REVIEW ARTICLE
Cardiopulmonary exercise testing as a risk assessment method in non cardio-pulmonary surgery: a systematic review

T. B. Smith,1 C. Stonell,2 S. Purkayastha3 and P. Paraskevas4
1 Medical Student, 2 Consultant, Department of Anaesthesia, 3 Clinical Lecturer, 4 Senior Lecturer and Consultant Surgeon, Departments of Biosurgery and Surgical Technology and Academic Surgical Unit, St Mary's Hospital, Imperial College Healthcare NHS Trust, London, UK

Summary
This study reviews the predictive value of maximum oxygen consumption (VO$_2$ max) and anaerobic threshold, obtained through cardiopulmonary exercise testing, in calculating peri-operative morbidity and mortality in non-cardiopulmonary thoraco-abdominal surgery. A literature review provided nine studies that investigated either one or both of these two variables across a wide range of surgical procedures. Six of the seven studies that reported sufficiently detailed results on peak oxygen consumption and four of the six studies that reported sufficiently detailed results on anaerobic threshold found them to be significant predictors. We conclude that peak oxygen consumption and possibly anaerobic threshold are valid predictors of peri-operative morbidity and mortality in non-cardiopulmonary thoraco-abdominal surgery. These indicators could potentially provide a means of allocating increased care to high-risk patients.

Correspondence to: Dr Christopher Stonell
E-mail: christopher.stonell@imperial.nhs.uk
Accepted: 1 April 2009

Cardiopulmonary exercise testing (CPET or CPX) is a non-invasive, integrated assessment of cardiovascular and pulmonary function both at rest and under stress. Among the purported benefits is its ability to determine the ability of the subject's physiological capacity to cope with the metabolic demands created by the trauma of major surgery. During the test subjects are exposed to incremental physical exercise up to their maximally tolerated level, dictated either by exhaustion or symptom related cessation (e.g., breathlessness or angina). Several physiological variables are recorded including ventilatory parameters, inspiratory and expiratory gases, blood pressure (BP) and electrocardiogram (ECG). From these are derived two key indicators: the body's maximum oxygen uptake (VO$_2$ max) and the point at which anaerobic metabolism exceeds aerobic metabolism (Ventilatory Anaerobic Threshold or VAT). Together, these broadly indicate the ability of the cardiovascular system to deliver oxygen to the peripheral tissues and the ability of the tissues to utilise that oxygen. It has already been demonstrated that measures such as VO$_2$ max are useful predictors of postoperative complications of pulmonary resection surgery [1–3] and assessing the timing of cardiac transplant surgery [4], while the VAT is a predictor of postoperative cardiac complications in abdominal surgery [5].

Physiology of the anaerobic threshold
Lactate accumulation in exercising muscle occurs when the muscle O$_2$ demand exceeds the supply. At this point the venous lactate concentration will begin to rise. This point is called the Lactate Anaerobic Threshold (LAT). The muscle lactate/pyruvate ratio increases at the LAT which supports the concept that the lactic acidosis results from relative muscle hypoxia [6–8].

As an individual commences an incremental exercise test, their expired minute volume ($\dot{V}_E$), oxygen consumption per minute (VO$_2$) and CO$_2$ production per minute (VCO$_2$) all increase linearly with respect to variables such as work rate or time. However a point is reached when $\dot{V}$CO$_2$ increases out of proportion to $\dot{V}$O$_2$...
This change is attributed to HCO$_3^-$ buffering the lactate produced and consequently generating a relative excess of CO$_2$. This transition is termed the Lactate Anaerobic Threshold (LAT). This phenomenon allows determination of the anaerobic threshold by respiratory measurements. Generally this is seen at 47–64% of VO$_{2\text{max}}$ in healthy untrained individuals [9] and is a marker of the maximum work rate that can be sustained for a prolonged period [10, 11]. As this point is less effort dependant than VO$_{2\text{max}}$ it is a clinically attractive measurement in a patient population where it may be difficult to achieve the maximal effort required for a true VO$_{2\text{max}}$ result. Failure to perform sufficient exercise during the test to reach VAT suggests either poor motivation or a non-cardiac (e.g. pulmonary or musculoskeletal) problem [12]. On reaching the AT a fall in the arterial pH caused by anaerobic respiration results in the carotid bodies stimulating an increase in minute ventilation [13]. This point is termed the Ventilatory Anaerobic Threshold (VAT) (Fig. 2). During this period the VE increases in proportion to the VCO$_2$ and the VE/VCO$_2$ remains constant. The rate of rise of VE exceeds the rate of rise of VO$_2$ and so VE/VO$_2$ starts to increase. This gives one method of determining the VAT from gas exchange measurements as AT is the point when the Ventilatory Equivalent for oxygen (VE/VO$_2$) increases relative to a constant Ventilatory Equivalent for carbon dioxide (VE/VCO$_2$) [14] (Fig. 3).

Another common method used to determine VAT is the ‘V-Slope’ method [15]. In order to determine the VAT VCO$_2$ is plotted against VO$_2$ during an incremental exercise test. Initially the slope is unitary, however at the VAT, when VCO$_2$ begins to increase, the slope increases. From this point of deflection of VCO$_2$ the VO$_2$ at which AT is reached can be determined (Fig. 4). Thus there
against transition between aerobic and anaerobic metabolism. Fig. 3 plotted in a different way in order to determine the transition between aerobic and anaerobic metabolism.

Add a degree of error to accurately assigning points such as actual physiological occurrence [27]. This has potential to variability is a by-product of the conversion method or an incomplete conversion method or an actual physiological occurrence [27]. This has potential to add a degree of error to accurately assigning points such as VAT and would benefit from further elucidation on the subject.

Notwithstanding these important issues needed to clarify the physiology underlying VAT and LAT, it is still possible to use either or both as a (perhaps relatively crude) non-invasive and useful marker of the onset of a significant anaerobic component in respiration.

Cardiopulmonary exercise testing in practice

There are several different protocols for achieving VO2max. Most use either an exercise bicycle (cycle ergometer) or a treadmill. Cycle ergometers tend to be used in peri-operative clinical settings as they provide a degree of safety in maximal testing, offset the work of walking in the obese and are less prone to artifacts in the ECG, airflow and pressure measurement. A higher peak VO2 can usually be generated on a treadmill [9, 12], however there is a learned technique to walking on a treadmill, which patients may not be familiar with, and should the subject hold on to any part of the treadmill the amount of work performed is altered as a consequence of static work being performed in the arms. This would interfere with the determination of an accurate VO2–work rate relationship as the static work performed is not accounted for by the ergometer [9]. In both methods, work rate is gradually and imperceptibly increased in a stepped or ramped manner until the subject is unable to continue. On a treadmill this can be achieved by increasing both the gradient and speed while on an exercise bicycle it is achieved by increasing the resistance of the pedals while the subject maintains a constant pedaling rate. If a subject is unable to cycle or walk then an arm crank may be employed.

The optimal duration of exercise is approximately 10 min [28]. Too short a duration of exercise may produce an underestimation of VO2max [28], while too long a duration may cause premature cessation due to the subject becoming demotivated or because of orthopaedic factors [29]. VAT seems to be independent of the duration of exercise [28]. The testing protocol can be adjusted according to predicted values based on age and height to achieve this [30]. In the United Kingdom the Consensus Protocol for Pre-Operative CPX Testing for York, Torbay and UCLH (Pre-operative Association, UK) is an accepted pre-operative testing protocol.

All physiological variables are recorded breath by breath and respiratory variables are then typically averaged over periods of 15–45 s [29].

The term VO2max is used for the highest VO2 recorded when a subject’s VO2 is observed to reach a plateau with increasing workload. If this plateau is not seen, as is often
the case, but the subject is believed to be near to their maximal effort it is termed their VO₂peak. In a clinical setting these are considered essentially similar.

**Current applications of CPET**

The American Thoracic Society/American College of Chest Physicians Statement on Cardiopulmonary Exercise Testing (2003) lists the specific indications for CPET testing [31] that include:

1. Evaluation of exercise tolerance, where the diagnosis is known, in order to objectively evaluate functional capacity, disability or response to treatment.
2. Evaluation of undiagnosed exercise intolerance where cardiac and respiratory aetiologies may coexist, the symptoms are disproportionate to the results of resting investigations or the investigations are non diagnostic.
3. Evaluation of patients with cardiovascular diseases.
4. Evaluation of patients with respiratory diseases/symptoms.
5. Pre-operative evaluation.
6. Exercise evaluation and prescription for pulmonary rehabilitation.

We explain some of these in greater detail:

**Evaluation of exercise tolerance:** As VO₂max measures maximal aerobic capacity it is a very useful indicator of the ‘cardiovascular fitness’ of an athlete, an important variable in endurance sports that have a high aerobic demand such as long distance running, cycling and skiing. However as most athletes at a specific level of a sport will have similar values it does not serve to predict their individual competitive performance (as other variables such as economy come into play) but serves instead to act as a personal indicator after a period of training or injury through comparison with previous values (e.g., akin to a golf handicap).

**Evaluation of impairment/disability:** Oxygen consumption can be used to indirectly assess caloric expenditure in subjects, since the two mirror each other during exercise. This results in the calculation of a Metabolic Equivalent (MET), with 1 MET corresponding to an oxygen usage of 3.5 ml.kg⁻¹.min⁻¹ (or the same resting intake as a healthy 70 kg man) [29]. It is therefore possible to establish the level of physical activity a subject should be able to obtain in reference to common activities. For example making a bed has a metabolic cost of 2 MET’s, painting a fence 4.5 MET’s, jogging 7 MET’s and moving furniture 9 METS [32]. With this reference data and the known VO₂max of an individual it is possible to determine those daily tasks that can be completed independently and those for which assistance is needed. VO₂max rather than VAT is more accurate in this context since individuals with a very low VO₂max may be producing lactate at rest [33].

**Cardiopulmonary testing as a pre-operative predictor of risk**

Cardiopulmonary exercise testing is most documented in cardiac and pulmonary medicine and surgery. A review of multiple lung tumour resection studies by Beckles et al. [34] provides data to suggest VO₂max can be used to predict peri-operative complication rates in tumour resection surgery:

1. VO₂peak > 20 ml.kg⁻¹.min⁻¹ = no increased risk of complications or death,
2. VO₂peak > 14 ml.kg⁻¹.min⁻¹ = increased risk of peri-operative complications,
3. VO₂peak < 10 ml.kg⁻¹.min⁻¹ = a very high risk of for post-operative complications.

Mancini et al. [4] used CPET to assess the urgency of cardiac transplantation by comparing the survival of heart failure patients who were not accepted for cardiac transplantation vs those who received a transplant. Of those not accepted for transplantation, a VO₂peak > 14 ml.kg⁻¹.min⁻¹ yielded survival comparable to those after transplantation; a VO₂peak < 14 ml.kg⁻¹.min⁻¹ yielded lower survival.

However few studies have investigated the role of cardiopulmonary exercise testing in the pre-operative assessment of patients undergoing non-cardiopulmonary thoraco-abdominal surgery. If CPET was able to provide a reliable indicator of peri-operative morbidity and mortality it could have a significant impact on the care of surgical patients more generally and on the cost of their healthcare. By identifying high risk candidates additional resources could be put in place ahead of time (e.g. an Intensive Care bed) to ensure appropriate postoperative care and thus potentially reduce morbidity and mortality. Furthermore if candidates identified as being at risk were placed on an exercise regime before elective surgery this could potentially improve their CPET measurements, and in turn their postoperative risk.

In cardiac surgery, and some forms of pulmonary surgery, it is expected that cardiac and respiratory function will improve postoperatively. In contrast, after non-cardiothoracic surgery cardio-respiratory function will remain, at best, unaffected and may well deteriorate postoperatively, at a time that the demands placed on the cardio-respiratory system are increased. We used the method of systematic review to establish the role of CPET testing in non-cardiopulmonary surgery.
Methods of systematic review

Study selection
Pubmed and ISI Web of Knowledge were searched for studies investigating pre-operative exercise testing as an indicator of postoperative outcome in surgery. The search terms used were ‘CPET/surgery’, ‘CPEX/surgery’, ‘cardiopulmonary/exercise testing/surgery’ and ‘VO2max/surgery’. With both search engines the ‘related articles’ or ‘related records’ function was used to broaden the search. The references of the studies of interest were reviewed as were any studies which were found to cite them.

Inclusion/exclusion criteria
In order to be included studies had to compare subjects pre-operative VO2max and/or AT with their postoperative outcome in terms of morbidity or mortality. Studies were excluded if they reported on (i) pulmonary or cardiac surgery, (ii) non-thoraco-abdominal surgery. The selected studies were then analysed for data investigating the relationship between pre-operative CPET values, such as VO2max or AT, to postoperative outcomes such as mortality or a defined morbidity. Where present this data was extracted and included in the review.

Results
In all, 322 studies were initially identified, 315 of which were excluded (Fig. 5). The reference lists of the seven remaining studies yielded two additional articles (Table 1).

All bar one study was in English [39] which was written in Japanese with an English abstract. The abstract was deemed to have sufficient data for the purposes of this review. The nine studies can be divided into four categories: abdominal aortic aneurysm repair (AAA; n = 2), hepatic transplantation (n = 1), upper GI surgery (n = 4) and major intra-abdominal surgery in the elderly (n = 2).

Abdominal aortic aneurysm repair
Carlisle et al. [35] measured AT, VO2peak, VE/VCO2 and V̇E/V̇CO2 in 130 patients before elective repair of an AAA. All variables correlated with survival (median follow up 35 months) in univariate analysis; VO2peak Hazard ratio (95% CI [standard error of the hazard ratio (SE)]) was 0.83 (0.74–0.94 [0.05]) p = 0.002 and AT 0.74 (0.63–0.87 [0.06]) p < 0.001. In multivariate analysis however V̇E/V̇CO2 correlated best with survival to both 30 days and the total observation period 1.13 (1.07–1.19 [0.03]) p < 0.001. AT was also associated with mid-term survival but to a lesser degree 0.84 (0.72–0.98 [0.07]) p = 0.033, VO2peak was not.

Nugent et al. [41] recorded the VO2peak of 30 patients prior to elective AAA repair and followed them up postoperatively for 12 months. He found no significant difference in the VO2peak in those who had postoperative complications compared to those who did not, though there was an underlying trend with 70% of the patients with complications having a VO2peak of < 20 ml.min⁻¹.kg⁻¹ as opposed to 50% in the non-complication group. However this study did not specify any power calculations for their analysis and compared to other studies in the review had the smallest number of subjects. Therefore findings from this group should be treated with caution.

These two studies suggest that there is a weak relationship between VO2peak and postoperative morbidity and mortality in patients undergoing elective AAA repair. Only one study investigates the role of AT but finds it to be a stronger predictor.

Hepatic transplantation
Fifty nine subjects who underwent hepatic transplantation [36] and were followed up 100 days postoperatively, were subsequently divided into two groups, survivors (n = 53) and non-survivors (n = 6). Non-survivors were more likely to have a peak VO2 < 60% predicted (p = 0.04) and a VO2-AT < 50% predicted (p = 0.03). They were also more likely to have a combined peak VO2 < 60% predicted and a VO2-AT < 50% predicted (p < 0.01). This study thus suggests that both VO2peak and AT provide an indication of short term mortality following hepatic transplantation.

Upper GI surgery
Seventy-eight patients underwent oesophagectomy [37], and comparison between patients who did or did not
experience postoperative cardiopulmonary complications showed significant difference in their $V\text{O}_2\text{peak}$ ($p = 0.04$) and although there was a trend in their AT, this was not significant ($p = 0.07$). $V\text{O}_2\text{peak}$ and AT were also analysed by receiver operating characteristic curves with values of 0.63 (95% CI 0.50–0.76) and 0.62 (95% CI 0.49–0.75) respectively. Although the authors state that these values were significant different from 0.5 ($p = 0.03$ and 0.02 respectively) the values are under 0.65 and so do not show strong predictive ability.

In the earlier of Nagamatsu et al.’s [39] two papers, subjects were divided into two groups depending on the presence ($n = 11$) or absence ($n = 41$) of cardiopulmonary complications. There were significant differences between the two groups in their pre-operative $V\text{O}_2\text{max} \cdot m^{-2}$ and AT $m^{-2}$ ($p < 0.001$). Although the authors state that these values were significant different from 0.5 ($p = 0.03$ and 0.02 respectively) the values are under 0.65 and so do not show strong predictive ability.

Nagamatsu et al.’s [40] later paper investigated CPET in 91 patients undergoing radical oesophagectomy with three-field lymphadenectomy. Subjects were divided into two groups postoperatively based on cardiopulmonary complications. It was observed that $V\text{O}_2\text{peak} \cdot m^{-2}$ was lower among patients having cardiopulmonary complications ($p < 0.001$), though this was not seen for AT $m^{-2}$ values. Subjects were also grouped postoperatively into one of seven groups by their $V\text{O}_2\text{max} \cdot m^{-2}$ and the cardiopulmonary complications for each group calculated. This was 86% with a $V\text{O}_2\text{max} \cdot m^{-2}$ of $699 \text{ml.min}^{-1} \cdot m^{-2}$, 44% with a $V\text{O}_2\text{max} \cdot m^{-2}$ of 700–799 ml.min$^{-1} \cdot m^{-2}$, 10% with a $V\text{O}_2\text{max} \cdot m^{-2}$ of 800–1099 ml.min$^{-1} \cdot m^{-2}$ and 0% with a $V\text{O}_2\text{max} \cdot m^{-2}$ of 1100 ml.min$^{-1} \cdot m^{-2}$ or more.

McCullough et al. [38] recorded the $V\text{O}_2\text{peak}$ and AT of 109 bariatric patients (i.e., body mass index (SD) 48.7 (7.2)) prior to laparoscopic Roux-en-Y gastric bypass surgery. Patients were divided into tertiles based on their $V\text{O}_2\text{peak}$ (1: 6.8–15.8 ml.kg$^{-1} \cdot \text{min}^{-1}$, 2: 15.9–18.4 ml.kg$^{-1} \cdot \text{min}^{-1}$ and 3: 18.5–27.7 ml.kg$^{-1} \cdot \text{min}^{-1}$) and their complication rates were compared between these groups. It was found that complications occurred in 16.6% of patients with a peak $V\text{O}_2 < 15.8 \text{ml.kg}^{-1} \cdot \text{min}^{-1}$ and in only 2.8% of patients with a peak $V\text{O}_2 > 15.8 \text{ml.kg}^{-1} \cdot \text{min}^{-1}$, univariate odds ratio 6.77 (95% CI 1.14–51.76) $p = 0.02$. Multivariate analysis using peak $V\text{O}_2$ as a continuous variable showed it was a significant predictor of complications, odds ratio of 1.61 per unit decrease (95% CI 1.19–2.18) $p = 0.002$. Hospital length of stay and 30-day readmission rates were highest in the lowest tertile, 3.8 days vs 2.8 days for all others ($p = 0.002$). The AT was not investigated as a postoperative indicator but showed a significant trend across the tertiles ($p < 0.001$).

The first three studies all show that patients who experience cardiopulmonary complications after oesophagectomy, as a group, have significantly different $V\text{O}_2\text{peak}$ from patients who do not experience cardiopulmonary complications. Nagamatsu et al.’s [40] later study chose to use $800 \text{ml.min}^{-1} \cdot m^{-2}$ as the minimal acceptable value for patients undergoing oesophagectomy. However only one of the three papers [39] saw a significant difference in AT between the complication and non-complication groups. The final study on bariatric patients [38] shows that their $V\text{O}_2\text{peak}$ played a predictive role on their postoperative complication rate, however this was only between the lowest tertile and the two higher ones.

### Intra-abdominal surgery in the elderly

Older et al. [42] recorded the pre-surgical anaerobic threshold of 187 elderly patients undergoing major abdominal surgery (abdominal surgery which would be ‘likely to cause a significantly increase in oxygen demand’ [42]). It was found that in patients with an AT of $<11 \text{ml.min}^{-1}. \text{kg}^{-1}$ ($n = 55$) 10 died of cardiovascular causes (18%), while in patients with an AT of $>11 \text{ml.min}^{-1}. \text{kg}^{-1}$ ($n = 132$) only one died of cardiovascular causes (0.8%), $p < 0.001$. When a low AT was combined with known pre-operative ischaemia ($n = 19$) the mortality rate was 42%, as opposed to 4% in those with a high AT and pre-operative ischaemia ($n = 25$). This established that AT and ischemia were important risk factors for surgery in the elderly and led the

### Table 1: Study characteristics.

<table>
<thead>
<tr>
<th>Author (References)</th>
<th>Type of surgery</th>
<th>Year</th>
<th>Country of origin</th>
<th>Design</th>
<th>Total patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carlisle et al. [35]</td>
<td>AAA repair</td>
<td>2007</td>
<td>UK</td>
<td>Bicycle</td>
<td>130</td>
</tr>
<tr>
<td>Epstein et al. [36]</td>
<td>Hepatic transplantation</td>
<td>2004</td>
<td>USA</td>
<td>Bicycle</td>
<td>59</td>
</tr>
<tr>
<td>Forshaw et al. [37]</td>
<td>Oesophagectomy</td>
<td>2008</td>
<td>UK</td>
<td>Bicycle</td>
<td>78</td>
</tr>
<tr>
<td>McCullough et al. [38]</td>
<td>Roux-en-Y gastric bypass</td>
<td>2006</td>
<td>USA</td>
<td>Treadmill</td>
<td>109</td>
</tr>
<tr>
<td>Nagamatsu et al. [39]</td>
<td>Oesophagectomy and lymphadenectomy.</td>
<td>2001</td>
<td>Japan</td>
<td>Not specified</td>
<td>52</td>
</tr>
<tr>
<td>Nagamatsu et al. [40]</td>
<td>Thoraco-lapartomy for oesophageal cancer</td>
<td>1994</td>
<td>Japan</td>
<td>Bicycle</td>
<td>91</td>
</tr>
<tr>
<td>Nugent et al. [41]</td>
<td>AAA repair</td>
<td>1998</td>
<td>Northern Ireland</td>
<td>Treadmill</td>
<td>36</td>
</tr>
<tr>
<td>Older et al. [42]</td>
<td>Elderly (intra-abdominal surgery)</td>
<td>1993</td>
<td>Australia</td>
<td>Bicycle</td>
<td>187</td>
</tr>
<tr>
<td>Older et al. [5]</td>
<td>Elderly (intra-abdominal surgery)</td>
<td>1999</td>
<td>Australia</td>
<td>Bicycle</td>
<td>548</td>
</tr>
</tbody>
</table>
authors to suggest that all patients scheduled for major intra-abdominal surgery with an AT < 11 ml.min⁻¹.kg⁻¹ should be admitted to ICU pre-operatively. Older et al. [5] then subsequently investigated the impact of triaging patients, based on this data, on their postoperative outcome.

In this latter study [5] 548 patients over the age of 60 (or younger if they had cardiopulmonary disease) were assigned into one of three groups pre-operatively: ICU, HDU or ward admission. This was based on their pre-operative AT and ECG ST segment changes, which were ascertained by CPET. If a patient had an AT < 11 ml.min⁻¹.kg⁻¹ they were assigned to ICU pre-operatively (additionally if a patient was undergoing a high risk surgery e.g. oesophagectomy or aortic they were also assigned to this group). This accounted for 28% of the patients. If a patient had an AT > 11 ml.min⁻¹.kg⁻¹ but either demonstrated myocardial ischemia (on CPET) or a VE/VO₂ of > 35 they were admitted to HDU post-operatively (21%). All other patients were sent to the general ward after surgery (51%). Of the nine patients who died postoperatively from cardiopulmonary complications, seven had an AT < 11 ml.min⁻¹.kg⁻¹ and the other two were in the HDU category. There were no deaths in the ward admissions from cardiopulmonary causes.

These two papers show that AT has a predictive value in determining postoperative complications across a range of surgical procedures in the elderly.

Results summary

Seven of the nine studies reviewed sufficiently investigated the predictive value of pre-operative VO₂max/peak for postoperative complications, of these six found VO₂max/peak to be a valid indicator. Fewer studies reported investigating the role of Anaerobic Threshold in this regard but in the six that did so adequately, four also found it to be a valid indicator (see Table 2).

Discussion

This review demonstrates that the majority of investigations of CPET as an indicator of peri-operative complications, in non-cardiopulmonary surgery, find the technique, and VO₂peak in particular, to be a valid marker. There were three studies where no significant relationship was found. In the first, Nugent et al. [41] did not find a significant relationship between VO₂peak and postoperative outcome, though there was an underlying trend. The authors suggested that this outcome may have been influenced by small sample sizes and a reduced tendency to operate on clinically borderline patients in AAA repair as opposed to other surgical indications. The other two instances were in oesophagectomy studies. Nagamatsu et al.'s [40] later study and Forshaw et al. [37] did not observe a relationship between AT and postoperative complication rate. In the former this was hypothesised to be due to oesophagectomy, with three-field lymphadenectomy, having a cardiopulmonary burden that exceeded the exercise workload determined by AT [40]. While in the latter it was proposed that the small proportion of patients with an AT of < 11 ml.min⁻¹.kg⁻¹ (16%), the non-distinction of cardiopulmonary complications from others such as anastomotic leak and the fact all subjects spent their first night postoperatively on ITU equivalent wards may have masked any underlying trend.

One possibility is that some types of surgery maybe more sensitive to one of these two indicators than the other, that the thresholds for different types of surgery are different (as suggested by Forshaw et al. [37]), and that the level of pre-existing morbidity accepted for surgery may affect the predictive value of the CPET. Thus while we might generally conclude that CPET provides a valid prognostic indicator, this needs to be validated prior to use in specific surgical procedures. The weaker support for AT as a predictor may be a reflection of the uncertainty surrounding the exact physiological significance of this measure.

The data that was used in this review has certain limitations. As with all studies reviewing published results publication bias has the potential to influence our conclusions. However there are other limiting factors specific to this analysis. These include the range of CPET values used as pre-operative indicators, the variability of the defined end point both in terms of time after operation and classification of morbidity and the heterogeneous nature of the subject groups. One study, Older et al. [5], investigated the package of care patients received postoperatively and did not specifically examine the predictive value of CPET data. However it was included in the review as the pre-operative CPET data made up a substantial component of the pre-operative assessment and it demonstrated the increased mortality in the high risk patients even with intensive postoperative support. As a result by demonstrating that patients defined as high risk by AT (a CPET variable) still had a higher mortality despite more intensive care provides further evidence for the ability of CPET to differentiate between levels of postoperative risk.

In Older et al.'s later study [5] elderly patients undergoing intra-abdominal surgery were assigned into one of three classifications of pre and postoperative care, based primarily on their AT. Of the nine cardiopulmonary related deaths, seven came from the most high risk category and two were in the middle category. This demonstrates that using indicators recorded during CPET it is possible to pre-emptively and accurately identify
Table 2 The ability of VO2peak and anaerobic threshold to predict postoperative cardiopulmonary complications in the studies reviewed. *Hazard ratios (95% CI) and [SE].

<table>
<thead>
<tr>
<th>Author</th>
<th>VO2peak/max</th>
<th>Outcome measure</th>
<th>Study result</th>
<th>Summary findings</th>
<th>Anaerobic threshold</th>
<th>Outcome measure</th>
<th>Study result</th>
<th>Summary findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carlisle [35]</td>
<td>Postoperative survival</td>
<td>Significant predictor</td>
<td>Univariate analysis: VO2peak 0.83 (0.74–0.94 [0.05]) p = 0.002* Multivariate analysis: VE/(\dot{V}CO_2): 1.13(1.07–1.19 [0.03]) p &lt; 0.001*</td>
<td>Postoperative survival</td>
<td>Significant predictor</td>
<td>Univariate analysis: AT: 0.74 (0.63–0.87 [0.06]) p &lt; 0.001* Multivariate analysis: AT: 0.84 (0.72–0.98 [0.07]) p = 0.033* Non-survivors more likely to have a VO2-AT &lt; 50% predicted (p = 0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epstein [36]</td>
<td>Postoperative survival</td>
<td>Significant predictor</td>
<td>Non-survivors more likely to have a VO2peak &lt; 60% predicted (p = 0.04). Non-survivors more likely to have a combined VO2peak &lt; 60% and a VO2-AT &lt; 50% predicted (p &lt; 0.01).</td>
<td>Postoperative survival</td>
<td>Significant predictor</td>
<td>Non-survivors more likely to have a VO2-AT &lt; 50% predicted (p = 0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forshaw [37]</td>
<td>Postoperative complications</td>
<td>Significant predictor</td>
<td>Significantly different VO2peak between patients who did or did not experience postoperative cardiopulmonary complications (p = 0.04)</td>
<td>Postoperative complications</td>
<td>Not a significant predictor</td>
<td>No significant difference between the AT of the complications and non-complications group (p = 0.07)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCullough [38]</td>
<td>Postoperative complications</td>
<td>Significant predictor</td>
<td>Complications occurred in 16.6% of patients with a peak VO2 &lt; 15.8 ml.kg(^{-1})min(^{-1}) and in only 2.8% of patients with a peak VO2 &gt; 15.8 ml.kg(^{-1})min(^{-1}) (univariate odds ratio 6.77 (95% CI 1.14–51.76) p = 0.002). VO2 peak as a continuous variable was a significant predictor of complications (odds ratio of 1.61 per 1 unit decrease (95% CI 1.19–2.18) p = 0.002.</td>
<td>Postoperative complications</td>
<td>Not documented/insufficient analysis</td>
<td>AT was not investigated comprehensively but showed a significant trend across the tertiles (p &lt; 0.001).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagamatsu [39]</td>
<td>Postoperative complications</td>
<td>Significant predictor</td>
<td>Significant difference between the VO2max.m(^{-2}) of the complications and no complications group (p &lt; 0.001).</td>
<td>Postoperative complications</td>
<td>Significant predictor</td>
<td>Significant difference between the AT.m(^{-2}) of the complications and no complications group (p &lt; 0.001). No significant difference between the AT.m(^{-2}) of patients who did or did not have cardiopulmonary complications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagamatsu [40]</td>
<td>Postoperative complications</td>
<td>Significant predictor</td>
<td>VO2max.m(^{-2}) was significantly lower in patients with cardiopulmonary complications (p &lt; 0.001).</td>
<td>Postoperative complications</td>
<td>Not a significant predictor</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nugent [41]</td>
<td>Postoperative complications</td>
<td>Not a significant predictor</td>
<td>No significant difference between the VO2peak of the complications group (18.6 ml.min(^{-1}).kg(^{-1})) and the VO2peak of the no complications group (21.8 ml.min(^{-1}).kg(^{-1}))</td>
<td>–</td>
<td>Not documented/insufficient analysis</td>
<td>–</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
patients requiring an increased level of postoperative care. VO2 max was not used in this study but if this was found to add greater discriminatory power it would further narrow down the high risk patient sub group and so minimise the unwarranted expense of patients unnecessarily receiving more intensive care.

The framework for utilising CPET data with other important predictors of peri-operative risk already exists and provides potential for individualised estimations [43]. However it only currently employs data on a subject’s MET (actual peak MET in comparison to their age predicted value). By further understanding the role of the other variables recorded in a CPET and obtaining procedure specific data it may be possible to increase its predictive capacity and improve risk prediction as a whole.

The impact of exercise training on VO2 peak in patients has been investigated by several studies. In one study [44] subjects underwent supervised endurance cycle ergometry at 60–100% VO2 peak 5 days a week for 4–6 weeks prior to their elective operation. Subjects who had an adherence of ≥ 80% increased their VO2 peak by 3.3 ml.kg⁻¹.min⁻¹ (95% CI 1.1–5.4) p = 0.006, representing an increase of 21.5%. This improvement in VO2 peak has also been shown using home exercise [45] for 10–12 weeks with weekly telephone monitoring (11% increase in VO2 peak, p < 0.05). This was less effective than the first study but as subjects followed a different regime it is not possible to allocate this difference to either the supervision or exercise program followed. Both these studies however demonstrate that in a relatively short pre-surgical period patients are able to significantly elevate their VO2 peak. It is yet to be investigated whether this improvement will translate into a lower complication rate from their baseline (pre-fitness regime). However if it did it could represent a powerful method of improving postoperative morbidity and mortality without significant delay to their pre-operative waiting period or adding substantial cost.

There is no firmly identified causal mechanism in the literature which would link a higher VO2 peak or AT with an improved risk profile. One possible explanation is that fitter patients are simply more able to mount the sustained increase in oxygen delivery induced by surgery without exceeding their aerobic physiological parameters. A more intriguing hypothesis is that regular exercise can induce a systemic effect similar to ischemic preconditioning which lessens the impact of any imbalance in oxygen delivery and demand associated with surgery, by increasing the body’s ability to extract oxygen and better tolerate ischemic conditions. There is both structural and biochemical evidence to support this second possible mechanism. Firstly endurance exercise in humans is known to increase mitochondrial mass [46], which may improve

<table>
<thead>
<tr>
<th>Author</th>
<th>VO2 peak</th>
<th>Anaerobic threshold</th>
<th>Anaerobic</th>
<th>Outcome measure</th>
<th>Study result</th>
<th>Summary findings</th>
<th>Study result</th>
<th>Summary findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Older [42]</td>
<td>–</td>
<td>–</td>
<td>Not documented/ insufficient analysis</td>
<td>Postoperative survival</td>
<td>Not documented/ insufficient analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Author</th>
<th>VO2 peak</th>
<th>Anaerobic threshold</th>
<th>Anaerobic</th>
<th>Outcome measure</th>
<th>Study result</th>
<th>Summary findings</th>
<th>Study result</th>
<th>Summary findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Older [5]</td>
<td>–</td>
<td>–</td>
<td>Not documented/ insufficient analysis</td>
<td>Postoperative survival</td>
<td>Not documented/ insufficient analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**This study formed patient groups based on AT, type of surgical procedure, ECG changes and VT/VO2. There was no statistical analysis in the paper of patient outcome based on AT alone though it is strongly suggestive of being a significant predictor.**
oxygen utilisation by mitochondria and thus delay the onset of anaerobic respiration and ischemia in the tissues. Secondly animal studies [47] have shown that exercise causes a number of physiological cardiac changes, in particular an increase in the antioxidant capacity and the number of sarcolemmal ATP-sensitive potassium channels. These are proposed to respectively reduce reactive oxygen species damage and cytosolic Ca2+ overload [47] thereby helping the tissues deal with ischemic conditions. These adaptations may make individuals better able to tolerate periods of anaerobic metabolism per-operatively.

To gain further evidence to support the use of CPET as a predictor of postoperative morbidity and mortality further studies need to be conducted. These studies need to focus on several key areas, particularly: (i) further investigation of the physiology of AT and the clinical relevance of this phenomenon, (ii) can other variables (O2 pulse, $\dot{V}_E/\dot{V}CO_2$ etc) independently or combined with $\dot{V}O_2_{max}/\text{peak}$ and AT provide a more powerful index, and if so would this be able to further narrow down the high risk patients without producing unacceptable type II error, as has been shown in other clinical applications [48]? (iii) Can the risk threshold of the variables be refined for different types of major surgery? (iv) Does increasing a subject’s $\dot{V}O_2_{max}$ or AT pre-operatively significantly improve their postoperative outcome? (v) What, if any, is the significance of the variables recorded during the recovery phase of the test?

Acknowledgements

The authors would like to thank Professor Brian Whipp for his comments on this manuscript. The authors are grateful for support from the NIHR Biomedical Research Centre Funding Scheme.

References


21 Gladden LB, Yates JW, Stremel RW, Stanford BA. Gas exchange and lactate anaerobic thresholds: inter- and


36 Epstein SK, Freeman RB, Khayat A, Unterborn JN, Pratt DS, Kaplan MM. Aerobic capacity is associated with 100-day outcome after hepatic transplantation. *Liver Transplantation* 2004; **10**: 418–24.


