©Journal of Sports Science and Medicine (2008) **7,** 15-22
http://www.jssm.org

Research article

Lack of agreement between gas exchange variables measured by two metabolic systems

DjordjeG. Jakovljevic [1]✉, David Nunan [1], Gay Donovan [1], Lynette D. Hodges[2], Gavin R.H. Sandercock [3] and David A. Brodie [1]

[1] Research Centre for Society and Health, Buckinghamshire New University, Buckinghamshire, UK
[2] Department of Sport and Exercise Science, University of Bedfordshire, Bedfordshire, UK
[3] Centre for Sports and Exercise Science, University of Essex, Colchester, UK

Abstract

The purpose of this study was to assess the agreement and consistency between gas exchange variables measured by two online metabolic systems during an incremental exercise test. After obtaining local ethics approval and informed consent, 15 healthy subjects performed an incremental exercise test to volitional fatigue using the Bruce protocol. The Innocor (Innovision, Denmark) and CardiO$_2$ (Medical Graphics, USA) systems were placed in series, with the Innocor mouthpiece attached to the pneumotach of the CardiO$_2$. Metabolic data were analysed during the last 30 seconds of each stage and at peak exercise. There were non-significant differences ($p > 0.05$) between the two systems in estimation of oxygen consumption (VO_2) and in minute ventilation (V_E). Mean Cronbach's alpha for VO_2 and V_E were 0.88 and 0.92. The Bland-Altman analysis revealed that limits of agreement were -0.52 to 0.55 l·min^{-1} for VO_2, and -8.74 to 10.66 l·min^{-1} for V_E. Carbon dioxide production (VCO_2) and consequently respiratory exchange ratio (RER) measured by the Innocor were significantly lower ($p < 0.05$) through all stages. The CardiO$_2$ measured fraction of expired carbon dioxide ($FeCO_2$) significantly higher ($p < 0.05$). The limits of agreement for VO_2 and V_E are wide and unacceptable in cardio-pulmonary exercise testing. The Innocor reported VCO_2 systematically lower. Therefore the Innocor and CardiO$_2$ metabolic systems cannot be used interchangeably without affecting the diagnosis of an individual patient. Results from the present study support previous suggestion that considerable care is needed when comparing metabolic data obtained from different automated metabolic systems.

Key words: Metabolic system, oxygen consumption, minute ventilation, carbon dioxide production, Bruce protocol.

Introduction

Exercise testing provides an integrative approach to different aspects of cardiovascular health (Wasserman et al., 1999). Physiologists have been interested in monitoring gas exchange variables such as oxygen consumption and carbon dioxide production over the last two centuries (Macfarlane, 2001). Technological advances today have contributed to the development of portable rapid-response breath-by-breath metabolic systems. These systems have become standard tools for diagnosing cardiorespiratory performance, not only in healthy and sporting populations, but also in those with different cardiovascular and ventilatory pathophysiological abnormalities (Hodges et al., 2005). Many automated metabolic measurement systems today present 'black boxes' which can generate many data without providing the user with sufficient information to evaluate exactly how the data were generated (Macfarlane, 2001). Considerable care is therefore needed when comparing metabolic data obtained from different automated metabolic systems (Hodges et al., 2005). There is general concern regarding the limited knowledge available about the accuracy of a number of commercially available systems (Carter and Jeukendrup, 2002).

Different suggestions may be found in the literature about accuracy and acceptable levels of agreement between different metabolic systems. It has been suggested that minute ventilation should be accurate within ± 5% (Gore, 2000) or within ±50 ml·min^{-1} (Thoden, 1991), while reported acceptable limits of agreement are -0.8 to 1.2 l·min^{-1} (Bassett et al., 2001). The accuracy of oxygen consumption measurements should have a technical error of measurement of less than 3% (Gore, 2000). For oxygen consumption it has been suggested that acceptable limits of agreement should be -0.08 to 0.11 l·min^{-1} (Bassett et al., 2001). For the purposes of the present study the limits of agreement of ± 1.2 l·min^{-1} for minute ventilation and ± 0.15 l·min^{-1} for oxygen consumption will be considered as acceptable.

The purpose of this study was to assess the agreement and consistency between gas exchange variables measured by recently introduced online metabolic measuring system (Innocor, Innovision, Denmark) and an alternative online metabolic system (CardiO$_2$, Medical Graphics, USA) during a standard incremental exercise test.

Methods

Subjects

Fifteen healthy adults (10 males and five females), staff and students of a South East UK University gave their signed informed consent to participate. The research was carried out in accordance with the Declaration of Helsinki and was approved by the Faculty Ethics Committee. All subjects were familiar with graded exercise testing. Subjects were asked to refrain from eating for a minimum of 2 h prior to the test and from vigorous exercise 24 hours prior to testing.

Equipment

Received: 10 September 2007 / Accepted: 30 October 2007 / Published (online): 01 March 2008

The Innocor (Innovision, Odense, Denmark) is a compact device intended to be used for non-invasive measurement of cardiac output using an inert gas (N_2O) rebreathing methodology. It also incorporates a breath-by-breath module to measure gas exchange variables (e.g. fractions of expired O_2 and CO_2 (FeO_2 and $FeCO_2$), minute ventilation (V_E)) and to calculate a number of derived variables (e.g. oxygen consumption (VO_2), carbon dioxide production (VCO_2), respiratory exchange ratio (RER)). The gas sample line and airflow umbilical are connected to the respiratory valve unit. Gas data analysis is performed and results presented by the Innocor software (version 5.05). Monitoring and presentation of the data is via the Innocor integrated computer with a Pentium MMX and Windows NT/XP embedded operating system. Measurement of airflow is performed by a pressure difference pneumotach. Carbon dioxide analysis is performed by using a photoacoustic gas analyser. Oxygen is analysed using an oxygen sensor (Oxigraf Inc., USA) based on the principle of laser diode absorption spectroscopy. Only the oxygen sensor needs 1-point calibration on a regular basis by the user while both oxygen sensor and photoacoustic gas analyser require multi-point calibration performed by manufacturer periodically (6-12 months). The manufacturer reported accuracy for measurements of ventilation ±1%, while for O_2 and CO_2 concentrations ±0.01%.

The $CardiO_2$ (Medical Graphics Corp., St Paul, MN, USA) is an automated breath-by-breath respiratory gas analysis system. The $CardiO_2$ measures the same variables as the Innocor. However, it provides the advantage of a single, light-weight, sample line and pressure transducer umbilical. Gas analysis is performed and results presented by the BreezeSuite (version 5.0) gas exchange testing software (Medical Graphics Corporation, St. Paul, Minnesota, USA). Monitoring and presentation of the data are via a personal computer (Dell Precision 340 Pentium 4, Dell Computer Corporation, Texas, USA). Gas analysis is performed by means of Zirconia electrochemical (O_2) and infrared (CO_2) analysers. Airflow is measured by means of a 'Prevent' pitot tube flowmeter. The manufacturer of the $CardiO_2$ reported accuracy for measurement of ventilation ±3% or 50 ml, while for O_2 and CO_2 concentrations ±0.03%.

The Innocor and $CardiO_2$ both have a gas drying sample circuit which ensures that only dry air comes into the analysers.

Procedure

Accuracy of the Innocor and $CardiO_2$ systems for measurements of ventilation, O_2 and CO_2 concentrations was checked by an engineer before the study was conducted. Certified gas samples of 14-16% O_2 and 3-5% CO_2 were introduced into the analysers in order to check their accuracy. Minute ventilation was checked using a certified volume syringe (MedGraphics, St Paul, Minn) of 2.5 litres at a rate of 50 strokes per minute.

In order to obtain simultaneous measurements the two systems were placed in series and subsequently calibrated according to the manufacturers' recommendations before the exercise test. The gas mixtures used for the $CardiO_2$ calibration were 5% CO_2, 12% O_2 in balanced N_2 (calibration gas) and 21% O_2 in balanced N_2 (references gas). The Innocor system only requires gas delay determination (specific breathing pattern performed by user) and O_2 adjustment to the ambient air prior to the test. The volume calibration was performed following the two systems were placed in series. The Innocor respiratory valve unit, with bacterial filter, was attached into the pneumotach of the $CardiO_2$ system using a five cm flexible tube. Both systems were attached to the subject's face mask, where the pneumotach of the $CardiO_2$ system was closer to the subject's mouth and the Innocor respiratory valve unit was 10 cm from the subject's mouth (see Figure 1). Such configuration increased dead space and the distance between the subject's mouth and the Innocor gas sensors.

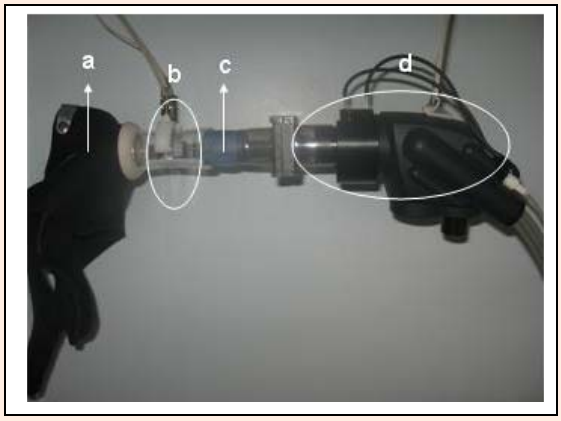


Figure 1. Graphical description of the installation of the two systems: a) face mask, b) $CardiO_2$ pneumotach with flow sensor and gas sample line, c) flexible tube, d) Innocor respiratory valve unit with bacterial filter, flow sensor and gas sample line.

Before exercise testing, weight was measured using Secca scale and height on a wall-mounted stadiometer. All subjects were instructed on the use of the Borg 6-20 Rating of Perceived Exertion (RPE) scale (Borg, 1982) to ascertain their perception of effort during each test.

Subjects performed an incremental exercise test to volitional fatigue on a motor driven treadmill (Cardio Control, Delft, Netherlands) using the protocol described by Bruce and colleagues (1973). Breath-by-breath metabolic data were analysed during the last 30 seconds of each stage and at peak exercise.

Data analysis

All statistical analysis was performed using SPSS version 13.0 (SPSS Inc., Chicago, IL, USA). A paired samples t-test and nonparametric Wilcoxon signed rank test were used. Statistical significance was indicated if $p < 0.05$. Bland-Altman plots (Bland and Altman, 1986) were constructed to assess agreement between gas exchange variables measured by the Innocor and $CardiO_2$ systems. Bland-Altman plots included in total 72 data points (from the end of each exercise stage and from peak exercise). Cronbach's alpha (Cronbach, 1951) was also calculated to demonstrate consistency of measures between the two systems at the end of each stage and from peak exercise. Values are expressed as means ± SD unless otherwise indicated.

Table 1. Gas exchange variables measured by the Innocor and CardiO$_2$ at peak exercise.

Gas exchange variables	Innocor mean (SD)	CardiO$_2$ mean (SD)	*p* value	Difference %	Cronbach's Alpha
VO$_2$ (l·min^{-1})	2.88 (1.02)	2.83 (.80)	.97	.15	.89
VCO$_2$ (l·min^{-1})	2.95 (1.14)	3.19 (1.21)	.19	7.52	.91
V$_E$ (l·min^{-1})	79.62 (28.72)	81.31 (32.76)	.55	2.08	.97
FeO$_2$ (%)	17.29 (0.54)	17.31 (.52)	.87	.12	.78
FeCO$_2$ (%)	3.69 (0.59)	3.97 (.39)	.11	7.05	.35
RER	1.02 (0.13)	1.13 (.14) *	.04	8.15	.53

Abbreviations: VO$_2$ - oxygen uptake, VCO$_2$ - carbon dioxide production, V$_E$ - minute ventilation, FeO$_2$ - fraction of expired oxygen, FeCO$_2$ - fraction of expired carbon dioxide, RER – respiratory exchange ratio. * significantly different ($p < 0.05$).

Results

Subjects were aged 34 ± 11.5 years; stature, 1.73 ± 0.11 m; weight, 71 ± 12.7 kg. All 15 subjects completed the first three stages of the Bruce protocol, while 12 subjects completed four. Table 1 presents the peak values of gas exchange variables measured by the Innocor and the CardiO$_2$ metabolic systems.

Stage by stage analysis revealed a range of differences in gas exchange variables. Oxygen consumption was not significantly different ($p > 0.05$) between the two systems ranging from 2% (stage one) to 6% (stage three). Cronbach's alpha ranged from 0.80 (stage two) to 0.99 (stage one), with a mean measurement of 0.88. The mean difference in VO$_2$ was 0.02 l·min^{-1} and the limits of agreement were -0.52 and 0.55 l·min^{-1} respectively (Figure 2).

Carbon dioxide production reported by the Innocor was significantly lower than the CardiO$_2$ throughout all exercise stages ($p < 0.05$, Figure 3). The differences for VCO$_2$ ranged from 9% (stage two) to 22% (stage four). Cronbach's alpha ranged from 0.84 (stage three) to 0.94 (stage four). The mean difference in VCO$_2$ was -0.23 l·min^{-1} and the limits of agreement were -1.01 and 0.56 l·min^{-1} (Figure 4).

Minute ventilation measurements recorded by the Innocor were higher compared with CardiO$_2$ throughout all exercise stages except at peak exercise, ranging from 2% (stage one) to 7% (stage three). These differences were non significant ($p>0.05$). Cronbach's alpha for measured V$_E$ ranged from 0.84 (stage one) to 0.98 (stage four), with a mean measurement of 0.92. Bland-Altman analysis which included data from the end of each exercise stage and from peak exercise showed that the Innocor reported higher V$_E$ values by a mean value of 0.96 l·min^{-1} compared with the CardiO$_2$ and the limits of agreement of -8.74 and 10.66 l·min^{-1} (Figure 5).

Respiratory exchange ratio estimated by the Innocor was significantly lower throughout all exercise stages ($p < 0.05$, Figure 6). The difference for RER ranged from

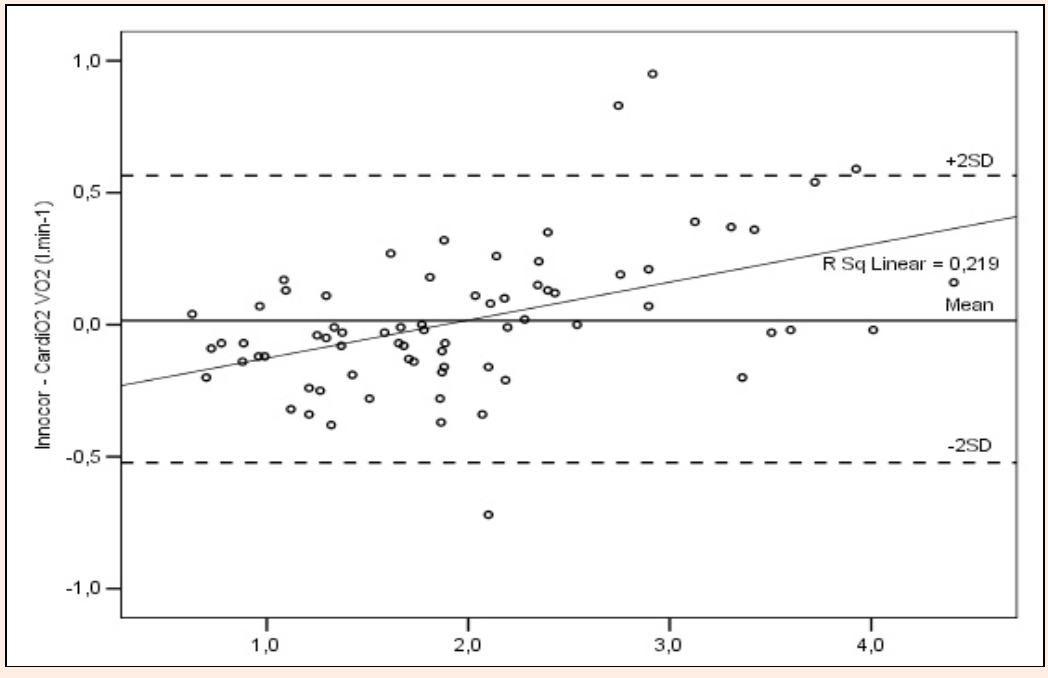
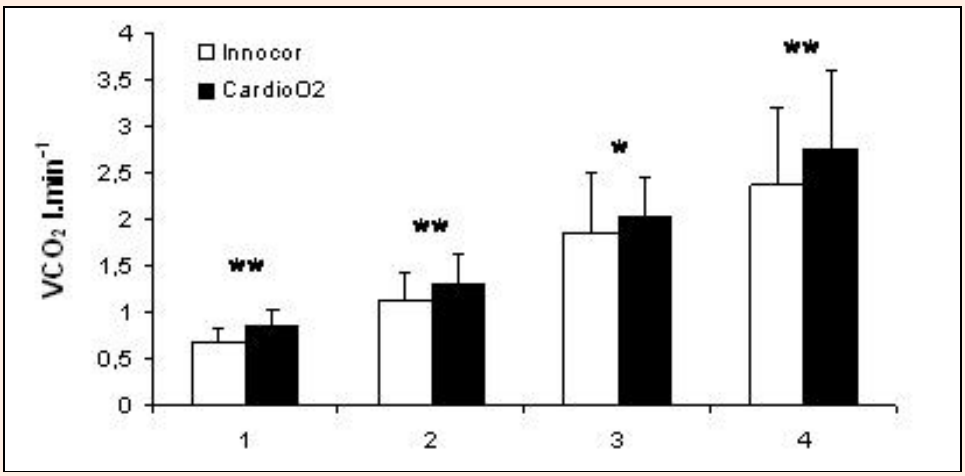


Figure 2. Bland-Altman plot of individual oxygen uptake (VO$_2$, l·min^{-1}) differences between the Innocor and the CardiO$_2$ metabolic systems. The solid line is at the bias (mean of the difference) and the dashed lines are at ±95% limits of agreement (SD of the differences multiplied by 2).



Figures 3. Mean differences in estimated carbon dioxide production (VCO_2, $l \cdot min^{-1}$) between the Innocor and the $CardiO_2$ through four stages of Bruce protocol. * p < 0.05, ** p < 0.01.

8% to 15% whilst Cronbach's alpha ranged from 0.22 to 0.61. The mean difference in RER was -0.11 and the limits of agreement were -0.18 and 0.02.

The mean differences in measured FeO_2 between the two systems ranged from 2% (stage one) to 7% (stage three), and were all non-significant ($p > 0.05$). Cronbach's alpha ranged from 0.64 to 0.90, dependent on the stage. The mean difference in measured FeO_2 was -0.19% and the limits of agreement were -1.77 and 1.38%.

$FeCO_2$ measured by the Innocor was significantly lower compared with the $CardiO_2$ ($p < 0.01$, $p < 0.05$) (Figure 7). The differences in measured $FeCO_2$ ranged from 12% (stage two) to 20% (stage four), respectively. Cronbach's alpha for $FeCO_2$ ranged from 0.20 to 0.77, dependent on the stage. The mean difference in measured $FeCO_2$ was -0.59% and the limits of agreement were -1.87 and 0.70%.

Discussion

The purpose of present study was to assess the agreement and consistency between gas exchange variables measured by the Innocor and the $CardiO_2$ metabolic systems during an incremental exercise test.

The Innocor yielded mean VO_2 and V_E that were not significantly different from those obtained by the $CardiO_2$ system.

The percentage VO_2 differences in the present study compare favourably with findings of Porszasz et al. (1994) who reported the difference of less than 5.9%

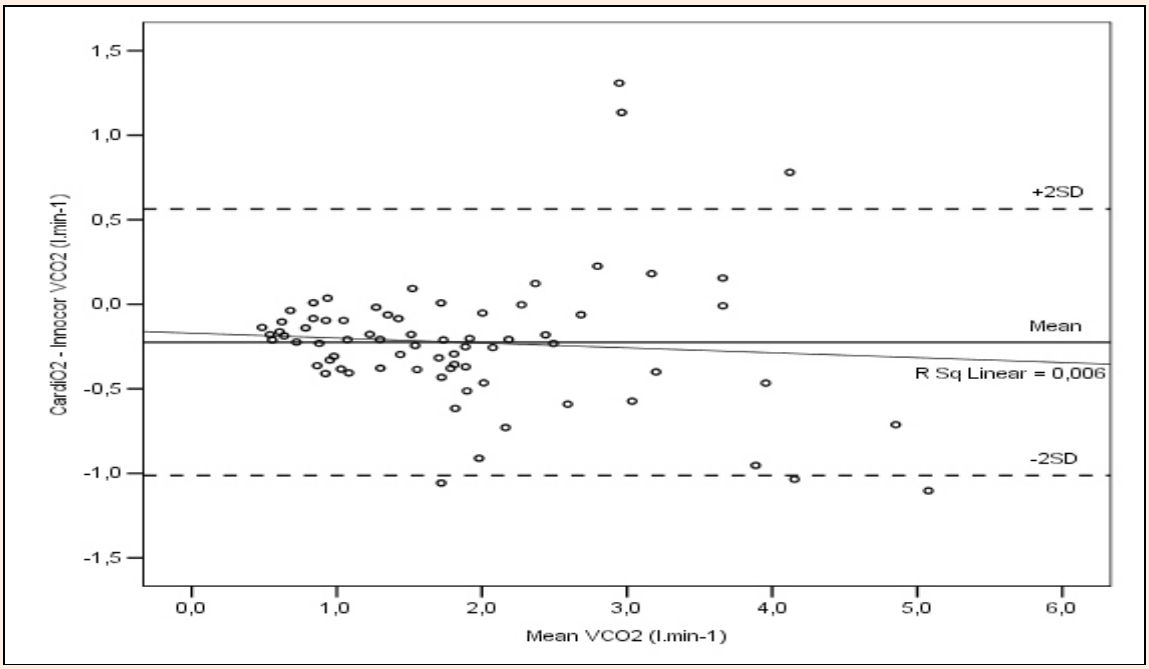


Figure 4. Bland-Altman plot of individual carbon dioxide production (VCO_2, $l \cdot min^{-1}$) differences between the Innocor and the $CardiO_2$ metabolic systems. The solid line is at the bias and the dashed lines are at ±95% limits of agreement.

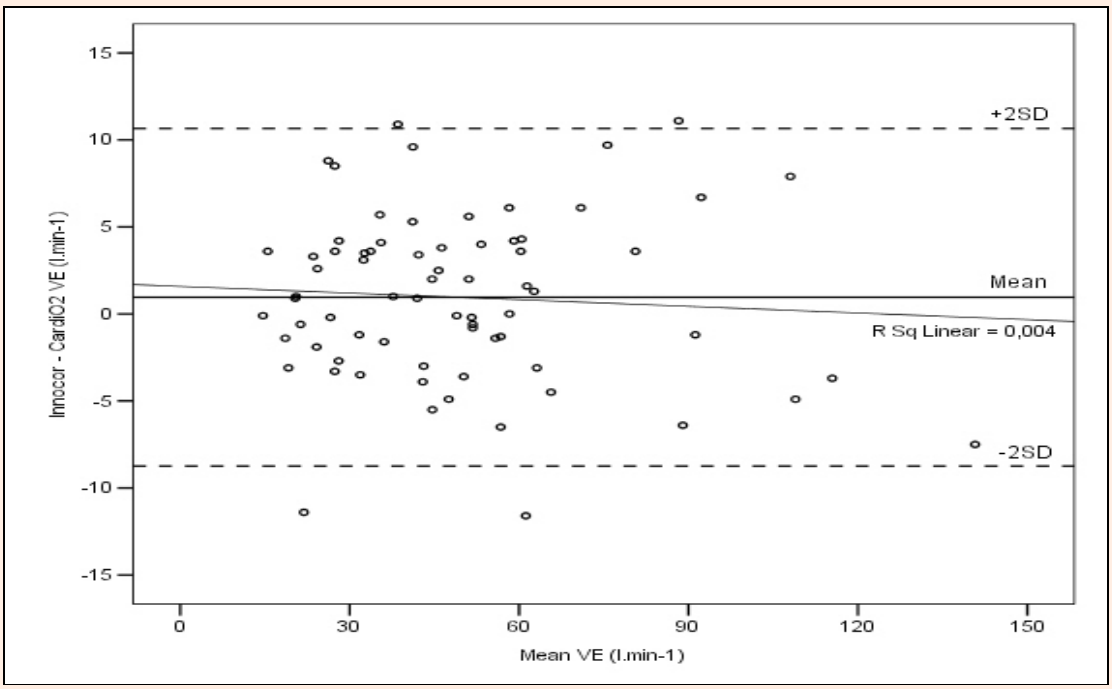


Figure 5. **Bland-Altman plot of individual minute ventilation (V_E, $l \cdot min^{-1}$) differences between the Innocor and the $CardiO_2$ metabolic systems. The solid line is at the bias and the dashed lines are at ±95% limits of agreement.**

between Medical Graphics system and the Douglas bag method. Engebretson (1998) found that the Medical Graphics system produced values of VO_2 that were within 3.6% compared with the Douglas bag. Non-significant differences in mean VO_2 in the present study were not surprising, although there were non-significant differences in measured FeO_2 between the Innocor and $CardiO_2$ systems.

In the present study, the limits of agreement reported for VO_2 of -0.52 to 0.55 $l \cdot min^{-1}$ are, however, wide and unacceptable in cardio-pulmonary exercise testing. Bassett et al. (2001) compared the ParvoMedics computerized system and Douglas bag method over a range of exercise intensities and reported that the limits of agreement of -0.08 to 0.11 $l \cdot min^{-1}$ for VO_2 are acceptable. McLaughlin et al. (2001) compared the Cosmed portable metabolic system with the Douglas bag and reported wider limits of agreement (-0.33 to 0.15 $l \cdot min^{-1}$). Surprisingly, the authors concluded that the portable metabolic system is acceptable for measuring oxygen consumption. This is in spite of these limits of agreement being ~10% of the reported mean peak VO_2 value of ~3.5 $l \cdot min^{-1}$.

The Innocor measured V_E slightly higher than the $CardiO_2$ through all exercise stages except at peak exercise. Miles et al. (1994) showed that the Medical Graphics automated system produced the lowest V_E measurement among four different metabolic systems. On the other hand Engebretson (1998) showed no significant

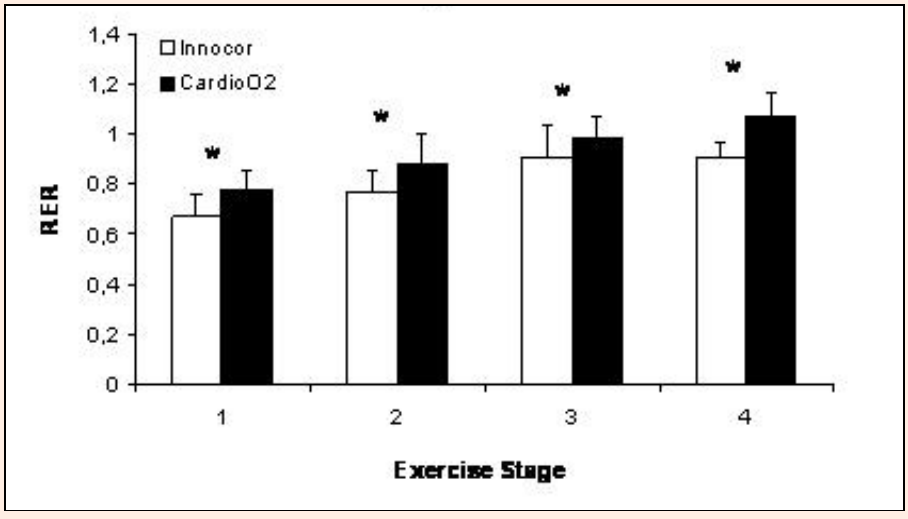


Figure 6. **Mean differences in estimated respiratory exchange ratio (RER) between the Innocor and the $CardiO_2$ through four stages of Bruce protocol. * $p < 0.05$.**

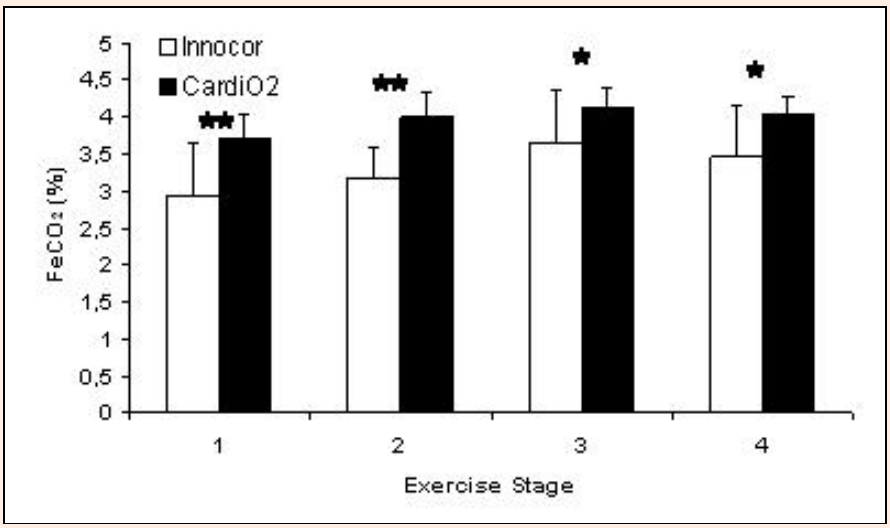


Figure 7. Mean differences in measured fraction of expired carbon dioxide ($FeCO_2$, %) between the Innocor and the $CardiO_2$ through four stages of Bruce protocol. * p < 0.05, ** p < 0.01.

differences between the Medical Graphics and the Douglas bag method in measured V_E, while La Mere et al. (1993) illustrated that the Medical Graphics system overestimated V_E by 3.1 $l \cdot min^{-1}$ compared with the Douglas bag.

Despite non-significant differences in measured V_E, it should be noted that the Bland-Altman analysis indicated that the limits of agreement for V_E are wide (-8.7 to 10.7 $l \cdot min^{-1}$). This is in contrast with finding of Bassett et al. (2001) who reported limits of agreement for V_E of -0.8 to 1.2 $l \cdot min^{-1}$, even at higher maximum V_E values than those reported in the present study (~100 $l \cdot min^{-1}$ vs. ~80 $l \cdot min^{-1}$). Individual differences in V_E never exceeded 1.6 $l \cdot min^{-1}$ (Bassett et al., 2001). On the other hand, McLaughlin et al. (2001) reported wider limits of agreement for V_E that were similar to those in the present study (~ -6 to 10 $l \cdot min^{-1}$). Also maximum V_E was ~80 $l \cdot min^{-1}$. However, we believe that limits of agreement for V_E reported in our study are wide and not acceptable in cardio-pulmonary exercise testing.

As suggested, the Cronbach's alpha coefficient should be used to indicate a degree of consistency between measurements (Cronbach, 1951). Bland and Altman (1997) reported that the Cronbach's alpha should be a minimum of 0.90, and 0.95 would be desirable for clinical application. The results of the present study demonstrate that high consistency exists in measured VO_2 and V_E between the two systems. However, the results of the present study are an obvious example that reporting only the Cronbach's alpha without calculating the limits of agreement may lead researchers to make wrong assumptions and draw inappropriate conclusions.

Both systems were calibrated and checked for their accuracy for measurements of V_E, VO_2 and VCO_2 by an engineer before study was conducted. The report demonstrated that both systems met manufacturers' recommendations regarding the accuracy. From manufacturers' specifications it seems that the Innocor has capability to measure V_E more accurate than the $CardiO_2$ (±1% vs. ±3%). In contrast, the $CardiO_2$ has capacity to measure O_2 and CO_2 concentrations with accuracy of $\pm$0.03% compared with ±0.01% by the Innocor. However the results from present study clearly indicate that differences in measured gas exchange variables between the two systems are higher than those suggested.

When Beaver et al. (1973) compared metabolic measurements between an on-line breath-by-breath computerised system and a standard method, they suggested that differences in measured V_E and VO_2 may be due to temporal alignment of a gas flow or analyser dynamic response. Further potential source of error in V_E and VO_2 may be the method and equation used by the Innocor and the $CardiO_2$ to estimate the BTPS factor (Hodges et al., 2005). When a subject exhales during a cardiopulmonary exercise test, the air leaves the lungs and enters the spirometer at 33-35°C (Cole, 1954). Most volume type spirometers assume instantaneous cooling of the air as it enters the spirometer, although errors can occur due to incorrect assumptions of instantaneous cooling of the air (Hodges et al., 2005). Depending on the environmental temperature, the BTPS correction factor could be as large as 10% (Crapo, 1994). As stated earlier, in order to obtain direct comparison, the two systems were placed in series. The distance of the subject to the metabolic analyser together with a bacterial filter through which air passed before reached the Innocor gas sensors may potentially affect the cooling of exhaled air. These possible differences in physical characteristics of the analysed air (e.g. temperature) may affect differences in V_E and VO_2 by the two systems. However, it is important to note that reversing the order of the systems, and subsequent measurement, was not possible due to equipment design.

Carbon dioxide production and consequently RER values reported by the Innocor were significantly lower than those of the $CardiO_2$. The limits of agreement for VCO_2 (-1.01 and 0.56 $l \cdot min^{-1}$) were wider than those previously reported as acceptable, (e.g. -0.08 to 0.08 $l \cdot min^{-1}$, as reported by Bassett et al., 2001). These differences in VCO_2 were due to lower measurements of $FeCO_2$ by the Innocor. Although the Innocor measured V_E slightly

higher during all exercise stages, this was not enough to compensate for the significantly lower $FeCO_2$ when calculating VCO_2. RER values reported at peak exercise by the $CardiO_2$ appear to be more valid than those reported by the Innocor. This indicates that the Innocor underestimated $FeCO_2$ and VCO_2 compared with the $CardiO_2$.

Miodownik et al. (2000), when comparing a newly developed semi-automated metabolic system based on a Douglas bag design and the Medical Graphics system, reported a non significant difference of 1.5% in VCO_2. Engebretson (1998) reported significantly lower VCO_2 by the Medical Graphics system compared to the Douglas bag method. By contrast, few studies have reported non significant differences in measured VCO_2 between the Medical Graphics systems and the Douglas bag method (Porszasz et al., 1994; Prieur et al., 1998), while Miles et al. (1994) reported that the Medical Graphics system measured a higher VCO_2 compared with three other metabolic systems.

It is accepted that an increased dead space and the distance between the Innocor sensors and the subject's mouth, may have contributed to the poor consistency and agreement in measured $FeCO_2$ and VCO_2. Analysing RER results at peak exercise it seems that this direct comparison study design may possibly affect more the Innocor CO_2 sensors than those in the $CardiO_2$. An additional problem is that of correcting water vapour pressure in the expired air, as this pressure may be quite different to that in the calibration gas (Davies et al., 1974). Although the gas analysers were adjusted automatically to ignore the contribution of water vapour (effectively measuring dry air), most CO_2 analysers are sensitive to the presence of water vapour (Macfarlane, 2001). As the dead space was increased in the present study, this could potentially enhance higher water condensation, and possibly effect CO_2 analysers and measurement of $FeCO_2$. Therefore the possible inability of infrared sensors to cope with the water vapour could have contributed to the discrepancies in measured $FeCO_2$. The possible measurement of the temperature of the sample near the flow detector followed by calculation and correction according to the absolute water vapour pressure could possible identify the source of error. However, the measurement of the temperature was not possible due to direct comparison study design and specific configuration of the Innocor respiratory valve unit where the sensors are located.

Lower $FeCO_2$ measured by the Innocor may be due to lower response time of the analyser. This particularly may be emphasized with higher breathing frequencies at higher intensity of exercise. Supporting this assumption Figure 3 demonstrates that the difference in VCO_2 was higher at the end of stage four than at the end of previous stages of the Bruce protocol. There is evidence to suggest the use of different algorithms for correcting response time (Ariely and Van Liew, 1981; Farmery and Hahn, 2000). Farmery and Hahn, using specific correcting methods, were able to reduce response times for measured gases almost fivefold. Therefore future investigations should evaluate the use of suggested algorithms for correcting the response time not only in the most commonly used metabolic analysers but also in those which have recently appeared.

Finally, the differences in measured CO_2 between the two systems may be explained by the technological factors. The Innocor uses a newly developed portable multigas analyser which uses the principle of photoacoustic spectroscopy with an infrared spectrum, while the $CardiO_2$ uses a standard non-dispersive infrared CO_2 sensor.

Conclusion

The results of this study revealed that the Innocor produces similar mean values of VO_2 and V_E to those obtained from the $CardiO_2$ system. However, the limits of agreement for VO_2 and V_E were wide and unacceptable in cardio-pulmonary exercise testing. Also, the data from the present study suggest that systematic bias exists in measured $FeCO_2$ and VCO_2. Therefore the Innocor and $CardiO_2$ metabolic systems cannot be used interchangeably without affecting the diagnosis of an individual patient. Results from the present study support previous suggestion that considerable care is needed when comparing metabolic data obtained from different automated metabolic systems.

Acknowledgments

The authors have no financial or business interests in any of the companies mentioned in this article.

References

Ariely R. and Van Liew, H. D. (1981) Correction for the response time and delay of mass spectrometers. *Journal of Applied Physiology* **51,** 1417-1422.

Bassett, D.R., Howley, E.T., Thompson, D.L., King, G.A., Strath, S.J., McLaughlin, J.E. and Parr, B.B. (2001) Validity of inspiratory and expiratory methods of measuring gas exchange with computerized system. *Journal of Applied Physiology* **91,** 218-224.

Beaver, W.L., Wasserman, K. and Whipp, B.J. (1973) On-line computer analysis and breath-by-breath graphical display of exercise function tests. *Journal of Applied Physiology* **34,** 128-132.

Bland, M.J. and Altman, D.G. (1986) Statistical methods for assessing agreement between two methods of clinical measurements. *Lancet* **1,** 307-310.

Bland, M.J. and Altman, D.G. (1997) Statistics notes: Cronbach's alpha. *British Medical Journal* **314,** 572-573.

Borg, G. (1982) Psychophysical bases of perceived exertion. *Medicine and Science in Sports Exercise* **14,** 377-381.

Bruce, R.A., Kusumi, F. and Hosmer, D. (1973) Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *American Heart Journal* **85,** 546-542.

Carter, J. and Jeukendrup A.E. (2002) Validity and reliability of three commercially available breath-by-breath respiratory systems. *European Journal Applied Physiolology* **86,** 435-441.

Cole, P. (1954) Recording of respiratory air temperature. *Journal of Laryngology* **68,** 295-307.

Crapo, R.O. (1994) Standardisaton of spirometry. Medical section of the lung Association. *American Journal of Respiratory Care* **152,** 1107-1136.

Cronbach, L.J. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika* **16,** 297-334.

Davies, E.E., Hahn, H.L., Spiro, S.G. and Edwards, R.H. (1974) New technique for recording respiratory transients at the start of exercise. *Respiratory Physiology* **20,** 69-79.

Farmery, A.D. and Hahn, C.E.W. (2000) Response-time enhancement of a clinical gas analyzer facilitates measurement of breath-by-breath gas exchange. *Journal of Applied Physiology* **89,** 581-598.

Engebretson, J.E. (1998) Validity of a breath by breath gas exchange analysis system. *Medicine and Science in Sports and Exercise* **30,** S330.

Gore, C.J. (2000) Quality assurance in exercise physiology laboratories. In: *Physiological Testing for Elite Athletes.* Ed. Gore, C.J. Champaign (IL): Human Kinetics. 3-11.

Hodges, L.D., Brodie, D.A. and Bromley, P.D. (2005) Validity and reliability of selected commercially available metabolic analyzer systems. *Scandinavian Journal of Medicine and Science in Sport* **15,** 271-279.

La Mere, V.J., Brown, K., Wigglesworth, J.K. and Edwards, J.E. (1993) Reproducibility between three metabolic systems and validation by Douglas bag method. *Medicine and Science in Sports and Exercise* **25,** S9.

Macfarlane, D.J. (2001) Automated metabolic gas analysis systems. *Sports Medicine* **31,** 841-861.

McLaughlin, J.E., King, G.A., Howley, E.T., Bassett, D.R. and Ainsworh, B.E. (2001) Validation of the Cosmed K4 b^2 portable metabolic system. *International Journal of Sports Medicine* **22,** 280-284.

Miles, D.S., Cox, M.H. and Verde T.J. (1994) Four commonly utilized metabolic systems fail to produce similar results during submaximal and maximal exercise. *Sports Medicine Training and Rehabilitation* **5,** 189-198.

Miodownik, S., Carlon, V.A., Ferri, E., Burda, B. and Melendez, J.A. (2000) System of automated gas-exchange analysis for the investigation of metabolic processes. *Journal of Applied Physiology* **89,** 373-378.

Porszasz, J., Barstow, J. and Wasserman, K. (1994) Evaluation of a symmetrically disposed Pitot tube flowmeter for measuring gas flow during exercise. *Journal of Applied Physiology* **77,** 2659-2665.

Prieur, F., Busso, T., Castells, J., Bonnefoy, R., Benoit, H., Geyssant, A. and Denis, C. (1998) Validity of oxygen uptake measurements during exercise under moderate hyperoxia. *Medicine and Science in Sports and Exercise* **36,** 958-962.

Thoden, J.S. (1991) Testing aerobic power. In: *Physiological Testing of the High Performance Athlete.* Ed: MacDougall, J.D., Wenger, H.A., Green, H.I. Champaign (IL): Human Kinetics. 107-173.

Wasserman, K., Hansen, J.E., Sue, D.Y., Casaburi, R. and Whipp, B.J. (1999) *Principles of Exercise Testing and Interpretation.* Williams and Wilkins, Baltimore.

Key points

- There is general concern regarding the limited knowledge available about the accuracy of a number of commercially available systems.
- Demonstrated limits of agreement between key gas exchange variables (oxygen consumption and minute ventilation) as measured by the two metabolic systems were wide and unacceptable in cardiopulmonary exercise testing.
- Considerable care is needed when comparing metabolic data obtained from different automated metabolic systems.

AUTHORS BIOGRAPHY

Djordje G. JAKOVLJEVIC
Employment
PhD student, Research Centre for Society and Health, Buckinghamshire New University, United Kingdom
Degrees
MSc, BSc
Research interests
Exercise physiology in health and disease (heart failure) particularly focusing on overall cardiac function as represented by cardiac power output.
E-mail: djordje.jakovljevic@bcuc.ac.uk

avid NUNAN
mployment
hD student, Research Centre for Society nd Health, Buckinghamshire New Uniersity, United Kingdom
egrees
ISc, BSc
Research interests
Cardiovascular and autonomic health in heart failure patients; physical activity and recovery; exercise physiology.
E-mail: dnunan01@bcuc.ac.uk

Gay DONOVAN
Employment
PhD student, Research Centre for Society and Health, Buckinghamshire New University, United Kingdom
Degrees
MSc, BSc
Research interests
Exercise physiology in health and disease (heart failure), particularly focusing on heart rate and metabolic recovery following exercise.
E-mail: gcarpu01@bcuc.ac.uk

Lynette D. HODGES
Employment
Senior Lab Manager, Department of Sport and Exercise Science, University of Bedfordshire, United Kingdom
Degree
PhD
Research interests
Erectile dysfunction; exercise benefits for people with cardiovascular disease, especially peripheral vascular disease.
E-mail: lynette.sigleton@beds.ac.uk

Gavin R.H. SANDERCOCK
Employment
Lecturer in Clinical Exercise Physiology, Centre for Sports and Exercise Science, University of Essex, United Kingdom
Degree
PhD
Research interests
Cardiac autonomic function; cadiovascular disease prevention and rehabilitation.
E-mail: gavins@essex.ac.uk

David A. BRODIE
Employment
Professor of Cardiovascular Health and Head of Research, Buckinghamshire New University, United Kingdom
Degree
PhD
Research interests
Cardiovascular health as applied to healthy and unhealthy populations (eg heart failure patients); cardiac rehabilitation; erectile dysfunction.
E-mail: david.brodie@bcuc.ac.uk

✉ **Djordje Jakovljevic,**
Research Centre for Society and Health, Buckinghamshire New University, Gorelands Lane, Chalfont St Giles, Buckinghamshire HP8 4AD, UK.