Received: 27 January 2023 | Accepted: 5 March 2023

DOI: 10.1111/eci.13981

ORIGINAL ARTICLE

WILEY

Assessing cardiorespiratory fitness relative to sex improves surgical risk stratification

George A. Rose[1] | Richard G. Davies[1,2] | Jared Torkington[3] | Ronan M. G. Berg[1,4,5,6] | Ian R. Appadurai[2] | David C. Poole[7] | Damian M. Bailey[1]

[1]Neurovascular Research Laboratory, Faculty of Life Sciences and Education, University of South Wales, Pontypridd, UK

[2]Department of Anaesthetics, University Hospital of Wales, Cardiff, UK

[3]Department of Surgery, University Hospital of Wales, Cardiff, UK

[4]Department of Biomedical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[5]Department of Clinical Physiology and Nuclear Medicine, University Hospital Copenhagen—Rigshospitalet, Copenhagen, Denmark

[6]Centre for Physical Activity Research, University Hospital Copenhagen—Rigshospitalet, Copenhagen, Denmark

[7]Departments of Kinesiology, Anatomy and Physiology, Kansas State University, Manhattan, Kansas, USA

Correspondence
Damian M. Bailey, Neurovascular Research Laboratory, Alfred Russel Wallace Building, Faculty of Life Sciences and Education, University of South Wales, Pontypridd CF37 4AT, UK.
Email: damian.bailey@southwales.ac.uk

Funding information
Higher Education Funding Council for Wales; Japan Society for the Promotion of Science, Grant/Award Number: #JSPS/OF317; Royal Society International Exchanges Award, Grant/Award Number: IES\R2\192137; Royal Society Wolfson Research Fellowship, Grant/Award Number: #WM170007

Abstract

Background: To what extent sex-related differences in cardiorespiratory fitness (CRF) impact postoperative patient mortality and corresponding implications for surgical risk stratification remains to be established.

Methods: To examine this, we recruited 640 patients (366 males vs. 274 females) who underwent cardiopulmonary exercise testing prior to elective colorectal surgery. Patients were defined high risk if peak oxygen uptake was $<14.3\ mL\,kg^{-1}\,min^{-1}$ and ventilatory equivalent for carbon dioxide at 'anaerobic threshold' >34. Between-sex CRF and mortality was assessed, and sex-specific CRF thresholds predictive of mortality was calculated.

Results: Seventeen percent of deaths were attributed to sub-threshold CRF, which was higher than established risk factors for cardiovascular disease (CVD). The group (independent of sex) exhibited a 5-fold higher mortality (high vs. low risk patients hazard ratio = 4.80, 95% confidence interval 2.73–8.45, $p < 0.001$). Females exhibited 39% lower CRF ($p < 0.001$) with more classified high risk than males (36 vs. 23%, $p = 0.001$), yet mortality was not different ($p = 0.544$). Upon reformulation of sex-specific CRF thresholds, lower cut-offs for mortality were observed in females, and consequently, fewer (20%) were stratified with sub-threshold CRF compared to the original 36% ($p < 0.001$).

Conclusions: Low CRF accounted for more deaths than traditional CVD risk factors, and when CRF was considered relative to sex, the disproportionate number of females stratified unfit was corrected. These findings support clinical

consideration of 'sex-specific' CRF thresholds to better inform postoperative mortality and improve surgical risk stratification.

KEYWORDS
cardiopulmonary exercise test, cardiorespiratory fitness, colorectal surgery, sex, survival

1 | INTRODUCTION

The routine determination of cardiorespiratory fitness (CRF) employing cardiopulmonary exercise testing (CPET) is increasingly accepted in patient populations undergoing major surgery, given its capacity to improve risk prediction.[1–3] Indeed, CRF is considered a vital clinical sign presenting health professionals with unique opportunities to improve patient management and care.[4] Preoperative CPET is used to identify patients with low CRF which is an established predictor of survival following major intra-abdominal surgery.[5–9] When comparing the sexes, female patients demonstrate lower preoperative CRF than males.[10] However, to what extent sex differences in CRF influence preoperative risk stratification and corresponding 'normalization' approaches to optimize the latter has not previously been investigated.

Evidence demonstrates that CRF is lower in females compared to males across the lifespan given smaller body size, skeletal muscle mass, peak cardiac output, haemoglobin concentration, conducting airways, greater oxygen cost of breathing, and accelerated vascular ageing.[11–14] Current practice frequently adopts thresholds of CRF to stratify patients as 'high risk' for colorectal surgery (combination of anaerobic threshold [AT] <11 mL oxygen [O_2] kg^{-1} min^{-1} and ventilatory equivalent for carbon dioxide [$\dot{V}_E/\dot{V}CO_2$-AT] >34)[9,15–17] and it is recommended that thresholds evolve, particularly for heterogenous surgical populations.[18,19] However, when considering patient sex, and if a simple dose–response relationship exists between CRF and postoperative survival, females would intuitively be expected to be at increased risk given the physiological constraints. Evidence suggests otherwise as differences in post-operative outcomes following major colorectal surgery remain equivocal, with limited evidence highlighting lower[20] or not different[21] mortality with males. Therefore, sex differences in CPET metrics should also be considered during surgical risk stratification, and the existence of different thresholds in relation to sex requires investigation.

To address this knowledge gap, we retrospectively analysed a cross-sectional population-based cohort of male and female patients who underwent CPET prior to major colorectal surgery. We tested the following hypotheses. First, that CRF would account for disproportionately more deaths compared to other, more traditional CVD risk factors reinforcing its significance as an independent risk factor. Second, compared to males, females would be characterized by lower CRF and corresponding higher mortality, justifying the need to reassess the current (indiscriminate one size fits all) CPET metrics to surgical risk stratification to further optimize patient management and care.

2 | METHODS

2.1 | Ethics

The University of South Wales Ethics Committee (LSE1636GREO), and Cardiff and Vale University Health Board (15/AIC/6352) approved the analysis. Written informed consent was waived as this study constituted a service evaluation. Procedures were conducted in accordance with guidelines set forth by the Declaration of Helsinki of the World Medical Association[22] with the exception of registration in a database.

2.2 | Design

This study constituted a retrospective cross-sectional population-based observational analysis of anonymized longitudinal hospital databases prospectively populated at a single NHS tertiary referral centre (University Hospital of Wales, Cardiff, UK), in accordance with the STROBE statement.[23] A central CPET database was collated with merged data from colorectal surgeons' databases, critical care databases, and longitudinal mortality records from the Office for National Statistics (ONS). Males and females were considered for inclusion based upon chronological order of visit date logged in the CPET database.

2.3 | Patients

2.3.1 | Inclusion/exclusion criteria

Consecutive patients (n = 1110, 618 males/492 females) who underwent CPET at a preoperative assessment clinic between January 2010 and December 2016 were considered for inclusion (Figure 1). Patients were excluded from final analysis if there was no record of surgery (n = 174) or if surgery was performed more than 365 days following

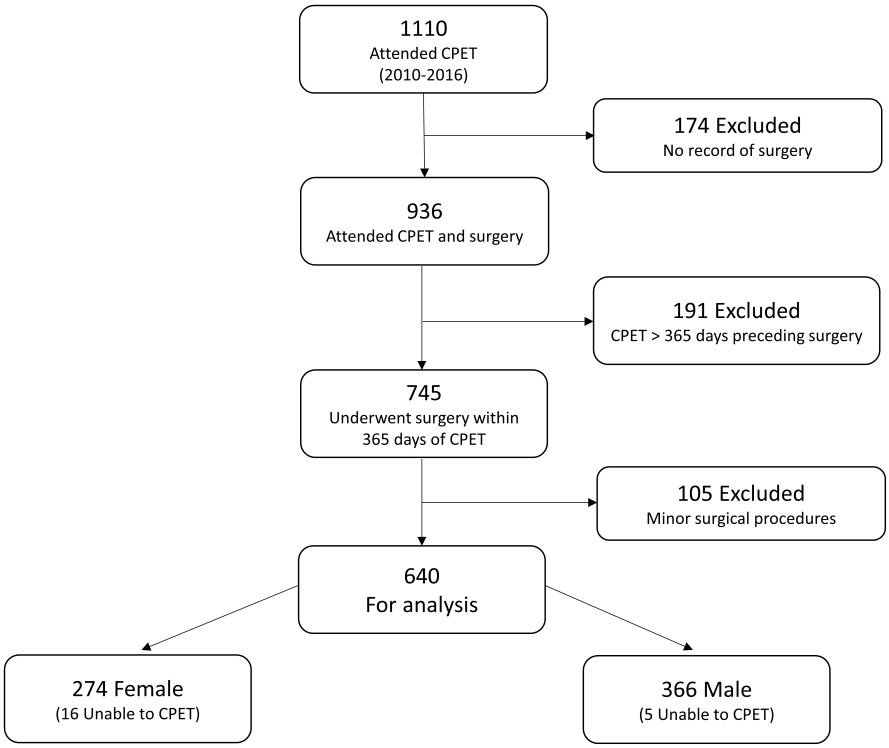

FIGURE 1 Patient flow. Consort diagram.[23] CPET, cardiopulmonary exercise test.

CPET ($n = 191$). In the case of multiple CPET visits, the visit closest to subsequent surgery was included. Patients undergoing minor or exploratory procedures were excluded ($n = 105$). Inability to undergo CPET was defined as having completed static pulmonary function tests but inability to perform cycle ergometry with expiratory gas analysis which was primarily due to musculoskeletal complications ($n = 21$; 6 males vs. 15 females).

2.3.2 | Demographics

Patient demographics including risk factors were gathered from medical notes at the preoperative evaluation by the consultant anaesthetist and included age, body mass index (BMI), cancer diagnosis, smoking history, hypertension (HTN), diabetes, ischemic heart disease (IHD), chronic obstructive pulmonary disease (COPD), cerebrovascular accident (CVA), statin use, renal disease, anaemia, and American Society of Anesthesiologists (ASA) physical status classification.[24] Only variables with less than 20% missing data were included.

2.4 | Measurements

2.4.1 | CRF

Following clinical examination and flow loop spirometry, preoperative CPET was conducted using an electromagnetically-braked cycle ergometer (Lode, Gronigen, The Netherlands) and a Medgraphics Ultima metabolic cart (MedGraphics™, Gloucester, UK) as previously described.[25–27] Calibration was conducted in accordance with the manufacturer's guidelines using a 3-litre volume syringe (Hans Rudolph) and reference calibration gases. During data collection, the middle five of seven breaths were averaged. Following 3 min of resting data collection, patients cycled at 60 revolutions per minute for 3 min in an unloaded freewheeling state. A progressively ramped period of exercise (5–15 W min^{-1} based on body mass, stature, age and sex) was then undertaken to volitional exhaustion or symptom limited termination, followed by a three-minute recovery period.[28] The Medgraphics Breeze™ software automatically determined peak oxygen uptake ($\dot{V}O_2$ peak) as the highest $\dot{V}O_2$ during the final 30 s of exercise. Oxygen uptake efficiency slope (OUES)[29] and peak oxygen pulse (O_2 pulse) were also automatically determined by the software. The AT was manually interpreted by the supervising consultant using the V-slope method,[30] also confirmed by a second consultant, and supported by comparison of ventilatory equivalents for oxygen ($\dot{V}_E/\dot{V}O_2$) and carbon dioxide ($\dot{V}_E/\dot{V}CO_2$) plots. Subcategories 'Unable to CPET' and 'AT not detected' were recorded if patients were unable to perform a CPET because of their clinical status, or if insufficient data was available for clear identification of the AT.

At this centre, patients were classified 'high risk' if CPET indicated an AT <11 mL O_2 kg^{-1} min^{-1} combined

with $\dot{V}_E/\dot{V}CO_2$ -AT >34 because patients with reduced AT or ventilatory deficiency are more likely to die after major intra-abdominal surgery, whether or not they have associated cardiovascular risk factors.[9,15–17] However, the AT is subject to high levels of natural (mostly biological) variation and requires caution during interpretation.[25] Therefore, optimized metrics of CRF were calculated to improve patient risk stratification. Patients unable to perform CPET, typically characterized by poor postoperative outcome,[6] were also classified as high risk.

2.4.2 | Surgical procedures

Procedures were obtained from the surgical database and included right or left hemicolectomy, transverse, sigmoid or subtotal colectomy, anterior or abdominoperineal resection and Hartmann's procedures. Open or laparoscopic technique were recorded. Major procedures outside these categories (such as panproctocolectomy) were classified as 'Other'. The approaches were classified to allow for comparison as surgery per se carries arguably the greatest risk.

2.4.3 | Postoperative morbidity

Postoperative morbidity was recorded by analysis of the surgical and critical care databases detailing length of stay for hospital, high dependency and intensive therapy unit, cardiopulmonary complications, all complications, postoperative destinations, and returns to theatre. The extent of complications were retrospectively used to construct Clavien-Dindo scores.[31]

2.4.4 | Postoperative mortality

Postoperative mortality was determined by review of ONS records calculated by comparison of surgery date with 30-day, 90-day, 1-year, and 2-year follow-up status, tracking National Bowel Cancer Audit reference points.[32] Mortality at 1 year was selected as the primary endpoint. Whilst 90-day mortality may be the preferred surgical perspective, sufficient events for meaningful survival analysis was unlikely. Furthermore, mortality relating to the surgical response may continue for 1 year.[33]

2.5 | Analysis

Statistical analysis was conducted using IBM SPSS Statistics for Windows (Version 28.0, Armonk, NY). Distribution normality was confirmed using repeated Shapiro–Wilk W tests. Continuous data are presented as mean (standard deviation), and dichotomous variables as number (percentage). Significance for all two-tailed tests was established at $p < 0.050$.

2.5.1 | Demographics

Between-sex differences were established using independent samples t-tests for continuous data, or χ^2 tests for frequency counts.

2.5.2 | Pooled data

Prognostic utility of CPET

First, we determined whether CRF was associated with postoperative mortality using Cox proportional hazards (PHs) regression models. The PH assumption was tested with Schoenfeld residuals.[34] Continuous and dichotomous variables, selected by clinical acumen and without missing values, were assessed using univariable Cox PH regression. This included age and BMI as demographic risk factors, CVA, IHD, HTN, COPD, statin therapy and diabetes as clinical risk factors, and $\dot{V}O_2$ peak, $\dot{V}O_2$-AT, $\dot{V}_E/\dot{V}CO_2$-AT, $\dot{V}_E/\dot{V}O_2$-AT, O_2 Pulse, OUES, Workload at AT, and peak Workload as CPET metrics. Subsequent multivariable Cox PH models were developed with variables included if $p < 0.20$ (from univariable analysis). Variables were removed if collinearity was identified by correlation coefficients >0.80.[35] A backward stepwise approach was employed with exclusion at $p < 0.10$ that yielded a final model with independent predictors of postoperative mortality. This process identified 'optimized' predictor variables for comparison with current practice.

We also determined the weighting that each risk factor contributed towards mortality using the concept of attributable fractions (AF), the estimated fraction of all deaths that would not have occurred had there been no exposure to that risk factor.[36] AF was calculated using the following equation[37]:

$$AF = P_c\,(1 - 1/HR)$$

where P_c = the prevalence of exposure among cases and HR = hazard ratio for the exposure.

HRs were calculated for 1-year postoperative mortality adjusted for age, BMI, smoking history, IHD, COPD, HTN, and CVA from a separate Cox Proportional Hazards (PH) model.

Long term survival was compared between dichotomized CRF strata (high risk vs. low risk defined from the

Cox PH model) and examined graphically using Kaplan–Meier plots with log-rank tests.

2.5.3 | Sex comparisons

CRF

Mean values for peak workload, oxygen uptake at AT, $\dot{V}O_2$ peak, $\dot{V}_E/\dot{V}CO_2$ –AT, O_2 pulse and OUES were compared between sexes using independent samples *t*-tests.

Postoperative outcome

A primary unadjusted Cox PH model was used to compare female and male survival with log-rank test for differences in survival curves. χ^2 tests were used to compare frequency counts for mortality, postoperative destinations, complications, and return to theatre between the sexes. Length of hospital stay was reported by median values with inter quartile range (IQR) assessed using Mann–Whitney *U* tests.

Defining sex-specific fitness thresholds

Area under receiver operating characteristic (AUROC) curves were constructed for markers of CRF identified as predictors of postoperative mortality. For a marker of CRF to be considered a valid predictor of mortality, an AUROC of >0.7 was used.[38] Threshold values for markers of CRF fulfilling this criterion were subsequently calculated for pooled, male, and female patients by examination of the minimal distance between AUROC plots and the upper left corner, optimizing sensitivity and specificity.

To assess the impact of sex-specific thresholds on the prediction of 1-year postoperative mortality, three risk stratification models were created. Each model used the Cox PH multivariable analysis previously described, but with scale metrics of CRF replaced with a single dichotomised value indicating patient risk (high or low). Model 1 reflected current practice (combination of AT <11 mL O_2 kg^{-1} min^{-1} and $\dot{V}_E/\dot{V}CO_2$-AT >34) for baseline comparison. Model 2 included optimized metrics of CRF defined from the original multivariable Cox PH survival model to dichotomise patient risk. Model 3 included optimized metrics of CRF from Model 2, but with threshold values (re)defined relative to sex.

3 | RESULTS

3.1 | Patient flow

For the final analysis, we included 640 out of 1110 patients examined (Figure 1). Females accounted for 43% of the sample (n = 274). Twenty-one patients were unable to perform CPET, including 16 females with two dying within 90 days of the CPET visit date. No serious cardiovascular complications were reported during any of the CPET visits.

3.2 | Demographics

Males were characterized by higher BMI, proportion of smokers and prevalence of diabetes, IHD and COPD, whereas ASA scores were not different (Table 1). More patients underwent laparoscopically assisted surgery compared to open surgery with equal distribution between the sexes. In contrast, more females underwent right hemicolectomy (39% vs. 27%, p = 0.002) and fewer underwent anterior resection (24% vs. 38%, p = <0.001).

3.3 | CRF and mortality

Forty-nine deaths occurred within 1 year of surgery (7.7%). The Cox PH multivariable model identified $\dot{V}O_2$ peak (HR 0.87 [0.80–0.95], p = 0.002) and $\dot{V}_E/\dot{V}CO_2$-AT (HR 1.05 [1.00–1.11], p = 0.037) as independently associated with mortality, alongside BMI (HR 0.91 [0.85–0.98], p = 0.006). Threshold values were defined at <14.3 mL kg^{-1} min^{-1} for $\dot{V}O_2$ peak and >34 for $\dot{V}_E/\dot{V}CO_2$-AT indicating patients with high risk. Using these optimized metrics, 149 (23%) of patients were considered high risk and exhibited 4.8-fold increased mortality at 1 year (Figure 2) relative to the low-risk group. Compared with traditional CVD risk factors, high risk patients who exhibited a combination of $\dot{V}O_2$ peak <14.3 mL kg^{-1} min^{-1} and $\dot{V}_E/\dot{V}CO_2$-AT >34 (adjusted for CVD risk factors) accounted for the largest proportion (17%) of deaths. They also exhibited more postoperative cardiopulmonary complications (HR 3.82 [95% CI 2.11–6.92], p < 0.001) and longer hospital stays (9 [IQR 8] vs. 7 [7] days, p = 0.011).

3.4 | Sex comparisons

3.4.1 | CRF

Females exhibited lower CRF (Figure 3) with reductions of 39% (70 vs. 115 W, p < 0.001) for work-load peak, 7% (10.5 vs.11.3 mL O_2 min^{-1} kg^{-1}, p = 0.001) for AT, 18% (14.5 vs. 17.6 mL kg^{-1} min^{-1}, p < 0.001) for $\dot{V}O_2$ peak, 6% (35 vs. 33, p = 0.007) for $\dot{V}_E/\dot{V}O_2$-AT, 34% (8 vs. 12 mL $beat^{-1}$, p < 0.001) for O_2 pulse, and 29% (1321 vs. 1864 [(mL min^{-1} O_2)/(L min^{-1} $\dot{V}_E$)], p < 0.001) for OUES.

TABLE 1 Demographics.

	Pooled	Male	Female	*p*-Value
Sample size, *n* (%)	640	366 (57)	274 (43)	
Age, mean (range), years	69 (23–90)	69 (23–90)	70 (24–89)	0.877
Risk factors				
BMI, mean (SD), $kg\,m^{-2}$	28.2 (5.6)	28.7 (5.4)	27.5 (5.9)	**0.009**
Cancer, *n* (%)	540 (84)	314 (86)	226 (83)	0.254
Smoker, *n* (%)	396 (62)	254 (70)	142 (52)	**<0.001**
Hypertension, *n* (%)	332 (52)	197 (54)	135 (50)	0.274
Diabetes, *n* (%)	114 (18)	79 (22)	35 (13)	**0.004**
IHD, *n* (%)	108 (17)	76 (21)	32 (12)	**0.002**
COPD, *n* (%)	139 (22)	91 (25)	48 (18)	**0.028**
CVA, *n* (%)	49 (8)	33 (9)	16 (6)	0.135
Statin, *n* (%)	265 (41)	166 (45)	99 (36)	**0.019**
Haemoglobin, mean (SD), $g\,L^{-1}$	128 (19)	133 (19)	122 (17)	**<0.001**
Creatinine, mean (SD), $\mu mol\,L^{-1}$	79 (25)	86 (28)	70 (18)	**<0.001**
ASA, *n* (%)				
I	19 (3)	9 (3)	10 (4)	0.481
II	382 (61)	215 (60)	167 (63)	0.625
III	219 (35)	132 (37)	87 (33)	0.274
IV	6 (1)	4 (1)	2 (1)	0.704
Surgery				
Method, *n* (%)				
Laparoscopic	336 (53)	185 (51)	151 (55)	0.264
Procedure, *n* (%)				
Right hemicolectomy	205 (32)	99 (27)	106 (39)	**0.002**
Transverse hemicolectomy	4 (1)	3 (1)	1 (0)	0.639
Left hemicolectomy	20 (3)	13 (4)	7 (3)	0.503
Subtotal colectomy	13 (2)	6 (2)	7 (3)	0.573
Anterior resection	205 (32)	140 (38)	65 (24)	**<0.001**
Hartmann's procedure	47 (7)	27 (7)	20 (7)	1.000
APR	47 (7)	24 (7)	23 (8)	0.444
Sigmoid colectomy	45 (7)	27 (7)	18 (7)	0.756
Other	54 (8)	27 (7)	27 (10)	0.315

Note: *p*-values refer to male vs. female comparisons.

Abbreviations: APR, abdominoperineal resection; ASA, American Society of Anesthesiologists Physical Status Classification; BMI, body mass index; COPD, chronic obstructive pulmonary disease; CVA, cerebrovascular accident; IHD, ischemic heart disease.

Significance for bold values = <0.050

3.4.2 | Postoperative outcome

Survival curve analysis demonstrated no difference ($p = 0.509$) in unadjusted postoperative mortality up to 5 years follow-up between the sexes. Similarly, Table 2 demonstrated no sex differences for all postoperative outcomes: 30-day, 90-day, 1-year, 2-year mortality, postoperative destination, complications, Clavien Dindo scores, returns to theatre, and length of stay ($p = > 0.050$ in all cases).

3.4.3 | Sex-specific fitness thresholds

AUROC curves (Table 3) indicated $\dot{V}O_2$ peak, $\dot{V}_E/\dot{V}CO_2$-AT, Peak workload, and OUES were predictive of 1-year mortality for pooled and respective sex-specific data (AUROC >0.7, $p < 0.050$ in all cases). Interestingly, cut-points indicative of increased mortality risk occurred at lower levels of CRF in females than males for all predictive markers ($\dot{V}_E/\dot{V}CO_2$-AT >35 vs. >33, $\dot{V}O_2$ peak 11.9 vs.

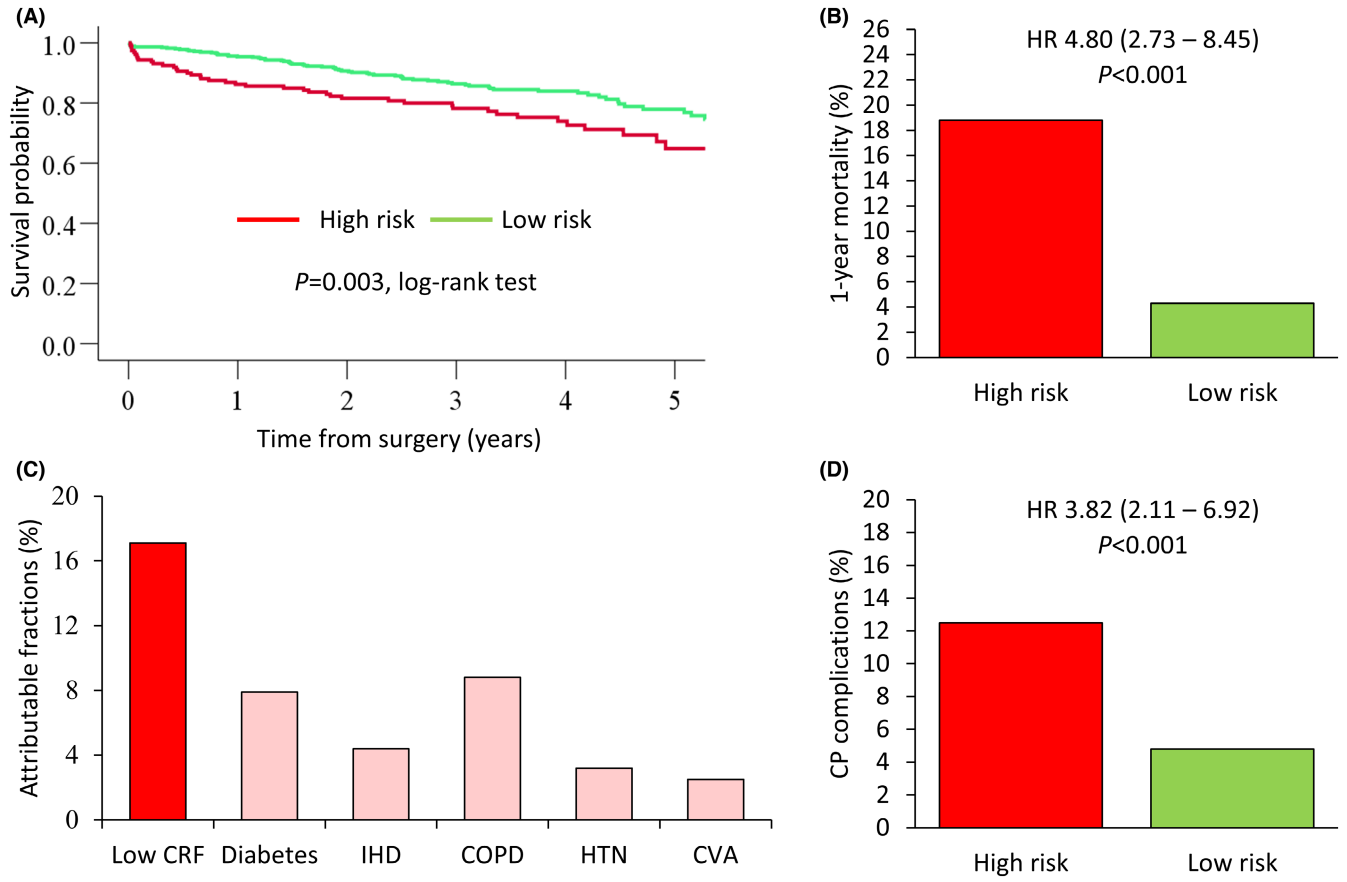

FIGURE 2 Cardiorespiratory fitness and postoperative outcome. Patients who were high risk (combination of peak oxygen uptake [$\dot{V}O_2$ peak] <14.3 mL kg^{-1} min^{-1} and ventilatory equivalent for carbon dioxide at the anaerobic threshold [$\dot{V}_E/\dot{V}CO_2$-AT] >34) for surgery were characterized by higher mortality rates. (A) Kaplan–Meier plot highlighting reduced survival in high risk (red) patients over a 5-year follow-up period after surgery; $p = 0.003$ log-rank test vs. low risk (green) patients. (B) Hazard ratio (HR) for 1-year postoperative mortality (95% confidence interval) demonstrating a 4.8-fold greater risk of mortality in high-risk patients. (C) Attributable fractions for 1-year postoperative mortality adjusted for age, body mass index, smoking history, and risk factors. COPD, chronic obstructive pulmonary disease; CVA, cerebrovascular accident; HTN, hypertension controlled by medication; IHD, ischemic heart disease. Low cardiorespiratory fitness (Low CRF) defined by combination of $\dot{V}O_2$ peak <14.3 mL kg^{-1} min^{-1} and $\dot{V}_E/\dot{V}CO_2$-AT >34, adjusted for all other CVD risk factors, accounted for the largest proportion (17%) of deaths. (D) The HR for postoperative cardiopulmonary (CP) complications demonstrated a 3.8-fold greater risk if patients were high risk.

14.9 mL kg^{-1} min^{-1}, Peak workload <50 vs. <100 W, and OUES <1148 vs. <1681).

Table 4 demonstrates current practice (Model 1) and newly optimized CPET metrics (Model 2) predictive of 1-year mortality, that more females were stratified high risk than males (36 vs. 23%, $p < 0.001$, and 32 vs. 17%, $p < 0.001$). Conversely, with fitness stratified by optimized sex-specific thresholds (Model 3), females and males were equally classified high risk (20% vs. 21%, $p = 0.702$), thus correcting the disparity. This occurred with improved specificity (74%–82%) albeit a small reduction in sensitivity (59%–55%). The predictive utility of each model improved AUROC curve values (Model 1: 0.67 [0.58–0.75], Model 2: 0.69 [0.61–0.78], and Model 3: 0.70 [0.61–0.78]) and the refined cut-points provided greater discrimination of risk with hazard ratios increasing (Model 1: HR 3.61 [2.05–6.36], Model 2: HR 4.80 [2.73–8.45], and Model 3: HR 5.13 [2.92–9.01]).

4 | DISCUSSION

This study has identified two novel findings. First, CRF was identified as the principal independent risk factor for postoperative mortality in patients undergoing major colorectal surgery with high risk classification accounting for disproportionately more deaths compared to other, more established CVD risk factors. Second, optimized threshold metrics of CRF predictive of postoperative mortality were identified to be sex specific highlighting the need to reassess the current indiscriminate approach to surgical risk stratification that overestimates high risk classification in

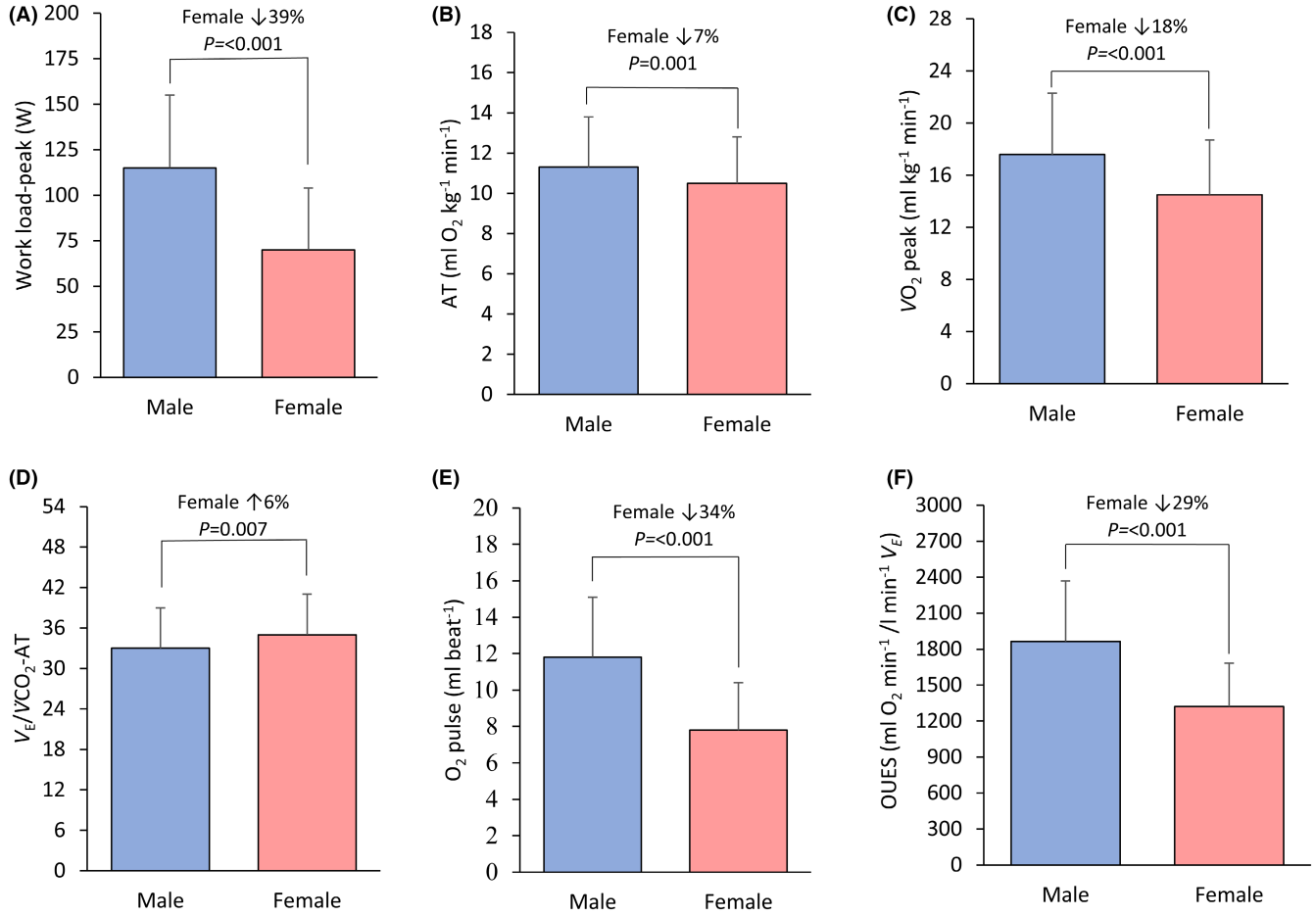

FIGURE 3 Sex-specific cardiopulmonary responses to exercise. Females were characterized by lower test values for: (A) Workload at peak exercise, (B) Oxygen uptake at anaerobic threshold (AT), (C) Peak oxygen uptake ($\dot{V}O_2$ peak), (D) Ventilatory equivalent for carbon dioxide at anaerobic threshold ($\dot{V}_E/\dot{V}CO_2$-AT), (E) Oxygen pulse at peak exercise (O_2 pulse), (F) Oxygen uptake efficiency slope (OUES).

females. Collectively, these findings reinforce the clinical significance of CRF and justify the need to personalize surgical risk stratification according to sex to further optimize patient management and care.

4.1 | CRF and mortality

High risk classification was associated with poor postoperative outcome following major colorectal surgery with lower survival at all postoperative timepoints, including a disproportionate 4.8-fold increased risk of 1-year mortality, a 3.8-fold greater risk of cardiopulmonary complications, and a 2-day increased hospital length of stay. This supports existing evidence in patients undergoing major colorectal surgery.[6,8,9,17] Novel application of attributable fractions highlighted that having low CRF was responsible for 17% of deaths, greater than any of the traditional CVD risk factors. This highlights low CRF as the most important independent risk factor for postoperative mortality, similar to its relationship with 'all-cause' mortality outside of the perioperative setting.[31] Importantly, we considered the confounding potential of advancing age and corresponding reduction in CRF that has an annual mortality HR of 1.06.[39] Our attributable fractions were thus age-adjusted, and furthermore, no age differences existed between the sexes.

When considering the CPET metrics employed, Wasserman and McIlroy's development of the AT was, in part, predicated on concerns that distinct patient populations either would not, or should not, be required to undergo the exhaustive exercise protocol requisite for measurement of 'maximal' oxygen uptake ($\dot{V}O_2$ max).[40] Subsequently it has become evident that good standard-of-practice for verifying that an accurate $\dot{V}O_2$ max has been achieved requires at least one additional exhaustive test, ideally at a higher work rate than achieved on the initial CPET.[41] Thus, what is typically measured by CPET is referred to as the $\dot{V}O_2$ peak; defined simply as the highest $\dot{V}O_2$ achieved on the test. Given the opportunity for underestimating the true $\dot{V}O_2$ max,[42] it is surprising that the coefficient of variation (CV) estimated for $\dot{V}O_2$ peak (~6%–7%) is less than AT (~7%–10%)

TABLE 2 Postoperative outcomes.

	Pooled ($n = 640$)	Male ($n = 366$)	Female ($n = 274$)	p-Value
Mortality, *n* (%)				
30-day	16 (2.5)	10 (2.7)	6 (2.2)	0.664
90-day	19 (3)	12 (3.3)	7 (2.6)	0.646
1-year	49 (7.7)	26 (7.1)	23 (8.4)	0.544
2-year	78 (12)	44 (12)	34 (12)	0.882
Destination, *n* (%)				
Ward	460 (72)	261 (71)	199 (73)	0.714
PACU	80 (13)	44 (12)	36 (13)	0.672
HDU	75 (12)	46 (13)	29 (11)	0.440
ITU	25 (4)	15 (4.1)	10 (3.6)	0.772
Complications, *n* (%)				
All	260 (41)	156 (43)	104 (38)	0.234
Cardiopulmonary	45 (7)	27 (7.4)	18 (6.6)	0.693
Clavien Dindo, *n* (%)				
0	371 (58)	204 (56)	167 (61)	0.186
I	35 (5.5)	20 (5.5)	15 (5.5)	0.996
II	151 (24)	90 (25)	61 (22)	0.493
IIIa	9 (1.4)	4 (1.1)	5 (1.8)	0.437
IIIb	42 (6.6)	25 (6.8)	17 (6.2)	0.752
IVa	11 (1.7)	9 (2.5)	2 (0.7)	0.096
IVb	2 (0.3)	1 (0.3)	1 (0.4)	0.839
V	19 (3)	12 (3.3)	7 (2.5)	0.593
RTT, *n* (%)	63 (10)	38 (10.4)	25 (9.1)	0.589
LoS, days (median, IQR)				
Hospital	8 (7)	8 (8)	8 (7)	0.427
HDU	2 (0)	2 (0)	2 (0)	0.873
ITU	3 (4.25)	3 (4)	2 (17)	0.937

Note: No between-sex differences were observed for any postoperative outcome. *p*-values refer to male versus female comparisons.

Abbreviations: HDU, high dependency unit; ITU, Intensive therapy unit; LoS, length of stay; PACU, post anaesthetic care unit; RTT, returns to theatre.

across patient cohorts undergoing colorectal surgery.[25] That we observed elevated Hazard Ratios for $\dot{V}_E/\dot{V}CO_2$-AT, despite higher CVs, may well reflect that this threshold occurs at a lower absolute $\dot{V}O_2$ (than $\dot{V}O_2$ peak) and thus that the CVs reported reflects a lower absolute CV in terms of mL O_2 min^{-1}. It is also pertinent that the $\dot{V}O_2$ max–$\dot{V}O_2$ peak difference, for the patient populations considered herein, has not been determined.

4.2 | Sex differences

Lower preoperative CRF in females was reflected across all CPET metrics and agrees with previous findings,[11,12] the likely consequence of comparatively smaller skeletal muscle mass and lower convective/diffusive O_2 transport. The smaller lung size of females accompanied by smaller conducting airways relative to males, predispose greater mechanical ventilatory constraints to exercise hyperpnea, and at a given ventilation, women have a higher oxygen cost of breathing.[14] Haemoglobin concentration was lower in females, closely agreeing with a reported reduction of 12% compared with males.[43] Peak cardiac output is lower in females who, even after indexing for body size, have smaller left ventricular chambers and accordingly lower stroke volumes.[13] Vascular ageing is also accelerated in females, characterized by greater arterial elastance, higher pulse pressure, smaller and stiffer aortic arches, and earlier wave reflection than men, independent of body size and heart rate.[44]

TABLE 3 Area under receiver operating characteristic (AUROC) curves for preoperative CPET variables predictive of 1-year mortality. AUROC curves demonstrated female cut-points occurred at lower levels of cardiorespiratory fitness than males. AUROC curves are presented with 95% confidence intervals. Cut-points were calculated for variables with AUROC >0.7 and $p < 0.05$. There were no significant differences in between-sex AUROC curves ($p > 0.05$ in all cases).

	Pooled			Male			Female		
	AUROC	Cut-point	Sn/Sp (%)	AUROC	Cut-point	Sn/Sp (%)	AUROC	Cut-point	Sn/Sp (%)
AT (mL O_2 kg^{-1} min^{-1})	0.62 (0.53–0.71)	–	–	0.59 (0.47–0.71)	–	–	0.66 (0.53–0.79)	–	–
$\dot{V}_E/\dot{V}CO_2$-AT	0.74 (0.66–0.82)	>34	71/63	0.76 (0.66–0.86)	>33	85/65	0.71 (0.59–0.83)	>35	74/66
$\dot{V}O_2$ peak (mL kg^{-1} min^{-1})	0.72 (0.64–0.79)	<14.3	67/63	0.70 (0.60–0.80)	<14.9	62/67	0.76 (0.65–0.86)	<11.9	65/75
Peak workload (W)	0.72 (0.65–0.79)	<91	72/54	0.75 (0.67–0.84)	<100	80/66	0.78 (0.67–0.89)	<50	62/76
O_2 pulse (mL $beat^{-1}$)	0.70 (0.62–0.79)	<8.5	55/78	0.68 (0.56–0.80)	–	–	0.80 (0.71–0.89)	<6.5	74/71
OUES (mL min^{-1} O_2)/(L min^{-1} $\dot{V}_E$)	0.73 (0.65–0.80)	<1461	72/60	0.74 (0.64–0.83)	<1681	85/63	0.82 (0.74–0.89)	<1148	91/65

Abbreviations: AT, anaerobic threshold; O_2 pulse, oxygen pulse at peak oxygen consumption; OUES, oxygen uptake efficiency slope; Sn, sensitivity; Sp, specificity; $\dot{V}_E/\dot{V}CO_2$, ventilatory equivalent for carbon dioxide; $\dot{V}O_2$ peak, peak oxygen consumption.

The lower CRF in females culminated in a greater proportion stratified high risk compared to males. This finding agreed with a previous study in major colorectal surgery.[9] If a presumed existence of a dose–response relationship between CRF and postoperative survival, we would have expected higher mortality in females given physiological constraints in convective/diffusive O_2 transport. However, to the contrary, our data indicate no difference in (i.e., not elevated) mortality (and morbidity) between the sexes, highlighting the need to consider other CVD risk factors that ultimately contribute to the 'collective' disease burden.

Females might experience increased sex-specific 'sensitivity' to CRF reflecting greater reliance on the systemic vascular protection afforded by CRF[45] given the lower prevalence of CVD risk factors compared to (more conditioned) male counterparts. Indeed, the collective CVD burden was higher in males and therefore may have counteracted the vascular protective benefits of elevated CRF, culminating in similar survival. However, it is important to emphasize that surgical intervention was less complicated in females given fewer anterior resections and more right hemicolectomies[8] that may have also contributed to improved postoperative outcome. Furthermore, statin-use was more prevalent in the male cohort which may also counteract 'traditional CVD risk factors' observed in males independent of, or in combination with, the effects of CRF. Albeit in a different cohort, the combination of high CRF and statin use can reduce the risk of events more than high CRF or statins alone.[46]

4.3 | Preoperative risk re-stratification

These findings justify a reappraisal of current CRF thresholds employed for preoperative risk stratification, highlighting the importance of defining sex-specific thresholds to better predict postoperative outcome. Consequently, when fitness was stratified using the newly defined sex-specific thresholds, fewer females were classified high risk compared with current practice, which has important implications for surgical risk stratification and perioperative care.

Indeed, this may result in an overly risk-adverse approach with excess resource utilization (e.g., HDU beds) and associated cost implications. Our postoperative destination data did not indicate any significant between-sex differences in HDU usage; however, we were not able to distinguish between electively prepared HDU provision and perioperative decision-based allocation. Therefore, we advise caution if HDU provision is allocated based on CPET alone and reinforce the view that CPET, should be used in combination with clinical review, individualized assessment and shared decision making,[19] not in isolation.

TABLE 4 Fitness thresholds predictive of 1-year postoperative mortality. Reformulation of sex-specific CRF thresholds with corresponding lower 'cut-offs' corrected the excess identification of high risk females (vs. males) in current practice whilst preserving mortality prediction.

Model	Thresholds indicating high risk	High risk patients identified *n* (%) Male/Female	AUROC (95% CI)	Sensitivity/ specificity (%)	Mortality (%) High risk/low risk	Hazard ratio (95% CI) versus low risk
1 Current practice	AT<11 mL O_2.kg^{-1}.min^{-1} $\dot{V}_E/\dot{V}CO_2$-AT>34	83 (23)/99 (36)[a]	0.67 (0.58–0.75)	59/74	16/4[a]	3.61 (2.05–6.36)
2 Optimized metrics	$\dot{V}O_2$ peak < 14.3 mL kg^{-1} min^{-1} $\dot{V}_E/\dot{V}CO_2$-AT > 34	61 (17)/88 (32)[a]	0.69 (0.61–0.78)	57/79	19/4[a]	4.80 (2.73–8.45)
3 Optimized metrics, sex-specific	Male $\dot{V}O_2$ peak<14.9 mL kg^{-1} min^{-1} $\dot{V}_E/\dot{V}CO_2$-AT > 33 Female $\dot{V}O_2$ peak<11.9 mL kg^{-1} min^{-1} $\dot{V}_E/\dot{V}CO_2$-AT > 35	78 (21)/55 (20)	0.70 (0.61–0.78)	55/82	20/4[a]	5.13 (2.92–9.01)

Abbreviations:AT, anaerobic threshold; AUROC, area under receiver operating characteristic curve; $\dot{V}_E/\dot{V}CO_2$-AT, ventilatory equivalent for carbon dioxide at anaerobic threshold; $\dot{V}O_2$ peak, peak oxygen consumption.

[a] $p < 0.05$ for between-sex or between-mortality risk differences.

Importantly, fewer females were stratified as high risk when reformulated CRF thresholds with corresponding lower 'cut-offs' were applied. These findings are consistent with females undergoing heart transplantation who, despite having lower CRF ($\dot{V}O_2$ peak), exhibited improved 1-year transplant-free survival,[47] further reinforcing the need for sex-specific CRF thresholds that extend to other major surgeries. Given the recent adoption of multimodal prehabilitation strategies with exercise training identified as the most fundamental component,[48] our findings redefine 'target thresholds' of CRF relative to sex and highlight the increased sensitivity and importance of CRF for females in whom lower levels of CRF may respond more favourably to exercise intervention. In clinical practice, age-specific considerations are well recognized and the preoperative CRF of a 60 year old would not be expected to equal that of an 80 year old. Sex differences should thus be considered in a similar manner.

5 | CONCLUSIONS

This study highlights CRF as the principal independent risk factor for postoperative mortality; being high risk accounted for disproportionately more deaths than other, more traditional CVD risk factors. These findings have facilitated the assignment of 'sex-specific' CRF thresholds that may better inform surgical risk stratification to further optimize clinical decision-making and patient care.

AUTHOR CONTRIBUTION

All authors contributed to study design. R.G.D and I.R.A performed CPET tests. Analysis was performed by G.A.R. The manuscript was drafted by G.A.R, R.M.G.B, and D.M.B. All authors provided revisions and approved the final version. D.M.B and G.A.R had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of data analysis.

ACKNOWLEDGEMENTS

D.M.B. was supported by a Royal Society Wolfson Research Fellowship (#WM170007), Royal Society International Exchanges Award (IES\R2\192137), the Japan Society for the Promotion of Science (#JSPS/OF317), and Higher Education Funding Council for Wales. D.M.B. is Editor-in-Chief of Experimental Physiology, Chair of the Life Sciences Working Group, member of the Human Spaceflight and Exploration Science Advisory Committee to the European Space Agency, member of the Space Exploration Advisory Committee to the UK Space Agency, member of the National Cardiovascular Network for Wales and South East Wales Vascular Network. D.M.B. is affiliated to the

companies FloTBI Inc., BrainEx Inc., and OrgEx Inc. focused on the technological development of novel biomarkers of brain injury in humans.

CONFLICT OF INTEREST STATEMENT

The authors declare no support or financial relationships with any organizations that might have an interest in the submitted work.

DATA AVAILABILITY STATEMENT

Most of the data are presented in the main manuscript and the tables and figures of the paper. The data that support the findings of this study are available upon reasonable request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

George A. Rose https://orcid.org/0000-0002-9598-6372

TWITTER

Richard G. Davies @rgdavies70
Ronan M. G. Berg @ronan_berg
Ian R. Appadurai @corticalsilence
Damian M. Bailey @USW_Oxygen

REFERENCES

1. Guazzi M, Arena R, Halle M, Piepoli MF, Myers J, Lavie CJ. Focused update: clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations. *Circulation*. 2016;133:e694-e711.
2. Rose GA, Davies RG, Appadurai IR, et al. 'Fit for surgery': the relationship between cardiorespiratory fitness and postoperative outcomes. *Exp Physiol*. 2022;107:787-799.
3. Balady GJ, Arena R, Sietsema K, et al. Clinician's guide to cardiopulmonary exercise testing in adults: a scientific statement from the American Heart Association. *Circulation*. 2010;122:191-225.
4. Ross R, Blair SN, Arena R, et al. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. *Circulation*. 2016;134:e653-e699.
5. Moran J, Wilson F, Guinan E, McCormick P, Hussey J, Moriarty J. Role of cardiopulmonary exercise testing as a risk-assessment method in patients undergoing intra-abdominal surgery: a systematic review. *Br J Anaesth*. 2016;116:177-191.
6. Lai CW, Minto G, Challand CP, 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;111:607-611.
7. West MA, Parry MG, Lythgoe D, et al. Cardiopulmonary exercise testing for the prediction of morbidity risk after rectal cancer surgery. *Br J Surg*. 2014;101:1166-1172.
8. West MA, Asher R, Browning M, et al. Validation of preoperative cardiopulmonary exercise testing-derived variables to predict in-hospital morbidity after major colorectal surgery. *Br J Surg*. 2016;103:744-752.
9. Wilson RJ, 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;105:297-303.
10. Thomas G, West MA, Browning M, et al. Why women are not small men: sex-related differences in perioperative cardiopulmonary exercise testing. *Perioper Med*. 2020;9:18-24.
11. Jackson AS, Sui X, Hébert JR, Church TS, Blair SN. Role of lifestyle and aging on the longitudinal change in cardiorespiratory fitness. *Arch Intern Med*. 2009;169:1781-1787.
12. Fleg JL, Morrell CH, Bos AG, et al. Accelerated longitudinal decline of aerobic capacity in healthy Older adults. *Circulation*. 2005;112:674-682.
13. Beale AL, Meyer P, Marwick TH, Lam CSP, Kaye DM. Sex differences in cardiovascular pathophysiology: why women are overrepresented in heart failure with preserved ejection fraction. *Circulation*. 2018;138:198-205.
14. Sheel AW, Dominelli PB, Molgat-Seon Y. Revisiting dysanapsis: sex-based differences in airways and the mechanics of breathing during exercise. *Exp Physiol*. 2016;101:213-218.
15. Older P, Hall A, Hader R. Cardiopulmonary exercise testing as a screening test for perioperative management of major surgery in the elderly. *Chest*. 1999;116:355-362.
16. Carlisle J, Swart M. Mid-term survival after abdominal aortic aneurysm surgery predicted by cardiopulmonary exercise testing. *Br J Surg*. 2007;94:966-969.
17. Wilson RJT, Yates DRA, Walkington JP, Davies SJ. Ventilatory inefficiency adversely affects outcomes and longer-term survival after planned colorectal cancer surgery. *Br J Anaesth*. 2019;123:238-245.
18. Older P, Levett DZH. Cardiopulmonary exercise testing and surgery. *Ann Am Thorac Soc*. 2017;14:S74-S83.
19. Levett DZH, Jack S, Swart M, et al. Perioperative cardiopulmonary exercise testing (CPET): consensus clinical guidelines on indications, organization, conduct, and physiological interpretation. *Br J Anaesth*. 2018;120:484-500.
20. Byrne BE, Mamidanna R, Vincent CA, Faiz O. Population-based cohort study comparing 30 and 90-day institutional mortality rates after colorectal surgery. *Br J Surg*. 2013;100:1810-1817.
21. Mayr FB, Rucker D, Warkentin LM, Huynh H, Khadaroo RG. Sex differences in the treatment and outcome of emergency general surgery. *PLoS One*. 2019;14:e0224278.
22. Williams JR. The declaration of Helsinki and public health. *Bull World Health Organ*. 2008;86:650-652.
23. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Int J Surg*. 2014;12:1495-1499.
24. Saklad M. Grading of patients for surgical procedures. *Anesthesiology*. 1941;2:281-284.
25. Rose GA, Davies RG, Davison GW, et al. The cardiopulmonary exercise test grey zone; optimising fitness stratification by application of critical difference. *Br J Anaesth*. 2018;120:1187-1194.
26. Rose GA, Adamson MJ, Davies RG, Appadurai IR, Bailey DM. High-intensity exercise training improves perioperative risk stratification in the high-risk patient. *Physiol Rep*. 2020;8:e14409.

27. Rose GA, Davies RG, Appadurai IR, et al. Cardiorespiratory fitness is impaired and predicts mid-term postoperative survival in patients with abdominal aortic aneurysm disease. *Exp Physiol*. 2018;103:1505-1512.
28. Wasserman K. *Principles of Exercise Testing and Interpretation: Including Pathophysiology and Clinical Applications*. 5th ed. Wolters Kluwer; 2012.
29. Hollenberg M, Tager IB. Oxygen uptake efficiency slope: an index of exercise performance and cardiopulmonary reserve requiring only submaximal exercise. *J Am Coll Cardiol*. 2000;36:194-201.
30. Beaver WL, Wassermen K, Whipp BJ. A new method for detecting anaerobic threshold by gas exchange. *J Appl Physiol*. 1986;60:2020-2027.
31. Dindo D, Demartines N, Clavien P-A. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg*. 2004;240:205-213.
32. NBOCA. National Bowel Cancer Audit Annual Report. 2017.
33. Dekker JW, Van Den Broek CB, Bastiaannet E, Van De Geest LG, Tollenaar RA, Lieffers GJ. Importance of the first postoperative year in the prognosis of elderly colorectal cancer patients. *Ann Surg Oncol*. 2011;18:1533-1539.
34. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*. 1994;81:515-526.
35. Grant SW, Hickey GL, Wisely NA, et al. Cardiopulmonary exercise testing and survival after elective abdominal aortic aneurysm repair. *Br J Anaesth*. 2015;114:430-436.
36. Blair SN. Physical inactivity: the biggest public health problem of the 21st century. *Br J Sports Med*. 2009;43:1-2.
37. Mansournia MA, Altman DG. Population attributable fraction. *Br Med J*. 2018;360:k757.
38. Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. 3rd ed. John Wiley & Sons; 2013.
39. Carlisle JB. Risk prediction models for major surgery: composing a new tune. *Anaesthesia*. 2019;74(Suppl 1):7-12.
40. Wasserman K, McIlroy MB. Detecting the threshold of anaerobic metabolism in cardiac patients during exercise. *Am J Cardiol*. 1964;14:844-852.
41. Poole DC, Jones AM. Measurement of the maximum oxygen uptake $\dot{V}O_2$max: $\dot{V}O_2$peak is no longer acceptable. *J Appl Physiol*. 2017;122:997-1002.
42. Poole DC, Wilkerson DP, Jones AM. Validity of criteria for establishing maximal O_2 uptake during ramp exercise tests. *Eur J Appl Physiol*. 2008;102:403-410.
43. Murphy WG. The sex difference in haemoglobin levels in adults — mechanisms, causes, and consequences. *Blood Rev*. 2014;28:41-47.
44. Redfield MM, Jacobsen SJ, Borlaug BA, Rodeheffer RJ, Kass DA. Age- and gender-related ventricular-vascular stiffening: a community-based study. *Circulation*. 2005;112:2254-2262.
45. Calverley TA, Ogoh S, Marley CJ, et al. HIITing the brain with exercise: mechanisms, consequences and practical recommendations. *J Physiol*. 2020;598:2513-2530.
46. Kokkinos PF, Faselis C, Myers J, Panagiotakos D, Doumas M. Interactive effects of fitness and statin treatment on mortality risk in veterans with dyslipidaemia: a cohort study. *Lancet*. 2013;9864:394-399.
47. Elmariah S, Goldberg LR, Allen MT, Kao A. Effects of gender on peak oxygen consumption and the timing of cardiac transplantation. *J Am Coll Cardiol*. 2006;47:2237-2242.
48. Scheede-Bergdahl C, Minnella EM, Carli F. Multi-modal prehabilitation: addressing the why, when, what, how, who and where next? *Anaesthesia*. 2019;74(Suppl 1):20-26.

How to cite this article: Rose GA, Davies RG, Torkington J, et al. Assessing cardiorespiratory fitness relative to sex improves surgical risk stratification. *Eur J Clin Invest*. 2023;00:e13981. doi:10.1111/eci.13981