Validity of the 6 min walk test in prediction of the anaerobic threshold before major non-cardiac surgery

R. C. F. Sinclair1*, A. M. Batterham2, S. Davies1, L. Cawthorn1 and G. R. Danjoux1

1 Department of Anaesthesia, The James Cook University Hospital, Middlesbrough, UK
2 Health and Social Care Institute, Teesside University, Middlesbrough, UK
* Corresponding author: Department of Anaesthesia, Royal Victoria Infirmary, Queen Victoria Road, Newcastle-upon-Tyne, NE1 4LP, UK. E-mail: rhona.sinclair@ncl.ac.uk

The assessment of exercise capacity before major non-cardiac surgery is recommended to help improve risk prediction perioperatively at the individual patient level.1,2 There are two principal methods utilized in clinical practice in the UK: a cardiopulmonary exercise test (CPET) and patient-reported metabolic equivalent (MET) scores. A CPET is generally regarded as the ‘gold standard’ assessment, providing objective rather than subjective analysis of exercise capacity. Specific measurements obtained during testing have been validated in the prediction of perioperative risk for major non-cardiac surgery.3–5 The anaerobic threshold (AT) currently has the largest evidence base with cut-off thresholds of <11 and <8 ml O2 kg−1 min−1 generally regarded as representing high and very high perioperative risk, respectively.4–7 A high-risk cut-off threshold of slope >34 for the ventilatory equivalent for carbon dioxide (VE/VECO2) has a more limited evidence base.4 In thoracic surgery, a cut-off of <15 ml O2 kg−1 min−1 for maximum oxygen consumption achieved (VO2 max) identifies high-risk cases.6,9

Service infrastructure costs may prohibit setting up a CPET service. Subjective functional assessment of METs, although a simpler alternative, has been shown to have user and physiological limitations.10–12 An alternative, simple, objective measure of exercise capacity may therefore more robustly aid risk stratification, where CPET is unavailable. Ideally, such a test should be validated against measured CPET parameters.

A review of the validity data supporting functional exercise tests revealed the 6 min walk test (6MWT) to be the most extensively researched and established test for use in clinical or research contexts in the cardiorespiratory domain.13 Previous studies have demonstrated a positive correlation between CPET measurements and distance walked in patients with cardiorespiratory disease.14–17 Although the 6MWT has been shown to predict outcome after pulmonary resection18 and lung volume reduction surgery,19 there is no literature pertaining to major non-cardiac surgery. We believe that based on this evidence and pilot data from...
our institution, the 6MWT might be suitable to provide the simple, objective assessment of exercise capacity outlined above.

The aim of this study was to assess the validity of the distance walked during the 6MWT in predicting the AT (and other parameters) derived from CPET.

Methods
The protocol for this concurrent validity study was approved by the National Research and Ethics Service in August 2008 (08/H1305/62). Trial registration: ISRCT 12656789.

Participants were recruited from the preoperative assessment clinics at the James Cook University Hospital between October 2008 and January 2010. After verbal explanation and a patient information sheet, written informed consent was obtained.

Participants included in the study were aged 50–85 yr and awaiting scheduled major non-cardiac surgery (Grade 3 or 4 surgery as defined by NICE guidance).20 Exclusion criteria comprised: medical contraindication to CPET21 or failure to complete a baseline CPET, lower limb claudication and inability to maintain a steady walking pace on level ground. After a medical screening examination, patients were invited to participate. For a desired precision of estimation of \( \pm 0.10 \) (95% confidence interval width) around a postulated validity correlation coefficient of \( r = 0.70 \) (for 6MWT distance in the prediction of AT) derived from pilot work, a sample size of 100 patients was estimated. Allowing for an attrition rate of 25%, a final sample size of 125 participants was required. A total of 186 individuals were screened for inclusion. Of these, 129 participants were enrolled. Characteristics, co-morbid diseases, surgical procedures undertaken, and medications prescribed for participants completing both CPET and 6MWT (119 participants) are shown in Supplementary Table S1.

Participants were asked to complete two exercise tests: CPET (on a cycle ergometer) and a 6MWT. The CPET was performed first, in order to screen for significant cardiovascular pathology, thus ensuring the safe conduct of the 6MWT. To minimize participant inconvenience, both tests were undertaken on the same day. After CPET, patients were provided with refreshments and allowed an appropriate rest interval between tests. The 6MWT was only undertaken once the participants had reported that they had no residual fatigue from CPET. To avoid study bias, the 6MWT was administered by an investigator blinded to the results of the CPET.

Cardiopulmonary exercise test
The CPET was performed using the Medgraphics Ultima system (Tewkesbury, Gloucestershire, UK) and a Lode Corival V2 cycle ergometer (BV Medical Technology, Groningen, The Netherlands). Flow and gas calibrations were performed before each test session, which was subsequently conducted to our standard protocol (available in Supplementary material). All usual patient medication was continued.

The test was terminated when the participant reached volitional exhaustion (\( \dot{V}_{O2} \) peak) or earlier if another termination criterion was fulfilled. The V-slope comparison plot was compiled using Breeze software (Medgraphics) and interpreted by two trained observers on completion of all study testing (G.R.D. and R.C.F.S.).

Six min walk test
After successful completion of CPET, participants performed the 6MWT as outlined in the guidance published by the American Thoracic Society (ATS).22 Individuals walked to their own maximum pace along a flat corridor, marked with a 30 m track, aiming to cover as much distance as possible in the timed 6 min. Participants wore a MIRoxi pulse oximeter (Medical International Research, Roma, Italy) to record heart rate response and oxygen saturations.

The ATS suggest that a practice test is not needed in most settings.22 Furthermore, data from our pilot study (unpublished observation) confirmed that the test was highly reproducible, with an intraclass correlation coefficient (ICC 3.1) of 0.94, and a non-substantial mean bias of 18 m greater on a second walk. Thus, a single 6MWT was performed in the current study.

Test outcome measures recorded
- CPET—oxygen consumption at the AT (using the V-slope technique),23 oxygen consumption at volitional exhaustion (\( \dot{V}_{O2} \) peak), the \( \dot{V}_{E}/\dot{V}_{CO2} \) recorded at AT, and maximum heart rate achieved (HR\(_{max}\))
- 6MWT—maximal distance walked and HR\(_{max}\)

Statistical analysis
Ordinary least-squares linear regression models were applied to obtain the validity coefficient (\( r \)) and the standard error of the estimate (SEE) — the typical error associated with the prediction of AT (or \( \dot{V}_{O2} \) peak or \( \dot{V}_{E}/\dot{V}_{CO2} \) slope) from 6MWT distance in an individual patient. Receiver operating characteristic (ROC) curve analysis was used to derive cut-points for 6MWT distance for the prediction of AT <11 ml O\(_2\) kg\(^{-1}\) min\(^{-1}\), AT <8 ml O\(_2\) kg\(^{-1}\) min\(^{-1}\), \( \dot{V}_{O2} \) peak <15 ml O\(_2\) kg\(^{-1}\) min\(^{-1}\), and a combination of AT <11 ml O\(_2\) kg\(^{-1}\) min\(^{-1}\) and \( \dot{V}_{E}/\dot{V}_{CO2} \) slope >3.4. The optimum cut-point was determined as the value corresponding with the greatest accuracy (highest sum of sensitivity plus specificity; i.e. with sensitivity and specificity weighted equally). When a test is to be used for screening purposes and risk stratification, however, a cut-off value with greater sensitivity (fewer false-negatives) may be desirable. Therefore, we derived an alternative cut-point by adopting a 2:1 weighting for sensitivity:specificity.

To refine the ROC-derived cut-offs, we used the obtained regression equation and SEE, to derive lower and upper cut-points for 6MWT distance that are predictive of an AT that is likely to be less than or greater than these prognostic AT thresholds. (A 6MWT distance falling between these two cut-points is assumed to be in an area of ‘clinical uncertainty’.)
Herein, ‘likely to be’ is defined as a probability of $\geq 0.75$ (odds of at least 3:1 in favour) of the patient’s ‘true’ AT being less than (lower cut-point for 6MWT distance) or greater than (upper cut-point) 11 ml O$_2$ kg$^{-1}$ min$^{-1}$, given the predicted AT from the regression equation and the observed prediction error (SEE). This probability is derived from the disposition of the confidence interval for the predicted value to the prognostic value of 11 ml O$_2$ kg$^{-1}$ min$^{-1}$, and is calculated using the Student t-distribution. The required t-value is derived as the prognostic value for AT minus the predicted value from the regression and divided by the obtained SEE: $(11 - 9.7)/1.9 = 0.68$. The area under the t-distribution to the left of this value with the appropriate degrees of freedom is 0.75, providing the probability that the patient’s true AT is $< 11$ ml O$_2$ kg$^{-1}$ min$^{-1}$ if their predicted value from the regression was 9.7 ml O$_2$ kg$^{-1}$ min$^{-1}$. Rearranging the derived linear regression equation gives the 6MWT distance predictive of an AT of 9.7 ml O$_2$ kg$^{-1}$ min$^{-1}$; this is the lower cut-point. The upper cut-point was calculated in an identical fashion. All analyses were conducted using StatsDirect (Altrincham, UK; v. 2.7.8) and Medcalc (Mariakerke, Belgium; v. 11.5) software packages.

We adopted an objective criterion to identify and remove outliers with a standardized residual of $> 3.6$ from the analysis. With the assumption of normality, this threshold identifies values that would occur only rarely (<5% of the time) with this sample size.

**Results**

In total 119 of 129 recruited participants completed both exercise tests. Of the 10 individuals not completing: two withdrew consent after CPET, one failed to reach AT during CPET, five were unable to complete the full 6 min of walking, and two individuals had no reason documented. Of the 119 participants, an additional seven participants were eliminated from the analysis due to a persistently elevated respiratory exchange ratio (RER) likely a consequence of hyperventilation due to anxiety, poor accommodation to the mouthpiece, or both. In such cases, an AT is still detectable but it will be a ‘pseudo-threshold’ occurring before the actual AT resulting in an underestimation. Screening for severe outliers resulted in the removal of one case for the AT analysis and one case for the V$_{O2}$ peak analysis, resulting in a data set of $n=110$ complete cases.

**Exercise test results**

The CPET and 6MWT results for study participants are presented in Table 1. The peak exercise challenge was comparable between the two tests as judged by the similar mean maximum heart rate. Figure 1 illustrates a scatter plot of AT vs 6MWT distance.

Linear regression analyses to predict the AT, V$_{O2}$ peak, and V$_E$/V$_{CO2}$ from the distance walked during the 6MWT are shown in Table 2, and the results of the ROC curve analyses are detailed in Table 3. The area under the ROC curve indicates that a randomly selected individual from the positive group (AT $< 11$ ml O$_2$ kg$^{-1}$ min$^{-1}$) has a 6MWT distance value (Y) smaller than that of a randomly chosen individual from the negative group (X) 85.2% of the time [$P(Y<X)=0.852$]. The likelihood ratios indicate that a 6MWT distance of $< 440$ m is obtained around 15 times as frequently in patients with an AT of $< 11$ ml O$_2$ kg$^{-1}$ than in those with an AT above this threshold and that a 6MWT distance of $\geq 440$ m is obtained approximately a third as frequently in patients with an AT of $< 11$ ml O$_2$ kg$^{-1}$ min$^{-1}$ than in those with an AT above this value. The ROC curve for this analysis is shown in Figure 2, illustrating that the area under the curve is substantially larger than that of ‘no

**Table 1. Exercise test results [mean (so)]. HR$_{max}$, maximum heart rate achieved; AT, anaerobic threshold; V$_{O2}$ peak, peak oxygen consumption**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Slope</th>
<th>Intercept</th>
<th>Correlation coefficient (r) (95% CI)</th>
<th>Standard error of the estimate (SEE) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>0.019</td>
<td>1.598</td>
<td>0.68 (0.56–0.77)</td>
<td>1.9 (1.7–2.2) ml O$_2$ kg$^{-1}$ min$^{-1}$</td>
</tr>
<tr>
<td>V$_{O2}$ peak</td>
<td>0.033</td>
<td>$-0.326$</td>
<td>0.75 (0.65–0.82)</td>
<td>2.7 (2.4–3.1) ml O$_2$ kg$^{-1}$ min$^{-1}$</td>
</tr>
<tr>
<td>V$<em>E$/V$</em>{CO2}$</td>
<td>$-0.028$</td>
<td>48.479</td>
<td>0.46 (0.30–0.60)</td>
<td>5.0 (4.4–5.8)</td>
</tr>
</tbody>
</table>

**Table 2. Linear regression analyses with 6MWT distance as the predictor. AT, anaerobic threshold; V$_E$/V$_{CO2}$, ventilatory equivalents for carbon dioxide; V$_{O2}$ peak, peak oxygen consumption**

![Fig 1 Scatter plot for 6MWT distance (m) vs AT (ml O$_2$ kg$^{-1}$ min$^{-1}$).](image-url)
discrimination’ (0.50) indicated by the diagonal. Figure 3 shows the dot plot for 6MWT distance in the two groups. Lower and upper cut-points for 6MWT distance derived from regression analysis

From the regression modelling, the lower and upper cut-points for 6MWT distance predictive of a true AT that is likely to be ($P \geq 0.75$) <11 or >11 ml O$_2$ kg$^{-1}$ min$^{-1}$, respectively, were below 427 m (positive test) or above 563 m (negative test). For the 8 ml O$_2$ kg$^{-1}$ threshold for the AT, the lower and upper cut-points were <269 and >405 m. For the VO$_2$ peak <15 ml O$_2$ kg$^{-1}$ min$^{-1}$, the lower and upper cut-points were <409 and >520 m.

Discussion

In this study, we have confirmed that the 6MWT may be a useful practical method for risk stratification, with a large effect size observed for the correlation between the distance walked during a 6MWT and oxygen consumption at both AT and VO$_2$ peak. AT, measured during CPET, is presently recognized as the most robust endpoint to inform perioperative risk stratification. For this reason, the majority of our analysis and inference focuses on the relationship between 6MWT distance and AT, rather than VO$_2$ peak.

The ROC curve analyses for both the <11 and <8 ml O$_2$ kg$^{-1}$ min$^{-1}$ thresholds for AT revealed that the 6MWT

<table>
<thead>
<tr>
<th>CPET measurement (ml O$_2$ kg$^{-1}$ min$^{-1}$)</th>
<th>AT &lt;11</th>
<th>AT &lt;8</th>
<th>AT &lt;11 and VE/VO$_2$ &gt;34</th>
<th>VO$_2$ peak &lt;15</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROC curve sensitivity: specificity weighting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevalence (%)</td>
<td>1:1</td>
<td>2:1</td>
<td>1:1</td>
<td>2:1</td>
</tr>
<tr>
<td>AUC</td>
<td>0.852</td>
<td>0.857</td>
<td>0.801</td>
<td>0.856</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.511–0.757</td>
<td>0.731–0.922</td>
<td>0.637–0.970</td>
<td>0.762–0.999</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.852–0.995</td>
<td>0.520–0.805</td>
<td>0.750–0.911</td>
<td>0.590–0.790</td>
</tr>
<tr>
<td>Positive likelihood ratio</td>
<td>14.73</td>
<td>2.59</td>
<td>5.45</td>
<td>3.14</td>
</tr>
<tr>
<td>Negative likelihood ratio</td>
<td>0.09–1.5</td>
<td>0.1–0.5</td>
<td>0.05–0.5</td>
<td>0.01–0.50</td>
</tr>
</tbody>
</table>

Sensitivity 0.641 0.844 0.857 0.952 0.725 0.825 0.702 0.895

Specificity 0.957 0.674 0.843 0.697 0.800 0.686 0.868 0.895

Positive likelihood ratio 14.73 2.59 5.45 3.14 3.63 2.62 5.31 2.06

Negative likelihood ratio 0.38 0.23 0.17 0.07 0.34 0.26 0.34 0.19

Table 3 ROC curve analyses. AT, anaerobic threshold; VE/VO$_2$, ventilatory equivalents for carbon dioxide; VO$_2$, peak oxygen consumption.
distance is an adequate discriminator between high and low
AT patient groups. However, deriving an optimum single cut-
point from ROC curve analysis is challenging and not ideally
suited to a clinical context. Sensitivity can be weighted to
reduce the false-negative rate, but we believe that the
regression-based analysis represents a refinement of single
ROC curve cut-points allowing for clinical variation and
uncertainty. Using the regression method, a patient with a
positive test (6MWT < 427 m) is likely to be at high periopera-
tive risk, and a patient with a negative test (6MWT > 563 m)
would be considered low risk. A patient completing a dis-
tance of ≥ 427 but ≤ 563 m is in a zone that we define as
‘clinical uncertainty’. We can usefully incorporate these
regression analysis-derived risk categories into clinical
practice (see below).

The current international guidance relies on subjective
assessment of functional capacity, in the form of METs, as
one of the three key variables in the decision-making
process of risk stratification before non-cardiac surgery. A
functional capacity of < 4 METs (inability to climb a flight of
stairs) represents the threshold to trigger the high-risk limb
of the risk stratification pathways. In the current study only
one of 101 individuals (1%) reported a functional capacity
of < 4 METs (14 ml O2 kg min−1), whereas 58.2% of our par-
ticipants had an objectively measured AT of < 11 ml O2 kg
min−1 during CPET. Interestingly, the individual reporting a
functional capacity of < 4 METs attained an AT and VO2
peak of 11.8 and 15.2 ml O2 kg min−1, respectively, thereby
representing low risk based on objective testing. We believe
that based on our data, the 6MWT represents a superior and
more robust technique for risk stratification than a self-
reported cut-point of < 4 METs.

The other two key variables utilized to determine pre-
operative risk in the current international guidance are
the number of clinical risk factors and nature of surgical
intervention. Utilizing this approach in combination
with 6MWT distance could help identify the most at risk
individuals before surgery. The major benefits would be
the ease with which it could be administered, minimal
staff training and equipment requirements, and simple
and quick to perform. In addition, the test is repeatable, is
safe to perform, and entails a minimal increase in patient
attendance time. Although not recommended as a replace-
ment for CPET, the 6MWT could, in effect, act as a surro-
gate ‘sieve’ in identifying high-risk individuals who may
require further assessment or optimization before surgery.
In hospitals where CPET would perhaps not be utilized fre-
quently, the 6MWT could be used as a cheap accurate
alternative enabling identification and referral for CPET via
loco-regional preoperative networked arrangements.

We believe that the upper and lower cut-points derived
from the regression analysis provide the ideal platform in
providing for such a model. For example, no further assess-
ment would be required in an individual walking > 563 m
during the 6MWT (upper cut-point, true AT likely to be
> 11 ml O2 kg min−1), whereas a patient walking < 427 m
(lower cut-point, true AT likely to be < 11 ml O2 kg min−1)
would be considered high risk and should be referred for
further functional assessment. In individuals walking a dis-
tance in the area of ‘clinical uncertainty’ (≥ 427 but ≤ 563
m), it would be important to incorporate the number of clin-
ical risk factors and magnitude of surgical intervention into
this clinical decision-making process, before consideration
of further investigation. Therefore, a patient walking, say,
500 m together with two to three clinical risk factors
should be further assessed, whereas an individual walking
the same distance with a good health profile would not.

Our study is unique in being the first to examine the use of
the 6MWT before operation in patients undergoing
non-cardio-thoracic surgery. However, we identified three
studies within cardio-respiratory medicine reporting similar
correlations between 6MWT distance and CPET measure-
ments to ours. These studies predominantly concen-
trated on correlations between peak oxygen consumption
and 6MWT distance, reporting validity coefficients from
correlation coefficients of 0.64 to 0.88. The observed correlations between 6MWT distance and AT and VO2 peak in the current study are sub-
stantially larger than those reported in patients with chronic obstructive pulmonary disease.

Our results appear to conflict with those reported from a
study of the validity of an intermittent shuttle walk test in
assessing fitness for surgery, with the authors concluding
that the discriminatory ability of the test was poor. How-
ever, a robust comparison of our findings with this
study is not possible, as the patient group was substantially
fitter (mean AT 12.7 vs 10.2 ml O2 kg min−1) than our
sample, and the authors do not detail the method used to
determine the single ROC cut-points, nor the sensitivity,
especificity, and likelihood ratios associated with the derived
cut-points.

It is important to acknowledge a number of limitations to
our study. First, we are utilizing a specific cut-off value for AT
to discriminate between high- and low-risk individuals
(a threshold value of 11 ml O2 kg min−1). However, this
threshold remains robust, despite being unchanged since
proposed originally. Indeed, Snowden and colleagues
reported a very similar AT cut-point (10.1 ml O2 kg min−1)
in prediction of increased postoperative morbidity. Similarly,
Wilson and colleagues reported that an AT of < 11 ml O2 kg
min−1 was a clinically significant predictor of mortality in
major non-cardiac surgery patients.

Second, using the 6MWT as a surrogate provides limited
diagnostic information on cardiorespiratory reserve, which
can be obtained with CPET. We are however in effect
suggesting the 6MWT as an improvement over a subjective
cut-point of < 4 METs in identifying high-risk individuals
and not in replacement of CPET. With the current financial
constraints on the National Health Service in the UK, we
believe that the 6MWT represents a robust pragmatic
improvement where CPET is unavailable. Indeed, identifi-
cation of high-risk individuals utilizing the 6MWT may
enable streamlined pathways of care at the loco-regional
level as outlined above. Such a tertiary referral service
would be more cost-effective and avoid unnecessary
duplication of tests. Third, the 6MWT is of limited utility in assessing patients with limb ischaemia or major limitation to exercise, for example, lower limb arthritis. Patients who cannot walk at a good pace have a resultant decreased 6MWT distance when compared with exercise results during non-weight-bearing cycle exercise. Finally, this study was designed to examine the prediction of CPET parameters from a 6MWT and not powered to predict perioperative outcome. We acknowledge that this might be considered an important direction for future research were the 6MWT to be adopted into regular clinical practice.

In conclusion, our results demonstrate that the 6MWT can be used robustly at preoperative assessment to assess exercise capacity. Where CPET is unavailable, we believe the regression analysis model presented provides an accurate, simple, and cheap way of clinically guiding further patient management as part of a preoperative screening process.

Supplementary material
Supplementary material is available at British Journal of Anaesthesia online.

Acknowledgements
The CPET equipment used in the study was purchased by James Cook University Hospital Volunteers services.

Declaration of interest
None declared.

Funding
The study was supported by a small research grant from The James Cook University Hospital.

References
7 Older P, Smith R, Courtney P. Pre-operative evaluation of cardiac failure and ischaemia in elderly patients by cardiopulmonary exercise testing. Chest 1993; 104: 701–4
8 Bayram A, Candan T, Gebitekin C. Preoperative maximal exercise oxygen consumption test predicts postoperative morbidity following major lung resection. Respiratology 2007; 12: 505–10
13 Solway S, Brooks D, Lacasse E, Thomas S. A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. Chest 2001; 119: 256–70
18 Holden DA, Rice TW, Stelmach K, Meeker DP. Exercise testing, 6-min walk and stair climb in the evaluation of patients at high risk for pulmonary resection. Chest 1992; 102: 1174–9
20 NICE guidelines: Preoperative tests: the use of routine preoperative tests for elective surgery. Appendices. 2003
21 Weisman I. ATS/ACCP statement on cardiopulmonary exercise testing. Am J Respir Crit Care Med 2003; 167: 211–77