Eur J Appl Physiol (2002) 86: 435–441
DOI 10.1007/s00421-001-0572-2

ORIGINAL ARTICLE

James Carter · Asker E. Jeukendrup

Validity and reliability of three commercially available breath-by-breath respiratory systems

Accepted: 6 December 2001 / Published online: 5 February 2002

Abstract Information concerning the validity and reliability of commercial on-line gas analysis systems is limited. The aim of this study was to provide a comparison of the validity and reliability of three on-line systems (Oxycon Alpha, Oxycon Pro and Pulmolab EX670) with that of Douglas bags. Two separate studies were conducted. In study 1, the three gas analysis systems were compared with Douglas bags using a metabolic simulator over four increases in ventilation. In study 2, ten subjects were split into equal groups exercising at 100 W or 150 W for 85 min on three separate occasions. Each system was used twice per visit. Study 1 demonstrated that the Oxycon Alpha and Douglas bags produced similar respiratory values over all levels of ventilation. The Oxycon Pro tended to slightly overestimate mean expiratory flow ($\dot{V}_E$), oxygen uptake ($\dot{V}O_2$), carbon dioxide production ($\dot{V}CO_2$) and respiratory exchange ratio (RER) at the higher ventilations. The Pulmolab produced large overestimations at all ventilations for $\dot{V}CO_2$ and RER (up to 26.3% away from expectations), whilst values for $\dot{V}_E$, and $\dot{V}O_2$ were slightly underestimated at higher ventilations (up to 7.5% from expectations). The results of study 2 support the findings of study 1, with the Oxycon Pro and Oxycon Alpha producing similar results compared to Douglas bags for $\dot{V}O_2$, $\dot{V}CO_2$ and RER. The coefficients of variation for $\dot{V}O_2$ and $\dot{V}CO_2$ measured using Douglas bags, Oxycon Pro and Oxycon Alpha were 3.3–5.1%, 4.7–7.0% and 4.5–6.3%, respectively, whilst that for the Pulmolab was highly variable (26.8–45.8%). The exercise study showed the Oxycon Pro and Oxycon Alpha to be both valid and reliable on-line systems for the measurement of parameters of respiration, at least at workloads up to 150 W.

Keywords Gas exchange · Exercise testing · Reproducibility · Breath-by-breath measurements · On-line systems

J. Carter · A.E. Jeukendrup (✉)
Human Performance Laboratory,
School of Sport and Exercise Sciences,
The University of Birmingham,
Edgbaston, B15 2TT, Birmingham, UK
E-mail: a.e.Jeukendrup@bham.ac.uk
Tel.: +44-121-4144124
Fax: +44-121-4144121

Introduction

The measurement of oxygen consumption ($\dot{V}O_2$) and carbon dioxide production ($\dot{V}CO_2$) are standard tools of exercise physiology that are used to assess aerobic capacity, exercise intensity and energy expenditure. In addition, measurement of $\dot{V}O_2$ and $\dot{V}CO_2$ allows indirect measures of substrate utilisation. For many years the technique for the collection and analysis of these gases has been the Douglas bag method. Although this technique is still considered the gold standard, it has several disadvantages and its own sources of error. Firstly, no breath-by-breath data can be obtained, and therefore rapid changes in ventilation or $\dot{V}O_2$ cannot be studied. Secondly, the method is time consuming due to the requirement of sampling and analysis after collection. In addition, the bags are made of PVC material, which is slightly permeable to the external air. Increasing technology, though, has seen the emergence of portable and automated on-line, breath-by-breath gas analysis systems. These systems allow the continuous measurement of gas volumes and concentrations and the immediate display of this information on-line, and therefore markedly increase the efficiency of the gas analysis procedure.

Little information is available about the validity and reliability of the measurements produced by these commercially available systems. The literature that does exist concentrates largely upon portable systems, such as the Cosmed K2 telemetry system (Kawakami et al. 1992; Lothian et al. 1993; Lucia et al. 1993; Peel and Utsey 1993), the Cosmed K4 metabolic system (McLaughlin et al. 2001) and the Aerosport TEEM 100 portable

metabolic measurement system (Wideman et al. 1996). Versteeg and Kippersluis (1989) compared the $\dot{V}O_2$ measured using three automated systems with that measured using Douglas bags, but unfortunately did not report any reliability data. More recently, Rietjens et al. (2001) compared the Oxycon Pro on-line system with that of Douglas bags but did not extend the study to encompass other on-line systems.

The purpose of the present investigation was to compare the validity and reliability of three commercially available automated on-line systems with that of Douglas bags and the Servomex 1400B4 analyser (Sussex, UK). The following systems were assessed: Oxycon Pro (Jaeger, Wuerzburg, Germany), Oxycon Alpha (Jaeger) and the Pulmolab EX670 (Morgan Medical, Kent, UK). To accomplish this aim, the investigation was split into two separate studies, study 1 using a portable metabolic simulator (Jaeger, Germany) and study 2 involving in vivo measurements in a randomised crossover design. The simulator enabled us to simulate metabolism with the aim of generating very precise and reproducible values for $\dot{V}O_2$ and $\dot{V}CO_2$. The accuracy of the respiratory gas analysis systems was determined by the extent of deviation of their values from those of the metabolic simulator and from within-subject and between-subject variation.

Methods

Study 1

Each gas analysis system was tested twice at four different levels of ventilation of the simulator over a period of 2 days. By varying the O_2 and CO_2 together with overall flow rate it was possible to control three respiratory parameters; pulmonary ventilation ($\dot{V}_E$), $\dot{V}O_2$, and $\dot{V}CO_2$. The different levels were achieved by controlling the injection of CO_2 and N_2 together with the simulator's mass flow controllers, valve system and pump. The values generated for $\dot{V}_E$ were 20, 40, 80 and 160 l/min for levels 1–4, respectively, and these corresponded to values for $\dot{V}O_2$ and $\dot{V}CO_2$ of 1.0, 2.0, 3.0 and 4.0 l/min, respectively, with a respiratory exchange ratio (RER) of 1.0. The accuracy of each system was determined from their ability to reproduce these known values of $\dot{V}_E$, $\dot{V}O_2$, $\dot{V}CO_2$ and RER generated by the simulator.

Protocol

The simulator was allowed to warm up for 30 min before two compressed gas bottles (99.99% N_2 and 99.99% CO_2, 200 kPa) were connected to it. The simulator's potentiometer was set at a respiratory frequency of 30 breaths/min and $\dot{V}O_2$ and $\dot{V}CO_2$ at zero. Prior to each test, the gas analysis systems of the Oxycon Pro and Oxycon Alpha were switched on after sufficient warm-up time was allowed (according to the manufacturer's instructions, 1 h for both the Oxycon Pro and Oxycon Alpha). The vacuum pumps of the Pulmolab mass spectrometer were running for several weeks previous to the commencement of the study. All three systems were then calibrated according to the specifications outlined by the respective manufacturers.

Douglas bags

The individual bags (Cranleigh, UK) were emptied using a vacuum pump, and the volume of air was measured with a dry gas meter (Harvard, Kent, UK). The Servomex 1400B4 analysers were subjected to a two-way calibration process; they were zeroed with a sample of 100% N_2 before the range was determined using two samples of calibration gas (BOC Gases, Surrey, UK); 80% N_2, 15% O_2, 5% CO_2 and 80% N_2, 18% O_2, 2% CO_2, respectively.

Pulmolab EX670

The Pulmolab was subjected to a three-way calibration process; a turbine calibration, a gas analyser calibration and a delay time calibration. The turbine calibration was determined using a 3-l syringe, in a ten-pump series. The acceptance criteria were 3 (0.1) l. The gas calibration was completed by attaching a mixed gas bottle (74.85% N_2, 5% CO_2, 15.07% O_2 and 5.08% Ar) to the system and initiating the automatic calibration program. The Pulmolab also provided a delay time calibration. The delay is the time between an instantaneous flow/volume event, and the corresponding gas analysis event as measured by the Pulmolab. This was calculated by the system during a brief period of sharp inhalation and exhalation through the mouthpiece by a human subject.

Oxycon Pro

This was also subject to a three-way calibration process, involving a flow-volume sensor and a gas analyser and delay time calibration. The flow-volume sensor calibration ensures that the measuring system of the Oxycon (consisting of the amplifier, Triple V, and the pressure transducer) is functioning correctly. A calibrated 3-l syringe connected to the Triple V assembly was used for this purpose. A series of six complete pumps of the syringe was repeated until the percent difference between the current and the previous volume calibration was less than 1%. The gas analyser and delay time calibration involved an automated calibration procedure, as provided by Jaeger, whereby a calibration gas at 180 kPa (16.25% O_2, 4.13% CO_2 and 79.62% N_2) was introduced to the Oxycon. The automated program was repeated until the current and previous data for gain, offset and delay time were within 1%.

Oxycon Alpha

Procedures for the Oxycon Alpha were identical to those for the Oxycon Pro.

Environmental temperature (Brannan wet and dry bulb thermometer, UK), relative humidity (Brannan wet and dry bulb thermometer) and barometric pressure (Fortin Barometer, accuracy within 0.1 Pa) were recorded and compared to the corresponding values produced by the gas analysis system being tested. Differences were corrected by entering the correct values into the relevant gas analysis system.

Once the described preparatory criteria were satisfied, the calibration test was performed. Each gas analysis system was connected in turn to the simulator, which was switched to mode 1, and a 10-min recording was completed. This was repeated a further three times for each system at modes 2, 3 and 4 in a stepwise fashion, with the total test time therefore being 40 min per system.

This protocol was completed twice for each system on 2 separate days and was identical between systems. The volume of the Douglas bags (200 l) limited their recording time to a maximum of 5 min.

Study 2

Subjects

Ten healthy active volunteers gave their informed consent to participate in the study, which was approved by the Local Ethics Committee. Subjects were aged 23 (3) years and weighed 71.8 (9.9) kg [mean(SD)]. All subjects had previously been involved in

studies or exercise involving cycle ergometry at the intensities used and were fully familiar with all of the experimental procedures.

All exercise was carried out on an electrically braked cycle ergometer (Lode Excalibur, Groningen, The Netherlands) set in the pedal-rate-independent mode. Each subject visited the laboratory on three occasions, the only difference between visits being the order of the gas analysis systems used to collect gas samples. The ten subjects were separated into two groups; the first comprised five subjects, all of who cycled at 100 W, while the second group of five subjects cycled at 150 W. These relatively low work rates were chosen because subjects would be able to maintain them for 85 min and, more importantly, negligible changes in $\dot{V}O_2$ and substrate utilisation (RER) were expected.

The cycle ergometer was fully adjustable and the subjects' preferred position was recorded and reproduced for each test. The subject cycled at the set workload for 10 min before measurements were made. The measure of respiratory gas was recorded for 5 min by each of the gas analysis systems, during which time subjects wore a nose clip. There was a 5-min break between each sampling period, during which the subject continued to exercise. Each system was measured twice in each trial. An example of the test procedure is provided in Fig. 1.

Heart rate was recorded continuously throughout the test (Polar Accurex Plus, Polar Electro, Oy, Finland) and water was available ad libitum. The test was terminated after all gas samples were collected (two per system).

Visits two and three were identical to visit one, except for the order of gas collection. The respiratory parameters recorded and compared between systems were $\dot{V}_E$, $\dot{V}O_2$, $\dot{V}CO_2$ and RER.

Data and statistical analysis

Data are reported as the mean (SD), unless otherwise stated. During the calibration test, the Oxycon Alpha, Oxycon Pro and the Pulmolab assumed that a human subject was being tested, and therefore generated all values in BTPS. As this was not the case, all recorded values from these three systems were converted to STPD for statistical analysis. A repeated-measures analysis of variance with a Latin Square Design contrast was used to distinguish any differences between mean heart rate over the three exercise visits. A Friedman *k*-related sample test was used to determine differences between the respiratory systems and the measured mean respiratory parameters. A further Wilcoxon two-related sample test was used to examine the differences between the means. Friedman and Wilcoxon tests were used to analyse day-to-day variation between respiratory systems. In all cases, the level of statistical significance was set at $P < 0.05$ and all tests were carried out on SPSS for windows (version 10.0, 1999; SPSS, Chicago, Ill., USA,). The Douglas bags were treated as the control and all tests of significance were plotted against the respective Douglas bag values.

Non-parametric tests were chosen because the skewness and kurtosis of the data sets exceeded 1.0 and 2.0, respectively, in the majority of cases.

Results

Study 1

Figure 2 shows the results for the four respiratory parameters measured by each of the gas analysis systems compared to the metabolic simulator over the four ventilation modes. The Oxycon Alpha and Oxycon Pro both produced similar values for $\dot{V}_E$ over all four modes when compared to the simulator. The Pulmolab, however, tended to underestimate $\dot{V}_E$; this became more apparent with increasing ventilation, deviating by 7.5% from the simulator (Fig. 2).

The Pulmolab, Douglas bag and Oxycon Alpha produced similar values to the simulator for $\dot{V}O_2$, the exception being at the highest ventilation mode, where the Pulmolab produced a slight underestimation of $\dot{V}O_2$ (6.6%). The Oxycon Pro consistently produced a slight overestimation of $\dot{V}O_2$ throughout all four ventilation modes of the calibration test, ranging between 5.8% and 10.5% from the expected figures (Fig. 2).

The results for $\dot{V}CO_2$ closely followed those for $\dot{V}O_2$, except for those of the Pulmolab, which consistently produced higher values for $\dot{V}CO_2$ over all modes of the simulator ($>20\%$ different from expectations). The Oxycon Pro once more produced an overestimation (10.5–11.7%), but only at the higher ventilation rates (modes 3–4; Fig. 2).

The Pulmolab was the only deviant from the simulator with respect to RER (Fig. 2).

Study 2

All ten subjects successfully completed the study. There was no difference between mean heart rates over these three visits at either workload. The mean ambient temperature was 15.1 (2.3)°C and the relative humidity was 48.3 (4.3)%.

Figure 3 illustrates the mean respiratory parameters as collected by each system for both workload groups. In both the 100 W and 150 W exercise groups all systems produced higher mean values for $\dot{V}_E$ compared to the Douglas bags; those of the Oxycon Pro and Oxycon Alpha were found to be significantly higher (Fig. 3).

In both exercise groups the Pulmolab produced significantly lower mean values for $\dot{V}O_2$ when compared to the Douglas bags. No differences for $\dot{V}O_2$ were reported between the Oxycon Pro, Oxycon Alpha and Douglas bags (Fig. 3).

No differences were reported between any of the systems for total mean $\dot{V}CO_2$ and RER in the 100 W exercise group. In the 150 W exercise group the Pulmolab produced significantly higher mean values for $\dot{V}CO_2$ and RER compared to those of the Douglas bags. As is indicated by the standard deviation, in some instances the Pulmolab values for $\dot{V}CO_2$ and RER in the 150 W exercise group far exceeded normal physiological values; >28.0 l/min and >30.0 l/min, respectively (Fig. 3).

Tables 1–4 illustrate the coefficient of variation (CV) for the mean respiratory parameters as collected by each gas analysis system for both exercise groups. The Pulmolab produced significantly higher CV figures for $\dot{V}O_2$, $\dot{V}CO_2$ and RER than the other systems when averaged for the five subjects in each exercise group (7.6–45.8%). The CV for total mean $\dot{V}_E$ of the Pulmolab (7.6–12.5%) was also the greatest for both exercise groups, but this was not significant (Tables 1, 2).

Unlike Tables 1 and 2, which show the day-to-day variation, Tables 3 and 4 show the within-system variation. The Pulmolab produced the greatest variation for

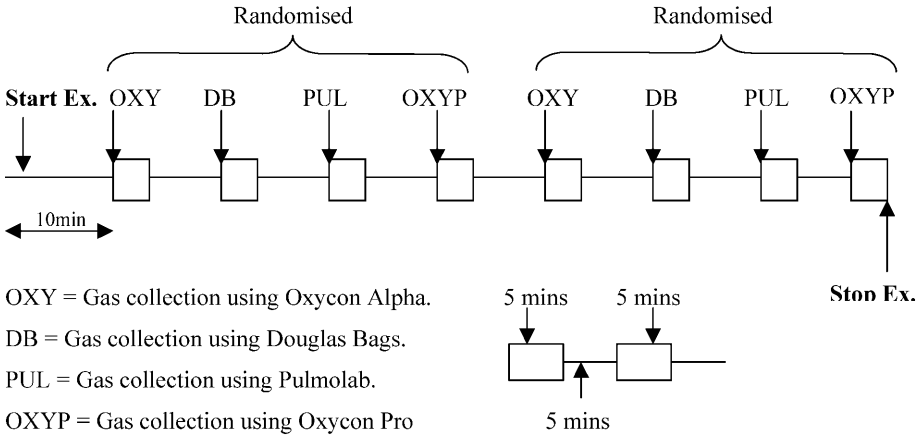

Fig. 1 Example of the test procedure used in the exercise study. (*OXY* Gas collection using the Oxycon Alpha, *DB* gas collection using Douglas bags, *PUL* gas collection using the Pulmolab, *OXYP* gas collection using the Oxycon Pro, *Start Ex.* start of the experiment, *Stop Ex.* end of the experiment)

each parameter, with values ranging from 7.3% to 22.6%. The Douglas bags, Oxycon Alpha and Oxycon Pro consistently produced small variations at both workloads, the range being 1.3–6.5%.

Discussion

Indirect calorimetry is one of the most common tools in exercise physiology. It is used for various purposes including the assessment of aerobic power, determination of exercise intensity and the measurement of energy expenditure. The Douglas bag method is considered the gold standard for indirect calorimetry, but automated systems have long been the standard means of measuring it. This study aimed to provide a direct comparison of the validity and reliability of three different automated gas analysis systems with that of Douglas bags.

The mean absolute values of $\dot{V}O_2$, $\dot{V}CO_2$ and RER achieved from the exercise testing were similar for the Oxycon Pro, Oxycon Alpha and Douglas bags.

These results suggest that both the Oxycon Alpha and Oxycon Pro are valid systems for generating accurate respiratory data for these three parameters during steady-state exercise up to 150 W. These findings are supported by Rietjens et al. (2001), who found no differences for similar respiratory variables throughout an incremental cycle test between the Oxycon Pro and Douglas bag system. In comparison, the Cosmed K2 telemetry system is consistently found to underestimate $\dot{V}O_2$ during sub-maximal exercise (Kawakami et al. 1992; Lothian et al. 1993; Peel and Utsey 1993). Peel and Utsey (1993) reported a 12.5–17% underestimation at all sub-maximal workloads, while Kawakami et al. (1992) described an underestimation at sub-maximal workloads. A 22% underestimation during maximal exercise has also been reported (Lothian et al. 1993). This consistent discrepancy has been attributed to the K2's calculation of $\dot{V}O_2$, in which it assumes, in the absence of a CO_2 electrode, that the RER is always 1.00 (Hausswirth et al. 1997). The latest Cosmed system, the portable K4 b^2, includes an infrared electrode to measure CO_2. In a recent study, McLaughlin et al. (2001) reported the K4 b^2 to be acceptable in measuring $\dot{V}O_2$ during exercise, but found it to overestimate this parameter (50–200 W) and to underestimate $\dot{V}_E$ and $\dot{V}CO_2$ (200–250 W). These small differences in $\dot{V}O_2$ and $\dot{V}CO_2$ resulted in slight differences in RER over all stages of an incremental cycle ergometer test.

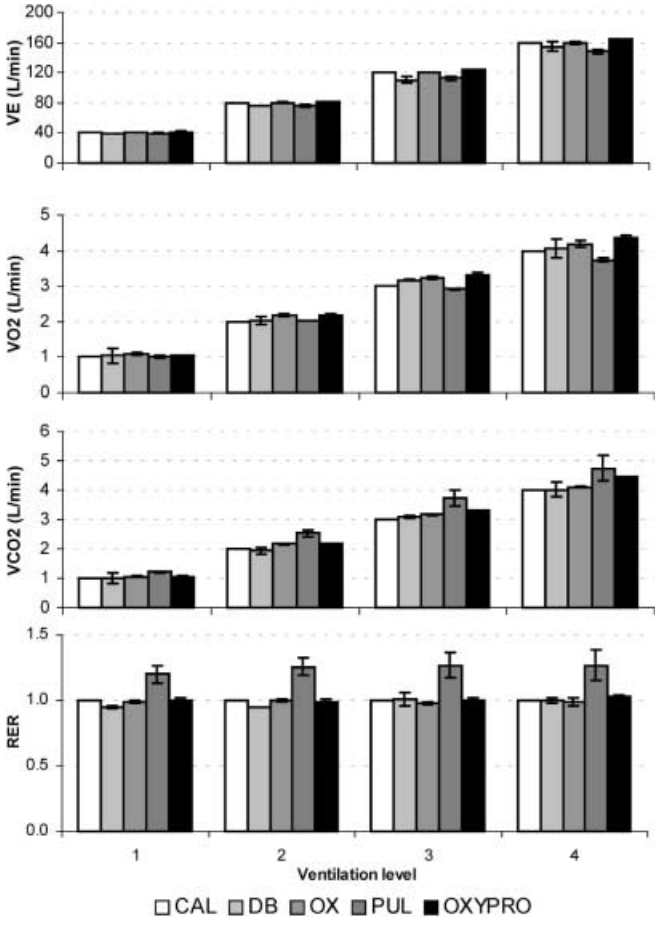

Fig. 2 Mean expiratory flow ($\dot{V}_E$), oxygen uptake ($\dot{V}O_2$), carbon dioxide production ($\dot{V}CO_2$) and respiratory exchange ratio (*RER*) for each system compared to the simulator. Data is mean ± the range. (*CAL* Calibration, *DB* Douglas bag, *OX* Oxycon Alpha, *PUL* Pulmolab, *OXYPRO* Oxycon Pro)

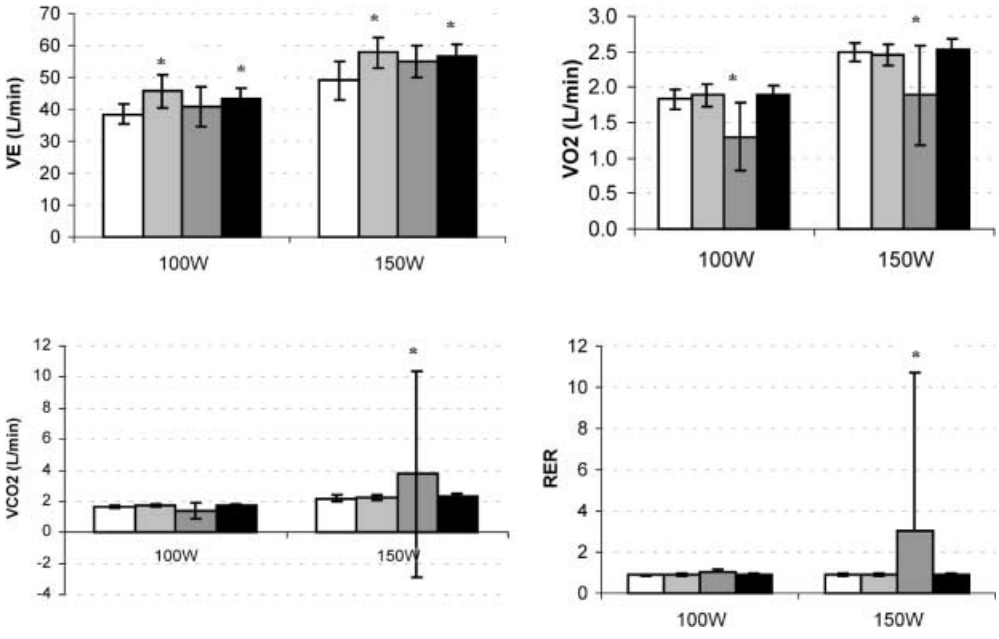

Fig. 3 Mean $\dot{V}_E$, $\dot{V}O_2$, $\dot{V}CO_2$ and RER for each system at both workloads. (*white bars* Douglas bags, *light grey bars* Oxycon Alpha, *dark grey bars* Pulmolab, *black bars* Oxycon Pro). *Significantly different from Douglas bag values

The only discrepancy in the current study was a significantly higher $\dot{V}_E$ value reported by the Oxycon Alpha and Oxycon Pro at both workloads studied. Interestingly, all systems tended to produce higher $\dot{V}_E$ values compared with those of the Douglas bags. In addition, when tested against the simulator, the Douglas bags produced lower values for $\dot{V}_E$. Although we have used Douglas bags as our control measurement, they are not without the potential for error. The bags are made of slightly permeable PVC, which can allow the diffusion of gases (Versteeg and Kippersluis 1989). The similarity between the Douglas bag $\dot{V}O_2$, $\dot{V}CO_2$ and RER data and that of both the Oxycon Pro and the Oxycon Alpha, however, can only lead us to speculate that the device used to empty and record the volumes within the bags was slightly inaccurate.

The Pulmolab proved less comparable with the Douglas bags. Despite producing similar $\dot{V}_E$ values, it produced significantly lower $\dot{V}O_2$ values and significantly higher $\dot{V}CO_2$ and RER values at both workloads. As mentioned in the results section, some of these values exceeded possible physiological limits.

The accuracy of the Pulmolab's $\dot{V}_E$ values suggests that the turbine and volume sensor were functioning correctly. However, the inaccuracy and high variability of the $\dot{V}O_2$ and $\dot{V}CO_2$ data indicates a delay time problem. On close inspection of the mouthpiece assembly it was discovered that the sample capillary tube, along which expired air was drawn into the system, was situated level to the saliva-collecting container. Therefore, a plausible explanation for the Pulmolab differences in data is that mid-test, a portion of the saliva generated by the subject entered the capillary tube causing partial blockage and thus, in some cases, irregular results.

The data concerning the day-to-day variation of the four gas analysis systems used in the present study compare favourably with data from previous studies (Table 5). Unfortunately, the majority of variability data relevant to this present study is limited to that of maximum $\dot{V}O_2$ ($\dot{V}O_{2max}$), although there is data available from time-trial studies. Hickey et al. (1992) investigated day-to-day 40-mile time-trial variation and reported the CV for mean $\dot{V}O_2$ as 3.0% and the percent $\dot{V}O_{2max}$ as 3.6%. The values generated for the CV of $\dot{V}O_2$ by our systems (excluding Pulmolab) are slightly higher than these, except for the Douglas bags at 150 W. Jensen and Johansen (1998) reported a range of CV values for test-retest parameters over a series of time trials. Although their CV value of 1.9% for $\dot{V}O_{2max}$ is markedly lower than any of our $\dot{V}O_2$ CV values, they also produced a CV for $\dot{V}O_2$ and $\dot{V}CO_2$ at 2 mM blood

Table 1 Mean (SD) coefficient of variation (CV) of all parameters for each system in the 100 W exercise group. ($\dot{V}_E$ Mean expiratory flow, $\dot{V}O_2$ oxygen uptake, $\dot{V}CO_2$ carbon dioxide production, *RER* respiratory exchange ratio)

System	$\dot{V}_E$	$\dot{V}O_2$	$\dot{V}CO_2$	RER
Douglas bags	5.1 (3.2)	5.1 (2.1)	3.9 (2.4)	3.2 (0.9)
Oxycon Alpha	7.3 (2.3)	6.3 (2.3)	4.8 (1.0)	4.1 (0.8)
Pulmolab	12.5 (10.3)	33.6 (16.0)*	33.1 (13.5)*	11.7 (5.5)*
Oxycon Pro	7.4 (1.0)	6.5 (2.0)	7.0 (1.0)	3.6 (0.6)

*Significant difference between that system and the Douglas bags

Table 2 Mean (SD) CV of all parameters for each system in the 150 W exercise group

System	$\dot{V}_E$	$\dot{V}O_2$	$\dot{V}CO_2$	RER
Douglas bags	5.7 (2.9)	3.3 (2.5)	5.0 (2.3)	3.5 (1.4)
Oxycon Alpha	6.1 (1.4)	4.5 (1.3)	5.3 (2.0)	4.3 (1.6)
Pulmolab	7.6 (1.5)	26.8 (21.3)*	45.8 (46.1)*	36.6 (53.6)*
Oxycon Pro	6.6 (1.2)	4.7 (1.2)	5.3 (1.0)	3.5 (1.4)

*Significant difference between that system and the Douglas bags

Table 3 Mean (SD) daily CV of all parameters for each system in the 100 W exercise group

System	$\dot{V}_E$	$\dot{V}O_2$	$\dot{V}CO_2$	RER
Douglas bags	2.5 (2.0)	2.5 (2.0)	2.5 (1.7)	1.3 (0.6)
Oxycon Alpha	5.9 (2.4)	4.5 (2.4)	4.2 (1.5)	2.5 (0.8)
Pulmolab	7.3 (2.6)	15.0 (9.4)	14.4 (9.0)	7.6 (5.3)
Oxycon Pro	6.5 (1.9)	5.7 (1.8)	5.6 (1.6)	2.5 (0.6)

Table 4 Mean (SD) daily CV (%) of all parameters for each system in the 150 W exercise group

System	$\dot{V}_E$	$\dot{V}O_2$	$\dot{V}CO_2$	RER
Douglas bags	4.0 (3.4)	2.3 (2.2)	1.8 (1.6)	2.1 (1.0)
Oxycon Alpha	5.5 (1.7)	4.0 (1.7)	4.1 (1.9)	2.6 (0.7)
Pulmolab	6.7 (1.9)	19.8 (20.3)	22.6 (23.7)	7.4 (5.2)
Oxycon Pro	6.1 (1.5)	4.4 (1.4)	4.8 (1.1)	2.7 (0.6)

lactate, which is of greater relevance to our experimental design. Their values of 7.4% and 7.7% for $\dot{V}O_2$ and $\dot{V}CO_2$, respectively, are distinctly higher than any values generated by our systems (bar the Pulmolab). Comparisons with data from time trials, however, are limited because work rate, and therefore respiratory responses, frequently change throughout the course of the exercise period.

Studies concerning the reliability of $\dot{V}O_2$ measures of the Cosmed K2 system are limited. Lothian et al. (1993) reported a range of 3.0–11.4% between trials, a value greater than that of the Douglas bags, Oxycon Pro and Oxycon Alpha in our current study. Unfortunately, a direct comparison cannot be made as the study by Lothian et al. (1993) incorporated both sub-maximal and maximal data, whereas we have only reported sub-maximal data.

Between systems, it is clear that the Douglas bag method possesses the lowest day-to-day variation, being consistent at both workloads. This is not surprising as Douglas bags do not have the breath-by-breath component and therefore do not have the variability factor associated with this. The Oxycon Pro and Oxycon Alpha systems are similar, but the measurements of the Oxycon Alpha resulted in slightly lower variation over the two workloads. The Pulmolab, however, with its absolute values, was widely variable in every parameter over both workloads.

The day-to-day variation was slightly larger than the within-day variation. For example the CV for $\dot{V}O_2$ measured by the Oxycon Alpha at 100 W was 6.34% day-to-day, but 4.53% for the within-day variation.

This reduction suggests that the subject as opposed to within-system variation caused the extra day-to-day variation. Kuipers et al. (1985) concluded that variability of physical performance has a physiological basis. They speculated that changes in skeletal muscle metabolic efficiency and/or changes in coordination of movements could in both cases change the energy required for a given workload. A decreased metabolic

Table 5 A summary of previous studies and their reported CV for $\dot{V}O_2$ and $\dot{V}CO_2$. ($\dot{V}O_{2\,max}$ Maximum oxygen uptake, *TT* time trial)

Author	Measure	CV (%)
Wright et al. (1978)	$\dot{V}O_{2max}$	5.1–6.8
Katch et al. (1982)	$\dot{V}O_{2max}$	3.7–7.3
Kuipers (1983)	$\dot{V}O_{2max}$	7.6
Kuipers et al. (1985)	$\dot{V}O_{2max}$	7.9
Hickey et al. (1992)	Mean $\dot{V}O_2$ 40 mile TT	3.0
Hickey et al. (1992)	% $\dot{V}O_{2max}$ 40 mile TT	3.6
Lothian et al. (1993)	Progressive $\dot{V}O_2$	3.0–11.4
Jensen and Johansen (1998)	$\dot{V}O_{2max}$	1.9
Jensen and Johansen (1998)	$\dot{V}O_2$ at 2 mM blood lactate	7.4
Jensen and Johansen (1998)	$\dot{V}CO_2$ at 2 mM blood lactate	7.7

efficiency and/or reduced coordination could, via increased stimulation of motor units, increase $\dot{V}O_2$, stimulation of the cardiorespiratory system and lactate production. This is supported by Armstrong and Costill (1985), who reported that the day-to-day variability of $\dot{V}O_2$ (4.0%) and $\dot{V}_E$ (3.6%) is the result of technological error and biological fluctuation.

The results of the simulator test support the findings of the exercise tests. The $\dot{V}_E$ values produced by the Douglas bags, as explained earlier, were slightly lower compared to those produced by the simulator. The Oxycon Alpha produced very similar values for $\dot{V}_E$, $\dot{V}O_2$ and $\dot{V}CO_2$ when compared with the simulator. The Pulmolab produced highly deviating figures for $\dot{V}_E$ (underestimation), $\dot{V}O_2$ (underestimation) and $\dot{V}CO_2$ (overestimation), suggesting that the blocked capillary tube theory can account for only part of the problem. The Oxycon Pro and the simulator produced very similar values for $\dot{V}_E$. Surprisingly, however, the Oxycon Pro produced slightly higher values for $\dot{V}O_2$ and $\dot{V}CO_2$ compared with the simulator, and these values became increasingly deviant at higher breath frequencies.

In conclusion, both the Oxycon Pro and Oxycon Alpha were similar to our reference system (Douglas bags) with respect to generating valid absolute values for $\dot{V}O_2$, $\dot{V}CO_2$ and RER at workloads of 100 W and 150 W. The results of the simulation test suggest that this will also be the case with the Oxycon Alpha at higher workloads, although the validity of the Oxycon Pro at higher workloads needs further validation. Rietjens et al. (2001) attempted this by calculating the limits of agreements for Oxycon Pro and Douglas bag incremental exercise data. They concluded that the Oxycon Pro is a valid system for measurement of $\dot{V}_E$, $\dot{V}O_2$, and $\dot{V}CO_2$.

The Oxycon Pro, Oxycon Alpha and Douglas bags produced similar (or lower) day-to-day variations than those reported previously. The Pulmolab EX670, did not compare favourably with respect to validity or reliability with the other systems. This could be explained, at least in part, by the arrangement of the mouthpiece allowing saliva to interfere with the delay time.

Acknowledgements The authors would like to thank Prof. David Jones and Dr. Lee Romer for their valuable input in the preparation of this manuscript.

References

Armstrong LE, Costill DL (1985) Variability of respiration and metabolism: responses to submaximal cycling and running. Res Q Exerc Sport 56:93–96

Hausswirth C, Bigard AX, Le Chevalier JM (1997) The cosmed k4 telemetry system as an accurate device for oxygen uptake measurements during exercise. Int J Sports Med 18:449–453

Hickey MS, Costill DL, McConnell GK, Widrick JJ, Tanaka H (1992) Day to day variation in time trial cycling performance. Int J Sports Med 13:467–470

Jensen K, Johansen L (1998) Reproducibility and validity of physiological parameters measured in cyclists riding on racing bikes placed on a stationary magnetic brake. Scand J Med Sci Sports 8:1–6

Katch VL, Sady SS, Freedson P (1982) Biological variability in maximum aerobic power. Med Sci Sports Exerc 14:21–25

Kawakami Y, Nozaki D, Matsuo A, Fukunaga T (1992) Reliability of measurement of oxygen uptake by a portable telemetric system. Eur J Appl Physiol 65:409–414

Kuipers H (1983) Variability of physiological responses to exercise. PhD Thesis, University of Limburg, Maastricht, The Netherlands. Di Vriesborch, Haarlem

Kuipers H, Verstappen FTJ, Keizer HA, Geurten P, Van Kranenburg G (1985) Variability of aerobic performance in the laboratory and its physiologic correlates. Int J Sports Med 6:197–201

Lothian F, Farrally MR, Mahoney C (1993) Validity and reliability of the cosmed k2 system to measure oxygen uptake. Can J Appl Physiol 18:197–206

Lucia A, Fleck SJ, Gotshall RW, Kearney JT (1993) Validity and reliability of the cosmed K2 instrument. Int J Sports Med 14:380–386

McLaughlin JE, King GA, Howley ET, Bassett DR Jr, Ainsworth BE (2001) Validation of the COSMED K4 b2 Portable Metabolic System. Int J Sports Med 22:280–284

Peel C, Utsey C (1993) Oxygen consumption using the k2 telemetry system and a metabolic cart. Med Sci Sports Exerc 25:396–400

Rietjens GJWM, Kuipers H, Kester ADM, Keizer HA (2001) Validation of a computerized metabolic measurement system (Oxycon-Pro) during low and high intensity exercise. Int J Sports Med 22:291–294

Versteeg PGA, Kippersluis GJ (1989) Automated systems for measurement of oxygen uptake during exercise testing. Int J Sports Test 10:107–112

Wideman L, Stoudemire NM, Pass KA, McGinnes CL, Gaesser GA, Weltman A (1996) Assessment of the aerosport TEEM 100 portable metabolic measurement system. Med Sci Sports Exerc 28:509–515

Wright GR, Sidney K, Shephard RJ (1978) Variance of direct and indirect measurements of aerobic power. J Sports Med 18:33–42