14</sup> Since most validation/calibration systems are not capable of fully simulating the changes during physiological exhalation, or replicate the effects of extreme breathing frequencies or breathing patterns, it is difficult to interpret the results from participants during different tests.<sup>7</sup> Inaccuracies using breath-by-breath indirect calorimetry devices have been shown previously<sup>7,8</sup> to underestimate VO<sub>2</sub> at levels as low as 2800 mL•min<sup>-1</sup>. Finally, a parameter that usually receives little attention is the discomfort subjects experience during a CPET. Breathing can be troubled by the relative small tubes that might cause some resistance and presumably lower outcome of gas exchange parameters during higher stages of a CPET.

In the present study, we aimed to investigate gas exchange parameters during maximal exercise testing using a total-capture indirect calorimeter, with the subject breathing without valves or restrictions in a grander stream of air open to the ambient at the inlet, that is, free-flow. The diluted exhalation of the subject is fully captured and analyzed, thereby bypassing limitations as mentioned above for breath-by-breath analysis. Importantly, a complete validation range by gas infusion and/or alcohol combustion can be performed on-site

for this type of calorimeter without the need for specialized equipment.

Therefore, the main purposes of the present study were (a) to assess reproducibility of VO<sub>2max</sub> using an on-site validated total-capture indirect calorimeter (Omnicol, Maastricht Instruments, Maastricht, The Netherlands) over a wide range of VO<sub>2</sub> values and (b) to compare the total-capture indirect calorimeter with breath-by-breath analysis.

Our hypothesis was that (a) a total-capture indirect calorimeter has high reproducibility for VO<sub>2max</sub> assessment and (b) can perform at least as well, and with validated accuracy, in comparison with a commercially available metabolic cart with breath-by-breath analysis.

As has been previously demonstrated,<sup>15</sup> the successful validation of an individual calorimeter should prove suitability for the full range of technical and biological variability expected.<sup>16</sup>

In brief, technical validation must be able to mimic the biological aspects as found in real life, for instance, breathing frequencies or energy expenditure at applicable physiological ranges. In contrast, a calorimeter that cannot be directly validated at the applicable physiological ranges and/or measurement mode may yet be validated by comparison of stable subject group results obtained with a validated second calorimeter.<sup>16</sup>

To assess the validity of our hypothesis, both a total-capture indirect calorimeter and a calorimeter using breath-by-breath mode were tested. The total-capture indirect calorimeter was validated for both technical accuracy and biological reproducibility. The breath-by-breath indirect calorimeter was tested for accuracy by comparison of biological subject group results as also obtained using the total-capture indirect calorimeter.

## 2 | METHODS

### 2.1 | Subjects and protocol

Twenty-seven healthy young men, recruited from local competition cyclists, (age 27 ± 5 years; BMI 23.0 ± 3.0 kg•m<sup>-2</sup>) performed two maximal CPETs on a calibrated cycle ergometer (Lode, Groningen, The Netherlands), with simultaneous gas analysis using an Omnicol total-capture indirect calorimeter (Omnicol, Maastricht Instruments, The Netherlands). The combination of trained cyclists and using a calibrated ergometer allowed assessment of workload independent of calorimetric measures. A subset of subjects, limited in number to the first 13 by subsequent unavailability of the Oxycon, performed a third maximal CPET test with simultaneous gas analysis using a breath-by-breath indirect calorimeter (Oxycon Pro, Carefusion, Höchberg, Germany). For one participant, no VO<sub>2</sub> values could be obtained during the final 2 minutes of the test, and hence, no VO<sub>2max</sub> value could be

obtained. The order of the three tests was always Omnical-Oxycon Pro-Omnical. A reason for this nonrandom order was that despite the fact that further tests in the order Omnical-Omnical-Oxycon Pro were intended these could not be completed with the Oxycon. For the reproducibility of  $VO_{2\max}$  with Omnical, the inclusion of a first visit CPET was deemed a "worst case" test for reproducibility, also allowing assessment of the impact of adaptation of subjects to the test protocol if any. All CPETs were separated by a  $7 \pm 2$  day interval.

Each participant was instructed to refrain from heavy exercise the day before a test. Participants were also asked to remain fasted from 22:00 hour the evening before a CPET. On the day of the test, participants came to the laboratory by car or by public transportation. After assessment of basal metabolic rate (BMR), participants consumed a standardized breakfast consisting of tea and two raisin buns, the same for all subjects and between trials, followed by a 45-60 minute rest before performing the maximum CPET. BMR was measured for all participants on their first visit following a previously described protocol to provide an individual baseline  $VO_{2\text{BMR}}$ .<sup>15</sup> Ten participants, the first 10 that volunteered to do so, were asked to participate in a second BMR measurement on the last visit to check if BMR reproducibility was in agreement with previously reported values.<sup>15</sup> Immediately after the CPET test, participants were asked to mark a visual analog scale (VAS) on the ease of breathing near maximal exertion. (10 cm line, measured as 0-10 score from "bad" to "excellent").

## 2.2 | Indirect calorimetry systems

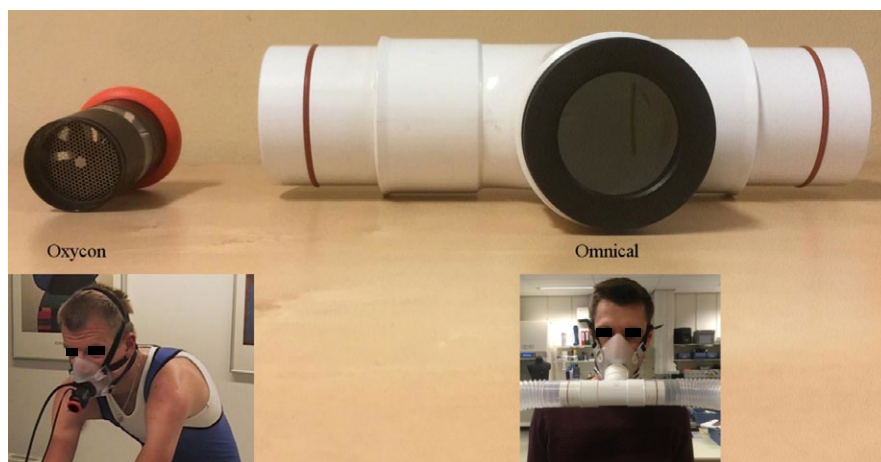
A face mask (Hans Rudolph Inc, Shawnee, KS, USA) was connected to a T-piece that was placed in a free airstream (Omnical) or turbine with sample tubing (Oxycon Pro), and

next, the face mask was placed over the participants' nose and mouth. Great care was taken to use masks that fit without detectable leakage to the participants' face. No gel-sealing of the total face mask was used.<sup>17</sup>

## 2.3 | Omnical

The Omnical is the fourth generation Omnical and is based on methods developed for whole-room calorimetry.<sup>18,19</sup> Analogous to a whole-room calorimeter, during all  $VO_{2\max}$  tests, the participants could breathe freely into a grand stream of air, for this study preset at  $450 \text{ L min}^{-1}$ , passing through the tubes. Whole-room calorimetry analysis techniques were used<sup>15,19</sup> so that no valves were required. In brief, the system measures total airflow passing the participant's face (Figure 1) and determines gas concentrations for inspired and expired air with a representative resolution<sup>19</sup> of  $\leq 0.001\%$ . Inspired air (environment) samples were taken every 2 minutes (analyzer 1), while expired air was analyzed every other 2 minutes (analyzer 1) as well as continuously (analyzer 2), allowing for a normalization between analyzers 1 and 2. Calculation was performed using gas exchange formulae as described previously.<sup>19,20</sup> Volume was set to nearly zero, and evaluation intervals set to seconds instead of multiple minutes or hours as for face mask or whole-room calorimetry, respectively. A calculation interval of 5 seconds was used, while samples were continuously evaluated by the analyzers and acquisition software at a rate up to 50 samples per second. The dependency on breathing frequency was that slow (supine, resting) breathing can be distinguished as individual peaks and valleys, while there is no upper limit for breathing frequency as this averages out as with fast changes in whole-room calorimetry.<sup>21</sup>

In breath-by-breath systems, measuring inhalation and exhalation in sequence through a single tube limits breathing



**FIGURE 1** The size of the pathway for breathing is visualized for both calorimeters, Oxycon on the left, Omnical on the right. The Oxycon triple-V sensor has an inner diameter of 2.6 cm, an in-line sputum catcher (metal mesh), and an in-line turbine behind the sputum catcher. It is estimated to provide  $400 \text{ mm}^2$  surface for breathing flow. The Omnical has an open pathway with an inner diameter of 3.5 cm, and it provides  $\pm 960 \text{ mm}^2$  surface for breathing flow. Both use common industry face mask or mouthpiece with nose-clip

frequency because of two main reasons; (a) the response time of the analyzers applied<sup>9</sup> and (b) the averaging of sample due to sequential transport of inhalation and exhalation through a single tube.<sup>11</sup>

In our study, the gas analyzers of the free-flow indirect calorimeter were calibrated automatically every 15-30 minutes. Nitrogen gas ( $\geq 99.999\%$ , Linde, Schiedam, The Netherlands) was used to set the zero, and a calibration gas with 18% O<sub>2</sub> and 0.8% CO<sub>2</sub> certified to 1% volumetric content (ie, 0.18% O<sub>2</sub>, 0.008% CO<sub>2</sub>) was used for calibration (HiQ specialty gas, Linde). The O<sub>2</sub> content of the calibration gas was further corrected based on the O<sub>2</sub> content of fresh air as described earlier for a whole-room calorimeter.<sup>19</sup> Sensors for flow, pressure, humidity, and temperature were precalibrated and checked at service intervals. Performance of calorimeters in our laboratories is checked in situ by methanol combustion and infusion tests before and during periods of use.<sup>15,19</sup>

## 2.4 | Oxycon Pro

The Oxycon Pro is an open circuit indirect calorimeter that measures flow at the participant's mouth using a bi-directional turbine (Triple V; Carefusion; Figure 1). It was calibrated before each test with a 3 L syringe (Carefusion). Gas concentrations VO<sub>2</sub> and VCO<sub>2</sub> were derived from undiluted samples taken at the turbine and thus alternate between inhaled and exhaled air, synchronously with breathing frequency. For optimal synchronization, the sample delay from turbine sample point into the Oxycon is determined with the help of a secondary tube equal to sample tube, used to infuse calibration gas pulses during calibration. VO<sub>2</sub> and VCO<sub>2</sub> results were stored every 30 seconds. The Oxycon Pro has been validated to achieve accurate results up to 5000 mL•min<sup>-1</sup> O<sub>2</sub> uptake.<sup>22</sup> The Oxycon requires users to calibrate its flow sensor before use and to perform a gas calibration. Gas calibration uses an ambient sample (room air) and a certified span gas (typically 5% CO<sub>2</sub> in N<sub>2</sub>) both for calibration of analyzer output and for determining sample synchronization from turbine-sample-point to analyzers. Sensors for pressure, humidity, and temperature were precalibrated similarly to the Omnicol.

## 2.5 | Instrumental validation

The Omnicol was validated both with methanol combustion (simulating BMR  $\sim 1$  kcal•min<sup>-1</sup> by setting wick height of burner) in a ventilated hood volume, and with gas infusions in the face mask connection up to the highest expected level of O<sub>2</sub> consumption and CO<sub>2</sub> production, range 225-6000 mL•min<sup>-1</sup> VO<sub>2</sub>. The principle of measuring total diluted gas flow allows for the use of combustion of known quantities of methanol or other fuels, as well as for straight-forward infusion of known quantities of gases (here: N<sub>2</sub> and

CO<sub>2</sub>). This allows simulation of the full range of O<sub>2</sub> consumption and CO<sub>2</sub> production, from resting to elite athletes' levels. Combustion tests provide for water vapor production similar to human breathing, while gas infusion tests are dry unless water vapor is specifically added. In the present study, handling of water vapor was validated with the combustion tests for Omnicol, and no humidity was added to infusion tests.

For validation in the BMR range, methanol (pro-analyse, 99.8%; Merck Millipore BV, Amsterdam, The Netherlands) was burned<sup>15,19</sup> at a target VO<sub>2</sub> rate of 225 mL•min<sup>-1</sup>. Infusions were performed for validation<sup>19</sup> in the higher ranges of CO<sub>2</sub> (99.99%, Linde) and N<sub>2</sub> ( $\geq 99.999\%$ ; Linde). All measurements were normalized to standard temperature and pressure dry (STPD) values by measuring temperature, humidity, and pressure.

The Oxycon Pro was serviced prior to the study and tested in 12 infusion tests during service using a proprietary dry gas sinusoidal breath simulator (Mijnhardt, Bunnik, The Netherlands) in the range of 1000-4000 mL•min<sup>-1</sup>. There was no available system to measure impact of humidity, or changes in breathing pattern, for the breath-by-breath mode. Gas infusion occurred inside the calibration unit's lung-simulation, that is, on the face-mask side of the flow sensor in a variable volume.

For both devices, the gases infused were N<sub>2</sub> and CO<sub>2</sub>. For CO<sub>2</sub>, the amount measured must equal the amount infused, for N<sub>2</sub>, the simulated O<sub>2</sub> uptake is derived from dilution of inlet O<sub>2</sub> air fraction (FiO<sub>2</sub>) as: O<sub>2</sub> uptake simulated = N<sub>2</sub> infused • FiO<sub>2</sub> • (1-FiO<sub>2</sub>)<sup>-1</sup>.

## 2.6 | VO<sub>2max</sub> tests

All VO<sub>2max</sub> tests were performed on a calibrated bicycle ergometer according to the protocol of Kuipers et al.<sup>23</sup> Subjects performed a 5-minutes warming-up at 100 W, after which the workload was increased by 50 W every 2.5 minutes. After the respiratory exchange quotient (RER) exceeded 1.0 or the heart rate exceeded 160 bpm, workload was increased by 25 W every 2.5 minutes. Subjects were verbally encouraged to continue until exhaustion by the researcher. The time to exhaustion was recorded. The maximal workload achieved ( $P_{\max}$ ) was calculated as the workload completed ( $W_{\text{completed}}$ ) plus time ( $t$ , in seconds) in the last stage divided by 150 (seconds ie, 2.5 minutes) and multiplied with the load increment of the final stage<sup>23</sup> ( $\Delta W$ ):  $P_{\max} = W_{\text{completed}} + \Delta W \cdot t \cdot 150^{-1}$ .

For the Omnicol, VO<sub>2max</sub> was calculated as the highest moving average VO<sub>2</sub> value obtained over 30 consecutive seconds (ie, six 5-second values). Additionally, the VO<sub>2peak</sub> was calculated as the single highest obtained VO<sub>2</sub> value (ie, one 5-second value). For the Oxycon, the highest registered 30 seconds VO<sub>2</sub> value was used as VO<sub>2max</sub>.

## 2.7 | Statistics

Agreement between the two  $\text{VO}_{2\text{max}}$  tests with the Omnicol, and between the Omnicol system and Oxycon Pro was examined using paired  $t$  tests and linear regression analysis to examine if the slope of the regression line was significantly different from 1 and the intercept significantly different from 0. In addition, Lin's concordance correlation coefficients were calculated. This coefficient is used to test reproducibility by taking into account the variation from the line of identity<sup>24</sup> (ie, slope 1, intercept 0). Bland-Altman plots were prepared to quantify systematic and random error, and paired  $t$  test's were used to test for differences in  $\text{VO}_{2\text{max}}$  between both Omnicol tests and between the Omnicol and Oxycon Pro tests, as well as for the VAS values marked by participants for ease of breathing.<sup>25</sup>

## 3 | RESULTS

### 3.1 | Instrumental validation

Validation data for the Omnicol were  $-0.05 \pm 0.70\%$  ( $n = 31$ ) for  $\text{VO}_2$  at  $225 \text{ mL}\cdot\text{min}^{-1}$  methanol combustion rate and  $0.04 \pm 0.75\%$  ( $n = 34$ ) over the  $500\text{-}6000 \text{ mL}\cdot\text{min}^{-1}$  for infusion rate.  $\text{CO}_2$  values were  $-0.23 \pm 0.80\%$  ( $n = 31$ ) at the  $150 \text{ mL}\cdot\text{min}^{-1}$  methanol combustion rate and  $-0.99 \pm 1.05\%$  ( $n = 24$ ) over the  $500\text{-}6000 \text{ mL}\cdot\text{min}^{-1}$  infusion range. The difference in levels of  $\text{O}_2$  (225) and  $\text{CO}_2$  (150) for alcohol combustion is caused by the respiratory quotient (RQ) of alcohol, that is, 0.6667. For the Oxycon Pro, results were  $-2.2 \pm 1.6\%$  for  $\text{VO}_2$  ( $n = 12$ ) and  $5.7 \pm 3.3\%$  for  $\text{CO}_2$  ( $n = 12$ ) over the  $1000\text{-}4000 \text{ mL}\cdot\text{min}^{-1}$  range of the breathing simulator. Note that alcohol combustion with the Oxycon is only possible in ventilated hood mode, effectively disabling the breath-by-breath mode. For that reason, no methanol combustion tests were performed with the Oxycon Pro; instead, the breathing simulator for validation of breath-by-breath mode was used.

### 3.2 | Basal metabolic rate analysis

Results for basal metabolic rate (BMR) were  $258 \pm 28 \text{ mL}\cdot\text{min}^{-1} \text{ O}_2$  ( $n = 27$ ) and repeatability for BMR between two visits was  $2.9 \pm 3.3\%$  coefficient of variation (CV) ( $n = 10$ ), when including substrate usage energy expenditure (EE) for BMR expressed in  $\text{kJ}\cdot\text{kg}^{-1}$  the CV was  $3.5 \pm 3.7\%$ .

### 3.3 | Repeatability $\text{VO}_{2\text{max}}$ Omnicol

The mean within-subject CV for  $\text{VO}_{2\text{max}}$  was  $1.2 \pm 0.9\%$  (range  $0.0\%\text{-}4.4\%$ ,  $n = 27$ ) and for  $P_{\text{max}}$   $1.3 \pm 1.3\%$  (range  $0.0\%\text{-}4.6\%$ ). Both maximum Power output ( $P_{\text{max}}$ ) ( $R^2 = 0.97$ ,

$P < 0.001$ ) and maximum  $\text{VO}_2$  ( $R^2 = 0.98$ ,  $P < 0.001$ ) were highly correlated (Figure 2). For  $\text{VO}_{2\text{max}}$ , during maximal exercise testing, the slope of the regression line was not significantly different from 1 ( $1.05$ ;  $95\% \text{ CI } 0.99\text{-}1.11$ ) and the intercept was not significantly different from 0 ( $-145 \text{ mL}\cdot\text{min}^{-1}$ ;  $95\% \text{ CI } -396 \text{ to } 107$ ). Lin's concordance correlation coefficient was 0.98. There was a small but significant difference in  $\text{VO}_{2\text{max}}$  between the two Omnicol tests ( $58 \pm 107 \text{ mL}\cdot\text{min}^{-1}$ ,  $P = 0.01$ ). Bland-Altman plots showed a significant and positive correlation between the difference and mean  $\text{VO}_{2\text{max}}$  of both tests ( $R^2 = 0.15$ ,  $P = 0.05$ ; Figure 3). The same was observed for the  $P_{\text{max}}$  ( $R^2 = 0.15$ ,  $P = 0.04$ ). When  $\text{VO}_{2\text{peak}}$  was used instead of  $\text{VO}_{2\text{max}}$ , the correlation between tests was comparable ( $R^2 = 0.97$ ,  $P < 0.001$ ), the slope of the regression line was not significantly different from 1 ( $1.01$ ;  $95\% \text{ CI } 0.94\text{-}1.08$ ), and the intercept was not significantly different from 0 ( $6 \text{ mL}\cdot\text{min}^{-1}$ ;  $95\% \text{ CI } -290 \text{ to } 301$ ). The Bland-Altman plot for  $\text{VO}_{2\text{peak}}$  did not show a systematic bias proportional to the measured value (Figure 3).

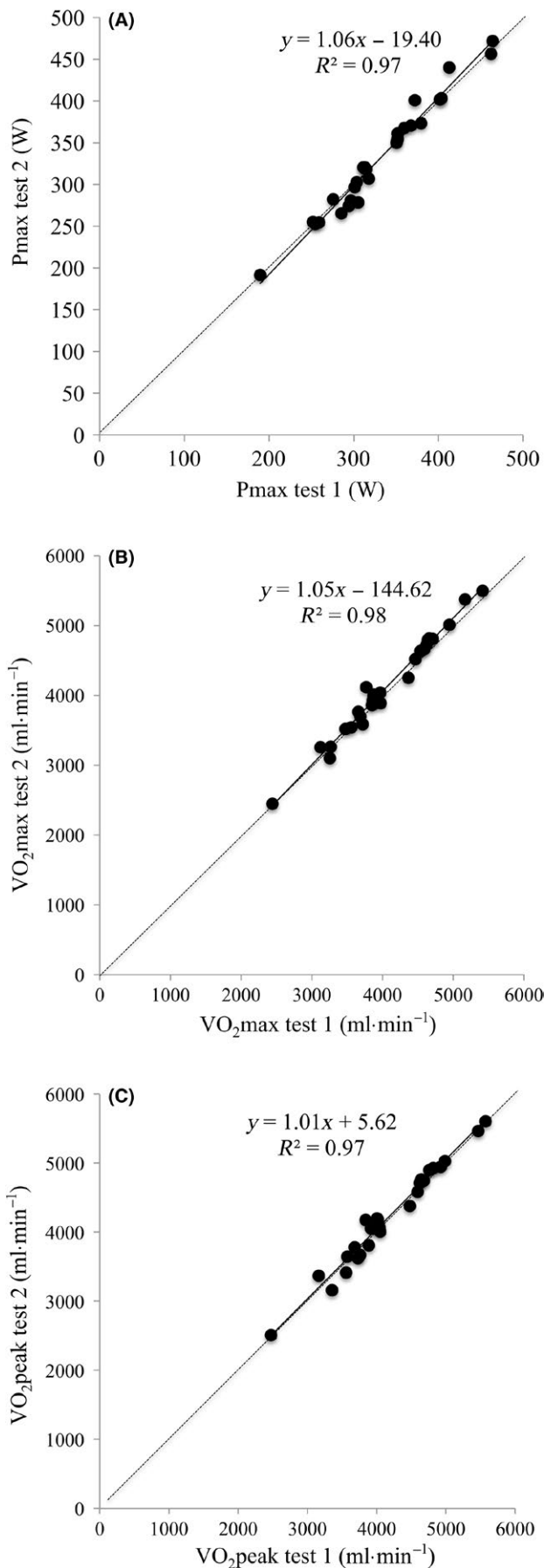
### 3.4 | Comparison Omnicol vs Oxycon Pro

The correlation between the  $\text{VO}_{2\text{max}}$  values as measured by the Oxycon Pro and Omnicol was 0.92 ( $R^2 = 0.83$ ;  $P < 0.001$ ). The slope ( $1.11$ ;  $95\% \text{ CI } 0.75\text{-}1.41$ ) and intercept ( $135 \text{ mL}\cdot\text{min}^{-1}$ ;  $95\% \text{ CI } -1052 \text{ to } 1323$ ) of the regression lines were not significantly different from 1 or 0, respectively (Figure 4A). Lin's concordance correlation coefficient ( $R_c$ ) was 0.77, and  $\text{VO}_{2\text{max}}$  values with the Oxycon Pro (second visit) were significantly lower with  $P < 0.001$  ( $t$  test) in comparison with Omnicol tests (third visit;  $P < 0.001$ ;  $t$  test) with mean  $3570$  vs  $4061$  and difference SD  $361 \text{ mL}\cdot\text{min}^{-1}$ . Omnicol visit 1 was not chosen for this comparison, yet values confirmed the lower Oxycon results with  $P < 0.003$ . No difference was observed in  $P_{\text{max}}$  or maximal heart rate ( $P_{\text{max}}$  mean difference  $2 \pm 5 \text{ W}$ ,  $\text{HR}_{\text{max}}$  mean difference  $2.5 \pm 3.3 \text{ BPM}$ ;  $P > 0.05$ ). The Bland-Altman plot showed a systematic bias that was proportional to the absolute  $\text{VO}_{2\text{max}}$  for the Oxycon Pro (Figure 4B). This is also visualized by plotting  $\text{VO}_{2\text{max}}$  vs  $P_{\text{max}}$  for both the Omnicol and Oxycon Pro (Figure 4C).

Visual analog scales (VAS) showed a significant difference in ease of breathing in advantage of the Omnicol, both overall and specifically at maximum exertion:

For breathing at maximum performance in the final stage, subjects reported mean  $\pm$  SD group scores for Omnicol of  $7.6 \pm 1.2$  vs Oxycon with  $5.1 \pm 2.7$  ( $P < 0.01$ ,  $n = 13$ ). Intra-individual values calculated with a paired  $t$  test were highly significant with  $P < 0.001$ ,  $n = 13$ .

For breathing over the complete duration of the CPET, subjects reported mean  $\pm$  SD group scores for Omnicol of  $7.7 \pm 1.2$  vs Oxycon with  $5.8 \pm 2.6$  ( $P < 0.03$ ,  $n = 13$ ).



**FIGURE 2** Scatterplot showing the test-retest reproducibility in (A)  $P_{\max}$ , (B)  $VO_{2\max}$  and (C)  $VO_{2\text{peak}}$  as measured with the Omnical. The full line represents the regression line, and the dotted line indicates the line of identity

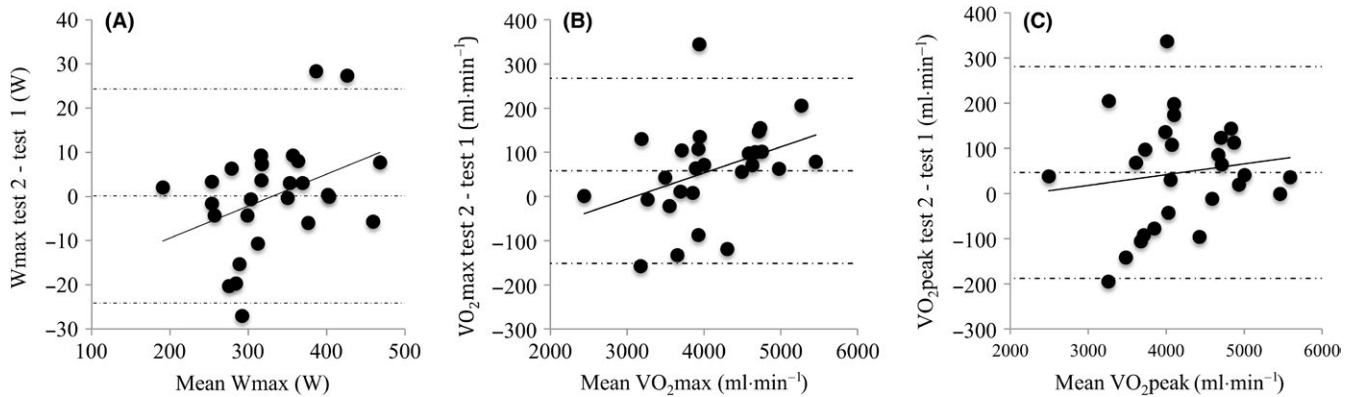
Intra-individual values calculated with a paired  $t$  test were also significant ( $P < 0.01$ ,  $n = 13$ ).

In Figure 1, the size of the pathway for breathing is visualized for both calorimeters. The Oxycon triple-V sensor has an inner diameter of 2.6 cm and an in-line sputum catcher (metal mesh) estimated at 25% of total surface, and it provides  $\pm 400 \text{ mm}^2$  surface for breathing flow. The Omnical has an open pathway with an inner diameter of 3.5 cm, and it provides  $\pm 960 \text{ mm}^2$  surface for breathing flow. Other aspects are the resistance and functionality of the turbine (Oxycon) and of any remaining resistance of the grander flow (Omnical).

## 4 | DISCUSSION

This is, to the best of our knowledge, the first study that shows that total-capture of exhalation diluted into a free-flow indirect calorimeter (Omnical) is capable of accurately measuring  $VO_2$  concentrations in athletes. We performed a technical and biological validation of an open circuit diluted flow calorimeter with total capture and analysis of participants' exhaled air until maximal CPETs. The stringent technical validation was deemed a requirement to prove suitability of the individual calorimeter for the full range of technical and biological variation expected, and this throughout the duration of the study. In contrast, no single application of calorimeter or calorimetric method should be considered a gold standard simply by reference to other units in literature; in that case, there is no proof that the application or individual apparatus performed as well. Our results showed consistent values during both validation tests, demonstrating that the Omnical was suitable for gas analysis at rest and during maximal intensity exercise testing in elite athletes and that it is preferable to devices in breath-by-breath mode, and specifically as researchers may validate the unit available on-site and fully independent. A suggestion for improvement would be to add water vapor to infusion tests as well,<sup>26</sup> though for diluted flow (Omnical) the vapor pressure should be kept below saturation as BTP conditions in reality are only applicable for approximately 50% of time (exhalation time).

The association between  $P_{\max}$  and  $VO_{2\max}$  showed a marked linearity across the entire range of  $P$ , indicating a low intra-individual variability for both parameters. At the level of basal metabolic rates (BMR), intra-individual variability was similar to previously reported results, acquired using a near-identical outpatient protocol and BMR measurement equipment (Omnical). Since the intra-individual BMR



**FIGURE 3** Bland-Altman plots showing the mean difference and 95% limits of agreement between test 1 and 2 on the Omnical. (A)  $P_{\max}$  (regression line  $y = 0.07x - 23.9$ ;  $R^2 = 0.15$ ;  $P < 0.05$ ), (B)  $VO_{2\max}$  (regression line  $y = 0.06x - 183$ ;  $R^2 = 0.15$ ;  $P < 0.05$ ), and (C)  $VO_2$  peak (regression line  $y = 0.02x - 54$ ;  $R^2 = 0.02$ ;  $P = 0.48$ )

variability as reported by Adriaens et al<sup>15</sup> was considered low,<sup>27</sup> the results of the present study illustrate strict adherence to the outpatient protocol by participants.

Our results furthermore indicated that the Omnical achieved a high degree of technically validated accuracy over the full range of human energy expenditure. The importance of reliable test and retest values of  $VO_{2\max}$  comes from a study by Vickers et al,<sup>28</sup> who studied errors in  $VO_{2\max}$  and proposed a guideline for corrections and some interesting future research recommendations. We have been able to address some of the issues and the fact that in our study, high accuracies have been achieved using participant generated data confirm the technical validation. More importantly, it implies that human maximal performance is very reproducible and almost perfectly linearly associated with  $P_{\max}$  across the entire range of  $P$  and  $VO_2$ .

As expected, complete capturing and analysis of exhaled air provide linear and full-range calorimetric performance. A challenging concern of our hypothesis is whether our findings are in line with the current scientific literature, which commonly expresses  $P_{\max} - VO_{2\max}$  as a single average. We found several studies describing the relationship between  $P_{\max}$  and  $VO_{2\max}$ , with  $P_{\max}$  values in the range from 100% to 133% of that observed in the present study.<sup>29-36</sup> Assuming  $P$ , expressed as Watts, is an independent and calibrated parameter, we used  $P_{\max}$  as a starting point for the comparison. In this study, we determined the mean  $VO_{2\text{BMR}}$  and  $P_{\text{BMR}}$  to be  $258 \text{ mL}\cdot\text{min}^{-1}$  and  $0.0 \text{ w}$  workload, and the mean  $VO_{2\max}$  and  $P_{\max}$  to be  $4112 \text{ mL}\cdot\text{min}^{-1}$  and  $331.4 \text{ w}$  ( $n = 27$ ) workload, respectively. A mean  $VO_2$  of  $4112 - 258 = 3854 \text{ mL}\cdot\text{min}^{-1}$  was required for  $331.4 \text{ w}$ , that is,  $11.63 \text{ mL}\cdot\text{min}^{-1}\cdot\text{w}^{-1}$ .

This study's curve is derived as:

$$VO_{2\max}(P_{\max}) [\text{mL}\cdot\text{min}^{-1}] = 258 [\text{mL}\cdot\text{min}^{-1}] + 11.63 \bullet P_{\max} [\text{w}].$$

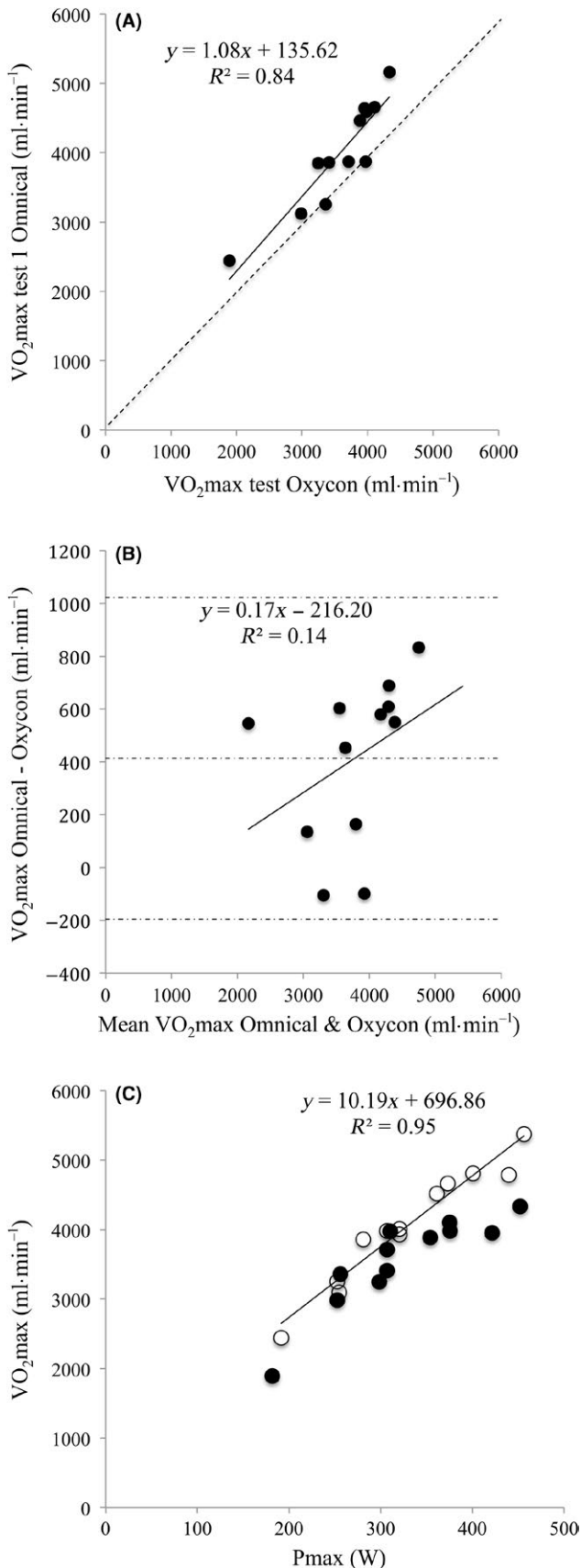
Using this curve with published mean  $P_{\max}$  values from the similar studies<sup>29-36</sup> provides us with predicted  $VO_{2\max}$

values as if measured in our laboratory. This allows comparison of predicted values with respective published mean  $VO_{2\max}$  values (Figure 5). Eight out of the nine studies used breath-by-breath analysis, the ninth used separate turbine and exhaled air sampling. None captured total subject exhaled air for analysis. We have to mention, however, that this comparison inherently includes dissimilarities in participants, laboratory conditions,  $P_{\max}$ -calibration, and other factors. We observed an average deviation of  $-316$  (range  $-59$  to  $-648$ )  $\text{mL}\cdot\text{min}^{-1}$ , that is,  $-6.4$  (range  $-1.4$  to  $-13.9$ ) %, between predicted and measured  $VO_{2\max}$ . In view of the successful technical validation over the full range of  $P$ , we conclude that  $VO_{2\max}$  has typically been underestimated in previous studies, and this is in agreement with earlier findings for specific calorimeters.<sup>7,8,37</sup>

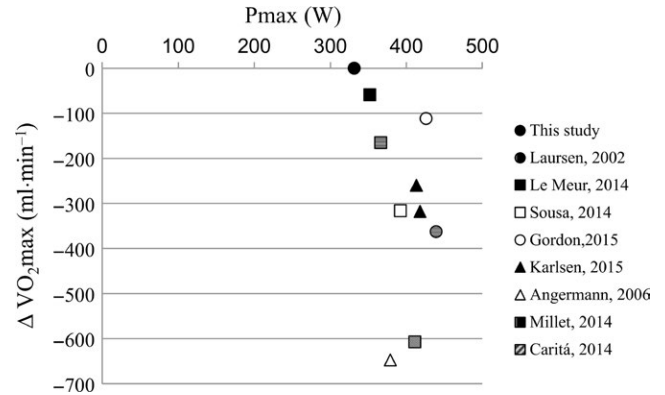
The low intra-individual variability in  $P_{\max}$  and  $VO_{2\max}$  that we observed, along with their co-linearity, implies that participants can be used to validate calorimetry equipment. In fact, we advocate that calorimetry equipment should not just show good technical validation results, but must also be able to reproduce this in a real-world participant-measuring setting. This is supported by the fact that results obtained using the Oxycon Pro suggest that this particular device, albeit serviced by the distributor, did not perform as well as the device tested previously by our group.<sup>22</sup> We surmise that not all calorimeters perform to the highest standards each and every day or with each and every unit; it is imperative to check performance of calorimeters on a regular basis.<sup>15</sup>

The Oxycon Pro and results found in literature (Figure 5) tend to underestimate  $VO_2$  in the higher range in comparison with results found by total capturing and analysis of exhaled air. These results are not surprising since at more intense efforts linearity is likely to disappear due to the increasing difficulty for analyzers in breath-by-breath mode to keep up with the increasing minute ventilation.

It is our experience that calorimetric carts are typically tested using alcohol combustion in ventilated hood mode.



**FIGURE 4** A, Scatterplot showing the test-retest variability in  $VO_{2max}$  between the Omnical and Oxycon Pro systems ( $n = 12$ ). The solid line represents the regression line and the dotted line the line of identity. B, Bland-Altman plot for  $VO_{2max}$  as measured with the Omnical vs the Oxycon Pro system. Mean bias and 95% limits of agreement are indicated with dotted lines. The regression line shows there is a systematic bias that is proportional to the measured  $VO_{2max}$ . C,  $VO_{2max}$  plotted against  $P_{max}$  for both the Omnical (open circles) and Oxycon Pro (closed circles)



**FIGURE 5** Scatterplot showing relation of average  $VO_{2max}$  vs  $P_{max}$  for 10 separate sources in literature including this article (Omnical) and using BMR for defining intercept (this article,  $VO_{2max} = 258 + 11.63 \times P_{max}$ ). The graph illustrates differences in average  $VO_{2max}$  between sources. From literature, no individual values for the relation  $VO_{2max} - W_{max}$  could be derived; thus, averages are compared only to the extent that sources have similar subjects with similar BMR and other variables

Alcohol combustion kits and descriptions for doing this are part of most product lines on offer, as an option. Yet, with respect to breath-by-breath mode, this only validates part of

the calorimeter, and specifically disregards the important response time of gas analysis in relation to required synchronous operation with participant in halation and exhalation. Performing a validation test in ventilated hood mode will provide basic feedback on operability in the absence of validation equipment for breath-by-breath mode. While it is better than no test at all, the presence of trained stable athletes with a validated performance on a calibrated workload should prove to be an important asset.

Possible limitations in this study originated from choices made and the reality of unavailability of the Oxycon for further testing. The choice to use cyclists may seem a limitation for other types of exercise, but the principle of reproducibility of known Oxygen uptake at a fixed workload will in general hold true for any type of exercise where the athlete is in a stable trained condition for the exercise of choice. This prerequisite of stable trained condition for the exercise of choice also circumvents problems with possible fatigue and training effects, other than getting used to a face mask, by having the athlete perform their routine type of exercise. In this study, choosing cyclists was deemed an advantage as the workload on the calibrated ergometer



provided a separate and validated value for the power of the workload over the range applied.

The choice in order of tests was Omnicol-Oxycon-Omnical, and if the Oxycon had been available in the second half of the study, this would have crossed over to Omnicol-Omnical-Oxycon but now was limited to Omnicol-Omnical. Despite the lack of randomization in order of testing, the comparison between devices did not include the first test in order to mostly eliminate adaptation of subjects to the test protocol. However, a small impact of lacking randomized order for Omnicol-Oxycon could still be present.

Reproducibility for  $\text{VO}_{2\text{max}}$  using Omnicol included a possible adaptation of subjects to the  $\text{VO}_{2\text{max}}$  protocol as the first test was always performed on Omnicol. The resulting reproducibility may therefore be considered a worst case and well achievable value or goal. This limitation in experiment order allowed investigating the adaptation of subjects to the test protocol if present, and indeed a small but significant difference was found. Eliminating this small difference by excluding first test in a series is expected to further improve physiological reproducibility.

Other practical limitations originated from limited availability of subjects for more tests, and lack of detailed data of similar tests in literature.

In conclusion, we report that the Omnicol as a free air-flow calorimeter capturing total exhalation in diluted mode is suitable for measuring human gas exchange ( $\text{O}_2$  and  $\text{CO}_2$ ) and hence energy expenditure over the full range from rest to maximal exertion and shows a significant increased ease of breathing near maximal exertion. Importantly, such a system can easily be validated by users on site over its full range and in the mode applicable for subjects. Finally, we conclude that healthy participants with a stable training status have a  $\text{VO}_{2\text{max}}$  that has similar reproducibility as technical validations.  $P_{\text{max}}$  and  $\text{VO}_{2\text{max}}$  show a linear relation over the full range of  $P$ . The latter challenges previously reported nonlinearity between  $P$  and  $\text{VO}_2$ .

## 5 | PERSPECTIVE

Cardiopulmonary exercise testing is an important tool to determine exercise capacity in patients with, for example, heart failure and cancer, as well as to determine training status and progress in athletes.<sup>33,38,39</sup> Since the introduction by Hill et al,<sup>39</sup> accurate assessment of the maximum plateau in  $\text{O}_2$  concentration has been the focus of many studies.<sup>40</sup> The concept of  $\text{VO}_{2\text{max}}$  originates from the notion that the  $\text{VO}_2$  curve flattens near exertion. Around 50% of studies nowadays assess  $\text{VO}_{2\text{max}}$  using breath-by-breath analysis of exhaled air using rapid-response gas analyzers.<sup>41</sup> In the present study, we provide evidence that open circuit indirect calorimetry is a valid approach to detect  $\text{VO}_{2\text{max}}$  and

is sensitive enough to reveal a linear association between  $\text{VO}_2$  and  $P$  over the entire range of  $P$  in both untrained and endurance-trained individuals. Our findings have implications for how results from studies reporting a plateau in  $\text{VO}_2$  near exhaustion should be interpreted. Finally, frequent validation of individual calorimeters in the modus operandi applied for subjects is deemed a historical and nowadays under-rated requirement.

## ORCID

Paul F. M. Schoffelen  <http://orcid.org/0000-0001-5680-4188>

## REFERENCES

1. Poole DC, Jones A. Measurement of the maximum oxygen uptake ( $\text{VO}_{2\text{max}}$ ):  $\text{VO}_{2\text{peak}}$  is no longer acceptable. *J Appl Physiol* (1985). 2017;122:997-1002.
2. Blair SN, Kohl HW 3rd, Paffenbarger RS Jr, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality. A prospective study of healthy men and women. *JAMA*. 1989;262(17):2395-2401.
3. Wei M, Kampert JB, Barlow CE, et al. Relationship between low cardiorespiratory fitness and mortality in normal-weight, overweight, and obese men. *JAMA*. 1999;282(16):1547-1553.
4. Totsikas C, Rohm J, Kantartzis K, et al. Cardiorespiratory fitness determines the reduction in blood pressure and insulin resistance during lifestyle intervention. *J Hypertens*. 2011;29(6):1220-1227.
5. Keren G, Magazanik A, Epstein Y. A comparison of various methods for the determination of  $\text{VO}_{2\text{max}}$ . *Eur J Appl Physiol Occup Physiol*. 1980;45(2-3):117-124.
6. Lundgren KM, Karlsen T, Sandbakk O, James PE, Tjonna AE. Sport-specific physiological adaptations in highly trained endurance athletes. *Med Sci Sports Exerc*. 2015;47(10):2150-2157.
7. Larsson PU, Wadell KM, Jakobsson EJ, Burlin LU, Henriksson-Larsen KB. Validation of the MetaMax II portable metabolic measurement system. *Int J Sports Med*. 2004;25(2):115-123.
8. Perret C, Mueller G. Validation of a new portable ergospirometric device (Oxycon Mobile) during exercise. *Int J Sports Med*. 2006;27(5):363-367.
9. Farmery AD, Hahn CE. Response-time enhancement of a clinical gas analyzer facilitates measurement of breath-by-breath gas exchange. *J Appl Physiol* (1985). 2000;89(2):581-589.
10. Prieur F, Busso T, Castells J, et al. A system to simulate gas exchange in humans to control quality of metabolic measurements. *Eur J Appl Physiol Occup Physiol*. 1998;78(6):549-554.
11. Yamamoto Y, Takei Y, Mokushi K, Morita H, Mutoh Y, Miyashita M. Breath-by-breath measurement of alveolar gas exchange with a slow-response gas analyser. *Med Biol Eng Comput*. 1987;25(2):141-146.
12. Aliverti A, Kayser B, Macklem PT. Breath-by-breath assessment of alveolar gas stores and exchange. *J Appl Physiol* (1985). 2004;96(4):1464-1469.
13. Potter CR, Childs DJ, Houghton W, Armstrong N. Breath-to-breath, "noise" in the ventilatory and gas exchange responses of children to exercise. *Eur J Appl Physiol Occup Physiol*. 1999;80(2):118-124.

14. Wessel HU, Stout RL, Bastanier CK, Paul MH. Breath-by-breath variation of FRC: effect on VO<sub>2</sub> and VCO<sub>2</sub> measured at the mouth. *J Appl Physiol Respir Environ Exerc Physiol*. 1979;46(6):1122-1126.
15. Adriaens MP, Schoffelen PF, Westerterp KR. Intra-individual variation of basal metabolic rate and the influence of daily habitual physical activity before testing. *Br J Nutr*. 2003;90(2):419-423.
16. Schoffelen P, Plasqui G. Classical experiments in whole-body metabolism: open-circuit respirometry-diluted flow chamber, hood, or facemask systems. *Eur J Appl Physiol*. 2018;118(1):33-49.
17. Prieur F, Busso T, Castells J, et al. Validity of oxygen uptake measurements during exercise under moderate hyperoxia. *Med Sci Sports Exerc*. 1998;30(6):958-962.
18. Webb P, Saris WH, Schoffelen PF, Van Ingen Schenau GJ, Ten Hoor F. The work of walking: a calorimetric study. *Med Sci Sports Exerc*. 1988;20(4):331-337.
19. Schoffelen PF, Westerterp KR, Saris WH, Ten Hoor F. A dual-respiration chamber system with automated calibration. *J Appl Physiol (1985)*. 1997;83(6):2064-2072.
20. Brown D, Cole TJ, Dauncey MJ, Marrs RW, Murgatroyd PR. Analysis of gaseous exchange in open-circuit indirect calorimetry. *Med Biol Eng Comput*. 1984;22(4):333-338.
21. Henning B, Lofgren R, Sjoström L. Chamber for indirect calorimetry with improved transient response. *Med Biol Eng Comput*. 1996;34(3):207-212.
22. Rietjens GJ, Kuipers H, Kester AD, Keizer HA. Validation of a computerized metabolic measurement system (Oxycon-Pro) during low and high intensity exercise. *Int J Sports Med*. 2001;22(4):291-294.
23. Kuipers H, Verstappen FT, Keizer HA, Geurten P, van Kranenburg G. Variability of aerobic performance in the laboratory and its physiologic correlates. *Int J Sports Med*. 1985;6(4):197-201.
24. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45(1):255-268.
25. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-310.
26. Prieur F, Castells J, Denis C. A methodology to assess the accuracy of a portable metabolic system (VmaxST). *Med Sci Sports Exerc*. 2003;35(5):879-885.
27. Donahoo WT, Levine JA, Melanson EL. Variability in energy expenditure and its components. *Curr Opin Clin Nutr Metab Care*. 2004;7(6):599-605.
28. Vickers RR Jr. Measurement error in maximal oxygen uptake tests. Naval health research center. 2003 (Report No. 04-03).
29. Angermann M, Hoppeler H, Wittwer M, Dapp C, Howald H, Vogt M. Effect of acute hypoxia on maximal oxygen uptake and maximal performance during leg and upper-body exercise in Nordic combined skiers. *Int J Sports Med*. 2006;27(4):301-306.
30. Carita RA, Greco CC, Denadai BS. The positive effects of priming exercise on oxygen uptake kinetics and high-intensity exercise performance are not magnified by a fast-start pacing strategy in trained cyclists. *PLoS ONE*. 2014;9(4):e95202.
31. Gordon D, Caddy O, Merzbach V, et al. Prior knowledge of trial number influences the incidence of plateau at VO<sub>2max</sub>. *J Sports Sci Med*. 2015;14(1):47-53.
32. Laursen PB, Shing CM, Peake JM, Coombes JS, Jenkins DG. Interval training program optimization in highly trained endurance cyclists. *Med Sci Sports Exerc*. 2002;34(11):1801-1807.
33. Le Meur Y, Louis J, Aubry A, et al. Maximal exercise limitation in functionally overreached triathletes: role of cardiac adrenergic stimulation. *J Appl Physiol (1985)*. 2014;117(3):214-222.
34. Millet G, Bentley DJ, Roels B, Mc Naughton LR, Mercier J, Cameron-Smith D. Effects of intermittent training on anaerobic performance and MCT transporters in athletes. *PLoS ONE*. 2014;9(5):e95092.
35. Sousa A, Ribeiro J, Sousa M, Vilas-Boas JP, Fernandes RJ. Influence of prior exercise on VO<sub>2</sub> kinetics subsequent exhaustive rowing performance. *PLoS ONE*. 2014;9(1):e84208.
36. Karlsen A, Racinais S, Jensen MV, Norgaard SJ, Bonne T, Nybo L. Heat acclimatization does not improve VO<sub>2max</sub> or cycling performance in a cool climate in trained cyclists. *Scand J Med Sci Sports*. 2015;25(Suppl 1):269-276.
37. Prieur F, Castells J, Denis C. A methodology to assess the accuracy of a 65 portable metabolic system (VmaxST). *Med Sci Sports Exerc*. 2003;35(5):879-885.
38. De Backer IC, Schep G, Hoogeveen A, Vreugdenhil G, Kester AD, van Breda E. Exercise testing and training in a cancer rehabilitation program: the advantage of the steep ramp test. *Arch Phys Med Rehabil*. 2007;88(5):610-616.
39. Rietjens GJ, Kuipers H, Adam JJ, et al. Physiological, biochemical and psychological markers of strenuous training-induced fatigue. *Int J Sports Med*. 2005;26(1):16-26.
40. Hill A. Muscular exercise, lactic acid, and the supply and utilization of oxygen. *Q J Med*. 1923;16:135-171.
41. Robergs RA, Dwyer D, Astorino T. Recommendations for improved data processing from expired gas analysis indirect calorimetry. *Sports Med*. 2010;40(2):95-111.

**How to cite this article:** Schoffelen PFM, den Hoed M, van Breda E, Plasqui G. Test-retest variability of VO<sub>2max</sub> using total-capture indirect calorimetry reveals linear relationship of VO<sub>2</sub> and Power. *Scand J Med Sci Sports*. 2019;29:213–222. <https://doi.org/10.1111/sms.13324>