The impact of decreasing cutoff values for maximal oxygen consumption (VO$_2$max) in the decision-making process for candidates to lung cancer surgery

Gaetano Rocco$^{1,2}$, Tindaro Gatani$^{1,2}$, Massimo Di Maio$^{1,2}$, Ilernando Meoli$^{1,2}$, Antonello La Rocca$^{1,2}$, Nicola Martucci$^{1,2}$, Carmine La Manna$^{1,2}$, Francesco Stefanelli$^{1,2}$

$^1$Department of Thoracic Surgery and Oncology, Division of Thoracic Surgery and Service of Physiopathology, National Cancer Institute, Naples, Italy; $^2$Division of Respiratory Physiopathology, Monaldi Hospital, Naples, Italy

ABSTRACT

Background: Maximal oxygen consumption (VO$_2$max) is considered a decisive test for risk prediction in patients with borderline cardiopulmonary reserve. Guidelines have adopted decreasing VO$_2$max cut-off values to define operability within acceptable mortality and morbidity limits. We wanted to investigate how the adoption of decreasing VO$_2$max cut-off-values assessment contributed to better select lung surgery candidates.

Methods: One hundred and nineteen consecutive surgical candidates have been prospectively analyzed as a sample population. Preoperative work-up included spirometry and transfer factor (DLco); irrespective of the spirometric values, these patients were subjected to VO$_2$max assessment. Surgical eligibility was decided by the same surgeon throughout the series. In the postoperative period, overall mortality and the occurrence of any, major or minor complications was recorded and graded according to the Common Terminology Criteria for Adverse Events v.4.3.

Results: Three arbitrary cut-offs were introduced at 15, 14 and 12 mL·kg$^{-1}$·min$^{-1}$. Notably, 15 and 12 mL·kg$^{-1}$·min$^{-1}$ correlated with percentage VO$_2$max values of 50% and 35% of predicted (P<0.0001 and 0.0079), respectively. Accordingly, the patients were subdivided into groups in which the prevalence of postoperative morbidity was recorded. The groups were homogeneous as to age, BMI, preoperative absolute and percentage FEV1 and DLco. In the Cox proportionate-hazards multivariate analysis, VO$_2$max less than 35% (P=0.0017) and CTCAE >2 (P=0.0457) emerged as significant predictors of survival after surgery. Conversely on logistic regression analysis, age over 70 years (P=0.03) and pneumonectomy (P=0.001), but not VO$_2$max cut-off values, were significant predictors of major (CTCAE >2) morbidity.

Conclusions: Since VO$_2$max is increasingly used to contribute to risk prediction for the individual patient, surgeons need to be advised that the concept of a definitive, generalized cut-off value for VO$_2$max is probably a contradiction in terms. Patient-specific VO$_2$max values are more likely to contribute to risk assessment since they may reflect the primarily affected component among the determinants of maximal oxygen consumption. Whether patient-specific VO$_2$max should be routinely used by surgeons to define operability for borderline patients needs further evaluation.

KEY WORDS

Maximal oxygen consumption; lung surgery; lung cancer


Introduction

In spite of the results from the analysis of multiinstitutional databases and the current guidelines focused on risk assessment for lung resection (1,2), the surgeon's perception of operability may at times not coincide with a borderline or adverse outcome of preoperative work-up (3). Selected patients have been offered surgery based on the surgeon's clinical instinct which still represents a useful adjunct to spirometric evaluation. Cardiopulmonary exercise testing is increasingly used to refine
the selection of surgical candidates (3-5). Decreasing cut-off values of maxymal oxygen consumption (VO\textsubscript{max}) have been proposed to avoid denying a surgical option to lung cancer patients (4). We investigated the possible correlation between the progressive reduction of VO\textsubscript{max} cut-offs and postoperative survival and morbidity after lung cancer surgery.

### Material and methods

Between January 2008 and December 2010, 119 consecutive candidates to pulmonary resection and operated on at the National Cancer Institute in Naples were subjected to VO\textsubscript{max} assessment as part of the routine preoperative work-up for lung cancer surgery. Cycle ergometry with incremental 10 watts RAMP (Cosmed Quark CPET, Italy) adaptable to the patient’s clinical condition, was used to determine VO\textsubscript{max} (maximal O\textsubscript{2} value at plateau despite increasing workload), peak VO\textsubscript{2} (maximal oxygen consumption at maximal workload), VCO\textsubscript{2} (quantity of expired CO\textsubscript{2} at a given time), RQ (VCO\textsubscript{2}/VO\textsubscript{2} or respiratory quotient), VO\textsubscript{2}/HR (oxygen consumption per systole), HRR (cardiac reserve) and the anaerobic threshold.

This patient population included 80 males and 39 females with a median age of 64 years (range, 22 to 80 years). Eighty-two percent of these surgical candidates were either former or current smokers. The median body mass index (BMI) was 26 kg/m\textsuperscript{2} (range, 17 to 50 kg/m\textsuperscript{2}). In particular, 66 of the patients in this series had a BMI of 25 or greater and were accordingly considered overweight or obese (5). Comorbidities included previous stroke in 4 patients (3.4%), concurrent significant peripheral vascular and coronary artery disease in 5 (4.2%) and 11 (9.2%) patients, respectively. In addition, 2 patients (1.7%) had atrial fibrillation whereas 61 (51.3%) had a history of hypertension under pharmacological treatment. Other comorbidities included diabetes in 17 patients (14.3%) and previous cancer history in 45 (37.8%). Thirty-eight patients (32%) had had chemotherapy or chemoradiotherapy as a part of their preoperative disease management. Of 86 patients undergoing lung resection for primary non small cell lung cancer, 11 (13%) had been subjected to neoadjuvant treatment. In the remaining patients, final pathology showed lung metastases (25 patients) and benign lesions (8 patients).

Of 119 patients, 11 (9.3%) had a preoperative FEV\textsubscript{1} less than 70% of predicted. The median FEV\textsubscript{1} absolute and percentage values were 2.29 L (range, 0.86-4.01 L) and 92.0% (range, 37.7-135.4%), respectively. In addition, the median DL\textsubscript{co} absolute and percentage values were 18.4 (range, 9.9-33.6) and 78.3% (range, 45-132%). Also, median VO\textsubscript{max} expressed in absolute and percentage values were 15.3 mL·kg\textsuperscript{-1}·min\textsuperscript{-1} (range, 8.3-28.8 mL·kg\textsuperscript{-1}·min\textsuperscript{-1}) and 62% (range, 31-129%).

Reasons for surgery included non small cell lung cancer in 80 patients (67.2%), metastases to the lung in 25 (21%) and benign lesions in the remaining 14 (11.8%). A pneumonectomy was performed in 20 patients (10.6%), a lobectomy in 54 patients (45.4%), a bilobectomy in 5 (4.2%), and, single or multiple segmentectomy/wedge resections in 40 (33.6%). A right-sided pulmonary resection was performed in 72 patients (60.5%). The standard surgical approach for primary lung cancer was a limited (8 to 10 cm) incision in the auscultatory triangle with minimal rib spreading and no muscle division with or without videoassistance (6). The decisions on operability was made in all instances by the same board-certified surgeon (G.R.) who also performed or directly supervised the surgical procedures. When final pathology was obtained in non small cell lung cancer patients, the disease was staged as pIA in 21 patients (26.2%), pIB in 44 (55%), pIIA in 5 (6.3%), pIIB in 4 (5%), pIIIA in 5 (6.3%), and, pIIIB in 1 (1.2%).

Major and minor postoperative morbidity were attributed according to the NCI Common Terminology Criteria for Adverse Events v.4.3 (7). In addition, to understand the relationship between morbidity and VO\textsubscript{max} and whether a cut-off value could be predictive of postoperative complications, we arbitrarily subdivided our patient population into three groups according to the VO\textsubscript{max} values currently reported in the literature (12 and 15 mL·kg\textsuperscript{-1}·min\textsuperscript{-1}) and a conservative (i.e., closer to the safest figure) intermediate value (14 mL·kg\textsuperscript{-1}·min\textsuperscript{-1}).

This routine use of cycloergometry was funded by the Italian National Health System (SSN) within the context of an institutional project aimed at improving risk management in oncologic surgery. Informed consent was obtained in all patients to undergo cardiopulmonary exercise testing prior to considering a surgical option.

### Statistical methods

Pearson’s correlation and t-test were used to define correlations between continuous variables. Categorical variables were compared by chi-square tests whereas Spearman’s tests were used to verify correlation. Multivariate analyses were conducted by logistic regression to define prognosticators for postoperative morbidity. In addition, Cox proportional hazard regression was conducted to evaluate survival predictors. Significance was set at P<0.05. For statistical analysis, SPSS version 19 and MedCalc version 12 softwares were used. Isometric graphs were constructed by using the JMP statistical software, version 9.

### Results

In this series of 119 patients, correlations between FEV\textsubscript{1}, FEV\textsubscript{1}/%DL\textsubscript{co} and VO\textsubscript{max} along with VO\textsubscript{max}% were studied. Expectedly, the correlation FEV\textsubscript{1}·DL\textsubscript{co} appeared statistically significant (P<0.0001). In addition, a definite correlation FEV\textsubscript{1}·VO\textsubscript{2}max (P=0.001) and DL\textsubscript{co}%·VO\textsubscript{max}%...
(P=0.0235) was also found. Moreover, in our series, 43 patients (36.1%) did have a disjunction between absolute and percentage VO$_2$max values. Absolute VO$_2$max values of 15 and 12 mL·kg$^{-1}$·min$^{-1}$ correlated with percentage VO$_2$max values of 50% (P=0.0001; rho=0.366, 95% CI for rho, 0.199 to 0.512) and 35% of predicted (P=0.0079; rho= 0.242, 95% CI for rho= 0.0652 to 0.405), respectively.

Overall, 4 patients died within 30 days of surgery (3.3%). Median FEV$_1$, FEV$_1$%, DLco and DLco% measured in these patients were 2.15 L (range, 1.29 to 2.8 L), 92.6% (range, 80% to 132%), 17.84 (range, 10.94 to 25.7), and, 83.2% (range, 64.3% to 107%), respectively. Absolute and percentage of predicted VO$_2$max values in these patients were 10.7, 12.2, 12.1, 19.3 mL·kg$^{-1}$·min$^{-1}$ and 65%, 56%, 57%, 103%, respectively. Causes of death included pulmonary embolism in 3 patients (2 after right pneumonectomy and 1 after right upper lobectomy) and sepsis following bronchopleural fistula in 1 following right lower lobectomy. In addition, postoperative complications occurred in 55 patients.

**Group 1 - cut-off: 15 mL·kg$^{-1}$·min$^{-1}$**

Fifty-three patients (44.5%) had a VO$_2$max less than 15 mL·kg$^{-1}$·min$^{-1}$. No statistically significant differences were noted as to age, BMI, percentage preoperative FEV$_1$, and DLco between patients with VO$_2$max greater or lesser than 15 mL·kg$^{-1}$·min$^{-1}$. Patients with less than 15 mL·kg$^{-1}$·min$^{-1}$ underwent pneumonectomy in 14 (26.4%) cases, lobectomy or bilobectomy in 30 (56.7%), and, segmentectomy/wedge in 9 (16.9%). Of these patients, 21 (40%) experienced some postoperative complications. On the other hand, the remaining 32 patients with VO$_2$max less than 15 mL·kg$^{-1}$·min$^{-1}$ (60%) did not experience any complications in the postoperative course. The difference with the 66 patients with VO$_2$max greater than 15 mL·kg$^{-1}$·min$^{-1}$ was not significant (P=0.26). If we distinguish between patients having minor and major complications in the postoperative period of patients with VO$_2$max less than 15 mL·kg$^{-1}$·min$^{-1}$, only 5 patients (9%) had major morbidity compared to 48 (91%) who did not. Again, no statistically significant difference was noted with the patients with a VO$_2$max greater than 15 mL·kg$^{-1}$·min$^{-1}$ (P=0.48).

If the FEV$_1$ less than 40% of predicted threshold is considered, no difference was noted in the severity of complications, if any, based on the 15 mL·kg$^{-1}$·min$^{-1}$ cutoff (P=0.21). In particular, this cutoff did not help stratify the patient population who would have developed major complications postoperatively (P=0.5).

**Group 2 - cut-off: 14 mL·kg$^{-1}$·min$^{-1}$**

Thirty-eight (26.8%) had a VO$_2$max less than 14 mL·kg$^{-1}$·min$^{-1}$. A pneumonectomy was performed in 8 (21%), lobectomy or bilobectomy in 21 (55.3%), and, segmentectomy/wedge in 9 (23.7%) patients. Of these patients, 15 (39%) experienced some postoperative complications. On the other hand, the remaining 23 patients with VO$_2$max less than 15 mL·kg$^{-1}$·min$^{-1}$ (61%) did not experience any complications in the postoperative course. The difference with the 81 patients with VO$_2$max greater than 14 mL·kg$^{-1}$·min$^{-1}$ was not significant (P=0.38). When we compared patients with VO$_2$max less than 14 mL·kg$^{-1}$·min$^{-1}$ developing minor and major complications postoperatively, only 3 patients (8%) had major complications while 35 (92%) did not. Again, no statistically significant difference was noted compared to the patients with a VO$_2$max greater than 14 mL·kg$^{-1}$·min$^{-1}$ (P=0.37).

**Group 3 - cut-off: 12 mL·kg$^{-1}$·min$^{-1}$**

Fifteen patients (12.6%) had a VO$_2$max less than 12 mL·kg$^{-1}$·min$^{-1}$; they underwent a pneumonectomy in 2 (13.3%) cases, lobectomy or bilobectomy in 7 (46.7%), and, segmentectomy/wedge in 6 (40%). Of these patients, 7 (47%) experienced some postoperative complications. On the other hand, the remaining 8 patients with VO$_2$max less than 12 mL·kg$^{-1}$·min$^{-1}$ (53%) did not experience any complications in the postoperative course. The difference with the 104 patients with VO$_2$max greater than 12 mL·kg$^{-1}$·min$^{-1}$ was not significant (P=0.91). If we distinguish between patients having minor and major complications in the postoperative period of patients with VO$_2$max less than 12 mL·kg$^{-1}$·min$^{-1}$, only 1 patient (7%) had major morbidity compared to 14 (93%) who did not. Again, no statistically significant difference was noted compared to patients with a VO$_2$max greater than 12 mL·kg$^{-1}$·min$^{-1}$ (P=0.51).

Pearson’s correlation tests were carried out by plotting decreasing values of FEV$_1$ and DLco against different cutoff levels of VO$_2$max. As expected, the categorical distinction (greater than or less than 15, 14 and 12 mL·kg$^{-1}$·min$^{-1}$) made it possible to separate our patient population in two risk groups according to each VO$_2$max cut-off value. If one looks at the diagrams contained in the figure in sequence (Figure 1), the upper parts of the diagrams show an increasing “osmosis” of high risk patients in the lower risk groups as we set the VO$_2$max cut-offs to decreasing values. Also, it clearly appears that a VO$_2$max less than 12 mL·kg$^{-1}$·min$^{-1}$ (or 35% of predicted) may contribute to the refinement of the stratification by clinical instinct albeit not all patients are finally allocated to newly determined risk class groups. In fact, the patients’ data falling into the low FEV$_1$-low DLco category (extreme sides of the scatterplots) become few and far between as the VO$_2$max cut-off values decreased to less than 12 mL·kg$^{-1}$·min$^{-1}$. When isometric graphs were constructed to study the correlation between FEV$_1$ and DLco at different absolute and percentage VO$_2$max values, the scattered distribution of VO$_2$max performance data was an obvious finding (Figure 2). Notably, patients with normal spirometry and/or DLco were found to have inconsistent
VO\(_2\)max values (Figure 2).

**Correlation between VO\(_2\)max and FEV\(_1\) and DLco in patients with VO\(_2\)max greater or less than 15 mL·kg\(^{-1}\)·min\(^{-1}\)**

Of the 66 patients with measured VO\(_2\)max greater than 15 mL·kg\(^{-1}\)·min\(^{-1}\) (i.e., normal preoperative risk), all (100%) had also an FEV\(_1\) >40% of predicted. In addition, 61 of these patients (92%) had also a DLco greater than 40% of predicted. Conversely, of the 53 patients with measured VO\(_2\)max less than 15 mL·kg\(^{-1}\)·min\(^{-1}\) (i.e., moderate to high preoperative risk), 52 (98%) had an FEV\(_1\) greater than 40% of predicted whereas only 47 (88%) had a preoperative DLco greater than 40% of predicted.

**Surgery for lung cancer**

Of the 119 patients in this series, 86 were resected for primary lung cancer. In this group, the patients subjected to wedge resection had a median VO\(_2\)max of 14.9 mL·kg\(^{-1}\)·min\(^{-1}\) (range, 10.8-20 mL·kg\(^{-1}\)·min\(^{-1}\)); when expressed in percentage, the median VO\(_2\)max was 64.6% (range, 35-115%). The patients subjected to lobectomy had a median VO\(_2\)max of 16 mL·kg\(^{-1}\)·min\(^{-1}\) (range, 8.32-28.89 mL·kg\(^{-1}\)·min\(^{-1}\)); the corresponding median percentage value was 68.3% (range, 42-129%). Pneumonectomy patients had a median VO\(_2\)max of 15.6 mL·kg\(^{-1}\)·min\(^{-1}\) (range, 11.5-22.9 mL·kg\(^{-1}\)·min\(^{-1}\)); the corresponding median percentage value was 61.1% (range, 39-82%). When the same analysis as the lung resection candidates was carried out to correlate decreasing cutoff values of VO\(_2\)max with postoperative morbidity in lung cancer patients, no differences were noted in all three subgroups with regard to either major or minor complications.

Variables entered in the multivariate analyses were age (70 years cut-off), gender, BMI (25 kg/m\(^2\) cut-off), hemoglobin levels,
smoking history, percent FEV₁ (70% cut-off), percent DLO₂ (50% cut-off), CPET expressed by VO₂max above mentioned cut-off values, including 35% and 50% cut-off, type of operation (pneumonectomy vs. others), morbidity expressed by CTCAE grading system (7) (CTCAE grade 3 cut-off).

In this setting, when preoperative indicators were considered, multivariate analysis demonstrated that age and cardiorespiratory performance (i.e., absolute VO₂max values) correlated with postoperative morbidity at large (Table 1). In addition, age over 70 years and type of operation emerged as significant predictors of major [CTCAE ≥ grade 3 (7)] postoperative complications (Table 2).

A 93.6% overall actuarial 5 year survival was computed with the Kaplan-Meier method. The Cox proportionate-hazards multivariate analysis demonstrated that both the percentage VO₂max value greater than 35% and CTCAE ≥3 emerged as significant predictors of postoperative 90-day survival (Table 3). Overall, the presence of a discrepancy between absolute and percentage VO₂max values was not a significant predictor of either 30 days mortality (P=0.509) or major morbidity (P=0.849).

## Discussion

In clinical practice, the quest for the perfect indicator of postoperative morbidity is far from over (8-11). In the meantime, the selection of surgical candidates is inspired by current guidelines which usually enable the surgeons to offer a surgical option to the majority of patients (1,2). In some circumstances, experienced surgeons tend to apply their “clinical instinct” to offer a surgical option despite potentially adverse spirometric figures (1,2). Concomitantly, in recent years, guidelines have been characterized by decreasing spirometric cutoff values and have introduced the concept of assessing the diffusing capacity and the cardiopulmonary performance in selected patients with borderline function (1,2,13). In this setting, cardiopulmonary exercise testing (CPET) has been extensively used and often referred to as the ultimate indicator of postoperative morbidity (14-16). Interestingly, high technology CPET has suffered the same fate as FEV₁ (and DLO₂) inasmuch as decreasing cutoffs have been proposed to define “the unacceptable surgical risk” (2). Recently, the ESTS-ERS guidelines have identified this cut-off in 10 mL·kg⁻¹·min⁻¹ or 35% of the predicted VO₂max value (2). Based on the review of the available literature, the limit for operability was lowered to 10 mL·kg⁻¹·min⁻¹ (or <35%) from the previous one (15 mL·kg⁻¹·min⁻¹) (2). Our series of surgical patients operated on prior to publication of these guidelines confirmed that the morbidity rates did not differ
significantly by altering the VO$_2$ max cutoff from previously accepted values. However, we observed that the 4 patients who died within 30 days from surgery had median spirometric values within the range of normality. Nevertheless, VO$_2$ max values were consistently around or below the accepted minimal cutoff for operability apart from one patient. However, there was a discrepancy between VO$_2$ max absolute and percentage values in our 4 patients who died within 30 days of the operation despite the significant correlation in the rest of our patient population. This finding seems to suggest that the pathophysiology of the limitation of performance during CPET may indeed result from failure of different systems and hence entail different prognostic significance. In this setting, the mortality in our series was due to pulmonary embolism and bronchopleural fistula complicated by sepsis. Predicting such adverse events despite an optimal prophylactic and postoperative management is a difficult task. In this setting, in order not to deny a potentially curative treatment to lung cancer patients, thoracic surgeons are faced from time to time with the occasional patient who seems to escape risk management guidelines. In those circumstances, the “clinical instinct” comes into play and surgical eligibility is warranted despite borderline - or even adverse - function tests. In this setting, anthropometric measurements, location of bronchial obstruction, different protocol criteria or technologies for CPET, origin of exercise limitation (i.e., deconditioning, primary heart or lung condition, peripheral arteriopathy), nutritional profile - all may impact the accuracy of VO$_2$ max estimation (17,18).

We believe that there are limitations to this study. This is a preliminary work which needs to be supported by a currently ongoing, prospective analysis comparing patients accepted for surgery to excluded, non surgical candidates. Also, the surgical risk should be stratified between central and peripheral lesions since type of operation and induction treatment play a major role in determining major morbidity. Nevertheless, this is a retrospective analysis of prospectively entered relatively limited sized series of patients operated on directly or under the supervision of a single surgeon. By spirometric data, the majority of these patients did present “normal” functional (spirometric) values - a concept of normality which needs careful revision since it does not entail total risk avoidance. Indeed, only a minority of high risk patients do require extraordinary clinical and objective evaluation. The aim of this preliminary report was to understand whether do we always adhere to guidelines for risk assessment in these patients. Our study is a “snapshot” of a homogeneous clinical conduct in a time when several guidelines were published and none accurate enough to be fully adopted (12). Certainly, guidelines play a crucial role in contributing to the education of young surgeons in training, whereas experienced surgeons tend to still rely on clinical judgment especially when dealing with borderline patients (12).

In this setting, Ferguson and colleagues have clearly raised the issue of further investigating the etiology, accuracy and utility of surgeon risk estimates (12). The interpretation of patient-specific VO$_2$ max values is more likely to contribute to risk assessment since these values may reflect the primarily affected component among the determinants of maximal oxygen consumption (18). Hence, should this concept be confirmed by larger studies, the adoption of generalized VO$_2$ max cut-off values in risk assessment for lung surgery may in the future become a contradiction in terms. Risk scores will have to be revised taking into account the primum movens of exercise limitation of individualized, patient-specific, VO$_2$ max cut-off values.

### Acknowledgements

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

### References


