

## Systematic Review

# Cardiopulmonary Exercise Testing for Predicting Early Outcomes after Major Cancer Resection: A Systematic Review

Stephen Lam<sup>1,2</sup>, and Andrew Hart<sup>1,2</sup>

## ABSTRACT

From the <sup>1</sup>Norfolk and Norwich University Hospitals NHS Foundation Trust, Colney Lane, Norwich, UK; <sup>2</sup>Norwich Medical School, University of East Anglia, Norwich, UK.

**Correspondence** to Dr. Stephen Lam at [stephen.lam@uea.ac.uk](mailto:stephen.lam@uea.ac.uk).

**Citation:** Stephen Lam, Andrew Hart. **Cardiopulmonary Exercise Testing for Predicting Early Outcomes after Major Cancer Resection: A Systematic Review.** *J Anesth Perioper Med* 2018; 5:136-48.

**doi:** 10.24015/JAPM.2018.0052

**Background:** Postoperative complications after major surgery are thought to be associated with reduced fitness. Surgical cancer patients are often malnourished, cachexic and subject to neoadjuvant chemotherapy resulting in low preoperative fitness levels. This review examined the associations between aerobic fitness, as determined objectively by preoperative cardiopulmonary exercise testing (CPEX), and short-term morbidity after cancer surgery.

**Methods:** A literature search using databases of PubMed, Medline, Embase, Cumulative Index to Nursing and Allied Health Literature (CINAHL) and the Cochrane Library for studies that examined associations between preoperative CPEX variables and postoperative complications following surgery for the ten commonest cancers.

**Results:** A total of 21 observational studies were identified with 4957 patients that underwent CPEX testing prior to lung, colorectal, liver, oesophagogastric, bladder and pancreas resections. The median sample size was 105 patients (range 64 - 1684). No studies were found for breast or brain cancers or lymphoma. In lung cancer patients undergoing thoracotomy, a  $VO_{2peak} \leq 15\text{ml/kg/min}$  was associated with an increased risk of respiratory complications and death. None of the studies in other cancer types had adequate sample sizes to report on mortality. CPEX testing had mostly poor to average discriminatory accuracy to predict postoperative morbidity in other cancer resection surgeries. Findings across studies were inconsistent, and detection and selective reporting biases were likely to be significant.

**Conclusion:** The utility of CPEX testing prior to cancer surgery is questionable and currently should not be used as a discriminatory tool, except in patients undergoing lung cancer resection by thoracotomy. Larger studies with more robust methodologies are currently required to determine the utility of CPEX.

Surgical stress response following major surgery results in muscle wasting and systemic inflammation, with a large increase in tissue oxygen demand and consumption (1); increasing the risk of ischaemic events (2). Furthermore, postoperative bed-rest and incisional pain inhibits normal lung mechanics, promoting shallow breathing, atelectasis and infective lung consolidation (3). These physiological challenges are in part met by a patient's cardiopulmonary reserves, or their ability to increase cardiac output and ventilation to meet increased demand. Such reserves are greater in physiologically 'fitter' patients. Cancer patients represent a specific population, more likely to have underlying malnutrition and cachexia than non-cancer patients, with depleted fitness levels (4, 5). Neoadjuvant chemotherapy has also been shown to independently reduce aerobic fitness (5). By measuring a patient's preoperative cardiopulmonary reserve, or functional capacity, we may hypothetically be able to discriminate patients that may or may not tolerate the physiological insult associated with surgery. Therefore, measurement of preoperative fitness may serve as a preoperative risk prediction tool for the development of complications prior to major cancer resection.

Cardiopulmonary exercise testing (CPEX) offers a quantitative and composite measure of functional reserve, where a low score indicates a patient's inability to effectively transfer oxygen to tissues on exertion. The test is carried out in laboratory conditions usually on a static exercise bike, where pedal resistance, or workload, is systematically increased until volitional termination. Two important CPEX variables are captured by analysis of gas exchange at the mouth,  $VO_{2peak}$  (the maximal oxygen consumed at the peak of exercise) and  $VO_2$  at estimated anaerobic threshold (AT), a measure of sustainable aerobic activity. Both parameters have shown promise in observational studies in predicting both morbidity and mortality after major elective surgery (6). In a seminal study from 1993, of 187 elderly patients undergoing major abdominal surgery, a preoperative AT cut-off of 11ml/kg/min had a sensitivity of 91% and specificity of 74% for predicting postoperative mortality after major abdominal surgery (7). Since then, multiple, often small, observational studies have been published in various cancer surgery specialties. However,

the results from these are inconsistent, possibly explained by biases inherent in the methodology such as performance and detection biases (due to unblinded clinical teams and outcome assessors, respectively) as well as selective reporting bias, which is likely to be substantial (8). Although 3 reviews of CPEX testing and major surgeries have been published (6, 9, 10), all included non-cancer patients and excluded common cancers such as lung and bladder. Furthermore, important sources of bias do not appear to have been adequately addressed previously. The aim of this article was to assess the association between CPEX testing and short-term postoperative morbidity and mortality after common major cancer resection surgery; with due consideration of potential biases both in terms of their magnitude and direction. If a convincing inverse association exists, it may not only support the preoperative use of CPEX to determine operability and postoperative monitoring and management, but also identify fitness as a modifiable risk factor for investigation in randomized controlled trials of surgical cancer patients.

## METHODS

The preferred reporting items for systematic reviews (PRISMA) guidelines were used to standardize the methods of conducting and reporting this review. The ten commonest causes of cancer deaths in the UK in 2014 were identified (lung, bowel, breast, prostate, pancreas, oesophagus, bladder, brain, liver, lymphoma) (11), and a literature search was conducted using PubMed, Medline, Embase, CINAHL and the Cochrane library (from commencement to 5th July 2017) for studies that examined associations between preoperative CPEX and postoperative complications following cancer resection surgery. The search terms used were: CPEX, CPET, exercise testing, anaerobic threshold,  $VCO_2$ , ventilatory inefficiency, oxygen consumption,  $VO_2$ , preoperative exercise, aerobic exercise. For each cancer, additional specific search terms were added (Supplementary Appendix 1). We included analytical studies (cohort, case-control, randomized controlled trials) that investigated the association between preoperative CPEX variables and short-term (up to 90 days)

morbidity and mortality. Studies were excluded with sample sizes <100 patients, unless there were <2 studies in that surgical population in which case the total sample size needed to be at least 60, which was considered the minimum size to determine associations with a moderate event rate. In order to select studies which examined only cancer populations, studies with a large proportion ( $\geq 25\%$ ) of non-cancer patients were also excluded (unless cancer patients were analysed separately) as were investigations that combined multiple surgical patient populations (e.g. colorectal, urological and upper gastrointestinal cancer patients), unless these groups were sub-analysed.

## RESULTS

A total of 21 hospital-based cohort studies (12 prospective, 9 retrospective) were included in this review with a total of 4957 patients that underwent CPEX testing prior to lung (12-17), colorectal (18-20), liver (21-23), oesophagogastric (24-26), bladder (27-29) and pancreas resection (30-32) (Figure 1). No studies were found in breast, brain or lymphoma cancers. No randomized controlled trials of any cancer site were identified. Data were extracted from each study (including study design, sample size, outcome measurement and effect sizes) and tabulated (Table 1).

### Lung Cancer

Four hundred and three studies were identified, with 63 potentially relevant papers by title. A review of these abstracts identified 23 potentially relevant papers. Six studies met the inclusion criteria and were included in this review (12-17). All were observational cohort studies (2 prospective, 4 retrospective), with a total of 2814 patients from hospitals within Europe and the USA. The largest study was a retrospective analysis of 1684 patients who had lung cancer surgery identified from the European Society of Thoracic Surgeons (ESTS) database (12), which is voluntarily contributed to by clinicians from 235 sites across Europe. The authors reported no association between  $VO_{2peak}$  and all-cause morbidity (for either video-assisted thoracoscopic surgery (VATS) or thoracotomy resections), but a significant association with mortality in pa-

tients undergoing thoracotomy resection with a  $VO_{2peak} < 15\text{ml/kg/min}$  ( $P = 0.008$ ). Whilst this study had several strengths including, a large sample size and use of propensity matching to reduce selection bias, it had significant limitations, which are acknowledged by the authors. The largest of these was the potential for measurement error (for complications) through the use of a voluntary multi-institutional database, where the accuracy of the data entry has not been validated (12). Inaccurate recording of outcome would lead to an attenuation of any associations between CPEX variables and outcome.

The second largest study was a multi-center (9 centers) prospective observational investigation of 346 patients from the USA who underwent thoracotomy and lung cancer resection surgery (13). In contrast to the results of the previous study, the authors reported that  $VO_{2peak}$  was significantly lower in the group with complications (15.2 ml/kg/min) compared to those without (16.7 ml/kg/min) although the mean difference was small: -1.47 ml/kg/min (95% CI 0.55-2.4),  $P = 0.002$ . The authors also undertook a further sub-analysis in patients with the outcomes of respiratory failure ( $n = 33$ ) and death ( $n = 15$ ). Both events were associated with a lower mean  $VO_{2peak}$  (14.7 ml/kg/min),  $P = 0.041$  compared to those without complications (mean difference 2 ml/kg/min), although the small number of events is noted. Whilst the large sample size, prospective design and multi-center participation increases power, reduces bias and enables generalisability, respectively, there were limitations. A pre-defined criterion for what constituted each complication was not established, nor were the post-operative outcomes measured in a blinded fashion, both of which could contribute to detection bias resulting in spurious over-estimation of the associations found. Furthermore, the authors defined postoperative morbidity as a composite outcome, which included complications lacking a clear plausible mechanistic relationship with aerobic fitness; such as red blood cell transfusions ( $n = 38$ ). Such an event is more likely to be associated with intra-operative blood loss (perhaps due to longer operating times in patients with underlying lung disease), suggesting that residual confounding may explain some of the associations found.

The other 4 studies in lung cancer patients were all hospital-based cohort studies, which reported inconsistent findings (14-17). Two of the investigations were by the same research group (15, 17), which reported an inverse association between  $VO_{2peak}$  and respiratory complications ( $P = 0.015$ ) in one of their studies (15), but were unable to subsequently replicate this finding ( $P = 0.5$ ) in the other despite similar methodologies and patient populations (17). Detection bias due to non-blinded assessment may have contributed to the variability in findings. The two other studies reported an inverse association between  $VO_{2peak}$  and cardiopulmonary complications, OR 0.05 (95% CI 0.01-0.58),  $P = 0.0216$  and OR 0.79, 95% CI 0.71-0.88,  $P \leq 0.0001$  (14). However, the large variation in the effect size estimates and large confidence intervals reflects the imprecision of the findings. Again, outcome assessment was not blinded, which may have spuriously inflated the reported effect sizes.

**Author's summary:** Observational studies have reported that  $VO_{2peak}$  appears to be associated with complications after lung cancer resection surgery. More specifically, the two largest studies reported that a  $VO_{2peak} < 15$  ml/kg/min was associated with an increased risk of respiratory failure (13) and death following lung cancer resection by thoracotomy (12, 13). However, only one study in lung cancer surgery examined outcomes after VATS and found no association between  $VO_{2peak}$  and morbidity or mortality (12). As VATS is increasingly becoming used for lung cancer resection, the utility of CPEX testing needs to be updated to determine if it is of value in less invasive surgeries for lung cancer resection.

### Colorectal Cancer

The literature search identified 431 studies, of which 30 were considered relevant based on their title. These were reduced to 13 after reading the abstracts, of which 3 met the inclusion criteria and were included in this review (18-20). The main reason for excluding other investigations was their inclusion of  $> 25\%$  of non-cancer patients. All 3 were observational studies from the same UK group. The largest was a retrospective multi-center (6 sites) UK investigation of 703 patients (20), most of which had a malignant disease (87%). In contrast to the studies in lung can-

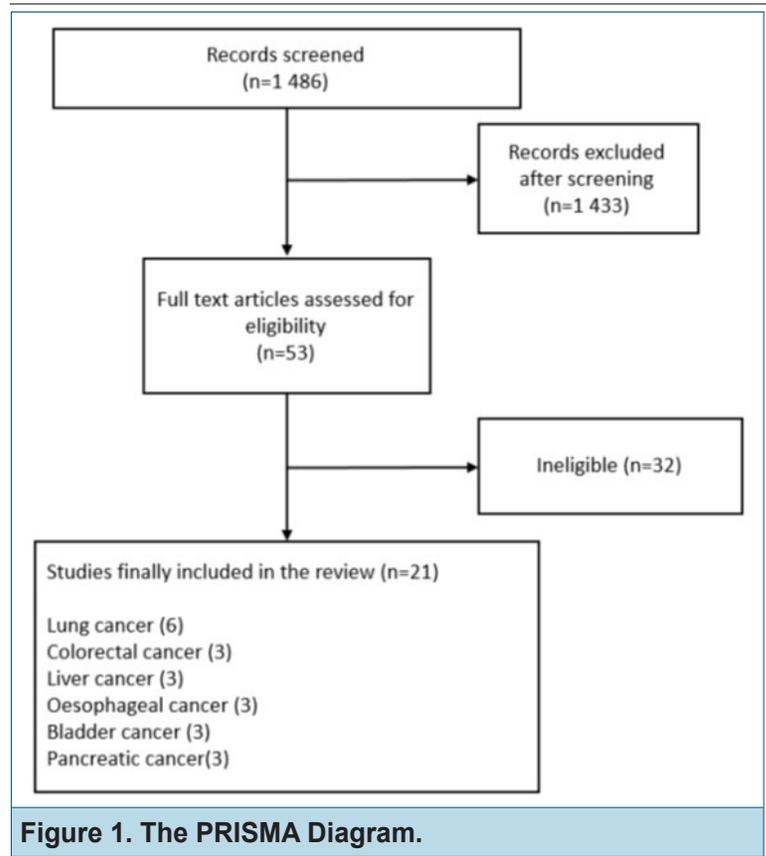


Figure 1. The PRISMA Diagram.

cer surgery, all-cause morbidity was measured using a validated PostOperative Morbidity Survey (POMS) at postoperative day 5 (33, 34). The severity of complications was graded using the system devised by Clavien and Dindo (34). This grading system is based upon the level of intervention required to treat a complication; from normal postoperative adjuncts such as supplementary oxygen, analgesia and anti-emetics (grade 1), to additional medicines above usual standard care, including antibiotics (grade 2). Grade 3 is a complication requiring operative intervention and grade 4 complications require organ support in critical care. The authors reported a significant difference in both median  $VO_{2peak}$  and AT in patients with and without all-cause morbidity of any Clavien-Dindo grade ( $P = 0.031$  and  $P = 0.002$ , respectively). Receiver Operator Characteristic (ROC) Curve analyses were also undertaken to show how sensitivity and specificity varied with changing thresholds,

**Table 1. Characteristics of Studies Investigating the Association between Preoperative CPEX and Postoperative Complications following Major Cancer Resection Surgery.**

Author, Year	Study Design	Sample Size (n)	Postoperative Complications Measured (rate of occurrence)	Diagnostic Criteria for Complications?	Complication Severity Classified?	Blinded Complication Assessment?	Association between VO <sub>2peak</sub> and Measured Complications	Association between AT and Measured Complications
Lung cancer								
Begum, 2015 (12)	Retrospective case-matched multi-centre cohort study	1684	In-hospital cardiopulmonary complications (30%)	No	No	No	No	Not measured
Brunelli, 2012 (17)	Retrospective hospital-based cohort study	225	30-day cardiopulmonary complications (23%)	Yes	No	No	No	Not measured
Licker, 2011 (14)	Retrospective hospital-based cohort study	210	In-hospital cardiopulmonary complications (22%)	Yes	No	No	Yes (OR = 0.79, 95% CI 0.71-0.88, P ≤ 0.0001)	Not measured
Torchio, 2010 (16)	Retrospective hospital-based cohort study	145	30-day cardiopulmonary complications (15%)	Yes	No	No	Yes (OR = 0.05, 95% CI 0.01-0.58, P = 0.017)	Not measured
Brunelli, 2009 (15)	Prospective hospital-based cohort study	204	In-hospital cardiopulmonary complications (23%)	Yes	No	No	No	No
Loewen, 2007 (13)	Prospective hospital-based multi-centre cohort study	346	30-day cardiopulmonary complications (40%)	No	No	No	Yes (No RR/OR/HR reported)	Not measured
Colorectal cancer								
West, 2016 (20)	Retrospective multi-centre cohort study	703	All cause morbidity at post-operative day 5 (36.7%)	POMS	ClavienDindo	No	Yes (≤ 18.2 ml/kg/min, OR 2.15, 95% CI 1.01-4.59, P = 0.05)	Yes (≤ 11.1ml/kg/min, OR 7.56, 95% CI 4.44-12.86, P ≤ 0.001)
West, 2014 (18)	Retrospective hospital-based cohort study	95	All cause morbidity at postoperative day 5 (48%)	POMS	ClavienDindo	Yes	Yes (≥ 18.8 ml/kg/min, OR 0.07, 95% CI 0.03-0.19, P ≤ 0.001)	Yes (≥ 11.2ml/kg/min, OR 0.07, 95% CI 0.03-0.19, P ≤ 0.001)

**Table 1. Characteristics of Studies Investigating the Association between Preoperative CPEX and Postoperative Complications following Major Cancer Resection Surgery (continued).**

Author, Year	Study Design	Sample Size (n)	Postoperative Complications Measured (rate of occurrence)	Diagnostic Criteria for Complications?	Complication Severity Classified?	Blinded Complication Assessment?	Association between VO <sub>2peak</sub> and Measured Complications	Association between AT and Measured Complications
West, 2014 (19)	Prospective hospital-based cohort study	136	All cause morbidity at postoperative day 5 (48%)	POMS	ClavienDindo	Yes	Yes (No RR/OR/HR reported)	Yes (OR 0.77, 95% CI 0.66-0.89, P = 0.0005)
Liver cancer								
Kasivisvanathan, 2015 (23)	Prospective hospital-based cohort study	108	3-day all cause morbidity (70%)	POMS	ClavienDindo	Yes	Yes (uninterpretable from data provided)	Yes (uninterpretable from data provided)
Dunne, 2014 (22)	Retrospective hospital-based cohort study	197	All cause in-hospital morbidity (44%)	No	ClavienDindo	No	No	No
Junejo, 2012 (21)	Prospective hospital-based cohort study	92	All cause in-hospital morbidity (51%)	POMS	No	No	No	No
Oesophageal cancer								
Moyes, 2013 (25)	Retrospective hospital-based cohort study	103	All cause in-hospital morbidity (55%)	CTCAE	No	No	No	Yes (No RR/OR/HR reported)
Forshaw, 2008 (24)	Retrospective hospital-based cohort study	78	All cause in-hospital morbidity (not stated)	CTCAE	No	No	Yes (No RR/OR/HR reported)	No
Nagamatsu, 2001 (26)	Retrospective hospital-based cohort study	91	In-hospital cardiopulmonary complications (19%)	No	No	No	Yes (No RR/OR/HR reported)	No

**Table 1. Characteristics of Studies Investigating the Association between Preoperative CPEX and Postoperative Complications following Major Cancer Resection Surgery (continued).**

Author, Year	Study Design	Sample Size (n)	Postoperative Complications Measured (rate of occurrence)	Diagnostic Criteria for Complications?	Complication Severity Classified?	Blinded Complication Assessment?	Association between VO <sub>peak</sub> and Measured Complications	Association between AT and Measured Complications
Bladder cancer								
Lamb, 2016 (29)	Prospective hospital-based cohort study	82	30-day Clavien-Dindo grade ≥3, all cause morbidity (12.6%)	No	Clavien-Dindo (grade ≥3 only)	No	No	No
Tolchard, 2014 (28)	Prospective hospital-based cohort study	105	90-day all cause morbidity (31%)	No	Clavien-Dindo	Yes	Yes (No RR/OR/HR reported)	Yes (AT ≥11.0ml / kg / min, OR 0.72, 95% CI 0.58-0.91)
Prentis, 2013 (27)	Prospective hospital-based cohort study	69	In-hospital all cause morbidity (56%)	No	Clavien-Dindo	Yes	No	No
Pancreatic cancer								
Chandrabalan, 2013 (31)	Retrospective hospital-based cohort study	100	Not stated	ISGPF	Clavien-Dindo	No	Not stated	No
Junejo, 2013 (32)	Prospective hospital-based cohort study	64	In-hospital all cause morbidity (64%)	ISGPF	Fistula by IS-GPF	No	No	No
Ausania, 2012 (30)	Retrospective hospital-based cohort study	124	In-hospital all cause morbidity (44%)	ISGPF & POMS	Fistula by IS-GPF	No	No	Yes (AT ≤10.1ml / kg / min, OR 5.79, 95% CI 1.62-20.63)

POMS = Postoperative Morbidity Survey (a validated measure of postoperative complication outcomes). Clavien-Dindo is a standardised therapy orientated grading system for the severity of postoperative complications in surgical practice.  
 CTCAE = Common Terminology for Adverse Events (a descriptive terminology used in Adverse Event (AE) reporting, usually in clinical drug trials).  
 ISGPF = International study group on Pancreatic Fistula (a definition derived by an international panel of pancreatic surgeons).

which was expressed as an area under the ROC curve (AUC). The AUC takes into consideration the accuracy of a diagnostic test (in terms of sensitivity and specificity) across a range of threshold values (35). In the context of CPEX testing, where the association may be inverse, the AUC is equal to the probability that if a pair of patients (one with a complication and one without) are selected at random, the patient with a complication will have a lower CPEX value than the complication-free patient (35). An AUC of 1.0 indicates a perfect test and 0.5 a completely uninformative test i.e. a result occurring by chance. An AUC of <0.7 would be indicative of a poor predictive test, 0.7-0.8 average accuracy and > 0.8 good accuracy as a diagnostic test across a range of thresholds (35). In this study, AT had average discrimination (i.e. 0.70-0.80) with an AUC of 0.79, 95% CI 0.76-0.83 with an optimal cut-point at 11.1 ml/kg/min (78% sensitivity and 71% specificity). The AUC for  $VO_{2peak}$  was 0.77, 95% CI 0.71-0.82 with an optimal cut-point of 18.2ml/kg/min (70% sensitivity and 72% specificity). However, similar to the previous studies in lung cancer, outcome assessors were not blinded to CPEX data, so detection bias could explain the associations found. Indeed, there were significant variations ( $P \leq 0.001$ ) in AUC values across recruited hospital sites (supplementary material), where the largest recruiting center ( $n = 239$ ) had more modest values for  $VO_{2peak}$  (AUC 0.73) and AT (AUC 0.68) compared to the above-pooled values quoted in the study. Complications without a clear plausible relationship with aerobic fitness were associated with CPEX values including; postoperative pain and gastrointestinal symptoms (such as ileus). Therefore, residual confounding may explain some inverse associations found. The same group previously published a prospective blinded observational study in 136 patients undergoing colonic surgery, most of whom (89%) had a malignant disease (19). With detection bias removed, the predictive performance of CPEX for day 5 morbidity was poor (i.e., AUC <0.70). For AT the AUC was 0.63, 95% CI 0.54-0.73, with a lower optimal cut-point at 10.1 ml/kg/min (68% sensitivity and 58% specificity). The AUC for  $VO_{2peak}$  was 0.63, 95% CI 0.53-0.73 with an optimal cut-point of 16.7 ml/kg/min (55% sensitivity and 69% specificity).

Furthermore, 14% of their sample who underwent CPEX testing and surgery were excluded as they “lacked complete data”. It is unclear whether this data was missing at random, and may therefore represent a source of selection bias.

The final study by the same group investigated 95 rectal cancer patients undergoing resection surgery, 68 of whom received neoadjuvant treatment (18). Morbidity and mortality were measured blinded, using the same methods reported in their other work (19, 20). Both  $VO_{2peak}$  and AT were associated with total morbidity at day postoperative day 5. For AT, the AUC showed good accuracy at 0.87 (95% CI 0.79-0.95) with an optimal cut-point of 10.6 ml/kg/min (84% sensitivity and 92% specificity).  $VO_{2peak}$  had an AUC of 0.85 (95% CI 0.77-0.93) and cut-point of 18.6 ml/kg/min (82% sensitivity and 80% specificity). However, the small sample size increases the risk of a chance finding.

**Author’s conclusion:** Overall the association between preoperative CPEX and postoperative outcome following colorectal cancer surgery is derived from observational studies by the same research group. There does appear to be inverse associations between  $VO_{2peak}$  and AT and all-cause morbidity at day 5 post surgery. However, detection bias and residual confounding cannot be excluded. Furthermore, 5-day POMS morbidity measured complications of a low severity (Clavien-Dindo  $\leq 2$ ) in the majority of patients, which makes the clinical usefulness of these findings questionable. POMS has also not been validated as an index of overall morbidity (33). The decision of whether or not to undergo surgery is unlikely to be meaningfully informed by this work. Larger multicentre studies are required to address whether improved fitness prior to colorectal surgery reduces the risk of major postoperative outcomes including death, with outcome assessment blinded to CPEX data.

### Liver Cancer

One hundred and eight studies were identified using the search terms, of which 7 were considered relevant based on their title. These were reduced to 4 after reading the abstracts and 3 met the inclusion criteria and were included in this review (21-23). All 3 were UK hospital-based cohort investigations (2 prospective and 1 retro-

spective) of patients that underwent both major and minor hepatectomies (21-23). The largest, a retrospective study of 197 patients, found no association between in-hospital morbidity and  $VO_{2peak}$  or AT (22). The study did not measure complications using a validated outcome measure or blind its assessors to CPEX values. However, the result of such bias may inflate, rather than reduce, the effect size reported. The second largest study, a UK prospective cohort of 104 patients, did report an association between both  $VO_{2peak}$ , AT and complications (23). This used a validated outcome measure (POMS) with assessors blinded to the CPEX scores. However, the authors chose to report complications at postoperative day 3. This timeframe was not defended in the study, and in the absence of a pre-defined protocol, selective reporting bias cannot be ruled out, which may have produced a false positive result. Furthermore, the high complication rate (70%) likely reflects routine, less severe, postoperative interventions (pain medication, urinary catheter, oxygen supplementation), which are likely to be clinically insignificant and rare beyond day 3 (33). When the authors graded complication severity according to Clavien-Dindo they found no associations between CPEX variables and complications of grade 3 (needing surgical intervention) or above. The final study was a UK prospective cohort investigation of 92 patients, which reported no association between  $VO_{2peak}$  or AT and 30-day morbidity (as measured by POMS) (21). The authors did document that  $VE/VCO_2$  (a CPEX measure of ventilatory efficiency) was associated with complications, but its predictive value was poor; where a value of 34.5 provided a sensitivity of 47% for complications.

**Author's summary:** The evidence base for CPEX and liver cancer surgery is derived from small observational studies with poor quality outcome reporting. From the evidence to date, there is insufficient data demonstrating an association between CPEX and outcome following hepatic resection. A large well designed multicentre study is needed to assess the role of preoperative CPEX in liver cancer surgery.

### Oesophageal Cancer

Four hundred and seventy-eight studies were

identified of which 11 were considered relevant based on their title, reduced to 5 after reading the abstracts. Of these, 3 met the inclusion criteria and were included in this review (24-26). All 3 were retrospective hospital-based cohort studies with small sample sizes from single institutions. The largest was a Japanese analysis of 91 patients who underwent McKeown oesophagectomy for squamous cell carcinoma (26). Only cardiopulmonary complications were measured and occurred in 19% of patients. The mean  $VO_{2peak}$  was found to be lower in those with, vs. those without cardiopulmonary complications (789 ml/min/m<sup>2</sup> vs. 966 ml/min/m<sup>2</sup>, t-test,  $P \leq 0.001$ ). These values approximate to 20.9 ml/kg/min vs. 25.6 ml/kg/min [conversion using the average height and weight of a Japanese male] (36). No association was found between AT and complications (t-test,  $P = 0.12$ ). The second largest study was a UK investigation of 78 patients, predominantly with adenocarcinoma (74%), undergoing oesophagectomy (64% with neoadjuvant chemotherapy) (24). Cardiopulmonary outcomes occurred in 42% of patients ( $n = 33$ ) and non-cardiopulmonary in 24% ( $n = 19$ ). Similar to the Japanese study, a low mean  $VO_{2peak}$  was associated with cardiopulmonary complications although the mean difference was small (19.2 ml/kg/min in those with complications vs. 21.4 ml/kg/min in those without, t-test,  $P = 0.04$ ). AT was also not associated with complications (13.2 ml/kg/min in those with complications vs. 14.4 ml/kg/min in those without, t-test,  $P = 0.07$ ). ROC curve analysis estimated the predictive value of both  $VO_{2peak}$  and AT to be poor (i.e.,  $< 70$ ), AUC 0.63 (95% CI 0.50-0.76,  $P = 0.02$ ) and 0.62 (95% CI 0.49-0.75,  $P = 0.03$ ), respectively. The same group subsequently published a further study of 103 patients with both oesophageal and gastric cancers that underwent CPEX testing prior to oesophagectomy (62%) and gastrectomy (25). The findings were the reverse of their previous work, in that, this time; a lower AT was associated with cardiopulmonary complications (9.9 ml/kg/min in those with complications vs. 11.2 ml/kg/min in those without,  $P = 0.05$ ), while  $VO_{2peak}$  was not (16.6 ml/kg/min in those with complications vs. 14.6 ml/kg/min in those without,  $P = 0.07$ ). ROC analysis again reported both AT and  $VO_{2peak}$  to

be poorly predictive of complications (AUC 0.62 (95% CI 0.50-0.74,  $P = 0.06$ ) and 0.60 (95% CI 0.48-0.72,  $P = 0.08$ , respectively). The most significant limitations of all three studies, apart from their small sample sizes and single institution design are the potential for detection bias for complications due to unblinded outcome assessments, particularly for complications which can be subjectively diagnosed (e.g. pneumonia). This could lead to an over-estimate of the association between CPEX variables and outcomes. Such methodological error may explain why the same group were unable to replicate their previous findings (24, 25).

**Author's conclusion:** Associations between CPEX variables and outcome after oesophagectomy are from small retrospective observational studies that did not use a validated measure of postoperative outcomes, or capture complications with blinding to CPEX values. The absence of blinding could result in an inflation of the association between CPEX values and postoperative outcomes. Further large studies where complications are strictly defined and measured by assessors blinded to CPEX values are needed.

### Bladder Cancer

Thirty-five studies were identified, with 9 potentially relevant papers identified by their title. A review of these 9 abstracts identified 5 potentially relevant papers, but only 3 met the inclusion criteria. All were prospective hospital-based cohort studies, with a total of 256 patients from hospitals in the UK (27-29). The largest was of 105 patients who underwent preoperative CPEX testing prior to either robot-assisted ( $n=38$ ) or open ( $n=67$ ) cystectomy (28). Complications were measured at day 90 by blinded assessors and were associated with a significantly lower median AT (10.6 vs 11.8, U-test,  $P = 0.007$ ) and  $VO_{2peak}$  (14.3 vs 15.4, U-test,  $P = 0.02$ ) compared to those without complications. Additionally, VE/VECO<sub>2</sub> was higher in the complication group than in those without (33.3 vs 30.3, U-test  $P=0.007$ ). Whilst these findings are convincing in that there was consistency of associations across 3 CPEX values, the small study sample from a single institution presents a significant limitation to a more generalized interpretation of the results. The second largest study was of

82 patients who underwent CPEX prior to intracorporal robotic-assisted radical cystectomy (29). There were no associations between any CPEX variable and outcome. However, the small sample size and small number of complications ( $n=14$ ) likely result in a lack of statistical power to detect a clinically meaningful associations with CPEX. The smallest prospective cohort study was of 69 patients undergoing radical cystectomy. Again, no CPEX values were predictive of complications when the patients were divided into two groups, composed of those with and without a complication. However, sub-analysis according to the presence of a Clavien-Dindo grade  $\geq 3$  complication ( $n = 13$  vs  $n = 56$ ) found an inverse association between AT and major complication risk, OR 0.74 (95% CI 0.57-0.97). Again, the results of this single institution study with a small sample size are difficult to interpret, particularly when post-hoc analyses according to different complication severities were undertaken, which increase the risk of a chance finding in a small sub-group.

**Author's summary:** The evidence of an association between CPEX and post-cystectomy outcome is from small single-institution studies, which report conflicting findings. Therefore larger studies which limit sources of bias are required to clarify whether an association exists.

### Pancreatic Cancer

Thirty-one studies were identified, with 8 potentially relevant papers by title. A review of these 8 abstracts identified 3 potentially relevant papers and all 3 met the inclusion criteria (30-32). These were UK hospital based-cohort studies (2 retrospective, 1 prospective) with a total of 288 patients. None of the studies used blinded outcome assessments. The largest was a retrospective study of 124 patients who had CPEX testing prior to pancreaticoduodenectomy (30). Complications occurred in 44% of patients and were defined using POMS (33) and the International Study Group definition of Pancreatic Fistula (ISGPF) (37). There were no associations between  $VO_{2peak}$  and complications, including pancreatic leak. AT was dichotomised using a cut-point of 10.1 ml/kg/min (a value derived from their previous work) (38), and included in a multivariable logistic regression model, which

estimated that AT <10.1ml/kg/min greatly increased the odds of a pancreatic leak, OR=5.79 (95% CI 1.62-20.63). The imprecision of this estimate likely reflects both the dichotomisation of a continuous variable and the small pancreatic leak events in the total sample (n=29). The second largest study was a retrospective analysis of 100 patients that underwent preoperative CPEX and major pancreatic surgery (98% pancreaticoduodenectomy) (31). Again, AT was dichotomized, rather than treated as a continuous variable. The point of dichotomization (10 ml/kg/min) was not justified in the report and unclear if chosen a priori. The results showed a greater incidence of pancreatic leak (occurring in 25 patients), when AT was < 10ml/kg/min (35.4% vs. 16%, P = 0.028). However, statistical significance was lost when leaks were graded according to the ISGPF classification (P = 0.091). Selective reporting bias cannot be excluded, particularly when a seemingly arbitrary threshold was used to dichotomize a continuous variable. Furthermore, dichotomization during analysis results in a loss of statistical power and increases the risk of a false positive result (39). The final investigation was a prospective cohort study of 64 patients undergoing pancreaticoduodenectomy, which included a per-protocol statistical analysis (32). The authors reported no association between AT or VO<sub>2peak</sub> and all complications, OR 1.07 (95% CI 0.83-1.39) and OR 1.00 (95% CI 0.86-1.18), respectively.

**Author's summary:** The studies in pancreatic cancer have small patient numbers and potential sources of detection and selective reporting biases which makes interpretation of their findings difficult. Similar to other cancer resection surgeries, large well-designed studies are needed to clarify whether there is an association between fitness, as determined by CPEX testing, and outcome.

## DISCUSSION

Overall, the evidence for associations between preoperative CPEX values and postoperative outcome after cancer resection surgery is mostly derived from small observational studies. Many were underpowered to report on mortality risk. However, in the largest study of its kind, a low VO<sub>2peak</sub> was associated with mortality after lung

cancer resection, but only in thoracotomy surgery (not in VATS) (12). For morbidity, preoperative CPEX testing has poor (19, 24, 25) to average (20) discriminatory accuracy to predict postoperative outcomes after cancer resection surgery, so has limited utility as an isolated preoperative screening tool. Furthermore, investigations often used composite outcomes and included low Clavien-Dindo graded complications (19, 20, 23), which reduces the clinical meaningfulness of associations found (40). Much research in CPEX testing has been a continuation of the seminal work of Older et al, in a paper published in 1993; reporting that AT may predict postoperative cardiac-related death after major surgery (7). However, there seems to have been incorrect interpretations of this original plausible hypothesis. Cardiac-related death has a relationship with aerobic fitness; mortality events are a reflection of how patients respond once complications have occurred (41). Patients that die as a result of such complications are likely to lack the necessary cardiopulmonary and musculoskeletal reserves, which are required when there are ongoing physiological stresses and prolonged ITU bed-rest (41). These outcomes are very different to those measured in some of the CPEX studies in this review, which include surgical wound infection on postoperative day 3 (23) or increased analgesia need due to postoperative pain (as measured by POMS) (20, 33). Ideally, observational studies investigating cardiac death are required, but pragmatically may be difficult due to the small number of such events. However, to justify the pre-operative use of CPEX in cancer surgery, it should be shown that CPEX can accurately identify patients at risk of significant postoperative complications.

This is the first review to assess the utility of preoperative CPEX to predict postoperative complications after major cancer resection surgery. And, whilst it is plausible that greater cardiopulmonary and skeletal muscle reserves (as measured by CPEX) could reduce postoperative complications following major cancer resection surgery, evidence from the reviewed observational studies suggests that the effect size, if present at all, is likely to be small. This unintuitive finding may be explained by the size of the physiological insult associated with major cancer resec-

tion surgery. Taking oesophagectomy as an example; resection and reconstruction of the upper gastrointestinal tract results in a complication profile reflective of the operative field rather than fitness, where the cardiopulmonary system is directly affected. However, cardiorespiratory and musculoskeletal reserves may be critical to the ability of a patient to respond once a complication has occurred (41). This review high-

lights the need for large multi-center studies to assess the association between CPEX variables and mortality in cancer resection surgery. Until this evidence is available, CPEX testing in isolation is unlikely to meaningfully inform cancer surgery practice.

The authors declare no other conflicts of competing interest for this work.

## References

- Hemmings HC Jr, Wlody D, Mahajan R, Webster NR. The 2014 BJA/PGA special issue: a selection of six educational reviews. *Br J Anaesth* 2014;113 Suppl 2:ii1-2.
- Desborough JP. The stress response to trauma and surgery. *Br J Anaesth* 2000;85(1):109-17.
- Wahba RW. Perioperative functional residual capacity. *Can J Anaesth* 1991;38(3):384-400.
- Harter J, Orlandi SP, Gonzalez MC. Nutritional and functional factors as prognostic of surgical cancer patients. *Support Care Cancer* 2017;25(8):2525-30.
- Jack S, West MA, Raw D, Marwood S, Ambler G, Cope TM, et al. The effect of neoadjuvant chemotherapy on physical fitness and survival in patients undergoing oesophagectomy cancer surgery. *Eur J Surg Oncol* 2014;40(10):1313-20.
- 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):883-93.
- Older P, Smith R, Courtney P, Hone R. Preoperative evaluation of cardiac failure and ischemia in elderly patients by cardiopulmonary exercise testing. *Chest* 1993;104(3):701-4.
- Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ* 2005;330(7494):753.
- Hennis PJ, Meale PM, Grocott MP. Cardiopulmonary exercise testing for the evaluation of perioperative risk in non-cardiopulmonary surgery. *Postgrad Med J* 2011;87(1030):550-7.
- Moran J, Wilson F, Guinan E, McCormick P, Hussey J, Moriarty J, et al. Role of cardiopulmonary exercise testing as a risk-assessment method in patients undergoing intra-abdominal surgery: a systematic review. *Brit J Anaesth* 2016;116(2):177-91.
- Cancer Research UK. Cancer incidence for common cancers. (Accessed June 1, 2017, at <http://www.cancer-researchuk.org/health-professional/cancer-statistics/incidence/common-cancers-compared#collapseZero>).
- Begum SS, Papagiannopoulos K, Falcoz PE, Decaluwe H, Salati M, Brunelli A. Outcome after video-assisted thoracoscopic surgery and open pulmonary lobectomy in patients with low VO2 max: a case-matched analysis from the ESTS databasedagger. *Eur J Cardiothorac Surg* 2016;49(4):1054-8.
- Loewen GM, Watson D, Kohman L, Herndon JE 2nd, Shennib H, Kernstine K, et al. Preoperative exercise Vo2 measurement for lung resection candidates: results of Cancer and Leukemia Group B Protocol 9238. *J Thorac Oncol* 2007;2(7):619-25.
- Licker M, Schnyder JM, Frey JG, Diaper J, Cartier V, Ivan C, et al. Impact of aerobic exercise capacity and procedure-related factors in lung cancer surgery. *Eur Respir J* 2011;37(5):1189-98.
- Brunelli A, Belardinelli R, Refai M, Salati M, Soccia L, Pompili C, et al. Peak oxygen consumption during cardiopulmonary exercise test improves risk stratification in candidates to major lung resection. *Chest* 2009;135(5):1260-7.
- Torchio R, Guglielmo M, Giardino R, Ardisson F, Ciacco C, Gulotta C, et al. Exercise ventilatory inefficiency and mortality in patients with chronic obstructive pulmonary disease undergoing surgery for non-small-cell lung cancer. *Eur J Cardiothorac Surg* 2010;38(1):14-9.
- Brunelli A, Belardinelli R, Pompili C, Xiumé F, Refai M, Salati M, et al. Minute ventilation-to-carbon dioxide output (VE/VCO2) slope is the strongest predictor of respiratory complications and death after pulmonary resection. *Ann Thorac Surg* 2012;93(6):1802-6.
- West MA, Parry MG, Lythgoe D, Barben CP, Kemp GJ, Grocott MP, et al. Cardiopulmonary exercise testing for the prediction of morbidity risk after rectal cancer surgery. *Br J Surg* 2014;101(9):1166-72.
- 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):665-71.
- West MA, Asher R, Browning M, Minto G, Swart M, Richardson K, et al. Validation of preoperative cardiopulmonary exercise testing-derived variables to predict in-hospital morbidity after major colorectal surgery. *Br J Surg* 2016;103(6):744-52.
- Junejo MA, Mason JM, Sheen AJ, Moore J, Foster P, Atkinson D, et al. Cardiopulmonary exercise testing for preoperative risk assessment before hepatic resection. *Br J Surg* 2012;99(8):1097-104.
- Dunne DF, Jones RP, Lythgoe DT, Pilkington FJ, Palmer DH, Malik HZ, et al. Cardiopulmonary exercise testing before liver surgery. *J Surg Oncol* 2014;110(4):439-44.
- Kasivisvanathan R, Abbassi-Ghadi N, McLeod AD, Oliver A, Rao Baikady R, Jhanji S, et al. Cardiopulmonary exercise testing for predicting postoperative morbidity in patients undergoing hepatic resection surgery. *HPB (Oxford)* 2015;17(7):637-43.
- Forshaw MJ, Strauss DC, Davies AR, Wilson D, Lams B, Pearce A, et al. Is cardiopulmonary exercise testing a useful test before esophagectomy? *Ann Thorac Surg* 2008;85(1):294-9.
- Moyes LH, McCaffery CJ, Carter RC, Fullarton GM, Mackay CK, Forshaw MJ. Cardiopulmonary exercise testing as a predictor of complications in oesophagectomy cancer surgery. *Ann R Coll Surg Engl* 2013;95(2):125-30.
- Nagamatsu Y, Shima I, Yamana H, Fujita H, Shirouzu K, Ishitake T. Preoperative evaluation of cardiopulmonary reserve with the use of expired gas analysis during exercise testing in patients with squamous cell carcinoma of the thoracic esophagus. *J Thorac Cardiovasc Surg* 2001;121(6):1064-8.
- Prentis JM, Trenell MI, Vasdev N, French R, Dines G, Thorpe A, et al. Impaired cardiopulmonary reserve in an elderly population is related to postoperative morbidity and length of hospital stay after radical cystectomy. *BJU Int* 2013;112(2):E13-9.
- Tolchard S, Angell J, Pyke M, Lewis S, Dodds N, Darweish A, et al. Cardiopulmonary reserve as determined by cardiopulmonary exercise testing correlates with length of stay and predicts complications after radical cystectomy. *BJU Int* 2015;115(4):554-61.
- Lamb BW, Tan WS, Eneje P, Bruce D, Jones A, Ahmad I, et al. Benefits of robotic cystectomy with intracorporeal diversion for patients with low cardiorespiratory fitness: A prospective cohort study. *Urol Oncol* 2016;34(9):417.e17-23.
- Ausania F, Snowden CP, Prentis JM, Holmes LR, Jaques BC, White SA, et al. Effects of low cardiopulmonary reserve on pancreatic leak following pancreaticoduodenectomy. *Br J Surg* 2012;99(9):1290-4.
- Chandrabalan VV, McMillan DC, Carter R, Kinsella J, McKay CJ, Carter CR, et al. Pre-operative cardiopulmonary exercise testing predicts adverse post-operative events and non-progression to adjuvant therapy after major pancreatic surgery. *HPB (Oxford)* 2013;15(11):899-907.
- Junejo MA, Mason JM, Sheen AJ, Bryan A, Moore J, Foster P, et al. Cardiopulmonary exercise testing for preoperative risk assessment before pancreaticoduodenectomy for cancer. *Ann Surg Oncol* 2014;21(6):1929-36.
- Grocott MP, Browne JP, Van der Meulen J, Matejowski C, Mutch M, Hamilton MA, et al. The Postoperative Morbidity Survey was validated and used to describe morbidity after major surgery. *J Clin Epidemiol* 2007;60(9):919-28.
- Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg* 2004;240(2):205-13.
- The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions (Version 5.1.0). (Accessed May 15, 2018, at <http://handbook-5-1.cochrane.org/>).
- Ministry of Education, Culture, Sports, Science and Technology of Japan. Official statistics of the average height and weight of Japanese males. (Accessed June 1, 2017, at [http://www.mext.go.jp/b\\_menu/toukei/001/022/2004/002.pdf](http://www.mext.go.jp/b_menu/toukei/001/022/2004/002.pdf)).
- Bassi C, Dervenis C, Butturini G, Fingerhut A, Yeo C, Izbicki J, et al. Postoperative pancreatic fistula: an international study group (ISGPF) definition. *Surgery* 2005;138(1):8-13.
- 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):535-41.
- Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ* 2006;332(7549):1080.
- Lim E, Beckles M, Warburton C, Baldwin D. Cardiopulmonary exercise testing for the selection of patients undergoing surgery for lung cancer: friend or foe? *Thorax* 2010;65(10):847-9.
- Hall A, Older P. Cardiopulmonary exercise testing accurately predicts risk of major surgery including esophageal resection: letter 1. *Ann Thorac Surg* 2009;87(2):670-1.

## Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Stephen Lam, Andrew Hart. Cardiopulmonary Exercise Testing for Predicting Early Outcomes after Major Cancer Resection: A Systematic Review. *J Anesth Perioper Med* 2018;5:136-48. doi: 10.24015/JAPM.2018.0052

### Search terms

Search terms for cardiopulmonary exercise testing:

“cardiopulmonary exercise”, “CPEX”, “CPET”, “exercise testing”, “anaerobic threshold”, “VCO<sub>2</sub>”, “ventilatory inefficiency”, “oxygen consumption”, “VO<sub>2</sub>”, “preoperative exercise”, “aerobic exercise”.

Search terms for outcomes: “morbidity”, “mortality”, “outcome”, “complication”.

Additional search terms for lung cancer: “lung cancer surgery”, “lung resection”, “lobectomy”, “pneumonectomy”.

Additional search terms for colorectal cancer: “colorectal”, “colon”, “colectomy”, “rectal”, “rectum”, “surgery”, “resection”.

Additional search terms for colorectal cancer: “hepatic”, “hepatectomy”, “liver surgery”.

Additional search terms for oesophageal cancer: “oesophagectomy”, “oesophagogastrectomy”.

Additional search terms for bladder cancer: “bladder cancer”, “cystectomy”.

Additional search terms for pancreatic cancer: “pancreatic surgery”, “pancreatic resection”, “pancreatoduodenectomy”.

An example for the whole search string for colorectal cancer: “(((cardiopulmonary exercise).ti,ab OR (cpex).ti,ab OR (cpet).ti,ab OR (exercise testing).ti,ab OR (anerobic threshold).ti,ab OR (VCO<sub>2</sub>).ti,ab OR (ventilatory inefficiency).ti,ab OR (oxygen consumption).ti,ab OR (VO<sub>2</sub>).ti,ab OR (preoperative exercise).ti,ab OR (aerobic exercise).ti,ab) AND (((colorectal).ti,ab OR (colon).ti,ab OR (colectomy).ti,ab OR (rectal).ti,ab OR (rectum).ti,ab) AND ((surgery).ti,ab OR (resection).ti,ab))) AND ((complications).ti,ab OR (outcome).ti,ab OR (morbidity).ti,ab OR (mortality).ti,ab)”

ti=title, ab=abstract