Received: 7 October 2018 | Revised: 30 September 2020 | Accepted: 22 February 2021

DOI: 10.1111/sms.13944

ORIGINAL ARTICLE

WILEY

Validity and reproducibility of VO_2max testing in a respiration chamber

Jordi P. D. Kleinloog | **Stephen P. G. A. van Laar** | **Paul F. M. Schoffelen** | **Guy Plasqui**

Department of Nutrition and Movement Sciences, NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University Medical Centre +, Maastricht, The Netherlands

Correspondence
Jordi Kleinloog, Department of Nutrition and Movement Sciences, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands.
Email: j.kleinloog@maastrichtuniversity.nl

The aim of this study was to investigate whether VO_2max can be accurately measured in a respiration chamber. Thirty participants aged 23.4 ± 3.9 years with a wide range in VO_2max were included. Participants performed four incremental cycle ergometer tests (VO_2max) with a minimum of 5 days between tests. These tests consisted of one familiarization test with face mask, followed by two VO_2max tests in the respiration chamber and one test with face mask in randomized order. Oxygen consumption and CO_2 production were measured continuously using Omnical (Maastricht University, the Netherlands) gas analysis system. The mean VO_2max was 3634 ± 766 ml, which resulted in mean VO_2max per lean body mass of 60.8 ± 8.0 ml/kg. Repeated respiration chamber tests showed a high concordance, and no significant differences were detected between tests (Lin's concordance correlation coefficient (Rc) = 0.99; Δ70 ± 302 ml/min; $p = .38$). There was high concordance between the mean VO_2max from both respiration chamber tests and the mean face mask tests, and no significant difference (Rc = 0.99; Δ41 ± 173 ml/min; $p = .22$) was observed. The Bland-Altman plots showed no proportional bias between different tests. In conclusion, the respiration chamber has been found to be a valid and reproducible method for measuring VO_2max. New research opportunities are possible in the respiration chamber, such as maximal exercise testing during 24-hour measurements.

KEYWORDS
cycle ergometer, exercise, maximal oxygen consumption, Omnical, respiration chamber, validity

1 | INTRODUCTION

Cardiovascular fitness is an important health parameter. It is usually measured by maximal oxygen uptake (VO_2max), which is considered the gold standard.[1] High VO_2max levels have been shown to be inversely related to cardiovascular disease and mortality [2,3] Professional athletes use the VO_2max as a key parameter to assess their training status and progress.[4] The definition of VO_2max is the maximal rate of pulmonary oxygen uptake during exercise engaging a sufficient muscle mass.[5] Using an incremental exercise protocol, usually on a bicycle ergometer or treadmill, VO_2max can be accurately measured under laboratory conditions.[6-8]

For precise measurement of oxygen consumption, sophisticated gas analysis equipment is required.[9] Both inhaled and exhaled air analyses, and airflow through the system are required to accurately assess oxygen consumption. Several methods are available to achieve this. The first sports gas analysis system captured all exhaled air undiluted by using a breathing valve, analyzing the resulting

volume and gas fractions.[10,11] These gas analysis systems have for the greater part been replaced with breath-by-breath analysis systems.[12-14] The breath-by-breath method uses a face mask or mouthpiece (combined with nose clip) to breathe through a bi-directional flow sensor and evaluate synchronized gas samples and flows for sequential inhalation and exhalation (eg, Oxycon Pro, Jaeger Oxycon Pro, VIASYS Healthcare, the Netherlands).[12] An alternative method is the Omnical system (Omnical, Maastricht University) where a face mask is connected without valves or flow sensors by large bore (50 mm in diameter) tubing to the gas analysis system. For breathing, a continuous flow of air passes the face mask, capturing all exhaled air in diluted mode, and the total airstream is analyzed for oxygen and CO_2 concentrations.[9,15] The Omnical system and the breath-by-breath system depend on a mouthpiece or face mask to connect the exercising participant to the gas analysis system.[9,12]

Testing for the maximal cardiorespiratory capacity in a respiration chamber (ie, gas analysis in a chamber) could provide new opportunities for maximal or high-intensity intermittent sub-maximal exercise testing. For example, environmental testing conditions can be strictly regulated and modulated, since these respiration chambers are also strictly controlled and programmable climate chambers (ie, temperature and humidity).[9,16] Additionally, participants do not have to leave the respiration chamber anymore to perform exercise testing, instead exercise testing could be incorporated in 24-hour studies. Besides, post-exercise thermogenesis (also called excess post-exercise oxygen consumption; EPOC) could be investigated inside the respiration chamber. Also, exercise testing including gas analysis could be performed on clinical populations (eg, patients who underwent a tracheotomy). Measuring the VO_2max in a respiration chamber is technically challenging, because of the enormous difference between a participant breathing into the volume of a face mask (±0.2 L) compared to the volume of the whole chamber (18 000 L).[9,15,17] However, with the current technological advancements of the Omnical system and whole chamber gas analysis ventilation systems, it may be possible to accurately measure VO_2max in a respiration chamber. Despite the current relative rarity and the costs of a respiration chamber, this could also add to the increased demand for respiration chambers due to the increased versatility.

The purpose of this study was to investigate whether VO_2max can be accurately measured in a respiration chamber. More specifically, first the test-retest variability (reproducibility) of a VO_2max test performed in a respiration chamber was investigated. Secondly, the validity of the respiration chamber was tested by comparing the results of the VO_2max test in the respiration chamber with the test using a face mask. The third objective was to assess the physical comfort of the participant for a test in the respiration chamber compared to the test with face mask, using a preferred test questionnaire.

2 | METHODS

2.1 | Participants

Thirty healthy participants consisting of twenty-one males and nine females with a mean age of 23.4 ± 3.9 were included in the study. The anthropometric measurements resulted in a mean BMI of 22.0 ± 1.8 kg/m^2, mean fat percentage of 15.1 ± 6.2, and a fat-free mass (FFM) of 59.5 ± 8.7 kg. Based on the screening by a medical questionnaire and blood pressure measurements, it was determined whether participants were healthy and eligible to participate safely in four maximal exercise tests. All participants signed informed consent. The study was approved by the Medical Ethics Committee of the Maastricht University Medical Centre + and monitored by the Clinical Trial Centre Maastricht.

2.2 | Protocol

Participants' presence was required five times. On the first day, the medical screening was performed. The second visit (test day 1), body composition was assessed and thereafter participants received breakfast and had 45 minutes to relax before conducting the first physical fitness test. The first (familiarization) test was conducted in the examination laboratory with face mask (FM1). The following test days (at least 5 days apart), three more VO_2max tests were performed. Two tests (RM1 and RM2) were performed in the respiration chamber and one more test in the examination laboratory with face mask (FM2). The test location of these three successive physical fitness tests was assigned randomly, generated by randomizer.org.

2.3 | Body composition (including anthropometry)

On test day 1, underwater weighing was performed in the morning after an overnight fast to determine body density. First dry weight was measured to the nearest 0.01 kg (Mettler Toledo IDT+). Thereafter, underwater weight was measured to the nearest 0.01 kg (Mettler Toledo EC240), while breathing through a tube and functional residual capacity (FRC) was measured to the nearest 0.1 L using the helium dilution technique (Mijnhardt Volugraph VG-2000). From body density, fat percentage was calculated using Siri's equation,[18] and fat-free mass was then calculated and used to determine the VO_2max relative to lean body mass.

2.4 | Fitness testing

Physical fitness was assessed using an incremental test on a calibrated bicycle ergometer (Lode Excalibur Sport 1000 W/1.5 V, Groningen, the Netherlands) with the protocol adopted from Kuipers et al.,[8] while oxygen consumption (VO_2) and CO_2 production were fully captured in a grander airstream and measured continuously (Omnical, Maastricht University, the Netherlands). At the same time, the heart rate (HR) was monitored using a chest strap (Sport-tester Polar RS800CX).

The VO_2max test started with a 5-minute warm-up at a load of 75 W for women and 100 W for men. Afterward, workload was increased with 50 W every 2.5 minutes. When the respiratory exchange ratio (RER = $\frac{VCO_2}{VO_2}$) reached 1 or above, workload was increased with 25 W every 2.5 minutes until exhaustion.[8] The maximal workload achieved (Pmax) was calculated as the workload completed (Pcompleted) plus time (t) in the last stage divided by 150 and multiplied with the load increment of the final stage (ΔW): Pmax = Pcompleted + t/150 × ΔW. Pedal frequency was free of choice, but when dropping below a pedal frequency of 70RPM for 10 seconds because of fatigue the test was stopped. Throughout the whole test, participants were verbally encouraged and energetic music was played.

The abovementioned protocol was used for all VO_2max tests. The exact same cycle ergometer and same type of gas analysis system (Omnical) were used for all tests. The climate setting in the respiration chamber and in the laboratory was conducted at a room temperature of 18°C without extra fan. In the respiration chamber, the ergometer was facing the window in the door. The researches were outside, in front of the window, with continuous visual and auditory contact to the participant (hands-free intercom). The ergometer controller for adjusting workload was positioned with the researchers outside the chamber. Before each test in the respiration chamber, participants were asked to sit and rest in the chamber for about 30 minutes before starting the physical fitness test, allowing to start the exertion test from a stable level for both energy expenditure (resting levels) and chamber conditions (climate, mixing). Subsequently, the participant started the test, while the gasses in the chamber were analyzed continuously (Omnical, Maastricht University, the Netherlands). After the cycling test, the participant remained inside the chamber for approximately 20 more minutes to allow O_2 and CO_2 concentrations to return to resting levels. A comfortable chair was available for the participant to sit during rest.

2.5 | Whole chamber gas analysis

The whole chamber gas analysis has been described before[19,20] and was upgraded since.[9,16] The gas analysis system has been named "Omnical" and is used in Maastricht from the mid-80 s with its automated frequent calibration sample scheme, which is described in more detail in the Appendix S1. The climate control compartment creates a recirculation airflow that is filtered, pressurized, dampened for sound, and adjusted in temperature and relative humidity for correction of climate at the position of the participant. In brief, a diagonal downward flow passes the participant, and main flow at floor level is divided into a direct return path and a sampling flow taken at the floor all around the walls. The latter flushes walls and ceiling, and this flow feeds the outflow of the chamber and sample points, optimized for a representative sample. The inlet flow to the chamber joins the direct return path and the sampling flow at the inlet of the climate control compartment. Effectively, this setup for airflow was designed to allow for fast and representative sampling which is deemed a requirement for the achieved fast response.

The airstreams inlet, outlet, and chamber are sampled and prepared for analysis as described before,[19] and the upgraded analysis now determines gas fractions in a continuous path alongside the pre-existing system. Output of continuous analyzers (1 sample·sec^{-1}) was continuously normalized with data from the frequently calibrated (96 day^{-1}) pre-existing analyzers (1 sample·min^{-1}); this method is described in more detail in the Appendix S1. Data were evaluated using recently deployed molar balance equations,[16] allowing evaluation at 5-sec intervals. The molar equations were derived to allow exact physics and decrease assumptions. The results of the continuous analysis are in summary identical to the results of the pre-existing and validated minute evaluation.[19]

2.6 | Preferred test questionnaire

After completing all VO_2max tests, participants were asked to fill in a preferred test questionnaire, representing the following six experiences: dry mouth, movement restriction, claustrophobic feeling, possibility to communicate, ability to perform, and general rating. The subjective rating of the participant's preference for the fitness test with face mask or in the respiration chamber was re-coded in a way that a lower scoring indicated a preference for the respiration chamber and a high score a preference for the face mask test (scale: 1–5; neutral score = 3).

2.7 | Statistics

Calculations were done using the following software: Microsoft Office Excel 2016, IBM SPSS statistics 23, GraphPad Prism 6, and Mac labview. All variables are expressed as mean ± SD. VO_2max was defined as the mean maximum oxygen consumption over 30 seconds

TABLE 1 Descriptive statistics of VO_2max and Pmax values

	VO_2max 1 (ml/min)	VO_2max 2 (ml/min)	Mean VO_2max (ml/min)
Face mask	3607 ± 773 (2330–4915)	3660 ± 763 (2471–4976)	3634 ± 766 (2415–4946)
Respiration chamber	3620 ± 844 (2238–4937)	3690 ± 827 (2348–5520)	3593 ± 805 (2310–5229)
	Pmax 1 (W)	**Pmax 2 (W)**	**Mean Pmax (W)**
Face mask	293 ± 63 (185–401)	297 ± 63 (182–406)	294 ± 62 (184–404)
Respiration chamber	297 ± 65 (166–400)	308 ± 61 (191–400)	299 ± 61 (179–400)

Note: Data are presented as mean ± SD (range).

based on 5-second data expressed in absolute values (ml/min). Variables were checked for normal distribution and were transformed if necessary. For statistical significance, alpha was set at .05. To determine reproducibility of the respiration chamber, VO_2max as measured during the two tests in the respiration chamber was compared by a paired-samples *t* test and linear regression to obtain the correlation, intercept, and slope. Furthermore, Lin's concordance coefficient was calculated to test concordance by taking into account the variation from the line of identity. Also, a Bland-Altman plot was used to quantify systematic and random error. To determine within-participant reproducibility, the coefficient of variation (CoV; ratio of the standard deviation to the mean) was calculated. By comparing the VO_2max consumption from the respiration chamber with that of the test with face mask the validity was determined. This was done using the same statistical analyses as was used for reproducibility. For statistical analyses of the respiration chamber in comparison with the face mask test, we averaged the data from both VO_2max tests in the respiration chamber; if only one measurement was available for a participant, this was used. To investigate which test environment participants favored, non-parametric one-sample Wilcoxon's signed-rank test was used with hypothesized median of 3 (neutral answer).

3 | RESULTS

3.1 | Participants

The mean VO_2max of the participants was 3634 ml ± 766 (ranging from 2415 to 4946 ml), which resulted in mean VO_2max per lean body mass of 60.8 ml/kg ± 8.0 (ranging from 45.4 to 78.1 ml/kg). VO_2max per total body mass was 51.9 ml/kg ± 9.1 (ranging from 36.5 to 70.3 ml/kg). All participants reached their VO_2max based on the requirement that RER exceeded 1.15 (mean RER 1.18 ± 0.05) except for three tests in the respiration chamber. Seven respiration chamber measurements were excluded prior to performing statistical analyses. Four measurements were excluded because of protocol violation; during one measurement, the door of the respiration chamber was opened too soon, resulting in absence of data at the end of the maximal exercise phase. Three consecutively performed tests were excluded because of wrong flow settings, which did not become apparent during the measurements. Additionally, three measurements were excluded due to problems arising from the fact that subjects entered the chamber shortly after another subject from a different experiment had left. The other experiment used temperature settings above 30°C, whereas the current experiment needed a cool environment. Cooling the metal chamber content instantly to 18°C revealed a periodic cycle of the air-conditioning as it was working to keep the air temperature at the constant level of 18°C, while also correcting for the heat still being radiated by the metal mass (±2 metric tons) that had been warmed up to 30°C. The cycling is an effect of the control system that regulates the temperature inside the chamber. This cyclic working of the air-conditioning was deemed to affect short-term shrinking and expanding of enclosed air-mass and posed an unintended confounder on time resolution. Given this lack of correct preparation before the start of the experiment, these three tests were excluded before further statistical analysis took place. Therefore, for the reproducibility analysis the included number of participants was 23 and for validity analysis of the respiration chamber versus face mask n = 30. On request of the reviewer, we performed the reproducibility and validity analysis of VO_2max in the respiration chamber including all measurements, and these results are described in the Appendix S1 and are shown in red in Figures S2 and S3. All data were normally distributed, mean, standard deviation (SD), minimum, and maximum were calculated (Table 1). Table S1 shows these measures of the data including all measurements.

3.2 | Reproducibility VO_2max respiration chamber

The correlation between the two maximal exertion tests performed in the respiration chamber was 0.94 ($p < .001$) for VO_2max and 0.99 ($p < .001$) for Pmax (Figure 1). Both

VO_2max (p = .38) and Pmax (p = .06) did not significantly differ from each other in both respiration chamber tests (Table 1). For VO_2max, the slope of the regression line was not significantly different from 1 (0.92; 95% CI 0.77 to 1.08) and the intercept was not significantly different from 0 (345; 95% CI −243 to 934) (Figure 1A). For Pmax, the slope was not significantly different from 1 (0.95; 95% CI 0.88 to 1.02), and the intercept was not significantly different from 0 (21; 95% CI −1 to 42) (Figure 1B). Lin's concordance correlation coefficient (Rc) for VO_2max between the two tests in the respiration chamber was 0.99. The Bland-Altman plots did not show a systematic bias proportional to the measured value (Figure 1C,D). Furthermore, the results showed high within-participant reproducibility with a CoV of 4.2% ± 3.4.

3.3 | Validity respiration chamber versus face mask VO_2max (Omnical 4)

There was no significant difference between the first VO_2max test with face mask in comparison with the other three VO_2max tests (p = .77). Besides, the within-participant reproducibility was high for the face mask tests with a CoV of 2.1% ± 1.6, indicating there was no learning effect between the first face mask test and the consecutive tests; therefore, also the mean VO_2max values of the two face mask tests were used.

The correlation between the VO_2max values as measured in the respiration chamber and with face mask was 0.98 (p < .001) (Figure 2A). The slope (0.93; 95% CI 0.85 to 1.01) was not significantly different from 1, but

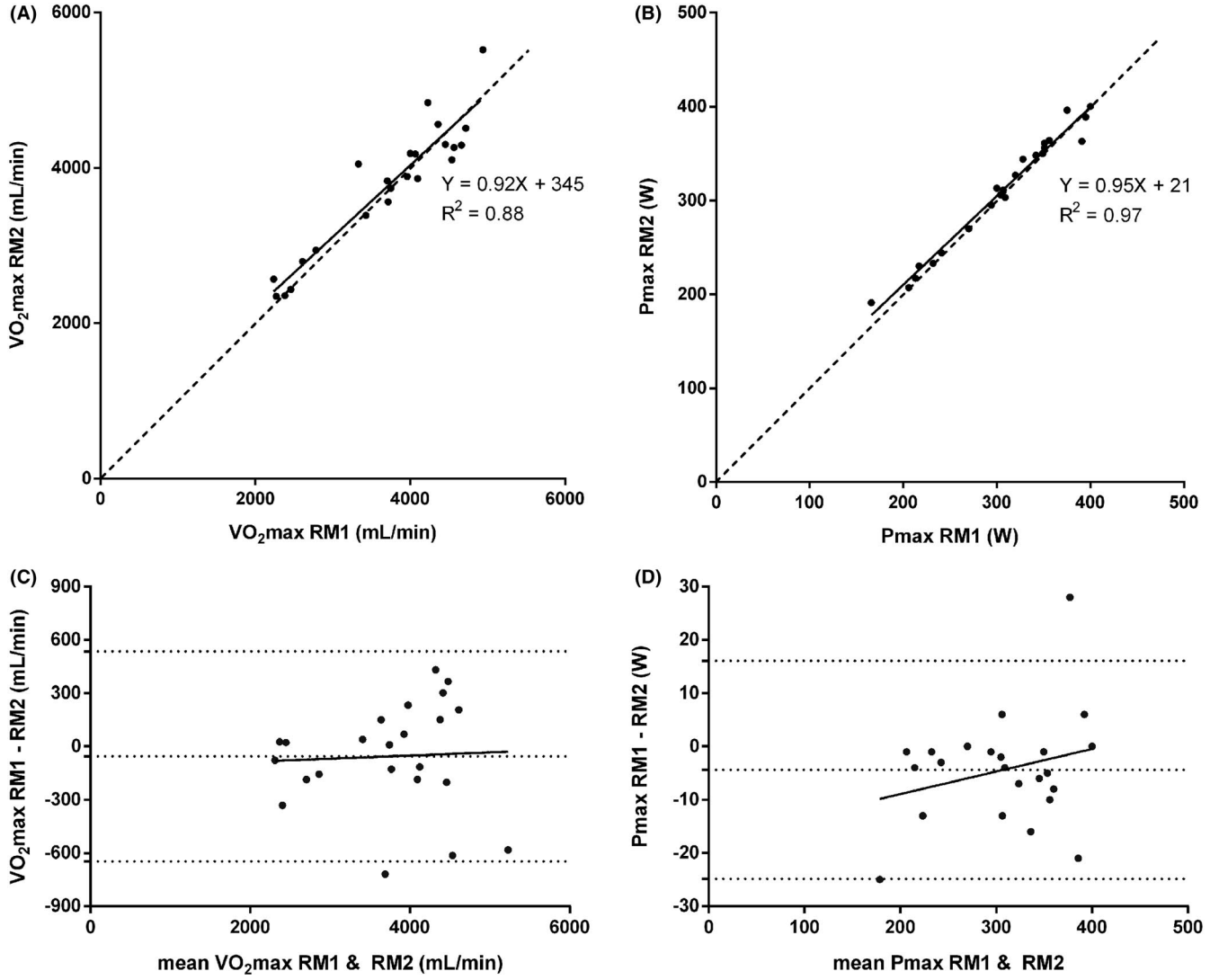

FIGURE 1 Reproducibility plots. (A) Scatterplot showing the test-retest of VO_2max as measured in the respiration chamber; (B) the Pmax as calculated for the respiration chamber tests. The solid line represents the regression line and the dotted line the line of identity. Bland-Altman plots of (C) the VO_2max as measured during the respiration chamber tests, mean bias 56 ml and 95% limit 647–535 ml; (D) the Pmax as calculated for the respiration chamber, mean bias 4 W and 95% limits 25–16 W. Mean bias and 95% limits of agreement are indicated with dotted lines. RM1: first test performed in the respiration chamber, RM2: second test performed in the respiration chamber

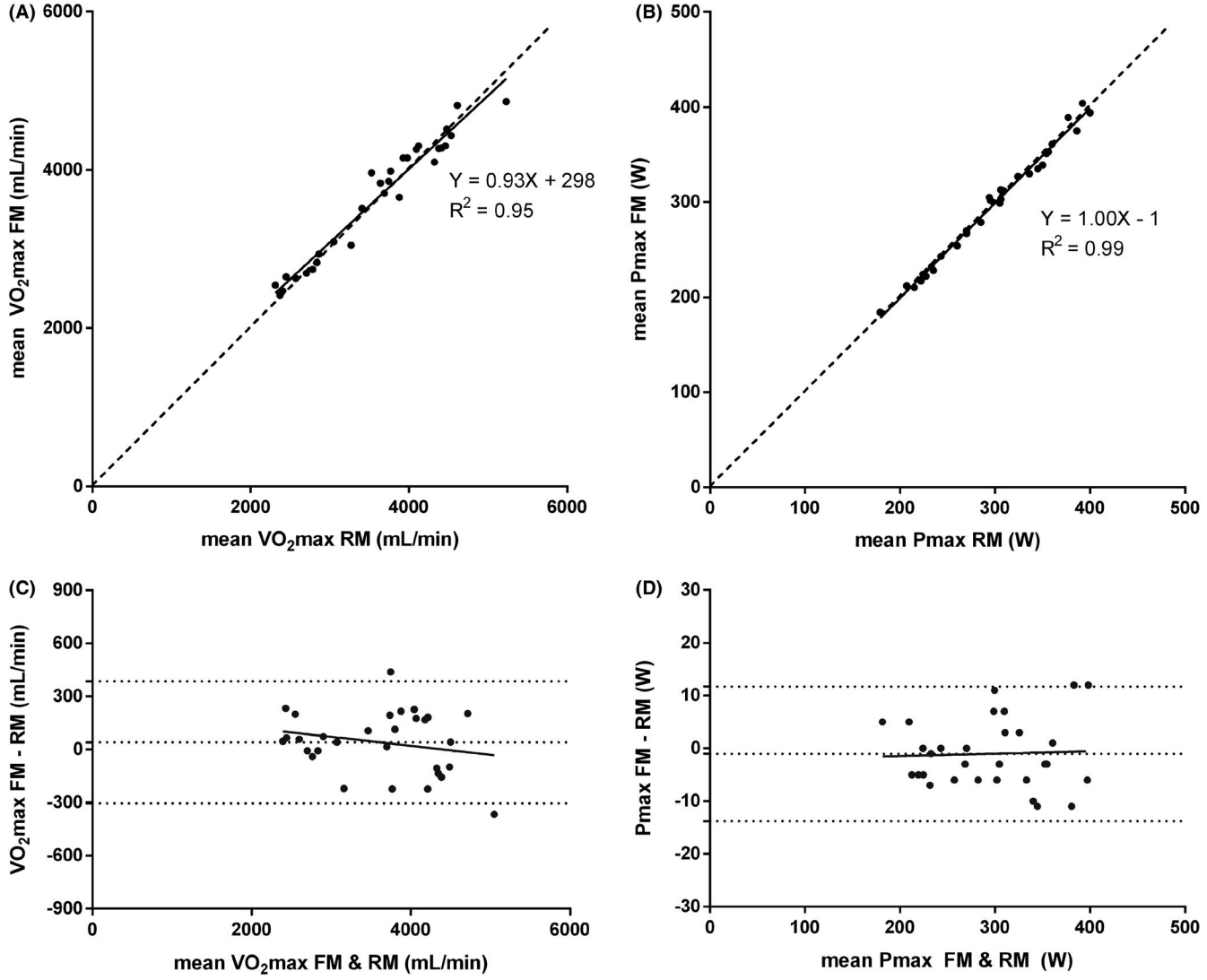

FIGURE 2 Validity plots. (A) Scatterplot showing the test-retest of VO_2max as measured in the respiration chamber versus with face mask; (B) the mean Pmax as calculated for the respiration chamber versus with face mask tests. The solid line represents the regression line and the dotted line the line of identity. Bland-Altman plots showing the mean (C) VO_2max of the respiration chamber versus test with face mask, mean bias 40 ml and 95% limits 305–385 ml; (D) Pmax as calculated for the respiration chamber tests versus the test with face mask, mean bias 1 W and 95% limits 14–12 W. Mean bias and 95% limits of agreement are indicated with dotted lines. FM, face mask; RM, respiration chamber

the intercept (298; 95% CI 4 to 592) was significantly different from 0 (Figure 2A). Lin's concordance correlation coefficient (Rc) was 0.99. There was no significant difference between the mean VO_2max of the tests in the respiration chamber and with face mask ($p = .22$). The same was found for the Pmax ($p = .23$). The correlation of the Pmax between the mean of the tests in the respiration chamber and with face mask was 0.99 ($p < .001$) (Figure 2B). The slope (1.00; 95% CI 0.96 to 1.04) and intercept (1; 95% CI −13 to 12) were not significantly different from 1 and 0 (Figure 2B), respectively. Just as with the reproducibility, the Bland-Altman plots did not show a systematic bias proportional to the measured value (Figure 2C,D).

3.4 | Preferred test questionnaire

After recoding, the results of the preferred test questionnaire were tested with the one sample Wilcoxon's signed-rank test equaled 3. Overall participants favored the respiration chamber on all tested parameters (combined score: 2.4; $p < .001$). Only Q4, representing the possibility to communicate, was rated equal to the face mask test (Table 2).

4 | DISCUSSION

This study was designed for testing validation and reproducibility of VO_2max testing in a respiration chamber.

TABLE 2 Descriptive statistics of preferred test questionnaire including resulting *p*-value of one sample Wilcoxon's signed-rank test (scale: 1–5; neutral score = 3)

	Mean score	*p*-Value
Q1 dry mouth	2.5 ± 0.9	.005
Q2 movement restriction	2.0 ± 0.6	<.001
Q3 claustrophobic	2.7 ± 0.6	.021
Q4 communication	2.9 ± 1.2	.524
Q5 performance experience	2.5 ± 1.2	.019
Q6 general rating	2.2 ± 0.8	<.001
Combined score	2.4 ± 0.5	<.001

Note: Data are presented as mean ± SD.

Results show that the respiration chamber achieved a high degree of reproducibility over the full range of oxygen consumption. There was no significant difference between the tests in the respiration chamber ($p = .38$; Lin's concordance correlation coefficient [Rc] = 0.99) and between the tests in the respiration chamber compared and the face mask tests for VO_2max ($p = .22$; Lin's concordance correlation coefficient [Rc] = 0.99). This suggests that the respiration chamber is a reproducible and valid method for determining VO_2max.

Several studies have examined different VO_2max testing instruments for reproducibility.[21-26] Typically found reproducibility for VO_2max is consistently mentioned as less than five percent (CoV < 5%). In the study by Crouter et al.,[22] three methods have been compared. The MedGraphics VO2000 showed the highest within-participant variability (CV 14.2–15.8%). The Douglas bag method (CV 5.3–6.0%) and the Parvo Medics TrueOne 2400 (CV 4.7–5.7%) showed a similar within-participant reproducibility. In comparison with these methods, the respiration chamber showed a lower within-participant variability (CV ~4.2%) indicating a high reproducibility.

In a recent study, the Omnical 4 was technically validated and tested for reproducibility as well as compared to an Oxycon Pro system.[9] Results revealed a linear relation for Pmax and VO_2max for the Omnical 4 with better reproducibility than the Oxycon Pro system. The Oxycon Pro tended to underestimate the VO_2max at higher levels of oxygen consumption. Besides, the Oxycon Pro showed a non-linear relation between Pmax and VO_2max. This particular Oxycon Pro system underestimated the VO_2max with 10.4% relative to the Omnical. In previous studies with various breath-by-breath systems, this non-linearity was also found, with an underestimation range of −1 to −30%.[9] This broad range of underestimation challenged the reality of underestimation, that is, of the non-linearity between Pmax and VO2max. The Omnical 4 system furthermore showed a high degree of technically validated accuracy over the full range of human energy expenditure (EE), as well as a near-identical result for VO_2peak and VO_2max illustrating correct timing at the end stage of exertion.[9] Additionally, the accuracy and reliability of the used Omnical system performed best compared with eleven other gas analysis systems.[27] Hence, the Omnical with face mask was a reliable method for validating VO_2max results against those obtained from the respiration chamber. The reliability of the methods applied using Omnical can be envisioned by looking at a classic Douglas bag application: All gasses are collected and analyzed by accurate flow measurement and gas analysis. The Omnical does exactly the same yet volume and gas analyses are performed in real-time. Distinctive aspect is the fact the accuracy of whole chamber gas analyzers allows for a dilution flow where a breathing valve becomes obsolete. Hence, the Omnical simply captures all exhalation diluted in a grander airstream or for gas diluted in a whole chamber.

The respiration chambers were ultimately designed for measuring intervals longer than 30 minutes, typically 24 hours, at relatively low energy expenditure levels (typical physical activity levels between 1.35 and 1.45 metabolic equivalent of task).[15] With the technological advancements of the system, the question arose whether it was possible to accurately measure and reproduce the same results within participants during a maximal exertion test that takes on average 15 minutes. The fact that similar results were achieved for VO_2max in the respiration chamber versus face mask with relatively constant Pmax once more confirmed the finding of a linear relation between Pmax and VO2max.[9] Besides, it validated the respiration chambers as being able to follow the step protocol correctly as otherwise the typical slow response (low-pass mechanical filter) of a whole chamber gas analysis [28] would have caused underestimation. Instead, it was able to follow the slope correctly to the peak-value. Additionally, the correlation and concordance between the respiration chamber and face mask test were substantial. Therefore, it can be concluded that the respiration chamber is a valid alternative for performing a maximal exertion test and can be used to evaluate gas analysis during sports.

Maximal exertion testing in the respiration chamber offers additional opportunities, such as measuring the influence of a different climate, lighting, or oxygen concentration on oxygen consumption during exercise. The respiration chambers are strictly controlled climate chambers, as a wide range of temperatures and humidity can be pre-set. Even oxygen pressure could be modified if necessary to simulate high altitude. Additionally, studies investigating circadian rhythm incorporating exercise testing within the respiration chamber could be designed. As a result, EPOC could be investigated in a habitual environment. Also, even though sub-maximal testing in a chamber is a known application,[29,30] the influence of eating and drinking during high-intensity exercise, while measuring oxygen consumption would be possible. The respiration chamber assures not only more comfort this way, but also allows for measuring the influence of free movement biological aspects like drinking or sputum removal (spit) during maximum exertion exercise, while oxygen consumption is still being measured continuously.

The preferred test questionnaire showed clear preference for the respiration chamber in terms of dry mouth, movement restriction, claustrophobic feeling, ability to performance, and general rating. The possibility to communicate was rated equal for the respiration chamber and test with face mask. While communication was possible with help of an intercom, the volume of the sound was not always sufficient for communication during a maximal cycling exertion test with loud music. This could have led to a decrease in motivation, because they could not hear our oral motivation although this was not visible in the Pmax.[31] The respiration chamber was designed to measure people over a broad range of temperature and level of EE, though for future maximal exertion measurements, the humidification should be increased in capacity. The system was set to 55% humidity and a temperature of 18°C, but due to low-temperature cooling capacity, the humidity dropped below 50%. For maximal exertion tests, a higher humidity is preferred in combination with the lower temperature, although it should not hinder VO_2max performance.[32,33] This deficit in humidification has since been remedied in the design of the respiration chamber.

Unfortunately, we had to exclude seven respiration chamber measurements prior to performing statistical analyses. These measurements were excluded because of protocol violation, either the flow settings were wrong or the door of the chamber was opened too soon. Three were excluded because our study required a temperature of 18°C, while another study preceding required high-temperature settings (32°C) and conforming to protocol should have been cooled down hours in advance. Although the room air temperature was 18°C after waiting an hour between experiments, remaining warmth contained in the dense walls resulted in increased cooling by the air-conditioning, which was expected to be detrimental for data in the subsequent VO_2max protocol and its time resolution. This can easily be solved by allowing more time between experiments. Some limitations should be taken into account when measuring VO_2max inside the respiration chamber. Firstly, ventilatory threshold cannot be measured without an additional validated spirometer. Ventilatory data are necessary to perform several analysis suggested by Wasserman et al. There is, however, a method available to determine ventilatory threshold which is based on the infliction point of plotting VCO_2 against VO_2.[34] Second, due to availability constraints the experiments ended after reaching maximal exertion, ascertaining elevated resting metabolic rate (RMR) after exertion would have increased protocol time with approximately a full hour for reaching stable results, as the participant would have to become rested for determining RMR.

4.1 | Perspectives

This study provides information about the reproducibility and validity of the respiration chamber for measuring maximum human oxygen consumption as a response to a stepwise protocol to maximal exertion. Respiration chambers have earlier been validated as the gold standard for measuring human EE from sleeping metabolic rate [35] up to high levels of exercise,[29] yet not for VO_2max. Despite the short-term character of this study as proof of principle, the employability of exercise testing during, for example, circadian rhythm studies opens new opportunities for innovative study designs using full size respiration/environmental chambers. As a consequence, new possibilities for measuring the cardiorespiratory system during exercise in different climatic condition and room-normal living conditions will arise. Finally, the respiration chamber holds advantages for participants regarding movement restriction, feeding, talking, spitting, among others. We surmise that the formerly well-known role of whole chamber gas analysis as a gold standard for testing smaller devices will resurface with the advent of more, smaller yet hard to validate devices. This is also reflected by the increasing number of chamber gas analysis systems around the world.

ACKNOWLEDGEMENTS

We would like to thank Marc Souren for the additional technical assistance, while conducting the VO_2max tests. The authors declare to have presented the results of this study clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

CONFLICT OF INTEREST

The authors declare to have no conflict of interest to report.

DATA AVAILABILITY STATEMENT

The datasets analyzed during the current study are available from last author on reasonable request.

ORCID

Jordi P. D. Kleinloog https://orcid.org/0000-0001-5861-9656
Paul F. M. Schoffelen https://orcid.org/0000-0001-5680-4188
Guy Plasqui https://orcid.org/0000-0003-4629-6479

REFERENCES

1. Vanhees L, Lefevre J, Philippaerts R, Martens M, Huygens W, Troosters T. How to assess physical activity? How to assess physical fitness? *Eur J Cardiovasc Prev Rehabil*. 2005;12(2):102-114.
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. *J Am Med Assoc*. 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. *J Am Med Assoc*. 1999;282(16):1547-1553.
4. Shephard RJ. Tests of maximum oxygen intake: a critical review. *Sports Med*. 1984;1:99-124.
5. Astorino TA, Robergs RA, Ghiasvand F, Marks D, Burns S. Incidence of the oxygen plateau at VO2max during exercise testing to volitional fatigue. *J Exerc Physiol*. 2000;3(4):1-12.

6. Arts FJ, Kuipers H, Jeukendrup AE, Saris WH. A short cycle ergometer test to predict maximal workload and maximal oxygen uptake. *Int J Sports Med*. 1993;14(8):460-464.
7. Eriksen L, Tolstrup JS, Larsen S, Gronbaek M, Helge JW. A maximal cycle test with good validity and high repeatability in adults of all ages. *Int J Sports Med*. 2014;35(14):1184-1189.
8. 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.
9. Schoffelen PFM, den Hoed M, van Breda E, Plasqui G. Test-retest variability of VO2max using total-capture indirect calorimetry reveals linear relationship of VO2 and Power. *Scand J Med Sci Sports*. 2019;29(2):213-222.
10. Douglas CG. A method for determining the total respiratory exchange in man. *Proc Physiological Society*. 1911;42:17-23.
11. Unnithan VB, Thomson KJ, Aitchison TC, Paton JY. Beta 2-agonists and running economy in prepubertal boys. *Pediatr Pulmonol*. 1994;17(6):378-382.
12. 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.
13. 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.
14. Perret C, Mueller G. Validation of a new portable ergospirometric device (Oxycon Mobile) during exercise. *Int J Sports Med*. 2006;27(5):363-367.
15. Schoffelen PFM, Plasqui G. *Human Energy Expenditure biological variability and technical validation: Utilization of different formulae to calculate energy expenditure from gas exchange and substrate oxidation; an updated equation for indirect calorimetry* [Doctoral dissertation]. Maastricht, The Netherlands: Datawyse / Universitaire Pers Maastricht: Human Biology and Human Movement Sciences. 2017;37-51.
16. Schoffelen PFM, Souren M, Plasqui G. *Human Energy Expenditure biological variability and technical validation: Measurement of respiratory gas exchange: derivation of equations for and comments on true molar balance in whole room calorimetry*. [Doctoral dissertation]. Maastricht, The Netherlands: Datawyse / Universitaire Pers Maastricht: Human Biology and Human Movement Sciences. 2017;141-167.
17. Schoffelen PFM, den Hoed M, van Breda E, Plasqui G. *Human Energy Expenditure biological variability and technical validation: Test-retest variability of VO2max using total-capture indirect calorimetry reveals linear relationship of VO2 and Power*. [Doctoral dissertation]. Maastricht, The Netherlands: Datawyse / Universitaire Pers Maastricht: Human Biology and Human Movement Sciences. 2017;119-138.
18. Siri WE. Body composition from fluid spaces and density: analysis of methods. 1961. *Nutrition* (Burbank, Los Angeles County, Calif). 1993;9(5):480-491; discussion 480, 492.
19. Schoffelen PF, Westerterp KR, Saris WH, Ten Hoor F. A dual-respiration chamber system with automated calibration. *J Appl Physiol*. 1997;83(6):2064-2072.
20. 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.
21. Babineau CL. Variability of maximum oxygen consumption measurement in various metabolic systems. *J Strength Conditioning Res*. 1999;13(4):318.
22. Crouter SA. Accuracy and reliability of the ParvoMedics TrueOne 2400 and MedGraphics VO2000 metabolic systems. *Eur J Appl Physiol*. 2006;98(2):139-151.
23. Katch VL, Sady SS, Freedson P. Biological variability in maximum aerobic power. *Med Sci Exerc Sports*. 1981;14(1):21-25.
24. MacFarlane WHL. Inter-unit variability in two ParvoMedics TrueOne 2400 automated metabolic gas analysis systems. *Eur J Appl Physiol*. 2013;113(3):753-762.
25. Miles DS, Cox MC, Verde TJ. Four commonly utilized metabolic carts fail to produce similar results during submaximal and maximal exercise: 584. *Med Sci Exerc Sports*. 1992;24(5):98.
26. Yule ET, Kaminisky LA, Sedlock DA, King BA, Whaley MH. Inter-laboratory reliability of VO2max and submaximal measurements. *Med Sci Exerc Sports*. 1996;28(5):15. https://journals.lww.com/acsmmsse/Fulltext/1996/05001/INTER_LABORATORY_RELIABILITY_OF_VO2MAX_AND.87.aspx
27. Kaviani S, Schoeller DA, Ravussin E, et al. Determining the accuracy and reliability of indirect calorimeters utilizing the methanol combustion technique. *Nutr Clin Pract*. 2018;33(2):206-216.
28. Henning B, Löfgren R, Sjöström L. Chamber for indirect calorimetry with improved transient response. *Med Biol Eng Compu*. 1996;34(3):207-212.
29. Brouns F, Saris WH, Stroecken J, et al. Eating, drinking, and cycling. A controlled Tour de France simulation study, Part I. *Int J Sports Med*. 1989;10(Suppl 1):S32-S40.
30. Brouns F, Saris WH, Stroecken J, et al. Eating, drinking, and cycling. A controlled Tour de France simulation study, Part II. Effect of diet manipulation. *Int J Sports Med*. 1989;10(Suppl 1):S41-S48.
31. Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol Bull*. 1999;125(6):627-668.
32. Stensrud T, Berntsen S, Carlsen KH. Humidity influences exercise capacity in subjects with exercise-induced bronchoconstriction (EIB). *Respir Med*. 2006;100(9):1633-1641.
33. Zhao J, Lorenzo S, An N, Feng W, Lai L, Cui S. Effects of heat and different humidity levels on aerobic and anaerobic exercise performance in athletes. *J Exerc Sci Fitness*. 2013;11(1):35-41.
34. Beaver WL, Wasserman K, Whipp BJ. A new method for detecting anaerobic threshold by gas exchange. *J Appl Physiol*. 1986;60(6):2020-2027.
35. Schoffelen PFM, Plasqui G. Classical experiments in whole-body metabolism: open-circuit respirometry—diluted flow chamber, hood, or facemask systems. *Eur J Appl Physiol*. 2018;118(1):33-49.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Kleinloog JP, Laar SP, Schoffelen PF, Plasqui G. Validity and reproducibility of VO_2max testing in a respiration chamber. *Scand J Med Sci Sports*. 2021;31:1259–1267. https://doi.org/10.1111/sms.13944