Original Research

Evaluation of a New Automated Pulmonary Gas Analysis System

TOVE HALLENSTÅL*, IDA SUNDQVIST*, MICHAEL SVENSSON‡, and JI-GUO YU‡

Department of Surgical and Perioperative Sciences, Sport Medicine Unit, Umeå University, Umeå, SWEDEN

*Denotes undergraduate student author, ‡Denotes professional author

ABSTRACT

International Journal of Exercise Science 8(3): 287-296, 2015. The study aimed to evaluate the validity and reliability of a new automated pulmonary gas analysis system - Exercise Physiology System (EPS). The Oxycon Pro, a well-established automated pulmonary gas analysis system was used as a reference system. Six well-trained cyclists were recruited and performed two identical submaximal exercises on a cycle ergometer over one week interval. During the exercises, pulmonary gas exchange: ventilation (VE), oxygen consumption (VO_2), elimination of carbon dioxide (VCO_2), and ratio between carbon dioxide and oxygen (RER) was measured using both systems in randomized order. The exercise was composed of three incremental work-loads (Watt) of low, medium, and high load corresponding to 40%, 60%, and 80% of individual maximal work-load. Each work-load was performed twice so that pulmonary gas exchange was analyzed twice on each work-load using both systems. EPS validity was assessed by comparing the measurements from the two systems. For measurements presenting agreement between the systems, Bland-Altman plot analysis was performed to identify outlier. Reliability was evaluated by comparisons of measurements from repeated tests through each system, and by calculations of intra-class correlation coefficient (ICC) and standard error of measurement (SEM). Validity evaluations revealed that the majority of the measurements were in disagreements between the systems and most of them appeared at lower work-loads. The results indicated that at lower work-loads, EPS had lower validity than Oxycon Pro. At higher work-loads, ESP had high and comparable validity with Oxycon Pro. Reliability assessments revealed that agreements between repeated tests appeared in almost all measurements through both systems. Excellent agreements (very high ICC) in measurements between repeated tests were observed in VE, VO_2, and VCO_2 from both systems, and very small measurement errors (SEM), close to zero were observed in VO_2, VCO_2, and RER. The results suggested that the two systems had high and comparable reliability in measurement of pulmonary gas exchange.

KEY WORDS: Exercise physiology system, validity, reliability, respiratory parameters

INTRODUCTION

The measurement of pulmonary gas exchange is a common way in exercise physiology to evaluate cardiopulmonary and metabolic fitness in both elite athletes and recreational sport enthusiasts. By collecting exhaled gas in large impermeable bags and subsequent analysis of gas fractions and expired volumes, the Douglas bag method has been deemed to be the golden standard method in measuring gas

exchange. However, the Douglas bag technique has several disadvantages and its own sources of error (14). More importantly, the Douglas bag method could not provide breath-by-breath data and is time consuming due to the requirement of sampling and post-collection analysis (3, 1).

Over 20 different automated metabolic gas analysis systems have been introduced into market over the last 40 years (6, 9). These systems are mostly used for diagnosing hospital patients, especially those with cardiorespiratory disease (10). Moreover, nowadays the measurement of maximal oxygen uptake (VO_{2max}) in sports to evaluate athletes' physical status becomes a routine in fitness laboratories (13). The automated analysis system makes it possible to quickly obtain an abundance of online respiratory values, which is far more advanced than the traditional golden standard Douglas bag method (1, 9, 12). In addition, the automated systems must perform high validity and reliability in order to facilitate a comparison between different studies.

The Exercise Physiology System (EPS ADInstruments, New Zealand) is a new automated pulmonary gas analysis system. PowerLab is a vital part of the system for data acquisition and is compatible with many different instruments in sports and clinic research settings (11, 15). However, to our knowledge, the validity of the instrument has so far never been evaluated. Only one study has evaluated the reliability of PowerLab 8M by comparing respiratory parameters measured during repeated upper body exercise (13). Oxycon Pro (Jaeger, Wuerburg, Germany) is a relatively well-established automated metabolic

analysis system with a measuring accuracy comparable to "the golden standard" of Douglas bag method (3, 5, 9). In this present study, we aimed to evaluate the validity and reliability of the EPS system using the Oxycon Pro as a reference system.

METHODS

Participants
Six well-trained cyclists (5 men and 1 female) were recruited for the study. The anthropological information and VO_{2max} values of the subjects are shown in Table 1.The participants were informed about the design of the study and informed consents were obtained from all participants. The study was approved by the Ethics Committee of Umeå University (EPN Nr 2011–236-31M).

Table 1. Anthropological information and VO_{2max} values of the subjects.

Subjects	Sex	Age (year)	Height (cm)	Weight (kg)	VO2max (ml/min/kg)
1	Male	28	183	67,6	68,1
2	Female	51	170	59	58,4
3	Male	41	190	82	65,1
4	Male	30	181	74	77,2
5	Male	47	171	72,5	58,2
6	Male	49	183	74	58,9
M ± SD		39.6 ± 9.9	179.6 ± 7.7	71.5 ± 7.7	64.3 ± 7.5

Protocol
The EPS consists mainly of a gas analyzer, gas mixing chamber, spirometer, and PowerLab 16/35. The system is driven by LabChart Pro software which supports the Metabolic Module in analyzing metabolic parameters such as ventilation (VE), oxygen consumption (VO_2), elimination of carbon dioxide (VCO_2), and ratio between carbon dioxide and oxygen (RER).

An air pump (2 liters) was used to simulate low, medium, and high respiratory flow. A

total of twelve pumps, four on each simulation were conducted to calibrate ventilation flow. The default of indoor air distribution is 20.93% oxygen and 0.04% carbon dioxide. Calibrations were also performed for gas and environment temperature. The EPS does not use a temperature probe, but setting temperature for exhale at 36.6° C. Data for VE was calibrated to body temperature, ambient pressure and saturated with water. Data for VO_2, VCO_2, and RER were calibrated to standard temperature, pressure, and dry from water.

Individual VO_{2max} was evaluated using the Oxycon Pro (Hans Rudolph, USA) in order to define individual submaximal exercise intensity. The VO_{2max} test was performed on a cycle ergometer with the start work-load of 150 Watts for males and 125 Watts for female. The load was increased for 25 Watts per minute until exhaustion. During the test, the parameters of pulmonary gas exchange: VE, VO_2, VCO_2, and RER were measured, and heart rate (Polar, Finland) as well as maximal work-load were recorded for each individual.

All the subjects were asked to refrain from any strenuous physical exercise, and to have ordinary meal, daily activity, and sleeping time the day before the tests. The tests were conducted in two separate occasions (Test 1 and Test 2) with one week interval, and each occasion lasted for three days. All the tests were performed at approximately the same time of the days (AM 8:00 to 12:00).

Before the tests, a standard warm-up was performed followed by ten minutes rest. For the tests, the subjects were asked to perform a submaximal exercise on a cycle ergometer while their pulmonary gas exchanges were measured using both the Oxycon Pro and the EPS systems. The exercise is composed of three incremental work-loads (Watts) of low load, medium load, and high load corresponding to 40%, 60%, and 80% of individual maximal work-load. The first two loads lasted three minutes and the last load two minutes. The pulmonary gas exchange in the last 60 seconds at each load was analyzed for estimation of VE, VO_2, VCO_2, and RER. Each load was performed twice with one minute interval for gas analysis systems switch; thus, under the same load, pulmonary gas exchange was analyzed twice using either of the two systems. The order of applying the gas analysis systems in each subject was randomized. One week after the first submaximal exercise (Test 1), the exercise was repeated (Test 2) as well as all the measurements.

The validity of the EPS system was firstly evaluated through comparison of all measurements obtained from the two systems to examine systematic difference. If a measurement did not present significant difference between the two systems, the analysis of the error between the two systems was conducted using the mean bias and agreement limits presented in a Bland-Altman plots (2).

The reliability of the systems was assessed by comparing all the measurements between Test 1 and Test 2 from each system. Reliability was also analyzed using the intra-class correlation coefficient (ICC), classifying ICC values as low (ICC < 0.4), good (0.4 < ICC < 0.75), and excellent (ICC > 0.75) (4, 16). Standard error of
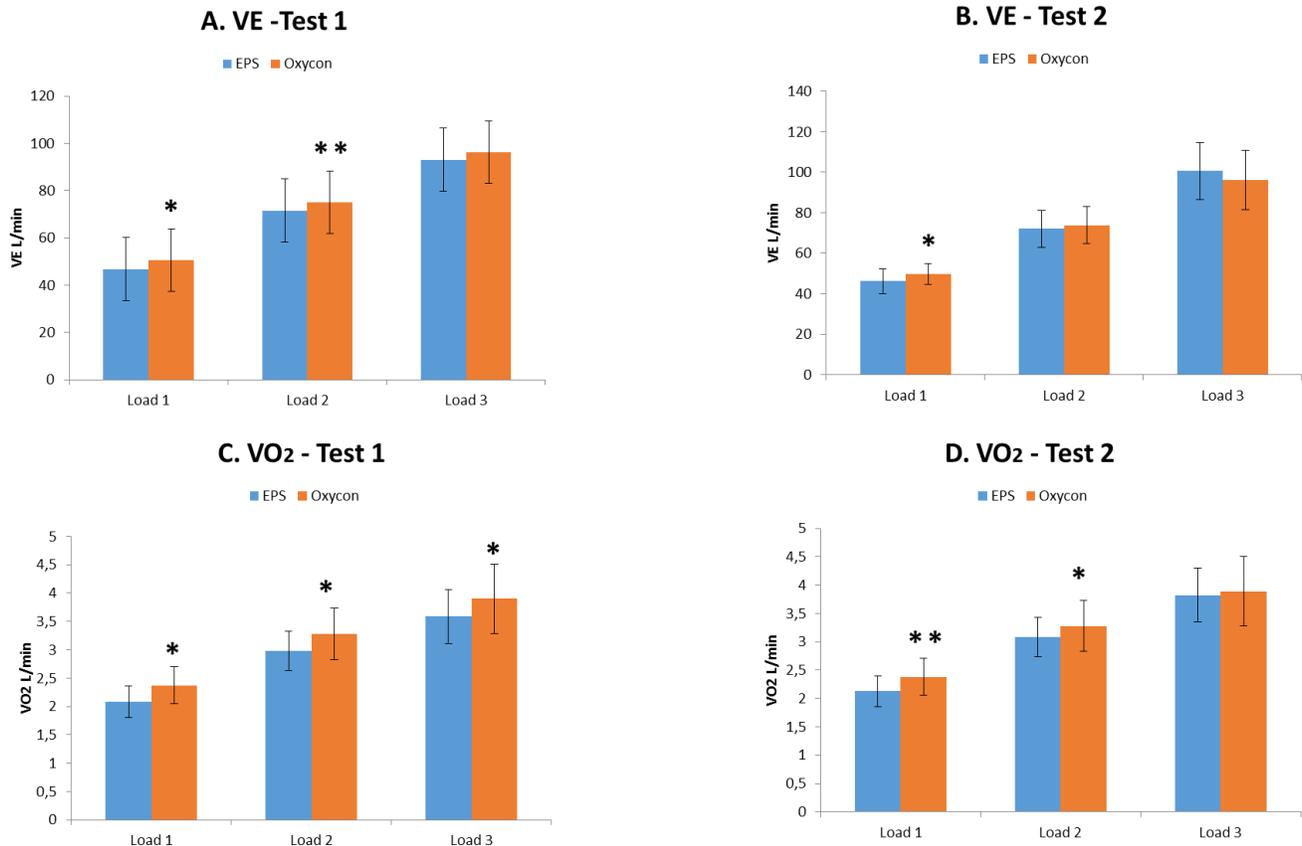
Figure 1. Comparisons of VE and VO_2 ofboth tests between the two systems. A, VE – Test 1; B, VE – Test 2; C, VO_2 – Test 1; D, VO_2 – Test 2. VE values of EPS were significantly lower than those of Oxycon Pro at load 1 and 2 in Test 1 (A), and at load 1 of Test 2 (B). VO_2 values of EPS were significantly lower at all the three loads in Test 1 than those of Oxycon Pro (C) and at load 1 and load 2 of Test 2 (D). * $p < 0.05$; ** $p < 0.001$.

measurement (SEM) were also calculated for reliability assessment.

Statistical Analysis
Statistical analysis was performed using SPSS (IBM SPSS statistics 22.0, CA, USA). Paired t-test was used for all comparisons, and one-way random of intra-class correlation with 95% confidence interval (CI) was applied for reliability assessment. Data were presented as mean ± standard deviation (M ± SD) and the level of significant difference was set at $P \leq 0.05$.

RESULTS

Comparisons of measurements in pulmonary gas exchange between EPS and Oxycon Pro were shown in Figure 1 and Figure 2. Most VE values obtained through EPS were lower than those through Oxycon Pro (Figure 1A, B), and significant differences were observed at load 1 ($p < 0.05$) and load 2 ($p < 0.001$) of Test 1 (Figure 1A), and at load 1 of Test 2 ($p < 0.05$; Figure 1B). Agreements between the two systems were observed in VE measurement at load 3 of Test 1, and at load 2 and load 3 of Test 2. Further analysis of the agreements using Bland-Altman Plot revealed that at load 3 of Test 1, one subject presented error value outside the 95% CI and had lower VE value
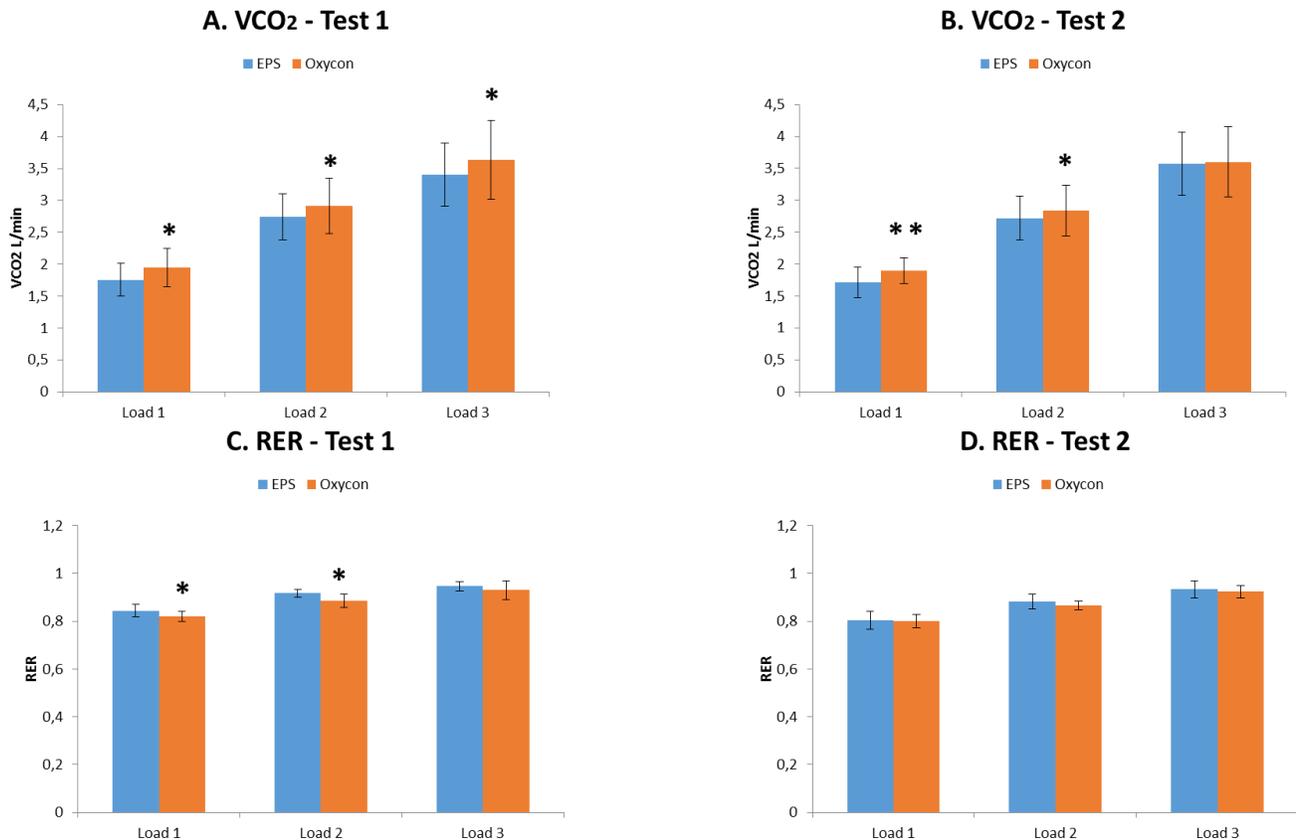
Figure 2. Comparisons of VCO_2 and RER of both tests between the two systems. A, VCO_2 – Test 1; B, VCO_2 – Test 2; C, RER – Test 1; D, RER – Test 2. All VCO_2 values of EPS were significantly lower than those of Oxycon Pro, except the load 3 of Test 2 (A, B). Significantly higher RER values of EPS than Oxocon Pro were observed at load 1 and load 2 of Test 1 (C). * $p < 0.05$; ** $p < 0.001$.

through EPS than through Oxyxon Pro. At load 2 or load 3 of Test 2, no outlier was identified.

All VO_2 values obtained through EPS were lower than those through Oxycon Pro, and significant differences were observed at all the three loads in Test 1 ($p < 0.05$; Figure 1C), and at load 1 and load 2 of Test 2 ($p < 0.001$ and $p < 0.05$, respectively; Figure 1D). Agreement in measurement of VO_2 between the two systems was only observed at load 3 of Test 2, and further analysis using Bland-Altman Plot did not reveal any outlier.

Paired t-tests of VCO_2 values of both Test 1 and Test 2 showed that the values obtained through EPS were significantly lower than those through Oxycon Pro at all the three loads of Test 1 ($p < 0.05$; Figure 2A), and at load 1 ($p < 0.001$) and load 2 ($p < 0.05$) of Test 2 (Figure 2B). Agreement in measurement of VCO_2 between the two systems was only observed at load 3 of Test 2, and further analysis using Bland-Altman Plot did not reveal any subject with error value outside the 95% CI.

RER values of EPS were significantly higher than those of Oxocon Pro at load 1 and load 2 ($p < 0.05$) of Test 1 (Figure 2C), but
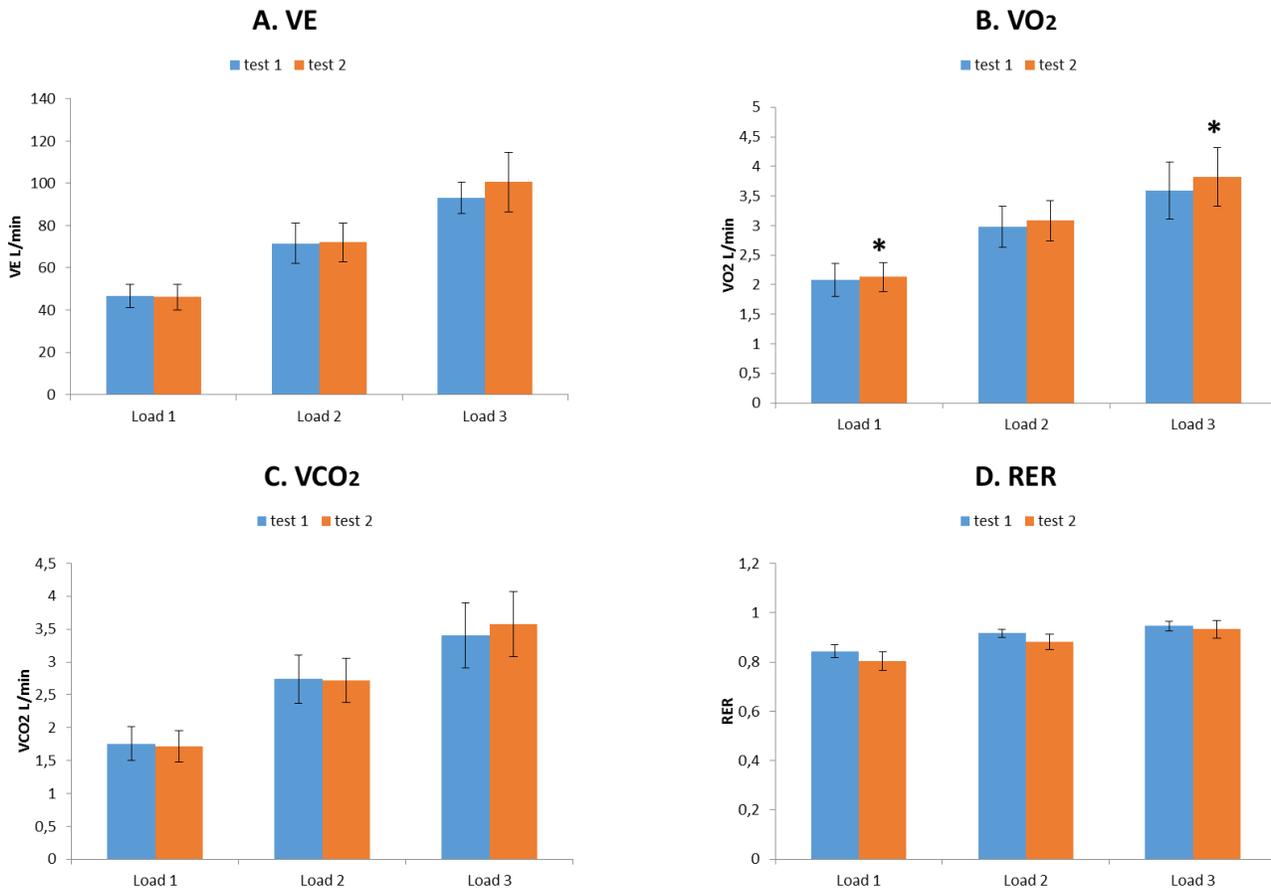
Figure 3. Comparisons of measurements in VE, VO_2, VCO_2 and RER between Test 1 and Test 2 obtained through EPS. A, VE; B, VO_2; C, VCO_2; D, RER. VO_2 values obtained at load 1 and load 3 of Test 1 were significantly lower than those at the corresponding loads of Test 2 (B). * $p < 0.05$.

comparable at all the three loads in Test 2 (Figure 2D). Agreements in measurements of RER between the two systems were observed at load 3 of Test 1, and at all the three loads of Test 2. Further analysis using Bland-Altman Plot did not reveal any outlier.

Comparisons of measurements between Test 1 and Test 2 from each system were shown in Figure 3 and Figure 4. For the EPS system, only measurements of VO_2 showed significantly lower values ($p < 0.05$) in Test 1 than Test 2 at load 1 and load 3 (Figure 3B). For Oxycon Pro system, no significant difference was observed in any parameter

between Test 1 and Test 2 at any load (Figure 4).

The ICC values of repeated measurements obtained from both systems were shown in Table 2. For both systems, excellent agreements between repeated tests were observed in measurements of VE (ICC = 0.79-0.99), VO_2 (ICC = 0.79-0.99), and VCO_2 (ICC = 0.93-0.98). However, the agreements in measurements of RER between repeated tests were low (0.07-0.35) through EPS, and low and good (ICC = 0.11-0.69) through Oxycon Pro. Visual inspection of SEM values revealed that VE presented in principle higher SEM values (0.39-3.67)
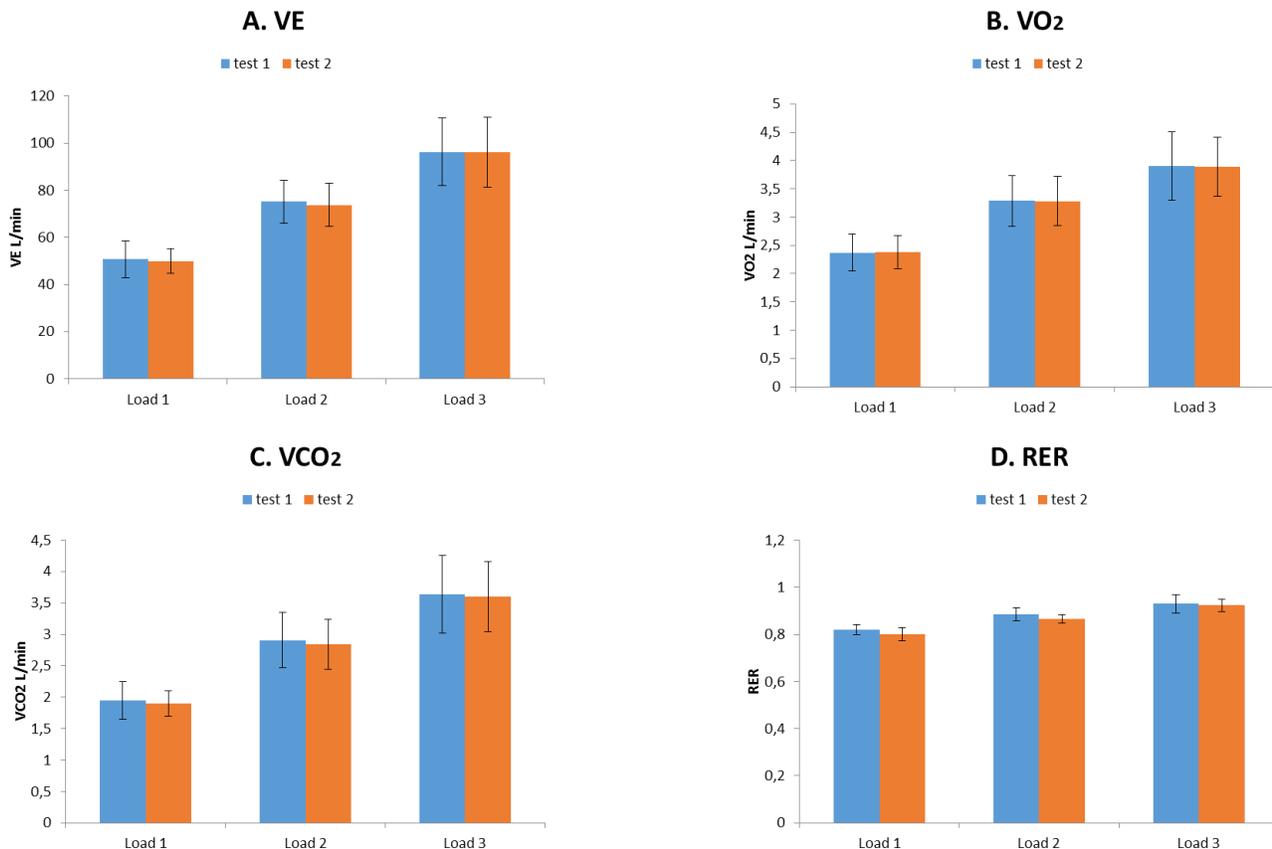
Figure 4. Comparisons of measurements in VE, VO_2, VCO_2 and RER obtained through Oxycon Pro between Test 1 and Test 2. A, VE; B, VO_2; C, VCO_2; D, RER. No significant difference was observed in any parameter between the two tests at any load.

than those of the other three parameters, which had SEM values close to zero (0.01-0.99; Table 2).

DISCUSSION

This present study evaluated the validity and reliability of a new automated pulmonary gas analysis system, the EPS system. The results showed that the EPS system had high and comparable reliability to the Oxycon Pro system. At low work-load, the EPS system had relatively lower validity than the Oxycon Pro system. At high work-load, the two systems had high and comparable validities.

The primary parameters, VE, VO_2, and VCO_2 presented lower values in measurements through EPS than through Oxycon Pro, indicating that the EPS system underestimated the respiratory measurements in comparison with those of Oxycon Pro. Comparisons of RER values between the two systems revealed that while most RER values of EPS were comparable with those of Oxycon Pro, two RER values of EPS were significantly higher than those of Oxycon Pro. As RER represents the ratio between carbon dioxide and oxygen (RER = VCO_2/VO_2), therefore higher/lower RER value only indicates higher/lower dioxide production (VCO_2) per unit of oxygen consumption (VO_2).

Table 2. Values of ICC (95% CI) and SEM of all parameters from both systems.

| | | EPS | | Oxycon | |
		ICC (p)	SEM	ICC (p)	SEM
VE	Load 1	0.99 (0)	0.39	0.84 (0.043)	2.12
	Load 2	0.94 (0.006)	1.95	0.96 (0.002)	1.51
	Load 3	0.79 (0.023)	3.44	0.91 (0.014)	3.67
VO_2	Load 1	0.99 (0)	0.02	0.98 (0)	0.04
	Load 2	0.95 (0.001)	0.06	0.98 (0)	0.05
	Load 3	0.93 (0)	0.04	0.99 (0)	0.06
VCO_2	Load 1	0.98 (0)	0.03	0.93 (0.007)	0.06
	Load 2	0.94 (0.006)	0.08	0.98 (0)	0.04
	Load 3	0.93 (0.003)	0.09	0.98 (0.001)	0.08
RER	Load 1	0.35 (0.23)	0.02	0.11 (0.44)	0.01
	Load 2	0.07 (0.45)	0.01	0.68 (0.052)	0.01
	Load 3	0.18 (0.42)	0.02	0.69 (0.132)	0.01

It is worth to notice that seven measurements were in agreement between the two systems, focusing mainly on higher work-load. Interestingly, further analyzing of the agreements using Bland-Altman Plot revealed that only one subject with error value outside the 95% CI. Taken together, the results suggested that the EPS had lower validity than Oxycon Pro at low work-load, but comparable validity to Oxycon Pro when work-load was high.

Oxycon Pro demonstrated to be a valid system in pulmonary gas analysis (3, 12, 5). However, Oxycon Pro has also been shown to present a slight tendency of overestimation for VO_2 and VCO_2 in breath-by-breath measurement. When breathing frequencies were increased, these values became more deviant (3). In contrast, other researchers have observed that Oxycon Pro slightly underestimated VO_2 and VCO_2 at high exercise intensities compared to that of Douglas Bags (12, 5). This present study showed that the EPS had lower validity than Oxycon Pro at low exercise intensities, but high and comparable validity to Oxycon Pro at high exercise intensities. It has been postulated that the differences in pulmonary metabolic parameters measured through different automated analysis systems are due to hardware or software, or both (6). Minor difference such as water vapor in analyzers may cause miscalculations in metabolic variables, leading to significant differences in results. Such concerns may also exist in this study, and need to be clarified in future study.

In previous study, validity evaluation of an instrument has been performed through

comparison of two instruments applied in two separate tests (9). However, this will result in not only technical error but also biological within-subject variation (9). To avoid the problem, in this present study, we evaluated the validity of EPS through consecutively applying both the EPS and the Oxycon Pro systems in a test. We used three submaximal progressive loads and each load was repeated so that pulmonary gas exchange could be measured at each load twice using both systems. The design of the study allowed sequential collection and analysis of the expired gases (7).

In this present study, the reliabilities of both systems were evaluated through comparisons of the measurements obtained from all the subjects in repeated tests through each individual system. While only two repeated measurements of VO_2 present disagreement through the EPS system, no measurement was observed to be in disagreement between repeated tests through the Oxycon Pro system.

Further analysis of reliability using intra-class correlation coefficient showed high ICC values in measurements of VE, VO_2, and VCO_2 through both systems. The results indicated that both systems could provide highly consistent results in estimation of VE, VO_2, and VCO_2. The ICC values of RER were relatively low, especially for EPS. However, RER of both systems presented very small measurement errors (close to zero of SEM values), at approximately the same levels of VO_2 and VCO_2. Taken all the reliability assessments into consideration, we concluded that the EPS system had high reliability in pulmonary gas exchange estimation, comparable to the Oxycon Pro system.

REFERENCES
1. Bassett DR, Howley ET, Thompson DL, King GA, Strath SJ, McLaughlin JE, Parr BB. Validity of inspiratory and expiratory methods of measuring gas exchange with a computerized system. J Appl Physiol 91: 218-224, 2001.

2. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327: 307-310, 1986.

3. Carter J, Jeukendrup AE. Validity and reliability of three commercially available breath-by-breath respiratory systems. Eur J Appl Physiol 86: 435-441, 2002.

4. Fleiss JL. The Design and Analysis of Clinical Experiments. New York, NY, USA: John Wiley, 1986.

5. Foss O, Hallen J. Validity and stability of a computerized metabolic system with mixing chamber. Int J Sports Med 26: 569-575, 2005.

6. Hodges LD, Brodie DA, Bromley PD. Validity and reliability of selected commercially available metabolic analyzer systems. Scand J Med Sci Sports 15: 271-279, 2005.

7. Kannagi T BR, Hossack KF. An evaluation of the Beckman Metabolic Cart for measuring ventilation and aerobic requirements during exercise. J Card Rehab 3: 38-53, 1983.

8. Macfarlane D. Automated Metabolic Gas Analysis Systems. Sports Med 31: 841-861, 2001.

9. Macfarlane DJ, Wong P. Validity, reliability and stability of the portable Cortex Metamax 3B gas analysis system. Eur J Appl Physiol 112: 2539-2547, 2012.

10. Makita K, Nunn JF, Royston B. Evaluation of metabolic measuring instruments for use in critically ill patients. Crit Care Med 18: 638-644, 1990.

11. Radespiel-Troger M, Rauh R, Mahlke C, Gottschalk T, Muck-Weymann M. Agreement of two different methods for measurement of heart rate variability. Clin Auton Res 13: 99-102, 2003.

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 22: 291-294, 2001.

13. Sealeya RM, Leichta AS, Spinksa W, Sinclair W. Reliability of two metabolic systems during sport-specific upper-body ergometry. Eur J Sport Sci 10(5): 305-309, 2010.

14. Shephard RJ. A critical examination of the Douglas bag technique. J Physiol 127: 515-524, 1955.

15. South AR, Somers SM, Jog MS. Gum chewing improves swallow frequency and latency in Parkinson patients: a preliminary study. Neurology 74: 1198-1202, 2010.

16. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res 19: 231-240, 2005.