Editor’s key points

- Clinical acumen and subjective assessment of a patient’s fitness for surgery have modest predictive utility.
- Cardiopulmonary exercise testing (CPET) evaluates the overall capacity of the cardiovascular and respiratory systems to work maximally.
- This study provides good evidence that CPET enhances risk stratification for patients undergoing major surgery.

Background. Postoperative complications are associated with reduced fitness. Cardiopulmonary exercise testing (CPET) has been used in risk stratification. We investigated the relationship between preoperative CPET and in-hospital morbidity in major colonic surgery.

Methods. We prospectively studied 198 patients undergoing major colonic surgery (excluding neoadjuvant cancer therapy), performing preoperative CPET (reported blind to clinical state), and recording morbidity (assessed blind to CPET), postoperative outcome, and length of stay.

Results. Of 198 patients, 62 were excluded: 11 had emergency surgery, 25 had no surgery, 23 had incomplete data, and three were unable to perform CPET. One hundred and thirty-six (89 males, 47 females) were available for analysis. The median age was 71 [inter-quartile range (IQR) 62–77] yr. Sixty-five patients (48%) had a complication at day 5 after operation. Measurements significantly lower in patients with complications than those without were \( \dot{V}_O_2 \) at estimated lactate threshold \( \hat{\lambda}_L \) \( [\text{median 9.9 (IQR 8.3–12.7)} \text{ vs 11.2 (9.5–14.2) ml kg}^{-1}\text{ min}^{-1}, P < 0.01] \), \( \dot{V}_O_2 \) at peak \( [15.2 (12.6–18.1) \text{ vs 17.2 (13.7–22.5) ml kg}^{-1}\text{ min}^{-1}, P = 0.01] \), and ventilatory equivalent for CO2 \( \dfrac{\dot{V}_E}{\dot{V}_CO_2} \) at \( \hat{\lambda}_L \) \( [31.3 (28.0–34.8) \text{ vs 33.9 (30.0–39.1), P < 0.01}] \). A final multivariable logistic regression model contained \( \dot{V}_O_2 \) at \( \hat{\lambda}_L \) [one-point change odds ratio (OR) 0.77 [95% confidence interval (CI) 0.66–0.89, \( P < 0.0005 \)], two-point change OR 0.61 (0.46–0.81) and gender [OR 4.42 (1.78–9.88), \( P = 0.001 \)], and was reasonably able to discriminate those with and without complications (AUC 0.71, CI 0.62–0.80, 68% sensitivity, 65% specificity).

Conclusions. CPET variables are associated with postoperative morbidity. A multivariable model with \( \dot{V}_O_2 \) at \( \hat{\lambda}_L \) and gender discriminates those with complications after colonic surgery.

Keywords: anaerobic threshold; cardiopulmonary exercise test; colorectal surgery; morbidity; postoperative complications

Accepted for publication: 10 September 2013
and enhances shared decision-making. Approaches to risk evaluation include clinical acumen, clinical prediction scores [e.g. ASA physical status (ASA-PS), Duke’s Activity Scores, POSSUM, CR-POSSUM], plasma biomarkers, measures of cardiac function, and shuttle walk tests, but their effectiveness in predicting complications is not well established.

Cardiopulmonary exercise testing (CPET), which has been used for risk stratification before thoracic and abdominal surgery, tests cardiorespiratory reserve (physical fitness) at rest and under the stress of maximal exercise (mimicking that of major surgery), and is the most objective and precise means of evaluating pre-surgical fitness.

This prospective, blinded, observational study tests the hypothesis that CPET variables are related to short-term in-hospital morbidity in patients undergoing major colonic surgery.

**Methods**

**Patients**

We included all patients aged > 18 yr considered for major colonic surgery (benign or malignant), except those with inflammatory bowel disease, patients undergoing neoadjuvant cancer therapy, or patients who were unable to perform CPET as part of their preoperative evaluation between February 2009 and December 2010. Patients were excluded on the basis of having no surgery performed or interim emergency surgery, lacking complete in-hospital morbidity data, or their inability to attain a definable lactate or anaerobic threshold (VO2 at \( \dot{a} \)). Discussions with Aintree University Hospitals NHS Foundation Trust and the North West Research Ethics Committee established that formal ethical approval was unnecessary, since CPET had been recently introduced as routine assessment in the hospital for major colorectal surgical patients, and results were not used by the multidisciplinary team (MDT) to alter clinical management as yet. We however adhered fully to Caldicott guidelines. All patients received an information sheet regarding CPET and written consent was obtained. No patient was refused surgery on the basis of information sheet regarding CPET and written consent was obtained. No patient was refused surgery on the basis of 

**Cardiopulmonary exercise testing**

CPET followed American Thoracic Society/American College of Chest Physicians recommendations. After resting spirometry (flow-volume loops), CPET on an electromagnetically braked cycle ergometer (Ergoline 2000) comprised 2 min rest, 2 min freewheel pedalling, ramped incremental pedalling until volitional termination, and 5 min recovery. Ventilation and gas exchange was measured using a metabolic cart [Geratherm Respiratory GmbH (Love Medical Ltd, Manchester, UK)]. Pulse, 12-lead ECG, arterial pressure, and pulse oximetry were monitored throughout. Ramp gradient was set to 10–25 W min⁻¹ based on a calculation using predicted freewheel oxygen uptake (\( \dot{V}O_2 \)), predicted \( \dot{V}O_2 \) at peak exercise, height, and age. No major adverse clinical events occurred during CPET.

**Measurements**

Patient characteristics recorded at CPET included age, gender, height, weight, diagnosis, staging (if malignancy), surgical procedure planned, WHO classification, and ASA-PS, and also diagnosis of diabetes, ischaemic heart disease, cerebrovascular disease, or heart failure. Resting flow-volume loops were used to derive forced expiratory volume over 1 s (FEV1) and forced vital capacity (FVC). Ventilation and gas exchange variables derived from CPET included \( \dot{V}O_2 \), ventilatory equivalents for oxygen and carbon dioxide (\( V_E/\dot{V}O_2 \), \( V_E/\dot{V}CO_2 \)), and oxygen pulse (\( \dot{V}O_2/\text{heart rate} \)), all measured at \( \dot{a} \) and at peak exercise. \( \dot{a} \) was estimated conventionally [breakpoint in the \( VCO_2 - \dot{V}O_2 \) relationship, with increases in \( V_E/\dot{V}O_2 \) and end-tidal (\( P_t \)) \( \dot{V}O_2 \) but no increase in \( V_E/\dot{V}CO_2 \) or decrease in \( P_t(\dot{V}CO_2) \). Peak \( \dot{V}O_2 \) was averaged over the last 30 s of exercise. CPETs were reported by two experienced assessors both blind to patient characteristics and outcome data.

Short-term surgical outcome was assessed as morbidity (by medical and nursing staff blind to any CPET data) using the nine domains listed in the Post-Operative Morbidity Survey on day 5, Clavien–Dindo Classification (highest grade for the most serious sustained in-hospital complication), and in-hospital mortality. Length of hospital stay (days) was recorded prospectively, and patients were followed for 30 days post-discharge for re-admission and mortality. The patients and the colorectal MDT (including anaesthetists) were blind to all CPET data. No perioperative management or decisions were influenced by CPET data.

The primary aim was to establish the relationship between postoperative complications (POMS present or absent on day 5) and \( \dot{V}O_2 \) at \( \dot{a} \); a secondary aim was to explore the multivariable relationship between CPET variables and other important prognostic variables with complications at day 5 after operation.

**Statistical methods**

Non-parametric receiver operator characteristic (ROC) curves were constructed for \( \dot{V}O_2 \) at \( \dot{a} \), \( \dot{V}O_2 \) at peak, \( O_2 \) pulse at \( \dot{a} \), and \( V_E/\dot{V}CO_2 \) at \( \dot{a} \) in order to assess their independence ability to discriminate between patients with and without day 5 morbidity. Optimal cut-points were obtained by minimizing the distance between points on the ROC curve and the upper left corner. Six variables (to satisfy the 10 events per variable rule) were identified as candidates for a multivariable logistic regression model: \( \dot{V}O_2 \) at \( \dot{a} \), \( \dot{V}O_2 \) at peak, gender, operation type (laparoscopic/open), and \( O_2 \) pulse at \( \dot{a} \) and \( V_E/\dot{V}CO_2 \) at \( \dot{a} \). A final multivariable model was obtained using forward stepwise selection [minimizing Akaike information criteria (AIC)]. Its sensitivity to variable exclusion and re-inclusion was also assessed using AIC. Model fit was assessed using the Hosmer–Lemeshow goodness-of-fit test. In order to explore the univariate relationship between CPET and length of stay, continuous CPET variables were dichotomized at the optimal cut-point for the ROC curve and the Kaplan–Meier curves were constructed. The log-rank test was used to compare survival curves; patients who died before discharge \( (n=2) \) were treated as right-censored.
Patients who left the study before day 5 (n = 14) were excluded from the analysis of length of stay. All analyses were conducted using Stata (StataCorp., 2011, Stata Statistical Software: Release 12. College Station, TX: StataCorp LP). Continuous variables are reported as mean and standard deviation (SD) or median and inter-quartile range (IQR) depending on the distribution. Categorical variables are presented as frequency (%). P-values in Tables 1 and 2 were obtained using univariate logistic regression, χ² tests, or Fisher’s exact tests where cell counts were insufficient. All percentages represent proportions within complications (yes/no classifications), except ‘n’ which refers to the proportion of patients with and without complications.

### Results

One hundred and ninety-eight patients consented; of whom, 25 had no surgery (15 due to patient choice, 10 due to irresectable metastasis), 11 needed an emergency procedure, and three were unable to perform a CPET. Of the 159 who had adequate CPET and underwent major elective surgery, 23 lacked complete data. Of the remaining 136 (89 males, 46 females), 41 had right and nine left hemicolectomy, 46 anterior resection, one subtotal colectomy, 13 abdomino-perineal resections, eight Hartman’s procedure, and 19 other major colonic resections. One patient developed a supraventricular tachycardia at peak exercise. The patient’s ECG was discussed at MDT and was subsequently referred to a cardiologist. Surgery on this patient proceeded as normal. Table 1 shows patients grouped by occurrence or not of in-hospital postoperative complications: these groups differed significantly in gender, age, and preoperative heart failure, but not in operation type, surgery, or presence of anastomosis/stoma. Table 2 shows grouped CPET data: patients with a complication had significantly lower VO₂ at $\dot{V}O_2$, VO₂ at peak, and higher $\dot{V}E$/ $\dot{V}CO_2$ at $\dot{V}O_2$. Three patients unable to attain $\dot{V}O_2$ sustained a complication and their discharge was delayed; these were excluded from analysis.
Sixty-five patients (48%) sustained a complication at day 5; of whom, two died in hospital (1.5% mortality) from myocardial infarction (at days 3 and 5) and eight (6.5%) suffered anastomotic leak at a median of 6 days (four anterior resection, three right, and one left hemicolectomy): of these, five were re-operated, three treated conservatively with radiologically inserted drains and i.v. antibiotics. A further two patients were re-operated at median 5 days (one patient suffered intestinal obstruction and another a necrotic stoma). All these suffered further complications and delayed hospital discharge.

Independently, VO₂ at $\dot{\text{V}}_\text{E}/\dot{\text{V}}$, VO₂ at peak, and $\dot{\text{V}}_\text{E}/\dot{\text{V}}$CO₂ were associated with day 5 morbidity ($P < 0.05$), whereas O₂ pulse at $\dot{\text{V}}$ and ASA were not ($P = 0.22$ and 0.11, respectively). For VO₂ at $\dot{\text{V}}$, [area under the curve (AUC) 0.63, confidence interval (CI) 0.54–0.73], the optimal cut-point was 10.1 ml kg⁻¹ min⁻¹, giving 68% sensitivity and 58% specificity. For VO₂ at peak (AUC 0.63, CI 0.53–0.73), the cut-point was 16.7 ml kg⁻¹ min⁻¹, giving 55% sensitivity and 69% specificity.

Only VO₂ at $\dot{\text{V}}$ and gender were retained in the final multivariable logistic regression model. In this model, the odds of complications are higher for a male than a female with the same VO₂ at $\dot{\text{V}}$ [odds ratio (OR) 4.19, CI 1.78–9.88, $P < 0.001$]; a 1.0 ml kg⁻¹ min⁻¹ increase in VO₂ at $\dot{\text{V}}$ is associated with ~20% reduction in the odds of complications (OR 0.77, CI...
0.66–0.89, P<0.0005) and a 2.0 ml kg⁻¹ min⁻¹ increase with ≏40% reduction (OR 0.60, CI 0.45–0.80, P=0.001), after adjustment for sex. The ability of this model to discriminate between patients with and without a complication was reasonable (AUC 0.71, CI 0.62–0.80, 68% sensitivity and 65% specificity at the optimal cut-point; positive predictive value = 62%, negative predictive value = 69%) (Fig. 2).

There is evidence to suggest that, independent of other predictive variables, patients with \( \dot{V}O_2 \) at \( \hat{u} \), or \( \dot{V}O_2 \) at peak (P = 0.003) (Fig. 3) or \( V_{E}/V_{CO_2} \) at \( \hat{u} \) (P < 0.0001) above the cut-point have a significantly reduced length of hospital stay.

## Discussion

### Main findings and comparison with other studies

This prospective, blinded, observational study provides novel evidence supporting CPET as an objective risk assessment tool before major colonic surgery. In this cohort, \( \dot{V}O_2 \) at \( \hat{u} \) and peak were significantly lower and \( V_{E}/V_{CO_2} \) at \( \hat{u} \) significantly higher, in patients encountering POMS-defined complication at day 5 after operation, and single-variable analysis confirms these associations, albeit with only moderate sensitivity and specificity. The poor predictive performance of CPET variables in our study when assessed independently is consistent with the literature and reflects the complex interactions between baseline physiology and elective surgical trauma on postoperative outcomes. However, multivariable analysis showed that \( \dot{V}O_2 \) at \( \hat{u} \) and gender were independent predictors of complications after surgery with moderate discrimination between patients with, and without, complications. ASA was not independently related to outcome and inclusion of this variable in the multivariable model had a negligible effect.

This study adds to the literature supporting objective measures of physical fitness for risk assessment in major abdominal surgery. The findings by Older and colleagues in 187 elderly patients undergoing major intra-abdominal surgery that preoperative \( \dot{V}O_2 \) ≤ 11.0 ml kg⁻¹ min⁻¹ was associated with increased cardiovascular mortality established CPET as a tool for preoperative risk assessment and stratification. In a later study and a review, Older investigated triaging: if a patient had \( \dot{V}O_2 \) < 11 ml kg⁻¹ min⁻¹, they were assigned to ICU before operation. Assessing 843 patients > 55 yr undergoing major colorectal surgery, radial nephrectomies, and cystectomies, Wilson and colleagues concluded that \( \dot{V}O_2 \) ≤ 10.9 ml kg⁻¹ min⁻¹ and \( V_{E}/V_{CO_2} \) at \( \hat{u} \) ≥ 34 had 88% sensitivity and 47% specificity for hospital mortality. Snowden and colleagues evaluated CPET in preoperative risk assessment in elderly (mean age 70 yr) patients undergoing major intra-abdominal surgery, finding that the \( \dot{V}O_2 \) optimal cut-point of 10.1 ml kg⁻¹ min⁻¹ gave 88% sensitivity and 79% specificity for discriminating postoperative complications (AUC 0.85; CI 0.78–0.91; P<0.001). In 32 patients...
undergoing major intra-abdominal surgery, Hightower and colleagues’ found that \( \hat{\phi}_1 < 75\% \) of predicted value predicted complications [AUC 0.72 (CI 0.57–0.87); sensitivity 88%; specificity 56%, \( P=0.016 \)]. A particular strength of our study, and those of Snowden and Hightower, is that clinicians were blinded to CPET results, which eliminates ‘confounding by indication’.

Junejo and colleagues’ evaluated preoperative CPET in predicting outcome after major hepatic resection in 108 patients: \( \hat{\phi}_1 < 9.9 \text{ ml kg}^{-1} \text{ min}^{-1} \) was 100% sensitive and 76% specific for in-hospital mortality, age and \( V_r/\dot{V}CO_2 \) at \( \hat{\phi}_1 \) (84% specificity and 47% sensitivity) were related to postoperative complications, and long-term survival with \( \hat{\phi}_1 < 9.9 \text{ ml kg}^{-1} \text{ min}^{-1} \) was significantly worse (hazard ratio 1.81, CI 1.04–3.17); however, only eight deaths out of 94 patients were recorded. Otto and colleagues’ retrospectively studied aerobic exercise capacity in inflammatory bowel disease patients, finding that adjusted \( \hat{\phi}_1 \) in Crohn’s disease was lower than in colorectal cancer [mean (SD): 11.4 (3.4) vs 13.2 (3.5) ml kg\(^{-1} \) min\(^{-1} \)]. This justifies our exclusion of inflammatory bowel disease patients as pathophysiologically distinct.

Our data further support CPET in perioperative risk assessment. Our best prognostic markers of postoperative complications were \( V_r/\dot{V}CO_2 \) at \( \hat{\phi}_1 \), with a cut-off at 32.9, \( \dot{V}O_2 \) at \( \hat{\phi}_1 \), with a cut-off at 10.1 ml kg\(^{-1} \) min\(^{-1} \), and \( \dot{V}O_2 \) at peak with a cut-off at 16.7 ml kg\(^{-1} \) min\(^{-1} \). These are similar to cut-off points found by Older and colleagues. Snowden and colleagues, Junejo and colleagues, although the sensitivity and specificity of individual variables were moderate when compared with other studies. However, our multivariable logistic regression model identifies gender and \( \hat{\phi}_1 \) as important predictors of day 5 complications, albeit with moderate AUC, specificity, and sensitivity. Interestingly, in upper gastrointestinal cancer surgery, bariatric surgery, and liver transplantation, surgery, the association with outcome and the cut-off point for \( \dot{V}O_2 \) at \( \hat{\phi}_1 \) is different. However, a recently published series by Colson and colleagues concludes that single-variable primary endpoints commonly derived from CPET (such as \( \hat{\phi}_1 \)) were not associated with 5 yr survival, although in a multivariable analysis, many CPET variables were important predictors.

Strengths and weaknesses

Strengths of this study include the homogeneous nature of the study population (only colonic surgical patients were included), rectal cancer patients who were undergoing cancer therapies were excluded, the blinded reporting of objectively measured CPET variables, the prospective nature of the study; the blinding to CPET results of caring clinicians and outcome data collectors, and the use of POMS (which has been validated in the UK) as a primary outcome measure and is currently used in National Institute for Health Research and Medical Research Council funded studies.

Potential weaknesses include the single-centre design which limits generalizability to other centres, as the ROC curve cut-off points are optimized for this local cohort and future validation work in similar cohorts should be performed.

Conclusion and further research

Our results indicate that when using CPET in patients awaiting colorectal surgery, clinicians should consider using \( \hat{\phi}_1 \) (ROC cut-off 10.1 ml kg\(^{-1} \) min\(^{-1} \)) and gender as a simple risk prediction tool. Of course, decisions regarding patient care or fitness for surgery should be made using the overall clinical and CPET picture. The identification of \( \hat{\phi}_1 \) and gender as a predictor for short-term outcome in colorectal surgery is novel; however, confirmation of these results in a larger colonic surgical cohort is encouraged to establish whether several preoperative risk assessment tools can be combined to predict risk more effectively.

Furthermore, we suggest that improving physical fitness, for example, \( \dot{V}O_2 \) at \( \hat{\phi}_1 \) and peak might improve surgical outcome in this population; to test this, we are currently investigating the role of a structured preoperative exercise training programme (prehabilitation) in this patient group (NCT01325909).

Authors’ contributions

M.A.W.: conception, study design, data acquisition, analysis, drafting article, revision, and final approval. D.L.: analysis and interpretation of data, drafting article, revising for intellectual content, and final approval. C.P.B.: study design, data acquisition, analysis, drafting article, revision, and final approval. L.N.: study design, data acquisition, analysis, drafting article, revision, and final approval. S.J.: conception, study design, revision, and final approval. G.J.K.: analysis and interpretation of data; critical revision of manuscript, and final approval. M.P.W.G.: conception, study design, revision, and final approval.

Acknowledgement

We would like to thank Aintree University Hospitals NHS Foundation Trust, Liverpool, Colorectal Multi-Disciplinary Team.

Declaration of interest

M.P.W.G. received honoraria for speaking for and/or travel expenses from: Edwards Lifescience, Fresenius-Kabi, BOC Medical (Linde Group), Ely-Lilly Critical Care, and Cortex GMBH. He has also received research grants from: National Institute of Health Research, Association of Anaesthetists of Great Britain and Ireland, Sir Halley Stuart Trust, Francis and Augustus Newman Foundation. He leads the Xtreme-Everest hypoxia research consortium, who have received unrestricted research grant funding from: BOC Medical (alinde Group) Ely-Lilly Critical Care, Smiths Medical, Deltex Medical, London Clinic, Rolex.

Funding

None.

References


2 Burns EM, Bottle A, Aylin P, Darzi A, Nicholls RJ, Foiz O. Variation in reoperation after colorectal surgery in England as an indicator of


11 Halm EA, Browner WS, Tubau JF, Tateo IM. Echocardiography for cated risk-adjustment scoring system for colorectal surgery (colo-


16 Schouten O, Box JJ, Poldermans D. Assessment of cardiac risk before non-cardiac general surgery. Heart 2006; 92: 1866–72


22 Smith TB, Stonell C, Purkayastha S, Paraskevas P. Cardiopulmonary exercise testing as a risk assessment method in non-cardio-


32 Older P, Smith R, Hone R. Preoperative evaluation of cardiac failure and ischemia in elderly patients by cardiopulmonary exercise testing. Chest 1993; 104: 701–4


39 Prentis JM, Manas D, Tre nell M. Submaximal cardiopulmonary exercise testing predicts 90-day survival after liver transplantation. Transplantation 2012; 18: 152–9


Handling editor: P. S. Myles