European Journal of Clinical Nutrition (1998) 52, 536–540

http://www.stockton-press.co.uk/ejcn

Precision and accuracy in a metabolic monitor for indirect calorimetry

JCK Wells[1,2] and NJ Fuller[1]

[1]*Dunn Clinical Nutrition Centre, Hills Road, Cambridge CB2 2DH; and* [2]*Childhood Nutrition Research Centre, Institute of Child Health, 30 Guildford Street, London WC1N 1EH*

Objective: To determine within-machine and between-machine precision (reproducibility) and accuracy, of the Deltatrac Mk 1 Metabolic Monitor.
Design: Within-machine and between-machine comparison for gas exchange (VO_2 and VCO_2), respiratory quotient (RQ) and energy expenditure (EE).
Subjects: 3 Deltatrac Mk 1 Metabolic Monitors.
Methods: Within-machine and between-machine reproducibility were assessed by five successive 10 min gas infusion tests in each machine. Accuracy was assessed by measuring independently the gas content of the infusion mixture. The Deltatrac flowmeters were evaluated by further infusion tests.
Results: Within-study reproducibility was < 1 ml/min for VCO_2, < 2.5 ml/min for VO_2, < 5 kcal/d for EE and < 0.01 for RQ. Between-study reproducibility was $< 0.2\%$ for RQ, $< 1\%$ for VCO_2 and $< 2\%$ for VO_2 and EE. Between-machine reproducibility was $< 0.1\%$ for RQ and $< 1\%$ for VO_2, VCO_2 and EE. Accuracy in all three machines was within 3% for VO_2, VCO_2 and EE, and within 0.2% for RQ.
Conclusions: The Deltatrac Mk 1 is a very precise metabolic monitor, and is accurate within 3% for gas exchange and EE. RQ is measured with greatest reproducibility and accuracy (within 0.2%), making the monitor particularly suitable for studies of substrate utilisation.
Descriptors: oxygen consumption; carbon dioxide production; energy expenditure; respiratory quotient

Introduction

Measurement of minimal metabolism in the form of basal (BMR), resting (RMR) or sleeping (SMR) metabolic rate is used increasingly in both research and clinical care. Although minimal metabolism can be predicted from anthropometry, predictive equations are inevitably subject to a degree of error, and are unsuitable for many patient groups and disease states. Consequently, there is a strong case for measuring minimal metabolism wherever possible (Kaplan *et al*, 1995; Journal of Pediatrics Editorial, 1995).

Several commercially-produced machines are now available for such measurements. Furthermore, these machines are able to measure separately oxygen (O_2) consumption, carbon dioxide (CO_2) production, and by derivation substrate utilisation, which are useful measurements in their own right. However, although commercial machines have been used for up to ten years, the success with which they measure respiratory gas exchange has not received adequate attention. We have therefore assessed precision (reproducibility) and accuracy, both within-machine and between-machine, in a metabolic monitor that is widely used in research and hospital environments.

Correspondence: Dr JCK Wells, Childhood Nutrition Research Centre, Institute of Child Health, 30 Guildford Street, London WC1N 1EH.
Received 10 February 1998; revised 10 April 1998; accepted 17 April 1998

Methods

Three Deltatrac Mk 1 Metabolic Monitors (Datex, Helsinki) were set up in close proximity to each other, in order to minimise environmental differences or ambient fluctuations. Carbon dioxide and nitrogen (N_2) were infused simultaneously to simulate physiological rates of CO_2 production and O_2 consumption respectively. N_2 alone was infused to simulate physiological rates of O_2 consumption in order to assess machine flow rate. For each series of tests the monitors were warmed up for over an hour before the infusions were undertaken. All machines were operated in canopy mode with the averaging and artefact suppression facilities set to 'off'.

Each machine received a gas mixture of nominal composition 20% CO_2 and balance N_2 (independent assessment of these values is described below) or N_2 alone. The infusion rate of either gas, set approximately to a rate of 1 l per minute, was measured accurately using an oil-filled flow meter (Model DM3A Alexander Wright, London) calibrated to 0.25%. The sites of the patient hood and the infusion line were not changed during the experiment. The same corrugated hose attached at one end to the canopy was connected at the other end to the flow generators of the different machines as required. Measurements of ambient temperature and pressure were recorded for each infusion. During each infusion, observed values of VCO_2 and VO_2 were allowed to reach a steady state plateau before the evaluation was carried out over a minimum ten minute period. The order in which the machines were tested was randomised.

Measured volumes of the gases infused, obtained from the Wright flowmeter, were corrected to standard tempera-

ture and pressure (STP) and expressed as volume per minute (V_{STP}). Simulated VO_2 and VCO_2 gas exchange rate was ascertained as follows:

$$VO_2 = \frac{F_t \times F_{oi} \times f_{nt}}{1 \quad f_{oi} \quad f_{ci}}$$

$$VCO_2 = \frac{F_t \times [f_{ct} \quad f_{ct} \times f_{oi} \quad f_{ci}]}{1 \quad f_{oi} \quad f_{ci}}$$

where F_t = flow rate of infusion gas
f_{ct} = proportion of CO_2 in infusion gas
f_{nt} = proportion of N_2 in infusion gas
f_{oi} = proportion of O_2 in ingoing air
f_{ci} = proportion of CO_2 in ingoing air

Simulated respiratory quotient (RQ) was calculated as VCO_2/VO_2. These equations apply to suction ventilated systems (Murgatroyd *et al*, 1987).

For the first part of this study, each machine was given five infusion tests over a period of 7 h, and was pressure and gas calibrated prior to every measurement. These tests were to assess accuracy and precision within and between machines. A second set of three infusions per machine was undertaken over a second 7 h period, with calibration not being repeated after the first time. The aim of these tests was to determine if accuracy of the monitor decreased if calibrations were not made immediately prior to the measurement. A third set of two tests per machine involved infusion of N_2 alone, to assess accuracy of the Deltatrac flowmeter, with calibrations performed before every infusion.

The composition of the infusion gas mixture was established as follows. The gas mixture, nominally 20% CO_2/80% N_2 (BOC, Guildford, Surrey, UK), was diluted 25 times (that is, to 4% of original CO_2 concentration) with CO_2-free air (BOC) using a Wösthoff pump (ED Gilbert & Co Ltd, Southampton, UK). This mixture was then analysed using an infra-red CO_2 analyser (ADC, Analytical Development Company Ltd, Hoddesdon, UK), calibrated with N_2 for zero and 0.8% CO_2 for span, in turn checked against a 0.75% (7500 ppm) alpha gas (BOC). The accuracy of the alpha gas was itself confirmed by a chemical (hyamine hydroxide) absorption and titration method (Fuller & Elia, 1989). Measured values for the main infusion mixture were 19.75% CO_2 and 80.25% N_2.

Evaluation of machine accuracy was dependent on the accuracy with which the composition of the infusion mixture was measured, and the rate at which the infused gases were delivered. We have therefore modelled the effect of error, in measurement of the composition of the gas infusion mixture, on the calculated gas infusions which form our main results, as described in the paragraph below. We have also evaluated the flowmeter of each Deltatrac separately, to determine whether variation in precision and accuracy of gas measurements can be attributed to the flowmeters.

Assuming the gas was analysed with a -1% error, that is the true composition was 19.95% CO_2 rather than 19.75%, VO_2 would have been over-estimated by 0.3% and VCO_2 under-estimated by 1.1%. Assuming the gas was analysed with a $+1\%$ error, namely the true composition was 19.55% CO_2, VO_2 would have been under-estimated by 0.2%, and VO_2 over-estimated by 1.0%. If the infusion mixture was assumed to have the manufacturer's specified content of 20% CO_2, V O_2 would have been over-estimated by 0.3% and VCO_2 under-estimated by 1.3%.

The effect of flowmeter accuracy is less complex. If the rate at which the gas mixture was infused was in error, then both VO_2 and VCO_2 would have been in error by the same relative amount, for example -1% flow rate error results in only 99% recovery of VCO_2 and VO_2. Finally, error of 1°C in temperature, or 1 mm in barometric pressure, would cause 0.4% or 0.1% errors in infusion rate respectively.

Statistics

Energy expenditure (EE) was calculated minute-by-minute using Weir's equation (Weir, 1949), and RQ as VCO_2/VO_2. Precision and accuracy were assessed for each of VO_2, VCO_2, RQ and EE. To evaluate within-machine within-study reproducibility, the standard deviation of each set of 10 one-minute measurements was calculated. The average reproducibility, from the five such 10 min tests undertaken for each machine, was then further calculated. These results were left in raw unit as their significance for measurement precision depends on the volumes of gas being infused.

To evaluate within-machine between-study reproducibility, and between-machine reproducibility, values of VCO_2 and VO_2 were expressed as percentage recovery of the infusion mixture, to adjust for variation in infusion volumes. Standard deviation values were expressed as an average of the mean. Measured RQ and EE were expressed as percentages of the expected value. Accuracy of the 3 machines was likewise assessed by percentage recovery of infused gases.

Results

Within-machine within-study reproducibility is given in Table 1. Values were consistent between machines, being < 1.0 ml/min for VCO_2, < 2.5 ml/min for VO_2, < 5 kcal/d for EE and < 0.01 for RQ. For a typical adult with a BMR of 1700 kcal/d, these results equate to precision of $< 0.5\%$ for VCO_2 and EE, and $< 1\%$ for VO_2 and RQ.

Within-machine between-study reproducibility is given in Table 2. Values were $< 0.2\%$ for RQ, $< 1\%$ for VCO_2

Table 1 Within-machine within-study precision (values are s.d.)

Deltatrac	*VO_2 (ml/min)*	*VCO_2 (ml/min)*	*RQ*	*EE (kcal/d)*
A	2.04	0.86	< 0.01	4.48
B	1.94	0.92	< 0.01	4.36
C	2.10	0.94	< 0.01	4.72

Values are mean of five precision values, each calculated as s.d. of 10 consecutive measurements.

Table 2 Accuracy and within-machine between-study precision

Deltatrac		*VO_2*	*VCO_2*	*RQ*	*EE*
A	error (%)	−1.0	−1.7	0.0	−1.1
	precision (%)	1.5	0.4	0.1	1.2
B	error (%)	−1.8	−2.4	−0.1	−1.9
	precision (%)	1.1	0.8	0.1	0.1
C	error (%)	−2.8	−1.7	0.0	−2.6
	precision (%)	1.9	0.6	0.1	1.5

Precision values for each machine are s.d. of five tests, expressed as a percentage of the mean percentage recovery of VO_2 and VCO_2, or s.d. of five tests expressed as a percentage of the observed mean for RQ and EE. Accuracy is given as error, that is 100 × [(observed − expected)/expected].

and < 2% for VO_2 and EE. Error was < 3% for VO_2 and EE, < 2% for VCO_2 and < 0.2% for RQ. Between-machine reproducibility values, given in Table 3 were < 0.1% for RQ, < 0.5 for VCO_2 and < 1% for VO_2 and EE. Average error values are also given in this table. Accuracy of the flowmeters is given in Table 4, with error being under-recovery of the infusion mixture by < 3% in all three machines.

The effect of measurement delay following a single calibration affected RQ but not EE. Recovery of CO_2 declined with increasing time after calibration, while recovery of oxygen was relatively unaffected. These patterns resulted in a decline on RQ in all 3 machines (Figure 1), but had a negligible and inconsistent effect on energy expenditure (Figure 2).

Discussion

Indirect calorimetry can serve a number of different functions, through its ability to measure O_2 consumption, CO_2 production, substrate utilisation and energy expenditure. These measurements can be carried out on the entire human age range from pre-term infants (Roberts *et al*, 1987) to the elderly (Fuller *et al*, 1996), and can be obtained during almost all states of disease.

Its main clinical use is in measuring energy expenditure, because of the limited success with which minimal metabolism can be predicted from anthropometry. Predictive equations are subject to error arising from several factors. Firstly, anthropometry can only provide an approximation for fat free mass, which is the best predictor of minimal metabolism (Ravussin & Bogardus, 1989). Secondly, predictive equations cannot take into account between-subject variation due to genetic factors, either familial (Bogardus *et al*, 1986) or ethnic (Case *et al*, 1997). Thirdly, predictive equations are frequently inappropriate for sub-populations whose minimal metabolism is altered by illness or abnormal body composition (Murphy *et al*, 1995; Dickerson *et al*, 1995; Schebendach *et al*, 1995).

However, indirect calorimetry can also serve a number of other functions. Separate measures of VCO_2 and VO_2 play an important role in ^{13}C breath tests (see Amarri *et al*, 1997) and determination of aerobic fitness (see Gutin *et al*, 1994) respectively. Measurement of RQ can be used to investigate substrate utilisation (see Sonko *et al*, 1994). Given these multiple uses, it is essential to understand the limitations of the technique. For example when designing studies, data on precision are required for estimating the ability to detect changes between and within subjects.

In this study, reproducibility was found to be high, whether within studies, between studies or between machines. Accuracy was found to be within 3% for gas exchange, likewise for EE, with recovery of both gases being marginally under-estimated. Given that this pattern was found in all three machines, it is important to consider the possible sources of such consistent error. Potential sources include error in the evaluation of the infusion mixture, error in the Deltatrac flowmeters and error in the Wright flowmeter.

Error in evaluation of the infusion gas mixture (CO_2/N_2) is unlikely to be responsible for the consistent error in gas recovery. The modelling described in the methods section indicates that the effects on VCO_2 and VO_2 would be in opposing directions, whereas our results show incomplete

Table 3 Between-machine accuracy and precision

	Error (%)	*Precision (%)*
VO_2	−1.9	0.9
VCO_2	−1.9	0.4
RQ	0.0	< 0.1
EE	−1.9	0.8

Accuracy is average error of the three machines. Precision is reproducibility between the three machines, calculated as s.d. expressed as percentage of mean percentage recovery for VO_2 and VCO_2, or s.d. as percentage of observed mean for RQ.

Table 4 Accuracy of flowmeters using nitrogen infusion

Deltatrac	*Factory setting (l/min)*	*Observed error (%)*
A	41.2	−2.1
B	39.0	−2.7
C	40.5	−1.9

Accuracy for each machine is given as error, that is 100 × [(observed − expected)/expected].

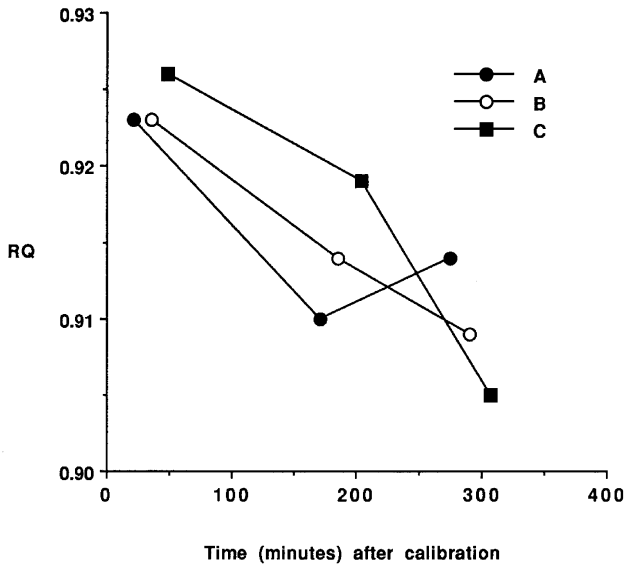

Figure 1 Change in respiratory quotient (RQ) (percentage of expected value) with time following a single calibration.

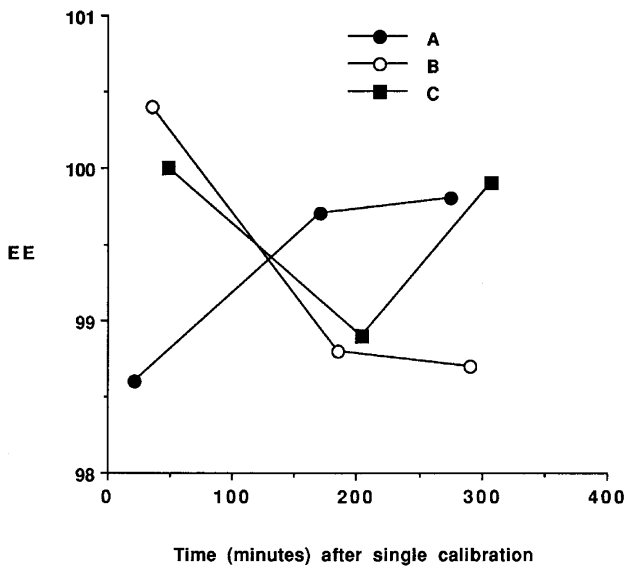

Figure 2 Change in energy expenditure (EE) (percentage of expected value) with time following a single calibration.

recovery of both gases. This makes an error in flow rate, which would act consistently on both gases, a more likely explanation.

Flow rate is notoriously difficult to measure, as the measurement procedure may interfere with the flow itself. If the Wright flow meter was inaccurate, the error would have to be considerably larger than the reported accuracy of $\pm 0.25\%$ to account for the observed error in VCO_2 and VO_2. A further possibility is that the flow rate of each machine is slightly low, which we investigated using infusion tests of N_2 alone which are not subject to errors involving the estimation of infusion gas content. Our results from these infusions reproduce the consistency of error in VCO_2 and VO_2 described above. Thus, assuming the oxygen analysers themselves to be accurate and precise, either all three Deltatracs have a consistent error in flow rate of $\sim 2\%$, or the Wright flowmeter has an error of similar magnitude. The Wright flowmeter was factory calibrated, but this procedure still cannot guarantee laboratory accuracy because the oil level, on which accuracy depends, changes with use. Use of a 1 l syringe indicated that the Wright flowmeter had less than 1% error. Therefore we conclude that worst case accuracy of the Deltatrac was an error of $\sim 2\%$ in VCO_2 and VO_2, but the actual errors is probably closer to $\sim 1\%$. Attributing all flow rate error to the Wright flowmeter would result in all Deltatrac measures of VCO_2 and VO_2 lying within $\pm 1.1\%$ of the expected value.

Previous studies have assessed the accuracy of the Deltatrac monitor, but our study is the first to be able to compare three machines using a single testing environment. Using a single machine, Phang *et al* (1990) reported errors of 1.5% and 1.9% for VCO_2 and VO_2 respectively, while Makita *et al* (1990) reported maximum errors of 2.9% and 4.0% respectively. Using the paediatric option in combination with a range of infusion rates Weyland *et al* (1994) reported mean errors of -2.4% and -3.2% for VCO_2 and VO_2 respectively. The study of Cooper *et al* (1991) compared three machines using different testing environments during a year-long evaluation. They found accuracy of the machines varied considerably, which they attributed principally to one machine having an inaccurate flow rate. However they did not take into account the gas composition of ingoing air in calculating their infusion volumes, which would be expected to influence gas recover by -1% for VO_2 and $+1.5\%$ for VCO_2. Our levels of error therefore compare favourably with other reports, with between-machine agreement being considerably higher than reported previously.

Previous studies have also considered precision of the Deltatrac Mk1 monitor, but again our study is the first to take into account both between-machine and between-study reproducibility in a stable testing environment. Our approach allows us to consider the precision of the monitor itself, rather than the canopy set-up which could introduce additional variation between testing occasions. Cooper *et al*, 1991 reported within-study reproducibility values of 5.2% and 3.2%, and between-study reproducibility values of 2.7% and 3.0%, for VCO_2 and VO_2 respectively. Our results show better precision both within study and between studies. The improved values for between-study reproducibility may be due to the stability of the testing environment. If this is so, it would demonstrate the importance of standardising the canopy set-up between measurements.

Our second series of tests demonstrates the effect of time elapsed following calibration on the Deltatrac's output. Energy expenditure and VO_2 were relatively unaffected, whereas VCO_2 and RQ decreased with increasing time. Such effects might be expected to be more marked if the interval were characterised by greater changes in environmental temperature. We therefore recommend that the monitor is calibrated according to the manufacturer's instructions immediately prior to every measurement.

Many potential uses for indirect calorimetry involve looking for differences in gas exchange and fuel utilisation either within or between subjects. The greater the accuracy and precision of the methodology, the more sensitive it is, enabling for example minor changes in substrate utilisation to be detected. Our study indicates that the Deltatrac Mk 1 Metabolic Monitor combines good accuracy with a high degree of precision. Most importantly, reproducibility was found to be high not only within studies and between studies in a single machine, but also between machines. These findings indicate that the monitor is suitable for longitudinal studies, and that results from different groups can be compared providing that measurement protocols are similar. However, all machines should be evaluated if possible to confirm accuracy and precision, particularly of the flow rate.

Acknowledgements—We are grateful to Peter Mugatroyd of the Dunn Clinical Nutrition Centre for his expert help in this study.

References

Amarri S, Harding M, Coward WA, Evans TJ & Weaver LT (1997): ^{13}C mixed triglyceride breath test and pancreatic enzyme supplementation in children with cystic fibrosis. *Arch. Dis. Child.* **76**, 349–351.

Bogardus C, Lillioja S, Ravussin E, Abbott W, Zawadzki JK, Young A, Knowler WC, Jacobowitz R & Moll PP (1986): Familial dependence of the resting metabolic rate. *N. Engl. J. Med.* **315**, 96–100.

Case KO, Brahler CJ & Heiss C (1997): Resting energy expenditure in Asian women measured by indirect calorimetry are lower than expenditures calculated from prediction equations. *J. Am. Diet. Assoc.* **97**, 1288–1292.

Cooper BG, McLean JA & Taylor R (1991): An evaluation of the Deltatrac indirect calorimeter by gravimetric injection and alcohol burning. *Clin. Phys. Physiol. Meas.* **12**, 333–341.

Dickerson RN, White KG, Curcillo PG, King SA & Mullen JL (1995): Resting energy expenditure of patients with gynecologic malignancies. *J. Am. Coll. Nutr.* **14**, 448–454.

Fuller NJ & Elia M (1989): Does mitochondrial compartmentation of CO_2 exist in man? *Clin. Physiol.* **9**, 345–352.

Fuller NJ, Sawyer MB, Coward WA, Paxton P & Elia M (1996): Components of total energy expenditure in free-living elderly men (over 75 years of age): measurement, predictability and relationship to quality of life indices. *Br. J. Nutr.* **75**, 161–173.

Gutin B, Islam S, Manos T, Cucuzzo N, Smith C & Stachura ME (1994): Relation of percentage of body fat and maximal aerobic capacity to risk factors for atherosclerosis and diabetes in black and white seven-to-eleven-year-old children. *J. Pediatr.* **125**, 847–852.

Kaplan AS, Zemel BS, Neiswender KM & Stallings VA (1995): Resting energy expenditure in clinical pediatrics: measured versus prediction equations. *J. Pediatr.* **127**, 200–205.

Makita K, Nunn JF & Royston B (1990): Evaluation of metabolic measuring instruments for use in critically ill patients. *Crit. Care Med.* **18**, 638–644.

Murgatroyd PR, Davies HL & Prentice AM (1987): Intra-individual variability and measurement noise in estimates of energy expenditure by whole body indirect calorimetry. *Br. J. Nutr.* **58**, 347–356.

Murphy MD, Ireton-Jones CS, Hilman BC, Gorman MA & Liepa GU (1995): Resting energy expenditures measured by indirect calorimetry are higher in preadolescent children with cystic fibrosis than expenditures calculated from prediction equations. *J. Am. Diet. Assoc.* **95**, 30–33.

Journal of Pediatrics Editorial (1995): Measuring resting energy expenditure in clinical practice. *J. Pediatr.* **127**, 269–271.

Phang PT, Rich T & Ronco J (1990): A validation and comparison study of two metabolic monitors. *JPEN* **14**, 259–261.

Ravussin E & Bogardus C (1989): Relationship of genetics, age, and physical fitness to daily energy expenditure and fuel utilisation. *Am. J. Clin. Nutr.* **49**, 968–975.

Roberts SB, Murgatroyd PR, Crisp JA, Nohria V, Schlingenseipen K-H & Lucas A (1987): Long-term variation in oxygen consumption rate in pre-term infants. *Biol. Neonate.* **52**, 1–8.

Schebendach J, Golden NH, Jacobson MS, Arden M, Pettei M, Hardoff D, Bauran N, Reichert P, Copperman N, Hertz S *et al* (1995): Indirect calorimetry in the management of eating disorders. *Int. J. Eat. Disord.* **17**, 59–66.

Sonko BJ, Prentice AM, Murgatroyd PR, Goldberg GR, van de Ven MLHM & Coward WA (1994): Effect of alcohol on postmeal fat storage. *Am. J. Clin. Nutr.* **59**, 619–625.

Weir JB de V (1949): New methods for calculating metabolic rate with special reference to protein metabolism. *J. Physiol.* **109**, 1–9.

Weyland W, Weyland A, Fritz U, Redecker K, Ensink F-B & Braun U (1994): A new paediatric metabolic monitor. *Inten. Care Med.* **20**, 51–57.