**Validation of preoperative cardiopulmonary exercise testing-derived variables to predict in-hospital morbidity after major colorectal surgery**

M. A. West1,2,3,4, R. Asher5, M. Browning6, G. Minto7,8, M. Swart9, K. Richardson10,11, L. McGarrity12, S Jack2,3,4 and M. P. W. Grocott2,3,4 for thePerioperative Exercise Testing and Training Society

1Academic Unit of Cancer Sciences, Faculty of Medicine, University of Southampton, and 2Critical Care Research Area, National Institute for Health Research Respiratory Biomedical Research Unit, 3Integrative Physiology and Critical Illness Group, Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, and 4Anaesthesia and Critical Care Research Unit, University Hospital Southampton NHS Foundation Trust, Southampton, 5Cancer Research UK Liverpool Cancer Trials Unit, Liverpool,  6Department of Anaesthesia, Maidstone and Tunbridge Wells NHS Trust, Maidstone, 7Directorate of Anaesthesia, Derriford Hospital and 8Plymouth University, Peninsula Schools of Medicine and Dentistry, Plymouth, 9Department of Anaesthesia and Critical Care Medicine, Torbay Hospital, Torquay, 10STRAPH Research Group, School of Sport and Exercise Sciences, University of Kent, Canterbury, 11Anaesthesia and Intensive Care Medicine, Medway Maritime Hospital, Gillingham, and 12Department of Anaesthesia, University Hospital Crosshouse, Kilmarnock, UK

*Correspondence to:*Mr M. A West, University of Southampton, Academic Unit of Cancer Sciences, South Academic Block, Tremona Road, Southampton SO16 6YD, UK (e-mail: m.west@soton.ac.uk)

Presented to a conference of the Association of Surgeons of Great Britain and Ireland, Harrogate, UK, May 2014

**Background:** In single-centre studies, postoperative complications are associated with reduced fitness. This study explored the relationship between cardiorespiratory fitness variables derived by cardiopulmonary exercise testing (CPET) and in-hospital morbidity after major elective colorectal surgery.

**Methods:** Patients underwent preoperative CPET with recording of in-hospital morbidity. Receiver operating characteristic (ROC) curves and logistic regression were used to assess the relationship between CPET variables and postoperative morbidity.

**Results:** Seven hundred and three patients from six centres in the UK were available for analysis (428 men, 275 women). ROC curve analysis of oxygen uptake at estimated lactate threshold (O2 at L) and at peak exercise (O2peak) gave an area under the ROC curve (AUROC) of 0.79 (95 per cent c.i. 0.76 to 0.83; *P* < 0.001; cut-off 11.1 ml per kg per min) and 0.77 (0.72 to 0.82; *P* < 0.001; cut-off 18.2 ml per kgpermin) respectively, indicating that they can identify patients at risk of postoperative morbidity. In a multivariable logistic regression model, selected CPET variables and body mass index (BMI) were associated significantly with increased odds of in-hospital morbidity (O2 at L 11.1 ml per kg permin or less: odds ratio (OR) 7.56, 95 per cent c.i. 4.44 to 12.86, *P* < 0.001; O2peak 18.2 ml per kg per min or less: OR 2.15, 1.01 to 4.57, *P =* 0.047; ventilatory equivalents for carbon dioxide at estimated lactate threshold (E/CO2 at L)more than 30.9: OR 1.38, 1.00 to 1.89, *P =* 0.047); BMI exceeding 27 kg/m2: OR 1.05, 1.03 to 1.03, *P* < 0.001). A laparoscopic procedure was associated with a decreased odds of complications (OR 0.30, 0.02 to 0.44; *P =* 0.033). This model was able to discriminate between patients with, and without in-hospital morbidity (AUROC 0.83, 95 per cent c.i. 0.79 to 0.87). No adverse clinical events occurred during CPET across the six centres.

**Conclusion:** These data provide further evidence that variables derived from preoperative CPET can be used assess risk before elective colorectal surgery.

**+A: Introduction**

Major colorectal surgery carries substantial morbidity1 and mortality, particularly in elderly patients and those with co-morbidities2. The 2014 UK colorectal cancer audit2 reported an overall 30-day mortality rate of 2.9 per cent for elective colorectal cancer surgery and a 90-day mortality rate of 3.2 per cent for major rectal cancer surgery. Outcome after major surgery depends on fixed variables (age and sex) as well as potentially modifiable factors, such as perioperative medical care and physiological tolerance of surgical trauma. Accurate risk stratification permits more effective collaborative decision-making, optimization of perioperative management and efficient use of hospital resources. Current approaches to risk prediction include clinical acumen, and use of risk prediction scores (for example American Society of Anesthesiologists physical score, Duke’s Activity Scores, POSSUM (CR-POSSUM in colorectal surgery)3,4, plasma biomarkers5, measures of cardiac function6 and shuttle walk tests7. Their effectiveness in predicting surgical morbidity remains poorly defined, especially in patients undergoing major surgery7,8.

Cardiopulmonary exercise testing (CPET) is now performed widely before surgery9,10 and is considered to be the most objective and precise means of evaluating presurgical physical fitness11–13. CPET evaluates physical fitness under stress, mimicking major surgery. This allows interrogation of the causes of exercise intolerance in individuals whose exercise capacity is reduced. CPET has also been shown to stratify risk accurately before major thoracic and abdominal surgery14–16. Single-centre studies have demonstrated a statistically significant association between CPET variables and in-hospital morbidity following major colonic and rectal cancer surgery17,18. The aim of this study was to confirm (or refute) the predictive value of selected CPET variables (identified in previous publications17,18) and their association with in-hospital morbidity in patients scheduled for major colorectal surgery in a multicentre setting.

**+A: Methods**

Six National Health Service (NHS) Trusts across the UK (Aintree University Hospitals NHS Foundation Trust, Liverpool; University Hospital Crosshouse – NHS Ayrshire and Arran, Kilmarnock; Medway NHS Foundation Trust, Gillingham; South Devon Healthcare NHS Foundation Trust, Torbay; Maidstone and Tunbridge Wells NHS Trust, Maidstone; Plymouth Hospitals NHS Trust, Plymouth) recruited consecutive adult patients between February 2011 and April 2014. Patients were referred for CPET by colorectal multidisciplinary teams (MDTs) and were assessed for suitability for inclusion. The predefined inclusion criterion was patients listed to undergo major elective colorectal surgery. Predefined exclusion criteria were: mechanical inability to perform CPET owing to lower limb dysfunction, inability to give informed consent for CPET, patients undergoing neoadjuvant cancer therapies, patients undergoing emergency surgery, patients diagnosed with distant metastasis and patients diagnosed with inflammatory bowel disease. Use of data for this validation study was approved by the NRES Committee East Midlands – Northampton (14/EM/1038) and the study was registered with clinicaltrials.gov (NCT02298907). All patients received an information sheet regarding CPET and written consent was obtained. No patient was refused surgery on the basis of gas exchange measurements, although any electrocardiographic abnormalities or oxygen transport variable patterns suggestive of clinically relevant myocardial ischaemia were raised at the colorectal MDT meetings and referred appropriately.

## **+B: Cardiopulmonary exercise testing**

CPET at all centres followed American Thoracic Society/American College of Chest Physicians recommendations19, previously reported from Torbay15, Plymouth20 and Aintree18. CPET was conducted on an electromagnetically braked cycle ergometer, and comprised 2–3 min resting (to allow gas exchange variables to stabilize), 3 min freewheel pedalling, then a ramped incremental protocol until volitional termination, and between 2–5 min recovery. Resting spirometry (flow–volume loops) was carried out before the CPET resting phase. Ventilation and gas exchange was measured using a metabolic cart (Aintree: Geratherm Respiratory ™, LoveMedical, Manchester, UK; Crosshouse: Carefusion™, San Diego, USA; Medway: Cortex™, Leipzig, Germany; Maidstone: Geratherm Respiratory™, Bad Kissingen, Germany; Plymouth: Zan nSpire™; Hertford, UK; Torbay, MedGraphics™, Gloucester, UK). Heart rate, full disclosure 12-lead electrocardiogram, blood pressure and pulse oximetry were monitored throughout. The ramp gradient was set to 10–25 W/min based on a calculation using predicted freewheel oxygen uptake (O2), predicted oxygen uptake at peak exercise (O2peak), height and age21.

## **+B: Patient characteristics and outcome measures**

Patient characteristics recorded at the initial CPET session included: age, sex, height, weight, and benign or malignant diagnosis (TNM stage if malignant). Resting flow–volume loops were used to derive forced expiratory volume over 1 s and forced vital capacity. Ventilation and gas exchange variables derived from CPET included O2, ventilatory equivalents for oxygen and carbon dioxide (E/O2 andE/CO2 respectively) and oxygen pulse (O2/heart rate), all measured at estimated lactate threshold (L) and at peak exercise21. TheO2 at L (anaerobic threshold, also sometimes referred to as the lactate threshold, ventilatory threshold, gas exchange threshold or lactic acidosis threshold) characterizes the upper limit of exercise intensity that can be accomplished almost wholly aerobically. The term O2 at L was used here. L was estimated conventionally (breakpoint in the CO2–O2 relationship22, with increases in E/ O2 and partial pressure of end-tidal oxygen but no increase in E/CO2 or decrease in partial pressure of end-tidal carbon dioxide23. O2peak was averaged over the last 30 s of exercise. CPET results were reported by experienced assessors, all with more than 5 years’ experience.

In-hospital surgical morbidity was recorded prospectively at day 5 after surgery by medical and nursing staff using the nine domains listed in the PostOperative Morbidity Survey (POMS)24; the Dindo classification25 (highest grade for the most serious sustained morbidity during the whole in-hospital stay) and 30-day mortality were also recorded. Postoperative in-hospital morbidity was defined as a POMS score of at least 1 or Dindo grade of I or higher. Patients who had been discharged by postoperative day 5 were presumed to have a POMS score of 0. Duration of hospital stay was also recorded and all patients were followed up to 1 year for mortality.

The primary variables of interest were O2 at L (ml per kg per min) and O2peak (ml per kg per min). Exploratory variables included O2 pulse at L (ml/beat) and E/CO2 at L. The primary aim was to establish the relationship between postoperative morbidity (present or absent when in hospital, assessed by POMS or Dindo classification) and O2 at L and O2peak. A secondary aim was to explore the multivariable relationship between patient demographics, CPET variables and postoperative in-hospital morbidity.

## **+B: Statistical analysis**

A sample size calculation was based on number needed to evaluate the area under the receiver operating characteristic (AUROC) curve. For a hypothesized AUROC of more than 0.63, based on a previous study18, 425 patients with complete data for both O2 at L and O2peak would be required to demonstrate that these variables are better than chance at discriminating between patients with, and without postoperative in-hospital morbidity. This estimate was based on a 5-day morbidity incidence of 48 per cent, 90 per cent power and a two-tailed 5 per cent significance level.

Continuous data are presented as median (i.q.r.) Univariable logistic regression with robust standard errors to take into account the clustered (centres) nature of the data was used to investigate the association between baseline demographics and postoperative complications.

ROC curves were constructed for O2 at L, O2peak, O2 at L and E/CO2 at L to assess their independent ability to discriminate between patients with, and without postoperative complications, taking into account the centre clustering. Optimal cut-off points were found by minimizing the distance between points on the ROC curve and the upper left corner. A variable was considered able to discriminate between patients with, and without postoperative complications if the AUROC and its 95 per cent c.i. were both greater than 0.7.

A multivariable logistic regression model with robust standard errors, taking into account the centre clustering, was fitted to the data and a stepwise selection procedure using Akaike’s information criteria was implemented. Variables with *P* < 0.250 in the univariable analysis were used as candidates for the final model (O2 at L, O2peak and E/CO2 at L, age at operation, body mass index (BMI) and operation type). O2 pulse at L was not included owing to a large proportion of missing data. The remaining CPET variables were dichotomized at their optimal cut-off point to improve model fit. The ability of the final model to discriminate between patients with, and without postoperative complications was investigated using ROC analysis, namely AUROC.

To explore the univariable association of CPET variables with duration of hospital stay, CPET variables were dichotomized at their optimal cut-off point and Kaplan–Meier curves were constructed. The log rank test was used to compare survival curves. DeLong’s test was used to investigate any differences in the discriminatory ability of each CPET between centres. All analyses were conducted using Stata release 12 (StataCorp LP, College Station, Texas, USA).

**+A: Results**

Seven hundred and three consecutive patients were recruited (428 men). All patients underwent CPET followed by major elective surgery and had complete outcome data on the primary outcome. Three of the six centres (Aintree, Torbay and Plymouth) recruited more than 85 per cent of the patients (*Table 1*). None of patients recruited from Aintree, Crosshouse, Medway and Maidstone were included in any other trial, and their data have never been published. Some patients recruited from Plymouth were part of the MIDAS (Microvascular Imaging During Abdominal Surgery) study (UKCRN ID 10093; ISRCTN21597243 – in press). Data for some patients recruited from Torbay had already been published in another case-controlled study26.

**+B: Safety**

No major adverse clinical events occurred during CPET. Fifteen patients (reported from 3 of 6 centres) developed supraventricular tachycardia at peak exercise, which resolved spontaneously during recovery; after review by a cardiologist, surgery proceeded as normal. Another patient developed CPET signs suggestive of severe myocardial ischaemia during early exercise. He was subsequently diagnosed with flow-limiting left mainstem coronary artery stenosis and underwent coronary revascularization before returning for low anterior resection 2 months later.

**+B: Postoperative morbidity**

A total of 258 patients (36.7 per cent) sustained in-hospital complications. Twenty-eight patients (4.0 per cent) had an anastomotic leak, and required relaparotomy, radiological-inserted drainage or conservative treatment with intravenous antibiotics. All of these patients developed further morbidity and their discharge from hospital was delayed. Increased BMI, age as well as method of operation (open *versus* laparoscopic) were associated with increased odds of in-hospital morbidity (*Table 1*). Analysis of grouped CPET data showed that lower O2 at L,O2peak and O2 pulse at L were associated with increased odds of in-hospital morbidity (*Table 2*).

There was a significant difference in the distribution of Dindo morbidity grade when the patient cohort was dichotomized at the optimal cut-off point for O2 at L; patients with O2 at L below or equal to 11.1 ml per kg per min had higher morbidity grades (I–V) (*Table 3*).

The median number of in-hospital morbidity events was 1 (0–2). Postoperative POMS-defined morbidity at day 5 dichotomized at the optimal cut-off for O2 at L also differed significantly between groups, except for neurological and haematological morbidities (*Table S1*, supporting information).

**+B: Duration of hospital stay**

The overall median duration of hospital stay was 7 (4–11) days. Patients with no POMS-defined morbidity had a median stay of 6 (4–7) days compared with 14 (9–19) days in patients with POMS-defined morbidity (*P* < 0.001) (*Fig. S1*, supporting information). One hundred and twenty-one patients were discharged within 5 days and were assumed to have a POMS score of 0. Kaplan–Meier analysis was conducted for each CPET variable (*Table S2* and *Fig. S2*, supporting information). Sixteen patients in whom no clear L could be identified sustained morbidity and their discharge was delayed.

**+B: Mortality**

Twelve patients (1.7 per cent) died within 30 days of surgery and 36 (5.1 per cent) had died by 1 year. All patients who were dead 1 year after surgery had a postoperative complication. The median O2 at L in this group was 10.1 (8.6–10.6) ml per kg per min.

**+B: Cardiopulmonary exercise testing variables and outcome prediction**

To investigate any differences between centres with regards to the discriminatory ability of each CPET variable, AUROCs were compared using DeLong’s test. The AUROCs differed significantly between centres forO2 at L,O2peak and O2 pulse at L (*P* < 0.001) but not for E/CO2 at L (*P =* 0.766) (*Table S3* and *Fig. S3*, supporting information).

The data from all centres were combined to find the optimal cut-off point for each variable. Both O2 at L and O2peak were able to discriminate between patients with, and without postoperative complications. For O2 at L (AUROC 0.79, 95 per cent c.i. 0.76 to 0.83) the optimal cut-off point was 11.1 ml per kgpermin, giving 78.2 per cent sensitivity and 71.4 per cent specificity (*Fig. 1a*). For O2peak (AUROC 0.77, 0.72 to 0.82) the optimal cut-off point was 18.2 ml per kgpermin, giving 70.3 per cent sensitivity and 72.0 per cent specificity (*Fig. 1b*). O2 pulse at L and E/CO2 at L did not discriminate between patients with, and without a postoperative complication (AUROC and/or 95 per cent c.i. 0.7 or less) (*Table 4*).

A total of 462 patients had complete observations and were included in the multivariable analysis; of these, 170 (36.8 per cent) had postoperative complications. Six variables were identified as candidates for a multivariable logistic regression model, all of which were retained in the final model (*Table 5*). In this model, the CPET variables, BMI and operation type were significantly associated with postoperative complications, with a O2 at L of 11.1 ml per kgpermin or less (odds ratio (OR) 7.56, 95 per cent c.i. 4.44 to 12.86; *P* < 0.001), O2peak of 18.2 ml per kgpermin or less (OR 2.15, 1.01 to 4.57; *P =* 0.047), E/CO2 at L over 30.9 (OR 1.38, 1.00 to 1.89; *P =* 0.047) and BMI exceeding 27 kg/m2 (OR 1.05, 1.03 to 1.08; *P* < 0.001) being associated with increased odds of complications, holding all other variables constant. A laparoscopic procedure was associated with a decreased odds of complications in comparison with an open procedure (OR 0.30, 0.02 to 0.44; *P =* 0.033) holding all other variables constant. The ability of this model to discriminate between patients with, and without complications was good (AUROC 0.83, 95 per cent c.i. 0.79 to 0.87; sensitivity 86.2 per cent, specificity 67.4 per cent, PPV 82.0 per cent, NPV 73.1 per cent) (*Fig. S4*, supporting information).

**+A: Discussion**

The findings in this study provide further evidence that preoperative CPET-derived variables can be used to risk assess patients before major elective colorectal surgery. Optimal cut-off points identified for O2 at L and O2peak were 11.1 and 18.2 ml per kgper min respectively. ROC analysis showed that both variables were effective at identifying patients at risk of postoperative morbidity.

The cut-off values established here for both primary outcome variables were similar to those found by others14,16,27–29, although individual variables from patients in this study showed higher sensitivity and specificity than reported from a previous single-centre study that included both benign and malignant colonic disease18. In contrast, in the present study O2 at L demonstrated lower sensitivity and specificity compared with a previous single-centre study on rectal cancer surgery17. In support of previous research in patientsundergoing colorectal surgery, E/CO2 at L did not discriminate between patients with, and without a postoperative complication. CPET-derived variables acting as prognostic variables for in-hospital morbidity here were similar to those identified in the authors’ previous two studies17,18. The prediction model generated in the present study, which included O2 at L, O2peak, E/CO2 at L, BMI and operation type, was found to predict in-hospital complications with better discrimination than another published model used for major colonic surgery18, and better still than a univariable model in two cohorts of patients undergoing major surgery16,29.

When comparison is made to the literature supporting objective assessment of physical fitness before major abdominal surgery, similar trends emerge. Older and colleagues28 showed that in elderly patients a preoperative O2 at L of 11.0 ml per kgperminor less was associated with increased cardiovascular postoperative mortality. Assessing 843 patients undergoing major colorectal surgery, radical nephrectomy or cystectomy, Wilson and colleagues14 concluded that a O2 at L of 10.9 ml per kg min or less and E/CO2 at L of 34 or more had 88 per cent sensitivity and 47 per cent specificity for identifying patients at risk of in-hospital mortality. Snowden and co-workers16 also evaluated CPET in elderly patients (mean age 70 years) undergoing major intra-abdominal surgery; they reported that the optimal cut-off point of O2 at L of 10.1 ml per kgper min gave 88 per cent sensitivity and 79 per cent specificity for discriminating postoperative complications (AUROC 0.85, 95 per cent c.i. 0.78 to 0.91; *P* < 0.001). Furthermore, Snowden and colleagues27 went on to evaluate 389 patients before major hepatobiliary surgery and identified a O2 at L of 10.0 ml per kgpermin as the most significant independent predictor of postoperative mortality among all other exercise variables.

The strengths of this study include the homogeneous study population, the clearly defined inclusion and exclusion criteria, the number of contributing centres, the wide ranging geographical distribution (in an attempt to reduce bias owing to socioeconomic status and geography), the robust reporting of objectively measured CPET variables, use of the POMS24 and Dindo25 classifications as primary outcome measures for morbidity, and the robust statistical analyses. The multivariable model can now be regarded as generalizable in patients undergoing major colorectal surgery. A final strength of this study is that 1-year postoperative mortality was recorded. Only one other study20 investigated mortality beyond 30 days after colorectal surgery, albeit in a smaller population. Also consistent with previous literature in other major surgical populations30, O2 at L as part of a multivariable model appeared to predict survival.

Potential limitations include the lack of blinding of colorectal MDTs (including anaesthetists) to CPET data; thus perioperative management might have been influenced31. Morbidity outcome collection was not performed in a blinded fashion either.CPETs were not recorded in a dual reporting style. Blinding and dual reporting were undertaken in previous publications17,18; however, the authors felt that this study should mimic real-life clinical practice. Finally, the AUROCs differed significantly between centres for three of the four CPET variables. This was an expected finding, as this study was not powered to assess intercentre variable discrimination; a large sample size in each centre would have been necessary to address this specific question.

The reliability and association of CPET variables with outcome following major surgery, especially major colorectal surgery, is established. CPET can identify high-risk patients by measuring O2 at L and O2peak before major colorectal surgery. The evidence supporting whether preoperative CPET is a useful test to alter clinician’s decisions in selected high-risk patients is, at present, incomplete. Studies are still required to assess whether CPET variables can be used with other non-CPET markers (Revised Cardiac Risk Index, Veterans Specific Activity Questionnaire (VSAQ), POSSUM, plasma biomarkers) to enhance perioperative risk assessment29. Meanwhile, decisions regarding perioperative care or fitness for surgery should be based on complete clinical and CPET assessment, and not on individual CPET variables or predictive models in isolation.

**+A: Acknowledgements**

The authors acknowledge the contribution of clinical colleagues, CPET services, research and development and administrative colleagues at all centres that contributed data to the Multicentre Colorectal CPET Collaborative.Particular mention goes to C. Challand, C. wan Lai (Directorate of Colorectal Surgery), R. Struthers and R. Sneyd (Plymouth Directorate of Anaesthesia) and E. Psarelli (Cancer Research UK Liverpool Clinical Trials Unit). This project was undertaken when M.A.W. was a National Institute of Health Research-funded Academic Clinical Fellow.

*Disclosure:* The authors declare no conflict of interest.

**References**

1. Morris EJ, Taylor EF, Thomas JD, Quirke P, Finan PJ, Coleman MP, et al. Thirty-day postoperative mortality after colorectal cancer surgery in England. *Gut* 2011;**60**(6):806–813.

2. Association of Coloproctology of Great Britain and Ireland. *National Bowel Cancer Audit Annual Report 2014*. http://www.hscic.gov.uk/bowel (accessed April 2015)

3. Hightower CE, Riedel BJ, Feig BW, Morris GS, Ensor JE, Woodruff VD, et al. A pilot study evaluating predictors of postoperative outcomes after major abdominal surgery: Physiological capacity compared with the ASA physical status classification system. *Br J Anaesth* 2010;**104**(4):465–471.

4. Tekkis PP, Prytherch DR, Kocher HM, Senapati A, Poloniecki JD, Stamatakis JD, et al. Development of a dedicated risk-adjustment scoring system for colorectal surgery (colorectal POSSUM). *Br J Surg* 2004;**91**(9):1174–1182.

5. Edwards M, Whittle J, Ackland GL. Biomarkers to guide perioperative management. *Postgrad Med J* 2011;**87**(1030):542–549.

6. Ethan A. Halm, Warren S. Browner, Julio F. Tubau, Ida M. Tateo and Mangano DT. Echocardiography for Assessing Cardiac Risk in Patients Having Noncardiac Surgery. *Ann Intern Med* 1996;**125**(6):433–441.

7. Struthers R, Erasmus P, Holmes K, Warman P, Collingwood A, Sneyd JR. Assessing fitness for surgery: a comparison of questionnaire, incremental shuttle walk, and cardiopulmonary exercise testing in general surgical patients. *Br J Anaesth* 2008;**101**(6):774–780.

8. Lee L, Schwartzman K, Carli F, Zavorsky GS, Li C, Charlebois P, et al. The association of the distance walked in 6 min with pre-operative peak oxygen consumption and complications 1 month after colorectal resection. *Anaesthesia* 2013;**68**(8):811–816.

9. Huddart S, Young EL, Smith R-L, Holt PJ, Prabhu PK. Preoperative cardiopulmonary exercise testing in England - a national survey. *Perioper Med* 2013;**2**(1):4.

10. Simpson JC, Sutton H, Grocott MPW. Cardiopulmonary exercise testing — a survey of current use in England. *J Int Care Soc* 2009;**10**(4):275–278.

11. Smith TB, Stonell C, Purkayastha S, Paraskevas P. Cardiopulmonary exercise testing as a risk assessment method in non cardio-pulmonary surgery: a systematic review. *Anaesthesia* 2009;**64**(8):883–893.

12. West M, Jack S, Grocott MPW. Perioperative cardiopulmonary exercise testing in the elderly. *Best Pract Res Clin Anaesthesiol* 2011;**25**(3):427–437.

13. Hennis PJ, Meale PM, Grocott MPW. Cardiopulmonary exercise testing for the evaluation of perioperative risk in non-cardiopulmonary surgery. *Postgr Med J* 2011;**87**(1030):550–557.

14. Wilson RJT, Davies S, Yates D, Redman J, Stone M. Impaired functional capacity is associated with all-cause mortality after major elective intra-abdominal surgery. *Br J Anaesth* 2010;**105**(3):297–303.

15. Carlisle J, Swart M. Mid-term survival after abdominal aortic aneurysm surgery predicted by cardiopulmonary exercise testing. *Br J Surg* 2007;**94**(8):966–969.

16. Snowden CP, Prentis JM, Anderson HL, Roberts DR, Randles D, Renton M, et al. Submaximal cardiopulmonary exercise testing predicts complications and hospital length of stay in patients undergoing major elective surgery. *Ann Surg* 2010;**251**(3):535–541.

17. West MA, Parry M, Lythgoe D, Barben C, Kemp G, Grocott M, et al. Cardiopulmonary exercise variables are associated with postoperative morbidity after rectal cancer surgery. *Br J Surg* 2014;**101**(9):1166-1172.

18. West MA, Lythgoe D, Barben CP, Noble L, Kemp GJ, Jack S, et al. Cardiopulmonary exercise variables are associated with postoperative morbidity after major colonic surgery : a prospective blinded observational study. *Br J Anaesth* 2014;**112**(4):665–671.

19. Weisman IM, Marciniuk D, Martinez FJ, Sciurba F, Sue D, Myers J, et al. ATS/ACCP Statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 2003;**167**(2):211–277.

20. Lai CW, Minto G, Challand CP, Hosie KB, Sneyd JR, Creanor S, et al. Patients’ inability to perform a preoperative cardiopulmonary exercise test or demonstrate an anaerobic threshold is associated with inferior outcomes after major colorectal surgery. *Br J Anaesth* 2013;**111**(4):607–611.

21. Wasserman K, Hansen JE, Sue DY SW and Whipp BJ. *Principles of Exercise Testing and Interpretation: Pathophysiology and Clinical Applications.* (4th Edition) . Baltimore, Maryland: Lippincott Williams & Wilkins; 2005.

22. William B, Wasserman K, Whipp BJ. A new method for detecting anaerobic threshold by gas exchange. *J Appl Physiol* 1986;**60**(6):2020–2027.

23. Sue D, Wasserman K, Moricca R, Casaburi R. Metabolic acidosis during exercise in patients with chronic obstructive pulmonary disease. Use of the V-slope method for anaerobic threshold determination. *Chest* 1988;**94**(5):931–938.

24. Grocott MPW, Browne JP, Van der Meulen J, Matejowsky C, Mutch M, Hamilton M a, et al. The Postoperative Morbidity Survey was validated and used to describe morbidity after major surgery. *J Clin Epidemiol* 2007;**60**(9):919–928.

25. Dindo D, Demartines N, Clavien P-A. Classification of Surgical Complications. A new proposal with evluation in a cohort of 6336 patients and results of a survey. *Ann Surg* 2004;**240**(2):205–213.

26. Swart M, Carlisle JB. Case-controlled study of critical care or surgical ward care after elective open colorectal surgery. *Br J Surg* 2012;**99**(2):295–299.

27. Snowden CP, Prentis J, Jacques B, Anderson H, Manas D, Jones D, et al. Cardiorespiratory fitness predicts mortality and hospital length of stay after major elective surgery in older people. *Ann Surg* 2013;**257**(6):999–1004.

28. Older P, Hall A, Hader R. Cardiopulmonary exercise testing as a screening test for perioperative management of major surgery in the elderly. *Chest* 1999;**116**(2):355–362.

29. James S, Jhanji S, Smith A, Brien GO, Fitzgibbon M, Pearse RM. Comparison of the prognostic accuracy of scoring systems, cardiopulmonary exercise testing , and plasma biomarkers : a single-centre observational pilot study. *Br J Anaesth* 2014;**112**(3):491–7.

30. Colson M, Baglin J, Bolsin S, Grocott MPW. Cardiopulmonary exercise testing predicts 5 yr survival after major surgery. *Br J Anaesth* 2012;**109**(5):735–741.

31. Grocott MPW, Pearse RM. Prognostic studies of perioperative risk: robust methodology is needed. *Br J Anaesth* 2010;**105**(3):243–245.

**Supporting information**

Additional supporting information may be found in the online version of this article:

**Fig. S1** Kaplan–Meier curve relating in-hospital complications to length of hospital stay (Word document)

**Fig. S2** Kaplan–Meier curves comparing hospital stay for patients grouped according to optimal cut-off point for cardiopulmonary exercise testing variables (Word document)

**Fig. S3** Receiver operating characteristic (ROC) curves for each cardiopulmonary testing variable according to centre (Word document)

**Fig. S4** Receiver operating characteristic (ROC) curve for multivariable logistic regression model. Final variables include oxygen uptake at estimated lactate threshold, peak oxygen uptake, ventilator equivalents for carbon dioxide at estimated lactate threshold, body mass index and operation type. Circle indicates optimal cut-off point obtained by minimizing the distance to the upper left corner. Area under ROC curve 0.83 (Word document)

**Table S1** In-hospital morbidity assessed for the nine domains of the PostOperative Morbidity Survey at day 5 after surgeryin patients grouped according to the optimal cut-off point for oxygen uptake at estimated lactate threshold (Word document)

**Table S2** Cardiopulmonary exercise testing variables dichotomized at their optimal cut-off point in relation to duration of hospital stay (Word document)

**Table S3** Area under receiver operating characteristic (ROC) curve for each cardiopulmonary exercise testing variable according to centre (Word document)

**Typesetter: please refer to marked-up figures**

**Fig. 1** Receiver operating characteristic (ROC) curves for **a** oxygen uptake at estimated lactate threshold (O2 at L) and **b** oxygen uptake at peak (O2peak). Symbols indicate optimal cut-off point obtained by minimizing the distance to the upper left corner. Area under ROC curve: **a** 0.79, **b** 0.77

**Table 1** Patient demographics and morbidity after elective colorectal surgery

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | Postoperative complications | |  |
|  | Overall  (*n* = 703) |  | No  (*n* = 445) | Yes  (*n* = 258) | *P*§ |
| Centre\*  Aintree  Crosshouse  Medway  Torbay  Maidstone  Plymouth | 192 (27.3)  26 (3.7)  8 (1.1)  180 (25.6)  58 (8.3)  239 (34.0) |  | 111 (57.8)  20 (77)  5 (63)  101 (56.1)  51 (88)  157 (65.7) | 81 (42.2)  6 (23)  3 (37)  79 (43.9)  7 (12)  82 (34.3) | < 0.001 |
| Age (years)† | 69 (61–76) |  | 67 (60–74) | 72 (64–78) | < 0.001 |
| Sex ratio (M : F) | 428 : 275 |  | 279 : 166 | 149 : 109 | 0.408 |
| Body mass index (kg/m2)† | 27 (24–31) |  | 27 (24–30) | 29 (25–32) | 0.007 |
| Method of surgery  Open  Laparoscopic | 435 (61.9)  268 (38.1) |  | 245 (55.1)  200 (44.9) | 190 (73.6)  68 (26.4) | 0.026 |
| Surgical procedure  Right hemicolectomy  Transverse hemicolectomy  Left hemicolectomy  Subtotal colectomy  Anterior resection  Hartman’s procedure  APR  Other | 159 (22.6)  4 (0.6)  37 (5.3)  25 (3.6)  374 (53.2)  30 (4.3)  47 (6.7)  27 (3.8) |  | 99 (22.2)  2 (0.4)  21 (4.7)  12 (2.7)  247 (55.5)  17 (3.8)  29 (6.5)  18 (4.0) | 60 (23.3)  2 (0.8)  16 (6.2)  13 (5.0)  127 (49.2)  13 (5.0)  18 (7.0)  9 (3.5) | 0.426 |
| Disease status  Benign  Malignant  Other | 85 (12.1)  617 (87.8)  1 (0.1) |  | 51 (11.5)  393 (88.3)  1 (0.2) | 34 (13.2)  224 (86.8)  0 (0) | 0.532 |
| Pathological tumour category‡  pT0  pT1  pT2  pT3  pT4  Unknown | 16 (2.3)  30 (4.3)  78 (11.1)  209 (29.7)  59 (8.4)  311 (44.2) |  | 10 (2.2)  17 (3.8)  46 (10.3)  130 (29.2)  28 (6.3)  214 (48.1) | 6 (2.3)  13 (5.0)  32 (12.4)  79 (30.6)  31 (12.0)  97 (37.6) | 0.600 |
| Pathological node category‡  pN0  pN1  pN2  pN3  Unknown | 222 (31.6)  122 (17.4)  44 (6.3)  2 (0.3)  313 (44.5) |  | 133 (29.9)  70 (15.7)  25 (5.6)  2 (0.4)  215 (48.3) | 89 (34.5)  52 (20.2)  19 (7.4)  0 (0)  98 (38.0) | 0.844 |

Values in parentheses are percentages by group size unless indicated otherwise; values are \*percentage by centre and †median (i.q.r.). APR, abdominoperineal excision of the rectum. ‡International Union Against Cancer TNM Classification of Malignant Tumours, seventh edition, LH. Sobin, MK. Gospodarowicz and C. Wittekind; Wiley-Blackwell. §Univariable logistic regression.

**Table 2** Cardiopulmonary exercise testing data in relation to postoperative complications

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | No. of patients tested |  |  | Postoperative complications | |  |
|  | Overall |  | No | Yes | *P* |
| O2 at L (ml per kg per min) | 703 | 11.9 (9.9–14.3) |  | 13 (11.3–15.5) | 9.9 (8.6–11.6) | 0.002 |
| O2peak (ml per kg per min) | 465 | 18.8 (15.4–22.9) |  | 20.4 (17.5–24.7) | 15.5 (12.8–18.6) | 0.031 |
| O2 pulse at L (ml/beat) | 284 | 9.1 (7.1–11.2) |  | 9.9 (7.8–12.3) | 7.5 (5.7–9.2) | < 0.001 |
| E/CO2 at L | 690 | 30.9 (27.5–34.2) |  | 30.2 (27.4–33.2) | 32.0 (27.9–36.0) | 0.111 |

Values are median (i.q.r.). O2 at L, oxygen uptake at estimated lactate threshold; O2peak, oxygen uptake at peak exercise; O2 pulse at L, oxygen pulse at estimated lactate threshold; E/CO2 at L, ventilatory equivalents for carbon dioxide at estimated lactate threshold. †Continuous univariable comparisons by logistic regression.



**Table 3** Postoperative morbidity by Dindo grade in patients grouped according to the optimal cut-off point for oxygen uptake at estimated lactate threshold

|  |  |  |
| --- | --- | --- |
| Grade\* | O2 at L (ml per kg per min) | |
| 11.1  (*n* = 284) | 11.1  (*n* = 419) |
| 0 | 100 (35.2) | 344 (82.1) |
| I | 20 (7.0) | 9 (2.1) |
| II | 125 (44.0) | 40 (9.5) |
| IIIa | 10 (3.5) | 11 (2.6) |
| IIIb | 20 (7.0) | 12 (2.9) |
| IVa | 4 (1.4) | 2 (0.5) |
| IVb | 1 (0.4) | 0 (0) |
| V | 4 (1.4) | 1 (0.2) |

Values in parentheses are percentages. \*Dindo classification25.O2 at L, oxygen uptake at estimated lactate threshold. *P* < 0.001 (Fisher’s exact test).

**Table 4** Summary of receiver operating characteristic (ROC) curve analysis for selected cardiopulmonary exercise testing variables

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | O2 at L  (ml per kg per min) | O2peak  (ml per kg per min) | O2 pulse at L  (ml/beat) | E/CO2 at L |
| No. of patients tested | 703 | 465 | 284 | 690 |
| Cut-off point | 11.1 | 18.2 | 8.7 | 30.9 |
| AUROC | 0.79 (0.76, 0.83) | 0.77 (0.72, 0.82) | 0.75 (0.69, 0.81) | 0.58 (0.54, 0.63) |
| Sensitivity (%) | 78.2 (73, 81) | 70.3 (65, 75) | 69.2 (62, 76) | 59.1 (52, 65) |
| Specificity (%) | 71.4 (65, 77) | 72.0 (65, 79) | 68.4 (58, 77) | 55.0 (50, 60) |
| PPV (%) | 82.0 (78, 86) | 81.4 (76, 86) | 81.0 (74, 86) | 42.4 (37, 48) |
| NPV (%) | 65.1 (59, 70) | 59.1 (52, 65) | 53.2 (44, 62) | 70.4 (65, 75) |

Values in parentheses are 95 per cent c.i. O2 at L, oxygen uptake at estimated lactate threshold; O2peak, oxygen uptake at peak exercise; O2 pulse at L, oxygen pulse at estimated lactate threshold; E/CO2 at L, ventilatory equivalents for carbon dioxide at estimated lactate threshold. AUROC, area under the receiver operating characteristic curve, PPV positive predictive value; NPV, positive predictive value.



**Table 5** Final multivariable logistic regression model for variables selected using a stepwise univariable analysis

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Odds ratio | *P* |
| O2 at L (≤ 11.1 ml per kg per min) | | 7.56 (4.44, 12.86) | < 0.001 |
| O2peak (≤ 18.2 ml per kg per min) | | 2.15 (1.01, 4.57) | 0.047 |
| E/CO2 at L (> 30.9) | | 1.38 (1.00, 1.89) | 0.047 |
| Age (per 5 years) | | 1.05 (0.92, 1.19) | 0.451 |
| BMI (> 27 kg/m2) | | 1.05 (1.03, 1.08) | < 0.001 |
| Operation (laparoscopic) | | 0.30 (0.02, 0.44) | 0.033 |

Values in parentheses are 95 per cent c.i. Cardiopulmonary exercise testing variables were dichotomized at their optimal cut-off point. O2 at L, oxygen uptake at estimated lactate threshold; O2peak, oxygen uptake at peak exercise; E/CO2 at L, ventilatory equivalents for carbon dioxide at estimated lactate threshold; BMI, body mass index.

