Measured versus predicted oxygen consumption in children with congenital heart disease

P O Laitinen, J Räsänen

Abstract

Objective—To compare measured and predicted oxygen consumption ($\dot{V}O_2$) in children with congenital heart disease.

Design—Retrospective study.

Setting—The cardiac catheterisation laboratory in a university hospital.

Patients—125 children undergoing preoperative cardiac catheterisation.

Interventions— $\dot{V}O_2$ was measured using indirect calorimetry; the predicted values were calculated from regression equations published by Lindahl, Wessel *et al*, and Lundell *et al*. Stepwise linear regression and analysis of variance were used to evaluate the influence of age, sex, weight, height, cardiac malformation, and heart failure on the bias and precision of predicted $\dot{V}O_2$. An artificial neural network was trained and used to produce an estimate of $\dot{V}O_2$ employing the same variables. The various estimates for $\dot{V}O_2$ were evaluated by calculating their bias and precision values.

Results—Lindahl's equation produced the highest precision ($\pm 42\%$) of the regression based estimates. The corresponding average bias of the predicted $\dot{V}O_2$ was 3% (range -66% to 43%). When $\dot{V}O_2$ was predicted according to regression equations by Wessel and Lundell, the bias and precision were 0% and $\pm 44\%$, and -16% and $\pm 51\%$, respectively. The neural network predicted $\dot{V}O_2$ from variables included in the regression equations with a bias of 6% and precision $\pm 29\%$; addition of further variables failed to improve this estimate.

Conclusions—Both regression based and artificial intelligence based techniques were inaccurate for predicting preoperative $\dot{V}O_2$ in patients with congenital heart disease. Measurement of $\dot{V}O_2$ is necessary in the preoperative evaluation of these patients.

(Heart 1998;80:601-605)

Keywords: oxygen consumption; congenital heart disease; neural networks

Oxygen consumption ($\dot{V}O_2$) must be known when haemodynamic indices are determined using the Fick principle. It can be measured using indirect calorimetry,^{1,2} but is often estimated from nomograms constructed according to the patient's characteristics.³⁻⁶ The regression equations are usually generated in a patient population under study, but they are seldom put to test in patients who were not part of the original group. Moreover, the relation

between $\dot{V}O_2$ and the variables commonly used to explain changes in $\dot{V}O_2$ may not be best described or modelled using linear methods.^{7,8} Several characteristics other than age, sex, and body size may influence $\dot{V}O_2$ in patients with congenital heart disease, but as the number of explaining variables increases, so does the complexity of prediction and probability of alinearity. For these reasons, haemodynamic results, determined using nomograms, may be misleading in evaluating the suitability of the patient for operation.

Our aim in this study was to test the value of the commonly used regression equations by comparing measured and predicted $\dot{V}O_2$ at cardiac catheterisation in a group of children with congenital heart disease who were not a part of the original dataset from which these equations were derived. To provide and evaluate an estimate of $\dot{V}O_2$ which is free from the constraints of known or assumed mathematical dependency, we trained an artificial neural network to predict $\dot{V}O_2$ from data available at the time of the procedure.

Methods

PATIENTS

Measured and predicted $\dot{V}O_2$ values were obtained at cardiac catheterisation from 125 patients with congenital heart disease. All patients fasted for at least four hours before induction of anaesthesia. Oral flunitrazepam (0.1 mg/kg, maximum dose 2 mg) was given as premedication. EMLA cream (Astra, Södertälje, Sweden) was applied on the groins and hands of the patients before arrival at the catheterisation laboratory. Glycopyrrolate (5 μ g/kg), was given to all the patients. Anaesthesia was induced with intravenous ketamine (1-5 mg/kg) and maintained with ketamine infusion (1-3 mg/kg/h) to assure spontaneous breathing and stable conditions for the investigation. Supplemental boluses of fentanyl were given as needed. The trachea was intubated with a cuffed tube if mechanical ventilation (Servo 900, Siemens, Solna, Sweden) was required. In these patients, neuromuscular blockade was provided with pancuronium bromide or atracurium. The groins of spontaneously breathing patients were infiltrated with a local anaesthetic.

PROCEDURES

An open circuit, indirect calorimetry device (Deltatrac, Datex-Engström, Helsinki, Finland) was used to measure $\dot{V}O_2$. With this device, the mean $\dot{V}O_2$ difference has been shown to average -3.2% and the precision $\pm 23\%$ when compared with mass spectrometry and

Department of
Anaesthesiology,
Hospital for Children
and Adolescents,
University of Helsinki,
Helsinki, Finland
P O Laitinen

Department of
Anesthesiology, Mayo
Clinic, Rochester,
Minnesota, USA
J Räsänen

Correspondence to:
Dr P O Laitinen, Hospital
for Children and
Adolescents, University of
Helsinki, Stenbäckinkatu 11,
FIN-00290, Helsinki,
Finland.

Accepted for publication
15 July 1998

wet gas spirometry in paediatric patients.² The calorimeter was calibrated before each measurement according to the instructions supplied by the manufacturer. The measurements were performed once every minute after a 10 minute stabilisation period, for 10 to 20 minutes. While the child was breathing room air spontaneously, a transparent canopy covered the child's head and upper chest. The expiratory gases were captured from inside the canopy by a continuous flow of gas into the calorimeter. Escape of expiratory gas was prevented by a soft plastic skirt attached to the rim of the canopy. If supplemental oxygen was used, the expiratory gases were fed through a mixing chamber. During mechanical ventilation, the calorimeter was connected to the outlet port of the ventilator and expired gas was collected through a mixing chamber into the metabolic monitor. An inspiratory oxygen fraction of 0.60 or lower was used during all measurements.

The predicted $\dot{V}_{O_2}$ values were calculated from the regression equations published by Lindahl,⁴ Wessel *et al.*,⁵ and Lundell *et al.*⁶

According to Lindahl, in children weighing less than 10 kg, $\dot{V}_{O_2}$ (ml/min) = $6.8 \times \text{weight (kg)} + 8.0$, and in those weighing more than 10 kg, $\dot{V}_{O_2}$ (ml/min) = $4.0 \times \text{weight (kg)} + 35.8$.⁴

In the equation published by Wessel, $\dot{V}_{O_2}$ (ml/min) = $144.8 \times \text{body surface area (m}^2\text{)} + 5.6$.⁵

According to Lundell, in children under 3 years of age, $\dot{V}_{O_2}$ (ml/min) = $0.40 \times \text{weight (kg)} + 1.91 \times \text{height (cm)} + 0.17 \times \text{heart rate (beat/min)} - 91.0$. In boys older than 3 years, $\dot{V}_{O_2}$ (ml/min) = $157.9 \times \text{body surface area (m}^2\text{)} + 0.79 \times \text{heart rate (beat/min)} - 61.8$. In girls older than 3 years, $\dot{V}_{O_2}$ (ml/min) = $159.0 \times \text{body surface area (m}^2\text{)} + 0.77 \times \text{heart rate (beat/min)} - 61.6$.⁶

The pulmonary vascular resistance index (PVRI) was determined according to a standard formula using both measured and predicted $\dot{V}_{O_2}$ values.

The age, sex, weight, and body surface area of the patients, type of cardiac malformation, treatment for heart failure, and decision of operability were recorded, as were heart rate, haemoglobin, systemic and pulmonary artery pressures, and pulmonary to systemic flow ratios obtained during the measurements. All catheterisations were performed at ambient temperature of 23–25°C.

STATISTICAL ANALYSIS

A power calculation for a 20% difference in PVRI between the methods, with the probability of type α error of 5% and a probability of type β error of 20%, yielded a sample size of 100 patients. The agreement between the measured and predicted $\dot{V}_{O_2}$ values was assessed by plotting the relative error of predicted $\dot{V}_{O_2}$ against the average of measured and predicted values.⁹ Stepwise linear regression and analysis of variance were used to evaluate the influence of age, weight, body surface area, type of cardiac malformation, heart failure, heart rate, haemoglobin, systemic and pulmonary artery pressures, and pulmonary to

systemic flow ratio on the precision of the predicted $\dot{V}_{O_2}$. The group means between canopy and ventilator measurements were compared with Student's unpaired t test. Results are presented as mean (SD) and range. Differences were considered statistically significant when the probability of type α error was less than 5%.

For the neural network training and testing,¹⁰ the dataset of 125 patient records was normalised so that all variables had values ranging from 0 to 1. Two separate datasets were then formed using random selection; 101 records were selected to train the neural networks, while the remaining 24 were used for testing the accuracy of the trained network. We first trained and tested a neural network with input consisting of age, sex, height, weight, and heart rate, the variables used by Lundell to estimate $\dot{V}_{O_2}$ with linear regression. To provide the neural network with additional information to improve its estimate of $\dot{V}_{O_2}$ we added, in a second analysis, nine further variables to describe the patient, the cardiac lesion, and the measurement of $\dot{V}_{O_2}$. These were body surface area, $\dot{V}_{O_2}$ calculated according to Lindahl's equation, mean pulmonary artery pressure, mean blood pressure, systolic pulmonary to systemic pressure ratio, cyanotic versus acyanotic nature of the cardiac malformation, haemoglobin concentration, the presence of heart failure, and whether the $\dot{V}_{O_2}$ measurement was performed with a canopy or a ventilator. The output requested from the neural network was a number representing the normalised estimate of $\dot{V}_{O_2}$. The absolute estimates of $\dot{V}_{O_2}$ were calculated by reversing the normalisation process. We used a back propagation neural network simulated with a software package (NNmodel; Neural Fusion, Middletown, New York, USA).

The neural network configuration was determined by repeated efforts to train networks of various configurations with the training dataset. The final network consisted of an input layer of five or 14 neurones, depending of the number of input variables, an intermediate layer of four neurones, and one output neurone. Before training, the neural network was initialised by setting its internal weights to random numbers. The training dataset was then fed into the neural network, allowing it to "learn" by adjusting its internal weights after each record was processed, according to the difference between the estimated and true $\dot{V}_{O_2}$. Observation during training indicated that 2000 passes of the training dataset were sufficient for the learning process. After learning was complete, the testing dataset was analysed with the network, and the resulting neural network estimates of $\dot{V}_{O_2}$ were compared with the corresponding measured values. The error of calculated $\dot{V}_{O_2}$ was compared with the error of the two neural network estimates with Student's t test for means of two samples.

Results

The demographic data of the patients are shown in table 1. Of the 125 measurements, 84 were performed in spontaneously breathing patients using the canopy.

Figure 1 Regression lines for measured oxygen consumption ($\dot{V}O_2$) values (ml/min), and predicted $\dot{V}O_2$ values (ml/min) according to Lindahl,⁴ plotted against weight (kg).

The regression lines for measured $\dot{V}O_2$ and $\dot{V}O_2$ predicted according to Lindahl,⁴ plotted against weight are shown in fig 1. The relative bias and precision of the $\dot{V}O_2$ values predicted

Figure 2 (A) The relative bias and precision of the predicted oxygen consumption ($\dot{V}O_2$) values (ml/min) according to regression equation by Lindahl.⁴ (B) The relative bias and precision of the predicted $\dot{V}O_2$ values (ml/min) according to regression equation by Wessel *et al.*⁵

Table 1 Demographic characteristics of the patients ($n = 125$)

Female/male	73/52
Age (years)	1.9 (2.1)
Weight (kg)	9.0 (4.2)
Body surface area (m ²)	0.43 (0.15)
Heart failure	109
Cyanosis	45
Qp:Qs >1.5	66
Inspired oxygen fraction	0.27 (0.08)

Values are given as mean (SD).

Qp:Qs, pulmonary to systemic flow ratio.

according to regression equation by Lindahl⁴ are presented in fig 2A and the corresponding values using the regression equation by Wessel *et al.*⁵ in fig 2B. The relative bias of the predicted $\dot{V}O_2$ was 3% (range -66% to 43%) and the precision $\pm 42\%$ when regression equation by Lindahl was used.⁴ According to Wessel *et al.*, the relative bias was 0% (range -69% to 39%) and the precision $\pm 44\%$.⁵ When the equation of Lundell *et al.* was used,⁶ the relative bias was -16% (range -94% to 100%) and the precision $\pm 51\%$. Since calculations based on the equation published by Wessel and Lundell resulted in poorer precision than those based on the equation published by Lindahl, the subsequent comparisons are made using Lindahl's equation.⁴ The relation between PVRI calculated using measured $\dot{V}O_2$ and predicted $\dot{V}O_2$ is shown in fig 3. In three patients with left to right shunts, two with Down's syndrome, and one patient with left heart downstream obstruction, the use of predicted $\dot{V}O_2$ in the calculation of PVRI would have resulted in overestimation of surgical risk. All these patients underwent surgical correction successfully. In four patients with Down's syndrome and left to right shunts, the use of predicted $\dot{V}O_2$ would have underestimated the calculated PVRI. One of these patients died of postoperative pulmonary hypertension; the others survived.

Age, weight, sex, body surface area, pulmonary or systemic artery pressure, heart rate, haemoglobin, and pulmonary to systemic flow index ratio had no statistically significant influence on the error of predicted $\dot{V}O_2$, nor did heart failure or the type of cardiac malformation.

When the variables found in the published regression equations were used as input to the neural network, the relative bias of the estimated $\dot{V}O_2$ was 6% (range -19% to 30%), and the precision was $\pm 29\%$. The addition of other input variables to the neural network did not produce a marked improvement in the estimation of $\dot{V}O_2$. Even though the neural network estimated $\dot{V}O_2$ with higher precision, there was no statistically significant difference in the error of $\dot{V}O_2$ calculated according to the equation published by Lindahl and the error of the neural network estimates of $\dot{V}O_2$.

There was a statistically significant difference in the bias of the predicted $\dot{V}O_2$ between measurements made with a canopy compared with those made in mechanically ventilated patients. The bias averaged 7% (range -50% to 43%) in spontaneously breathing patients, and -5% (range -66% to 30%) in mechanically ventilated patients ($p < 0.03$).

Figure 3 The relation between measured and predicted pulmonary vascular resistance index (PVRI; Wood units $\times$ m^2) according to Lindahl's regression equation.⁴

Discussion

A raised PVRI may be a contraindication to corrective surgery in patients with congenital heart disease. Therefore the evaluation of the operability and the assessment of risks of surgery can only be based on precise haemodynamic measurements. Our study shows that there is a lack of agreement between measured $\dot{V}O_2$ and $\dot{V}O_2$ predicted using regression based or artificial intelligence based methods.

Equipment for routine measurement of $\dot{V}O_2$ in children has only recently become widely available. Therefore it has been necessary to estimate $\dot{V}O_2$ on the basis of relatively small published studies conducted under variable clinical conditions and with variable types of equipment. Lindahl generated predictive regression equations for $\dot{V}O_2$ by making measurements with a pneumotachograph, a Douglas bag, an infrared carbon dioxide meter, and a mass spectrometer in 34 anaesthetised, healthy, spontaneously breathing children and four mechanically ventilated children with congenital heart disease. In that study, $\dot{V}O_2$ was related to weight, and there were no differences in $\dot{V}O_2$ between children with or without congenital heart disease.⁴ The patient population of that study, as well as the measurement technique, was different from the one employed in the present study. Nevertheless, we found that the regression line describing the relation between measured $\dot{V}O_2$ and weight is surprisingly similar to Lindahl's original equation (fig 1). However, neither of the presented regression equations was sufficiently accurate in predicting $\dot{V}O_2$ —despite the small relative bias of the predicted $\dot{V}O_2$, the agreement between the measured and predicted $\dot{V}O_2$ values was poor (fig 2A).⁹

Wessel *et al* determined oxygen uptake at cardiac catheterisation in 98 sedated spontaneously breathing patients with congenital heart disease using a flow-through technique with a paramagnetic oxygen analyser and an infrared carbon dioxide analyser. Their results showed a linear relation between oxygen uptake and body surface area.⁵ In the present study, the degree of agreement between the measured $\dot{V}O_2$ and the $\dot{V}O_2$ predicted according to Wessel was

no higher than between the measured $\dot{V}O_2$ and the $\dot{V}O_2$ calculated using equation published by Lindahl.

Lundell *et al* measured $\dot{V}O_2$ in 504 children with cardiac abnormalities using a hood system with a pneumotachograph and a paramagnetic oxygen analyser.⁶ The patients were sedated and were breathing room air spontaneously. In that study, in children under 3 years of age, the body dimensions and heart rate were found to have significant influence on $\dot{V}O_2$. In children older than 3 years, sex, body surface area, and heart rate affected the measured $\dot{V}O_2$ significantly. The investigators emphasised that the predictive nomograms, if employed, should be used taking into account the wide confidence intervals of the predicted $\dot{V}O_2$ values. We found that the degree of agreement between the measured and predicted $\dot{V}O_2$ values, calculated according to Lundell, was no higher than when the equations of Lindahl and Wessel were used.

FACTORS AFFECTING THE VARIATION IN $\dot{V}O_2$

Considerable care must be taken when models estimating the behaviour of a biological variable are adopted for clinical use. The sample used in the generation of the model must be sufficiently large and representative of the variation in the population in which it is used. Any model probably gives better predictions for those cases that were included its generation than for new cases which the model has not "seen." Therefore, an understanding of the "goodness" of a model can only come from its application to a separate test set of data.

In this respect, the tests of the regression equations in the current study revealed considerable problems. The large variability between measured $\dot{V}O_2$ values and values produced with the equations calls into question their routine clinical use.

We found that the use of a neural network, a pattern recognition technique that does not assume or require a mathematical relation between the independent and dependent variables, produced less variability in its estimate of $\dot{V}O_2$ than a linear regression technique, even when the same variables were used. This suggests that models that assume linear effects of these variables on $\dot{V}O_2$ are at a disadvantage. Nevertheless, we were not able to produce sufficiently accurate estimates of $\dot{V}O_2$ even with an artificial neural network. The addition of input variables to the neural network did not appear to make its prediction any more accurate from that based on age, sex, height, weight, and heart rate—variables found relevant by Lundell *et al*. We did not experiment further by including or excluding individual variables in the hope of finding an ideal set; rather, we provided an abundance of available variables for the second neural network and relied on the learning process to determine the significance of the individual variables to the output. The fact that very little was gained from the addition of variables indicates that their influence on $\dot{V}O_2$ was small. This finding was supported by the lack of effect of these variables in the stepwise linear regression analysis.

The large relative error of predicted $\dot{V}O_2$ values represents the large interindividual variation in the patients of the present study. Differences in the degree of heart failure and metabolic rate of the patients may have produced variation in $\dot{V}O_2$, even though the tested variables had no significant influence on the error of predicted $\dot{V}O_2$. The determination of heart failure may not have been sufficiently accurate in the evaluation of circulatory compromise in these patients. Variation in body temperature also may have introduced variation in $\dot{V}O_2$.¹¹

The most likely explanation for the observed difference in $\dot{V}O_2$ between spontaneously breathing and mechanically ventilated patients is the effect of anaesthetic depth and neuromuscular blockade on $\dot{V}O_2$. Even though stable conditions were achieved during the investigation, light anaesthesia may have augmented $\dot{V}O_2$ in the spontaneously breathing patients. In contrast, mechanically ventilated patients received full ventilatory support with induced neuromuscular blockade. These factors probably decreased the cardiopulmonary work and metabolic rate, resulting in the observed difference in $\dot{V}O_2$.

CONCLUSIONS

Modelling of the large interindividual biological variation in patients with congenital heart disease in order to predict $\dot{V}O_2$ is complicated. Our results show that the nomograms employed in this study, as well as the methods

based on artificial intelligence, do not accurately predict preoperative $\dot{V}O_2$ in patients with congenital heart disease. Though the decision to correct a cardiac malformation surgically does not solely depend on any one haemodynamic variable, we believe that measurement of $\dot{V}O_2$ is necessary in the preoperative evaluation of these patients.

This study was supported by a grant from the Foundation for Pediatric Research, Helsinki, Finland.

- 1 Takala J, Keinänen O, Väisänen P, *et al.* Measurement of gas exchange in intensive care: laboratory and clinical validation of a new device. *Crit Care Med* 1989;17:1041–7.
- 2 Weyland W, Weyland A, Fritz U, *et al.* A new paediatric metabolic monitor. *Intensive Care Med* 1994;20:51–7.
- 3 LaFarge CG, Miettinen OS. The estimation of oxygen consumption. *Cardiovasc Res* 1970;4:23–30.
- 4 Lindahl SGE. Oxygen consumption and carbon dioxide elimination in infants and children during anaesthesia and surgery. *Br J Anaesth* 1989;62:70–6.
- 5 Wessel HU, Rorem D, Muster AJ, *et al.* Continuous determination of oxygen uptake in sedated infants and children during cardiac catheterization. *Am J Cardiol* 1969;24:376–85.
- 6 Lundell BPW, Casas ML, Wallgren CG. Oxygen consumption in infants and children during heart catheterization. *Pediatr Cardiol* 1996;17:207–13.
- 7 Dehmer GJ, Firth BG, Hillis LD. Oxygen consumption in adult patients during cardiac catheterization. *Clin Cardiol* 1982;5:436–40.
- 8 Sherman MS, Kosinski R, Paz HL, *et al.* Measuring cardiac output in critically ill patients: disagreement between thermodilution-, calculated-, expired gas-, and oxygen consumption-based methods. *Cardiology* 1997;88:19–25.
- 9 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;i:307–10.
- 10 Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. *Lancet* 1995;346:1075–9.
- 11 Shapiro H, Stoner EK. Body temperature and oxygen uptake in man. *Annals of Physical Medicine* 1966;8:250–5.