Patients’ inability to perform a preoperative cardiopulmonary exercise test or demonstrate an anaerobic threshold is associated with inferior outcomes after major colorectal surgery

C. W. Lai1,2, G. Minto2,3, C. P. Challand1,2, K. B. Hosie1, J. R. Sneyd2,3, S. Creanor2 and R. A. Struthers2,3*

1 Department of Colorectal Surgery, Derriford Hospital, Plymouth, UK
2 Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK
3 Department of Anaesthesia, Derriford Hospital, Plymouth, UK
* Corresponding author: Department of Anaesthetic, Plymouth Hospitals NHS Trust, Derriford Hospital, Derriford Road, Derriford, Plymouth PL6 8DH, UK. E-mail: richardstruthers@nhs.net

Background. Surgical patients with poor functional capacity, determined by oxygen consumption at anaerobic threshold (AT) during cardiopulmonary exercise testing (CPET), experience longer hospital stays and worse short- and medium-term survival. However, previous studies excluded patients who were unable to perform a CPET or who failed to demonstrate an AT. We hypothesized that such patients are at risk of inferior outcomes after elective surgery.

Methods. All patients undergoing major colorectal surgery attempted CPET to assist in the planning of care. Patients were stratified by their test results into Fit (AT ≥ 11.0 ml O2 kg⁻¹ min⁻¹), Unfit (AT < 11.0 ml O2 kg⁻¹ min⁻¹), or Unable to CPET groups (failed to pedal or demonstrate an AT). For each group, we determined hospital stay and mortality.

Results. Between March 2009 and April 2010, 269 consecutive patients were screened, and proceeded to bowel resection. Median hospital stay was 8 days (IQR 5.1–13.4) and there were 44 deaths (16%) at 2 yr; 26 (9.7%) patients were categorized as Unable to CPET, 69 (25.7%) Unfit and 174 (64.7%) Fit. There were statistically significant differences between the three groups in hospital stay [median (IQR) 14.0 (10.5–23.8) vs 9.9 (5.5–15) vs 7.1 (4.9–10.8) days, P < 0.01] and mortality at 2 yr [11/26 (42%) vs 14/69 (20%) vs 19/174 (11%), respectively (P < 0.01)] although the differences between Unable and Unfit were not statistically different.

Conclusions. Patients’ inability to perform CPET is associated with inferior outcomes after major colorectal surgery. Future studies evaluating CPET in risk assessment for major surgery should report outcomes for this subgroup.

Keywords: exercise test; general surgery; length of stay; mortality; preoperative care; risk assessment

Accepted for publication: 13 April 2013

In the UK, elective surgical procedures have an overall in-hospital mortality of 0.4%. However, the figures conceal that a large proportion of the mortality is concentrated in a group of high-risk patients undergoing major surgery.1

Recent prominent reports have highlighted deficiencies in the identification of the high-risk surgical patient and appropriate planning of their perioperative care. Patients may be undergoing major procedures without having a full appreciation of the risk of mortality and severe complications that they face.2

The Royal College of Surgeons’ report, ‘Peri-operative Care of the Higher Risk General Surgical Patient’, recommends the use of scoring systems to estimate patients’ risk of mortality and guide escalation of care for those with an expected mortality > 5%.3

Major surgery triggers a strong systemic inflammatory response, which leads to an increase in oxygen demand. This is met by an increase in global oxygen delivery (DO2), through an increase in cardiac output (CO) and increase in tissue oxygen extraction. While most patients are able to mount an
Cardiopulmonary exercise testing (CPET) examines the ability of the cardio-respiratory system to deliver oxygen to tissues under stress and provides an objective determination of functional capacity and impairment. Anaerobic threshold (AT) is the point during exercise when anaerobic metabolism supplements aerobic with additional CO₂ production, creating an inflection point on a plot of pulmonary CO₂ efflux vs O₂ uptake. Lower AT indicates lower aerobic fitness and is associated with increased postoperative mortality and severe morbidity. However, CPET studies to date have not reported the perioperative outcomes of those unable to perform a test or in whom an AT was not determined. Patients with heart failure who do not demonstrate an AT are at higher risk of major cardiac events. This may be attributable to a short duration of tests or abnormal ventilatory patterns which attenuate the inflection point or alterations in cardiovascular physiology, which conceal AT in the early part of the test. We hypothesized that patients undergoing surgery who do not demonstrate an AT may similarly be at high risk of adverse outcomes.

Methods

All patients undergoing elective major colorectal surgery at our Colorectal Specialist Unit attempt preoperative CPET to assist in the planning of perioperative care. With Ethics Committee approval all patients undergoing bowel resection between March 2009 and April 2010 were screened for inclusion into a trial of intraoperative goal-directed fluid therapy and those with an AT > 8 ml O₂ kg⁻¹ min⁻¹ were approached to be enrolled and randomized. The testing methods and results of this study have been previously reported. CPET was performed on a stationary bicycle (Zan, nSpire, CO, USA) using a ramped exercise protocol with AT determined by V slope and ventilatory equivalents.

All patients tested during this period were asked for permission to audit their outcomes. Patients were stratified by their CPET results into Fit (AT ≥ 11.0 ml O₂ kg⁻¹ min⁻¹), Unfit (AT < 11.0 ml O₂ kg⁻¹ min⁻¹), or Unable to CPET if they failed to pedal the cycle or demonstrate an AT.

Patient characteristics and ‘end of surgery’ colorectal POSSUM (CR-POSSUM) predicted percentage mortality were recorded, along with type of surgery (laparoscopic or open), use of critical care, length of hospital stay, and 30-day-, 90-day- and 2-yr mortality. All patients were cared for on our Specialist Colorectal surgical unit.

Analysis

Data were collated into Microsoft Excel and statistical analysis performed using Minitab (Minitab® Statistical software v16, USA). Statistical significance was set at the 5% level. Data are presented as mean [standard deviation (SD)] where normally distributed and median (interquartile range) where not normally distributed. Between fitness group differences in total length of hospital stay were tested using one way analysis of variance (ANOVA) after log transformation of non-parametric data. Categorical data were compared between fitness groups using χ² tests. Where there was evidence of overall differences between the three fitness groups, Bonferroni-corrected follow-up multiple comparisons were used to identify between which pairs of groups had statistically significant differences. A survival plot, stratified by fitness groups, was performed using the Kaplan–Meier method with survival distributions compared using the log-rank test. Multivariable binary logistic regression was also undertaken to examine whether fitness group had a significant influence on the probability of death by 2-yr post-surgery, after allowing for other important variables (age, gender, and Dukes staging of malignancy).

Results

Between March 2009 and April 2010, 287 patients attended clinic and attempted CPET; 269 proceeded to surgery with follow-up, five formally declined to be followed up, and 13 did not proceed to surgery. One hundred and seventy-nine patients were recruited to the trial of intraoperative fluid management. All consenting survivors were followed to 2 yr. A summary of the results, stratified by CPET fitness group, is presented in Table 1.

A summary of the results, stratified by CPET fitness group, is presented in Table 1.

Thirteen patients did not proceed to bowel resection—five who would have been characterized as Fit, five Unfit and three as Unable to CPET. A summary of their results and reasons for not proceeding are presented as a web-based Supplementary Table S1.

For the cohort proceeding to surgery median hospital stay was 8 days (IQR 5.1–13.4) and there were 44 deaths at 2 yr (16%).

Twenty-six (9.7%) patients were categorized as Unable to CPET, 69 (25.7%) Unfit and 174 (64.7%) as Fit. Of the Unable patients 15 had a measured VO₂ peak but there was no AT demonstrable on V slope or ventilatory equivalents. The other 11 did not produce a test—seven because of arthritis, three because of frailty and one because of perineal pain.

The Unable patients were, on average, older than Unfit and Fit patients and typically had higher CR-POSSUM percentage predicted mortality although there was insufficient evidence to support a statistically significant difference between Fit and Unfit groups. The Unable group had higher mortality at each time point, and in particular at 2 yr (11/26 (42%) vs 14/69 (20%) vs 19/174 (11%), P < 0.01), although again there was insufficient evidence to support a statistically significant difference between the Unfit and Unable groups.

Kaplan–Meier survival curves are presented in Figure 1; there was evidence of significant differences in the survival distributions of the three groups (P < 0.01).

Binary logistic regression analysis was performed looking at main effects on probability of death at 2 yr after surgery. Results are presented in Table 2. Age at surgery was associated with an odds ratio of 1.09 per year (95% confidence interval (CI) 1.04–1.14) and CR-POSSUM Score at the end of surgery with an odds ratio of 1.17 (1.08–1.28) per percentage increase in predicted mortality. Male gender and Dukes C or D cancer staging (compared with non-malignant disease) were also
significantly associated with increased probability of mortality. After allowing for these variables, an inability to perform CPET/demonstrate an AT, compared with being in the Fit group, was associated with an increase in probability of mortality at 2 yr [odds ratio of 3.98 (95% CI 1.04–11.73)]. Interestingly, 20% of females who could not demonstrate an AT or perform CPET had died by 2 yr compared with 73% of males in the same risk group.

**Discussion**

Previous studies have investigated the relationship between aerobic fitness as determined by CPET and outcomes after major surgery. However, these have typically focused on patients who were both able to complete CPET and generate an AT. To our knowledge, no studies have reported on the outcomes of those who were unable to do so.

Our data demonstrate that such patients typically have a longer length of hospital stay and higher 30-day mortality rate and that a substantial proportion (42%) is no longer alive 2 yr after surgery. This has important implications for planning perioperative care and for the process of informed consent, including the decision whether or not to attempt curative resection.

Current consensus opinion is that patients should be explicitly provided with an estimation of the risk of perioperative death from the planned procedure and that this should be documented in the patient notes and on consent forms.\(^3\) For many procedures, data are available on unit-specific 30-day mortality rates for all comers; however, it is likely that the risk varies greatly for subgroups. More individualized risk assessment facilitates better clinical decision-making.

Information about medium and long term mortality after surgery conveys a more realistic picture of the burden of risk faced: mortality at 90 days is consistently double that at 30 days.\(^8\)\(^12\)\(^13\)

Scoring systems are potentially useful for individualized risk assessment. These indices are generally derived using logistic
regression in a large patient cohort, with subsequent validation of the score in a separate patient cohort. POSSUM and CR-POSSUM are specific examples which have been validated in a surgical population. These can assign patients into expected population risk categories although they are not necessarily indicative of risk for the individual. Interestingly, the Royal College of Surgeons’ report suggests that patients with a predicted operative mortality > 10% should be automatically triaged to receive postoperative critical care admission. The mean CR-POSSUM predicted mortalities for the Unable group in this trial exceeded 10%, whereas predicted (and observed) mortalities for those patients who could generate an AT were much lower.

Despite the high predicted mortality, only 10 of the 26 patients (38%) in the Unable group had planned postoperative critical care admission. This suggests that the attending clinicians did not appreciate the association of an inability to successfully complete a CPET with a poor clinical outcome. A primary intention of this paper is to highlight the association.

A higher level of planned postoperative critical care might have mitigated against mortality risk for this group. ‘Prehabilitation’, incorporating a program of physical exercise, smoking cessation, and optimization of medical conditions may improve outcome for some patients.

In contrast to risk scoring indices, which are based on a tally of patient characteristic data, co-morbidities and intraoperative events, CPET objectively quantifies the aerobic fitness of an individual. Thus, it is possible to identify patients who are at risk of perioperative functional heart failure, many of whom will not have previously had a formal diagnosis of cardiovascular or respiratory disease. We speculate that the inability to generate meaningful oxygen consumption measurements during the test is a marker of particularly poor functional capacity—the test becomes one not of cardio-respiratory reserve but of frailty.

The study has several limitations. The combined total of patients unable to pedal the cycle or not demonstrating an AT is small and prevents further modelling. Analysis of a larger database is advisable to confirm these findings. In particular, there was not a statistically significant difference between Unfit and Unable groups in length of stay or mortality when performing Bonferroni corrected comparisons—although being in the Unfit group was associated with statistically significant higher probability of mortality at 2 yr when multivariate analysis was performed.

Fitness is not a dichotomous measure and the use of a cutoff in AT of 11 ml O2 kg$^{-1}$ min$^{-1}$ is somewhat arbitrary. The Unfit patients are likely to be at the lower end of a continuum of fitness and as such at high risk of being unable to meet the metabolic demands of healing after surgery. As such the inability to demonstrate a statistically significant difference between outcomes in the Unfit and Unable patients might not be too surprising.

Surgical and anaesthetic techniques were not standardized, nor were access to planned critical care. Attending clinicians were not blinded to the results of any preoperative investigations. In the regression model, we have made statistical adjustment for differences in cancer staging between groups but in any small study such as this, co-morbidity imbalances and differences in care between groups may have an effect on comparative outcome.

We did not measure the incidence nor grade the severity of postoperative complications. It is likely that the longer length of hospital stay and increased mortality observed in the Unable group is linked to an increased adverse event rate but we cannot specifically comment on this. Similarly, we did not use quality of life indicators to gauge the longer term recovery from surgery. This is important in evaluating whether surgical resection represents the best treatment option for an individual. When a patient has limited life expectancy (because of frailty or co-morbidities) then it may not be in their best interest that they endure a substantial portion of their remaining days trying to regain an acceptable quality of life after a surgical procedure. Humanity and patient dignity is an important aspect in patient care highlighted in the first (N)ICEPOD report in 1987 and echoed in the most recent. We suggest that the benefits

### Table 2. Probability of death at 2 yr post-surgery, using binary logistic regression

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coefficient</th>
<th>se coefficient</th>
<th>Z</th>
<th>P-value</th>
<th>OR</th>
<th>95% CI for OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>−8.545</td>
<td>1.732</td>
<td>4.93</td>
<td>&lt;0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.086</td>
<td>0.024</td>
<td>3.65</td>
<td>&lt;0.01</td>
<td>1.09</td>
<td>1.04–1.14</td>
</tr>
<tr>
<td>Predicted CR-POSSUM % mortality</td>
<td>0.158</td>
<td>0.043</td>
<td>3.66</td>
<td>&lt;0.01</td>
<td>1.17</td>
<td>1.08–1.28</td>
</tr>
<tr>
<td>Fitness group (vs Fit)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfit</td>
<td>0.545</td>
<td>0.443</td>
<td>1.23</td>
<td>0.22</td>
<td>1.72</td>
<td>0.72–4.10</td>
</tr>
<tr>
<td>Unable</td>
<td>1.381</td>
<td>0.562</td>
<td>2.46</td>
<td>0.01</td>
<td>3.98</td>
<td>1.32–11.97</td>
</tr>
<tr>
<td>Female gender (vs male)</td>
<td>−0.837</td>
<td>0.413</td>
<td>−2.03</td>
<td>0.04</td>
<td>0.43</td>
<td>0.19–0.97</td>
</tr>
<tr>
<td>Malignancy stage (vs non-malignant)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dukes A</td>
<td>−0.541</td>
<td>0.913</td>
<td>−0.59</td>
<td>0.55</td>
<td>0.58</td>
<td>0.10–3.48</td>
</tr>
<tr>
<td>Dukes B</td>
<td>0.272</td>
<td>0.623</td>
<td>0.44</td>
<td>0.66</td>
<td>1.31</td>
<td>0.39–4.45</td>
</tr>
<tr>
<td>Dukes C and D</td>
<td>1.833</td>
<td>0.574</td>
<td>3.19</td>
<td>&lt;0.01</td>
<td>6.25</td>
<td>2.03–19.26</td>
</tr>
</tbody>
</table>
of operative over non-surgical options must be clear in order to justify proceeding to resection in this group of patients.

In summary, our results imply that poor preoperative functional capacity is associated with inferior postoperative outcomes. Failure to report the outcomes of those unable to perform the test may distort the clinical relevance of CPET studies.

The small proportion of patients who are unable to complete a cardiopulmonary exercise test have the longest lengths of hospital stay and higher early and medium-term mortality and as a consequence their care represents a substantial burden on healthcare resources. Further studies on a larger cohort are required to confirm our findings and to identify perioperative interventions to improve their outcomes.

**Authors’ contributions**

G.M., R.A.S., J.R.S., K.B.H. and C.P.C. conceived the idea for this audit. C.W.L., C.P.C. and R.A.S. performed data collection and C.W.L., R.A.S., G.M. and S.C. performed data analysis. C.W.L., R.A.S. and G.M. drafted the manuscript and all authors were actively involved in revising the manuscript.

**Supplementary material**

Supplementary material is available at British Journal of Anaesthesia online.

**Declaration of interest**

None declared.

**Funding**

This work was supported by an Anaesthesia/AABGI Departmental Research grant (Round 1, 2008) through the National Institute of Academic Anaesthesia. Service support and administrative costs were met by the National Institute of Healthcare Research.

**References**


*Handling editor: P. S. Myles*