Bariatric surgery leads to weight loss, reduces the prevalence of obesity-related illnesses (type 2 diabetes, hypertension), and increases short-term quality of life. Current guidelines from the UK National Institute for Clinical Excellence (NICE) recommend bariatric surgery as a treatment option for those with BMI > 40 kg m⁻², or 35–40 kg m⁻² in the presence of other risk factors such as hypertension or type 2 diabetes. However, the benefits of bariatric surgery should be weighed against the risks of the procedure, which include a 30 day mortality rate of 0.3% and a serious adverse surgical outcome rate of 4.3%. There is no established method for predicting such postoperative complications.

For several types of non-cardiac surgery, cardiopulmonary exercise testing (CPET) is used to assess patients' physical fitness and to identify individuals at increased risk of adverse perioperative outcome. The peak oxygen consumption (VO₂ peak) has been reported to have a moderate capacity to predict complications after bariatric surgery. However, this study used treadmill exercise to conduct
CPET rather than cycle ergometry, which is more widely used in the preoperative setting.\textsuperscript{2} Further, these data have yet to be replicated, or extended to other measures of functional capacity.

We aimed to assess whether CPET-derived physical fitness variables determined during cycle ergometry were predictive of adverse outcome after gastric bypass surgery.

**Methods**

**Patient population**

A prospective evaluation of the utility of the CPET service in predicting surgical outcome was approved as a service evaluation by the local research and development departments and by the local ethics committee, and permitted combination of CPET data with our on-going audit of surgical morbidity. All patients provided informed consent to perform CPET. Over an 18 month period (September 01, 2009–February 25, 2011), all patients aged \( \geq 18 \) yr undergoing elective gastric bypass surgery at the Whittington Hospital NHS Trust who had a preoperative CPET were included. Patients with a body mass over 190 kg were not eligible for CPET due to weight restrictions of equipment used to transfer patients in the event of medical emergencies.

**Cardiopulmonary exercise testing**

CPET was conducted at the Whittington Hospital NHS Trust in accordance with the American Thoracic Society/American College of Chest Physicians guidelines.\textsuperscript{7} In brief, patients performed a symptom-limited, incremental ramp protocol to volitional exhaustion using a semi-recumbent electromagnetically braked cycle ergometer (Corival Recumbent, Lode, Groningen, The Netherlands), during which on-line breath-by-breath gas exchange analysis and 12-lead ECG monitoring were conducted (Cortex Metalyzer 3B, Biophysik, Leipzig, Germany), and peripheral oxygen saturation (Sp\textsubscript{O2}) and arterial pressure were measured (Cortex MultiLyser, Biophysik).

After CPET, data were analysed to derive the following variables: ventilatory/anaerobic threshold (AT), peak oxygen consumption (\( \dot{V}O_2 \) peak), and ventilatory equivalent for carbon dioxide (\( V\dot{E}/VCO_2 \)). The \( \dot{V}O_2 \) peak was defined as the average of the highest exertional oxygen uptake achieved over the last 30 s of exercise.\textsuperscript{8} The AT was determined by an exercise physiologist and a consultant physician experienced in CPET interpretation, using the modified V-slope method,\textsuperscript{9} confirmed by patterns of change in ventilatory equivalent and end-tidal gas measurements.\textsuperscript{10} The \( \dot{V}O_2 \) and AT values were adjusted for body mass (ml kg\textsuperscript{-1} min\textsuperscript{-1}), and normalized for predicted adipose tissue free mass (PATFM) (ml kg\textsuperscript{-1} min\textsuperscript{-1}).\textsuperscript{11} The \( V\dot{E}/VCO_2 \) was recorded as the value measured at AT.\textsuperscript{12} Results from CPET were available to the surgical team before operation.

Other variables recorded were age, sex, height, and weight. Patients’ medical histories were used to complete the Lee’s revised cardiac risk index (RCRI),\textsuperscript{13} and the ASA physical status score\textsuperscript{14} was recorded by the attending anaesthetist before surgery.

**Outcome measures**

For the duration of the study, all patients who were referred for CPET and underwent gastric bypass surgery were followed-up after operation. The primary outcome variables were postoperative morbidity (PostOperative Morbidity Survey; POMS) recorded as present (\( + \)) or absent (\( – \)) on postoperative day 5\textsuperscript{15} and postoperative length of stay (LOS). The secondary outcome variables were postoperative length of critical care unit (CCU) stay, in-hospital mortality, and readmission to hospital. Those patients who left hospital before postoperative day 5 were assumed to have no POMS-defined complications.\textsuperscript{15}

**Statistical analysis**

SPSS 17.0 (SPSS Inc., Chicago, IL, USA) was used for all statistical analyses. Statistical significance was set at \( P<0.05 \). Normal (Gaussian) distribution of data was verified using the Kolmogorov–Smirnov test and visual inspection of histogram charts. Non-normally distributed data were transformed and analysed using parametric statistics where possible, and using appropriate non-parametric tests when not possible. Normally distributed data are presented as mean and 95% confidence intervals (CIs) or standard deviation (so), and non-normally distributed data as median and inter-quartile (IQ) range.

Patients were separated into groups according to surgical outcome [POMS\( + \) vs POMS\( – \), LOS\( <\) median (3 days) vs LOS\( >\) median (3 days), CCU LOS\( <\) median (1 day) vs CCU LOS\( >\) median (1 day), readmitted vs not readmitted]. Differences in CPET and non-CPET variables between these groups were compared using independent samples t-test for continuous, normally distributed data and Fisher’s exact test for categorical data.

The capacity of CPET- and non-CPET-derived markers to predict outcome variables was determined using receiver operating characteristic (ROC) curve analysis, and the area under the curve (AUC) was compared with a value of 0.50. The optimum sensitivity and specificity of the predictor variables was selected as the value closest to the top left of the ROC curve chart.

The G’power 3 power analysis program and Fisher’s exact test protocol was used to determine the population sample size required to evaluate the utility of CPET in our practice.\textsuperscript{16} Using the odds ratio (OR) for postoperative morbidity between high and low fitness groups determined by McCullough and colleagues\textsuperscript{5} (OR=6.77), an \( \alpha \)-value of 0.05, a power of 0.80, and a group allocation ratio of 1, the required total sample size was calculated to be 68. We decided that OR values below 6.77 would be clinically important and adjusted the OR used in the sample size calculation to reflect this. Using an OR of 4.00 and the same input data as above, the calculated sample size required 104 patients.
Results

Over the period of data collection, 121 patients with a body mass <190 kg were referred for CPET and underwent elective gastric bypass surgery. A total of 106 of these patients successfully completed the CPET and were included in the main study population. The patients not included in the main study population failed to attend their appointment (n=8), could not exercise due to mobility problems (n=6), or had a contraindication to CPET (unstable angina, n=1).

Excluded patients were not different from those included with regard to gender, age, or height, but did have a greater body mass (P=0.034) (Table 1).

The mean VO₂ peak of the 106 patients was 15.7 (15.1–16.4) ml kg⁻¹ min⁻¹. The associations between the VO₂ peak and surgical outcome were assessed using these 106 patients. Eleven patients ceased exercise before an AT could be determined. In the remaining 95 patients, the mean AT was 11.0 ml kg⁻¹ min⁻¹ (10.6–11.3) and the Vₑ/VO₂ at AT was 25.9 (25.5–26.3). The median ASA score of 106 patients was II (IQ range=1), and 47 of the 106 patients had a score above II. The median RCRI was 1 (IQ range=1), and 29 patients had a value above the median.

Surgical outcome

After surgery, 83 patients were electively admitted to a CCU as part of routine care, while 23 patients were admitted to a ward. Patients admitted to the CCU did not differ from ward patients in terms of body mass (P=0.448), height (P=0.636), risk indices (RCRI, P=0.602; ASA, P=0.349), or CPET variables (AT, P=0.930; peak VO₂, P=0.294; Vₑ/VO₂, P=0.975). Patients admitted to the CCU tended to be older (mean difference, 4.6 yr; 95% CI, −0.69 to 9.27; P=0.053). The frequency of patients POM+ on day 5 was 11% (n=12). The median postoperative LOS was 3 days (IQ range=2) and the median length of CCU stay was 1 day (IQ range=1). Nine (8.5%) patients were readmitted to hospital within 30 days of discharge. No patients died in hospital.

Morbidity

The AT was significantly lower in patients with postoperative complications than in those without (P=0.049); while VO₂ peak and Vₑ/VO₂ were not different between the two groups (P=0.217 and 0.381, respectively) (Table 2). The area under the ROC curve for AT as a predictor of postoperative morbidity was 0.693 (0.510–0.875) (P=0.058). The optimal AT to predict morbidity was 11.0 ml kg⁻¹ min⁻¹, which had a sensitivity of 78% and a specificity of 54%. The area under the ROC curve was not significantly different from 0.50 for the VO₂ peak (AUC, 0.592; 95% CI, 0.397–0.787; P=0.300) or Vₑ/VO₂ (AUC, 0.587; 95% CI, 0.355–0.818; P=0.395). Body mass and BMI at the time of CPET did not differ between those with or without morbidity on postoperative day 5 (P=0.416 and 0.667, respectively). Similarly, body mass and BMI at the time of surgery did not differ between the two groups (P=0.404 and 0.714, respectively). The risk of postoperative morbidity was not predicted by ASA score (OR, 1.131; 95% CI, 0.335–3.821; P=1.00) or RCRI score (OR, 2.121; 95% CI, 0.436–10.31; P=0.502).

Postoperative LOS

The AT of patients with a LOS>3 days was lower than those with a LOS<3 days (P=0.023) (Table 2). The AT was a significant predictor of LOS>3 day, with an area under the ROC curve of 0.640 (0.524–0.756) (P=0.030). The optimal AT to identify individuals with increased likelihood of having a LOS>3 days was 11.4 ml kg⁻¹ min⁻¹ (sensitivity, 76%; specificity, 53%). The VO₂ peak and Vₑ/VO₂ of patients with a LOS>3 days were not different from those who had a LOS<3 days (P=0.094 and 0.170, respectively). The ORs for increased LOS for ASA grade and RCRI score were 2.37 (95% CI, 1.04–5.39; P=0.042) and 1.444 (95% CI, 0.601–3.469; P=0.496), respectively. Body mass and BMI, both at the time of CPET and on the day of surgery, were not different according to postoperative LOS category.

CCU LOS and readmission to hospital

AT, VO₂ peak, and Vₑ/VO₂ were not different between those with a CCU LOS>1 and those with a CCU LOS≤1 day or between those readmitted to hospital and those not. ASA and RCRI scores did not influence CCU LOS (P=0.490 and 1.000, respectively) or readmission to hospital (P=0.074 and 0.710, respectively). Body mass and BMI, both at the time of CPET and on the day of surgery, were not different according to CCU LOS category or in patients readmitted to hospital.

CPET variables normalized to PATFM

The mean AT and median peak VO₂, expressed as a function of PATFM, were 28.8 and 40.8 ml kg⁻¹ min⁻¹, respectively. The AT and the peak VO₂ did not have the capacity to predict any of the outcome variables when expressed as a function of PATFM.

Undeterminable AT and surgical outcome

Eleven patients had a CPET where the AT could not be determined. The gender, age, height, and weight of these patients were not different from the rest of the study population.

---

**Table 1** Physical characteristics of the populations

<table>
<thead>
<tr>
<th>Patients who completed CPET (n=106)</th>
<th>Patients referred for CPET but not included in analysis (n=15)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (male/female)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (yr)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height (m)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body mass (kg)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---
operative morbidity (POMS). However, the surgery, preoperative AT determined during CPET predicted substantial event rate (LOS and/or mobility. A high ASA score was also associated with a greater risk of increased LOS and increased the precision of risk prediction when used in conjunction with AT.

Comparison with other studies

This study adds to the growing literature supporting the use of CPET as a preoperative risk stratification tool for general surgical procedures, and is only the second study assessing the association between CPET and surgical outcome after bariatric surgery. McCullough and colleagues published the first data to suggest a role for CPET in the risk assessment of patients undergoing gastric bypass surgery and using ROC analysis identified a VO2 peak of 15.8 ml kg\(^{-1}\) min\(^{-1}\) as the optimal risk prediction threshold. Our data agree with this earlier study in that both studies found objective cardiorespiratory fitness variables-predicted surgical outcome. However, the two studies differ regarding the specific CPET variable that identifies risk, with the current study agreeing better with the bulk of the general surgery perioperative literature in suggesting that AT is the better variable to predict outcome, rather than the VO2 peak. Our results demonstrate that AT was associated with both morbidity (POMS) and postoperative LOS, and the optimal AT values found to predict these outcomes were similar (morbidity, 11.0 ml kg\(^{-1}\) min\(^{-1}\); LOS, 11.4 ml kg\(^{-1}\) min\(^{-1}\)) to the thresholds identified in previous studies.

It is surprising, given that our study population comprised morbidly obese patients, that our results are in such close agreement with perioperative CPET studies conducted on 'normal' weight patients which have identified an AT of 11 ml kg\(^{-1}\) min\(^{-1}\) to be the optimal predictor of risk. Nonetheless, from a clinical perspective, it is helpful to have a single threshold AT value to categorize patients into a higher risk group. It is also interesting that the current study found an association between surgical outcome and AT when adjusted according to body mass but not when it was normalized to PATFM. McCullough and colleagues found the association between surgical outcome and the VO2 peak was strongest.

### Discussion

#### Main findings

In a cohort of bariatric patients undergoing gastric bypass surgery, preoperative AT determined during CPET predicted prolonged hospital LOS, and was associated with postoperative morbidity (POMS). However, the VO2 peak and \(V_E/VCO_2\) at AT, which predict operative risk for a range of other surgery types, were not associated with postoperative outcome after gastric bypass surgery. Further, the risk of having extended LOS was increased when AT was undeterminable, which occurs due to reduced effort, compliance, and/or mobility. A high ASA score was also associated with a greater risk of increased LOS and increased the precision of risk prediction when used in conjunction with AT.

### Table 2

<table>
<thead>
<tr>
<th>POMS day 5</th>
<th>POMS – (n=94)*</th>
<th>POMS+ (n=12)*</th>
<th>Mean difference (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT (ml kg(^{-1}) min(^{-1}))</td>
<td>11.1 (1.7)</td>
<td>9.9 (1.5)</td>
<td>1.2 (0.0 – 2.3)</td>
<td>0.049</td>
</tr>
<tr>
<td>Peak VO2 (ml kg(^{-1}) min(^{-1}))</td>
<td>15.9 (3.3)</td>
<td>14.6 (3.8)</td>
<td>1.2 (0.7 – 3.3)</td>
<td>0.217</td>
</tr>
<tr>
<td>(V_E/VCO_2) at AT</td>
<td>25.9 (2.0)</td>
<td>26.5 (2.7)</td>
<td>0.8 (0.8 – 2.1)</td>
<td>0.381</td>
</tr>
<tr>
<td>Body mass at CPET</td>
<td>126.0 (18.0)</td>
<td>130.6 (20.9)</td>
<td>4.6 (6.4 – 15.8)</td>
<td>0.416</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Postoperative LOS</th>
<th>LOS≤3 days (n=70)*</th>
<th>LOS&gt;3 days (n=36)*</th>
<th>Mean difference (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT (ml kg(^{-1}) min(^{-1}))</td>
<td>11.3 (1.8)</td>
<td>10.4 (1.4)</td>
<td>0.9 (0.1 – 1.6)</td>
<td>0.023</td>
</tr>
<tr>
<td>Peak VO2 (ml kg(^{-1}) min(^{-1}))</td>
<td>16.1 (3.2)</td>
<td>15.0 (3.6)</td>
<td>1.1 (0.2 – 2.5)</td>
<td>0.094</td>
</tr>
<tr>
<td>(V_E/VCO_2) at AT</td>
<td>25.7 (2.0)</td>
<td>26.4 (2.1)</td>
<td>0.6 (0.3 – 1.5)</td>
<td>0.170</td>
</tr>
<tr>
<td>Body mass at CPET</td>
<td>124.8 (19.5)</td>
<td>130.0 (15.6)</td>
<td>5.2 (2.2 – 12.7)</td>
<td>0.165</td>
</tr>
</tbody>
</table>

\(P=1.00, 0.930, 0.787, \text{ and } 0.207\), respectively). The OR for increased LOS in patients with an undeterminable AT was 3.984 (1.081 – 14.669) \(P=0.042\). The odds of having morbidity, increased CCU LOS, or being readmitted to hospital did not differ between the groups \(P=0.109, 1.00, \text{ and } 0.235\), respectively.

#### Combining variables

The effect of combining variables on surgical outcome was only tested for variables that were significantly associated with the outcome when analysed independently (AT and ASA grade) and on surgical outcome variables that had a substantial event rate (LOS>3 days, \(n=36\)). The odds of individuals having a LOS>3 days was higher in those with a low AT combined with an ASA grade ≥II when compared with those with a high AT and an ASA grade ≤II (OR, 5.890; 95% CI, 1.644 – 21.10) \(P=0.005\). The sensitivity and specificity to predict those with an increased LOS was 77% and 64%, respectively.
when expressed according to body mass, rather than absolute values. Together, these results suggest that CPET variables should be adjusted according to body mass when used to predict outcome after bariatric surgery. Our study also found that bariatric patients unable or unwilling to perform exercise to a moderate intensity and achieve a determinable AT were at greater risk of having increased LOS. While this finding requires further verification, it suggests that patients in this group should be treated in the same way as those identified as high risk based on AT criteria.

The current study and that of McCullough and colleagues differ in several ways that may explain why we identified the AT as the best predictor of gastric bypass surgery outcome as opposed to the VO₂ peak, which they found to have the strongest association. Our patients completed a cycle ergometric ramp test compared with a Bruce treadmill protocol conducted by McCullough and colleagues. This may influence the outcome in two ways. First, treadmill exercise is more familiar to patients and engages a larger muscle mass; therefore, the VO₂ peak values may be more reflective of patient’s cardiovascular capacity than during cycle ergometry. Secondly, AT determination is likely to be more sensitive using cycle ramp test than a Bruce protocol treadmill test as the latter uses fairly large stepped increases in work rate. Therefore, the current study is likely to have identified AT with more precision, improving the association between AT and surgical outcome. Interestingly, the earlier study’s ROC curve analysis of the AT’s capacity to predict outcome found an AUC of 0.69, which is comparable with the current study. However, the likelihood of this occurring due to chance is not known as probability statistic test results were not stated.

Weaknesses of the study
This study suffers from several weaknesses. First, although AT was a predictor of postoperative outcome variables, performance in this respect was moderate (morbidity, ROC AUC = 0.694; postoperative LOS, ROC AUC = 0.640). This level of diagnostic performance is consistent with the majority of literature in this area and reflects the multiple complex influences on surgical outcomes. It may also reflect the lack of binding of clinicians to the CPET data, which would tend to reduce the strength of association between predictor and outcome variable due to the natural tendency of clinicians to focus attention on patients perceived to be at highest risk (confounding by indication). Secondly, the number of subjects (n=106) and the number of positive outcome events (POMS+, n=12) were insufficient for the valid use of multivariate models to explore determinants of risk. Thirdly, the designation of high and low LOS groups was determined after data collection based on the median value, rather than being defined a priori based on clinical criteria. Fourthly, the absence of a control group is an inherent limitation of a service evaluation. Finally, this is a single-centre evaluation and generalizability to other centres is uncertain.

Implications for clinicians or policymakers
Decisions made regarding patient care should always be made in the context of the overall clinical picture. The current study indicates that when cycle ergometry is used to perform CPET for patients awaiting gastric bypass surgery, clinicians should consider using an AT of 11 ml kg⁻¹ min⁻¹ as the threshold to indicate increased risk. However, when treadmill exercise is used, it would be prudent to use results from McCullough and colleagues to guide care and use a VO₂ peak <15.8 ml kg⁻¹ min⁻¹ to indicate increased risk.

Future studies
Although already highlighted in various other surgical populations, the identification of AT as a predictor of outcome after gastric bypass surgery is new. Therefore, confirmation of results in similar populations is required and would increase confidence in results. To extend these results, a larger population of gastric bypass patients should be studied to establish whether several preoperative risk stratifying tools can be combined to predict risk more effectively. Finally, investigating whether preoperative training regimes designed to improve CPET variables improve surgical outcome is warranted.

Conclusion
In this service evaluation of the predictive utility of CPET variables in obese and morbidly obese patients undergoing gastric bypass surgery, we show that the AT is associated with postoperative morbidity and predicts postoperative LOS.

Declaration of interest
P.J.H, A.F.O’D., and J.O. are supported by the UCL/UCLH Comprehensive Biomedical Research Centre, which receives some of its funding through the UK Department of Health’s National Institute of Health Research Biomedical Research Centre/Unit Funding Scheme. M.P.W.G. is the British Oxygen Professor of Anaesthesia at the Royal College of Anaesthetists. M.P.W.G. is also supported by the Southampton Biomedical Research Centre, which receives some of its funding through the UK Department of Health’s National Institute of Health Research Biomedical Research Centre/Unit Funding Scheme. M.P.W.G. received funding from Cortex GmBH for lecturing and travel support to those lectures. This funding was paid directly into departmental funds.

References
5 Hennis PJ, Meale PM, Grocott MP. Cardiopulmonary exercise testing for the evaluation of perioperative risk in non-cardiopulmonary surgery. Postgrad Med J 2011; 87: 550–7
7 ATS/ACCP Statement on cardiopulmonary exercise testing. Am J Respir Crit Care Med 2003; 167: 211–77
18 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: 701–4