Differences between remaining ability and loss of capacity in maximum aerobic impairment

J.A. Neder, L.E. Nery, E. Bagatini, S.R. Lucas, M.S. Anção and D.Y. Sue

Abstract

In the evaluation of exercise intolerance of patients with respiratory diseases the American Medical Association (AMA) and the American Thoracic Society (ATS) have proposed similar classifications for rating aerobic impairment using maximum oxygen uptake (VO₂max) normalized for total body weight (ml min⁻¹ kg⁻¹). However, subjects with the same VO₂max weight-corrected values may have considerably different losses of aerobic performance (VO₂max expressed as % predicted). We have proposed a new, specific method for rating loss of aerobic capacity (VO₂max, % predicted) and we have compared the two classifications in a prospective study involving 75 silicotic claimants. Logistic regression analysis showed that the disagreement between rating systems (higher dysfunction by the AMA/ATS classification) was associated with age >50 years (P<0.005) and overweight (P = 0.04). Interestingly, clinical (dyspnea score) and spirometric (FEV₁) normality were only associated with the VO₂max, % predicted, normal values (P<0.01); therefore, in older and obese subjects the AMA/ATS classification tended to overestimate the aerobic dysfunction. We conclude that in the evaluation of aerobic impairment in patients with respiratory diseases, the loss of aerobic capacity (VO₂max, % predicted) should be used instead of the traditional method (remaining aerobic ability, VO₂max, in ml min⁻¹ kg⁻¹).

Introduction

Relating the loss of pulmonary function to the degree of exercise intolerance (impairment) (1,2) remains a controversial aspect of occupational medicine (1,6). Complaints of dyspnea in patients with occupation-related pulmonary disorders often exceed objective findings, especially in claimants, and it is well recognized that accurate assessment of work capacity must consider that subjects may be limited by non-respiratory problems (7). Analysis of metabolic and ventilatory responses during exercise has significantly improved the quality of both occupational and non-occupational respiratory disease evaluation (8). The American Thoracic Society (ATS; 1,2) and the American Medical

Correspondence
L.E. Nery
Disciplina de Pneumologia
EPM, UNIFESP
Rua Botucatu, 740, 3º andar
04023-062 São Paulo, SP
Brasil
Fax: 55 (011) 570-2127

Research partially supported by CNPq and FAPESP. Part of a Doctoral thesis presented by J.A. Neder to the Disciplina de Pneumologia, UNIFESP-EPM.

Received May 21, 1997
Accepted February 13, 1998

Key words
• Disability evaluation
• Work capacity
• Exercise tolerance
• Oxygen consumption
• Occupational diseases
Association (AMA; 3-5) have proposed similar criteria for rating “remaining aerobic ability”, based on maximum oxygen uptake (VO2max) attained in the laboratory. The observed values are normalized by actual weight (ml min⁻¹ kg⁻¹) allowing comparisons with the estimated work metabolic stress. For example, if an individual’s VO2max is found to be greater than 25 ml min⁻¹ kg⁻¹, remaining aerobic ability is judged normal using this scheme (Table 1).

Nevertheless, patients with identical measured VO2max (expressed as l/min) may demonstrate marked differences between remaining ability (VO2max, ml min⁻¹ kg⁻¹) and loss of capacity (VO2max, % predicted). A younger, non-obese individual with a high predicted normal VO2max may have sustained a considerable loss of capacity and yet have a “normal” remaining capacity (VO2max >25 ml min⁻¹ kg⁻¹). On the other hand, an older and especially overweight individual may have an abnormally low VO2max expressed ml min⁻¹ kg⁻¹ even though the loss of aerobic capacity expressed as VO2max, % predicted, may be slight. Therefore, rating aerobic dysfunction using only remaining ability (VO2max, ml min⁻¹ kg⁻¹) could lead to major discrepancies when compared to a scheme expressing loss of capacity (VO2max, % predicted).

The aim of the present study was to compare conclusions reached by two systems of rating aerobic dysfunction based on VO2max measured in a heterogeneous group of silicotic claimants: 1) a system similar to the AMA/ATS classification (1-5)(VO2max, ml min⁻¹ kg⁻¹) and 2) a new system based on reduction of VO2max values compared to predicted (VO2max, % predicted; Table 1). We hypothesized that there would be significant differences between the two systems that could lead to potentially different conclusions about impairment in individual subjects.

### Table 1 - Aerobic dysfunction according to a classification for rating remaining ability (AMA/ATS) and a new proposal for rating loss of capacity.

<table>
<thead>
<tr>
<th></th>
<th>None</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMA/ATS VO2max (ml min⁻¹ kg⁻¹)</td>
<td>&gt;25</td>
<td>20-25</td>
<td>15-20</td>
<td>&lt;15</td>
</tr>
<tr>
<td>New proposal VO2max (% predicted)</td>
<td>&gt;70</td>
<td>60-69</td>
<td>40-59</td>
<td>&lt;40</td>
</tr>
</tbody>
</table>

### Material and Methods

#### Subjects and clinical evaluation

A group of subjects undergoing evaluation was available for comparison of aerobic performance. These claimants with occupational exposure to free silica in the ceramics industry generally had some degree of dyspnea, and had radiographic abnormalities suggestive of silicosis. No patient was excluded due to a co-morbidity factor if the latter was compensated (systemic arterial hypertension, diabetes mellitus). Any medications utilized by the patients were maintained. Informed consent to participate in the study was obtained from all subjects. The protocol was submitted to and approved by the Ethics Committee of UNIFESP/EPM-HSP.

Clinical, occupational and smoking histories were obtained; physical examination and anthropometric evaluation were performed in all subjects. A dyspnea score was recorded using a Cotes modified scale (9) in which a normal degree of respiratory effort was defined as breathlessness occurring while walking uphill or climbing stairs, but not with a lesser degree of exertion. Posteroanterior chest radiographs were interpreted by two experienced readers according to the International Labor Organization (10) classification of radiographs of pneumoconiosis.

#### Spirometry

Pulmonary function tests were performed
with a CPF-System (Medical Graphics Corporation (MGC), St. Paul, MN) utilizing a Fleisch-type No. 3 pneumotachograph. Measurements were performed according to standard protocols following ATS guidelines (11), pre- and post-200 µg of salbutamol administered with a metered-dosed inhaler. Forced expiratory volume in 1 s (FEV₁) and forced vital capacity (FVC) were recorded in all subjects. Predicted normal values were those of Knudson et al. (12). Maximal voluntary ventilation (MVV) was estimated as the product of FEV₁ x 40 (13).

**Cardiopulmonary exercise testing**

The exercise study was performed with the subject sitting on a calibrated, electromagnetically braked cycle ergometer (CPE-2000, MGC). The cycle seat was adjusted for the subjects’ comfort. The cycle ergometer was interfaced with a digital computer-based exercise system (CPX-System, MGC) which increased the work rate at a constant rate (ramp protocol). The increase in work rate (10 to 20 watts/min) was chosen in such a way that the symptom-limited maximum was obtained between 8 and 12 min. Subjects breathed through a 115-ml deadspace, low-resistance Hans-Rudolph valve. The expiratory arm of the valve was connected to a Fleisch-type No. 3 pneumotachograph and the flow signal was electronically integrated to obtain volume variables. A 1-ml/s sample of expired gas from the mouthpiece was drawn for determination of expired O₂ and CO₂ with a zirconium cell and rapid infrared analyzer, respectively. Gas measurements and flow were analyzed breath-by-breath after phase delay correction (13). All subjects had their heart rate measured and electrocardiogram obtained each minute; subjects were also monitored by pulse oximetry (OxyShuttle™ SensorMedics, SensorMedics Corp., Anaheim, CA).

The exercise test consisted of measurements of gas exchange and flow at rest for 2 min, 2 min while the subject was pedaling without a load, and then while the work rate was increased until the subject indicated that he could not continue pedaling at least at 50 rpm, or the test was stopped by the attending physician. All subjects were encouraged to perform as long as possible. At the end of exercise, the subject was asked to describe his symptoms and to rate the severity of dyspnea, fatigue and leg pain (modified Borg scale; 14). Exercise could be terminated by the physician if a subject had a fall in systolic blood pressure (BP) of more than 15 mmHg, systolic BP >250 mmHg, diastolic BP >120 mmHg, progressive horizontal or downsloping ST-segment depression, T-wave inversion or the appearance of Q waves, increased frequency of premature ventricular contraction, ventricular tachycardia (3 or more consecutive ectopic beats), onset of atrial fibrillation or atrial tachycardia (15), or pulse oximeter oxyhemoglobin saturation (SaO₂) <80%.

The following variables were measured during the test: breath-by-breath O₂ uptake (VO₂, ml/min, standard temperature and pressure, dry; STPD), CO₂ output (VCO₂, ml/min STPD), respiratory exchange ratio (R), minute ventilation (VE, l/min, body temperature, ambient pressure, saturated BTPS), breathing frequency (f, bpm), tidal volume (VT, ml/min BTPS), ventilatory equivalents for O₂ and CO₂ (VE/VO₂, VE/VCO₂), end-tidal PO₂ and PCO₂ (PETO₂, PETCO₂, mmHg), dead space/tidal volume ratio (VD/VT), heart rate (HR, b/min) and oxygen pulse (VO₂/HR, ml/beat).

The VO₂ achieved by the subjects at maximal effort, limited by symptoms or by medical reasons, was considered to be the VO₂max (peak VO₂), and was defined as the highest value obtained in the analysis of 8-breath moving average. VO₂max was expressed as % predicted according to the reference values of Hansen and coworkers (16): i) in eutrophic and underweight males, the predicted VO₂max (ml/min) was calculated us-
ing the actual weight: weight (kg) x 50.72 - 0.372 x age (years) and, ii) in overweight subjects, the ideal weight normalized for height was used: VO₂max = (0.79 x height (cm) - 60.7 x 50.72 - 0.372 x age (years).

The anaerobic threshold (VO₂AT) was estimated by gas exchange at the point at which VCO₂ increased out of proportion to VO₂ (V-slope) and when the VE/VO₂ ratio and the PETO₂ increased while the ratio VE/VCO₂ and PETCO₂ remained constant (13). The lower limit of normality for VO₂AT was set at 40% of predicted VO₂max (13). The reading was performed by two observers who always agreed in their analysis. The predicted maximal HR (HRmax) was calculated as 220 - age (13). The chronotropic reserve was calculated by the equation: 1 - HRmax/HRpred, and the breathing reserve using the standard equation 1 - VEmax/MVV.

Statistical analysis

The Kappa reliability test was used to determine the presence of significant concordance between the classifications of aerobic dysfunction. Logistic regression analysis was performed to correlate the demographic and anthropometric variables with the discordance between the classifications. The Mann-Whitney U-test was carried out to contrast excess weight values among groups. To evaluate the associations among some studied variables, we used chi-square analysis (Fisher’s exact test). For all tests, a P value <0.05 was considered to be statistically significant.

Receiver operating characteristic (ROC) analysis was used to choose an appropriate VO₂max % predicted cutoff value for comparison to the VO₂max ml min⁻¹ kg⁻¹ criterion and to test the relative accuracy of the two systems (17). A ROC plot consists of a statistical method designed to evaluate the overall relationship between sensitivity and specificity of a test against a selected “gold standard”. Assuming a necessary inverse relationship between sensitivity and specificity, the best decision threshold corresponds to the best possible combination of sensitivity and specificity and, therefore, the area under the curve is closely related to the overall accuracy of the test. In this experiment, the selected gold standard was the VO₂AT, a variable that provides a global assessment of aerobic performance independent of effort or of achieving a true VO₂max (13). Thus, a ROC plot was constructed by plotting the true-positive fraction (sensitivity) against the false-positive fraction (1-specificity) for multiple VO₂max decision thresholds using discrete test values (17).

Table 2 - Subject characteristics, pulmonary function and exercise performance.

<table>
<thead>
<tr>
<th></th>
<th>Total (N = 75)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>52.2 ± 9.0</td>
</tr>
<tr>
<td>Actual weight (kg)</td>
<td>71.9 ± 12.9</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>166.2 ± 7.4</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>67</td>
</tr>
<tr>
<td>Female</td>
<td>8</td>
</tr>
<tr>
<td>Smoking history</td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>34</td>
</tr>
<tr>
<td>Former or current</td>
<td>41</td>
</tr>
<tr>
<td>Pack-years of smoking</td>
<td>20.0 ± 14.1</td>
</tr>
<tr>
<td>Free-silica exposition</td>
<td>23.9 ± 8.2</td>
</tr>
<tr>
<td>Spirometry</td>
<td></td>
</tr>
<tr>
<td>FEV₁ (l)</td>
<td>2.78 ± 0.82</td>
</tr>
<tr>
<td>FEV₁ (% predicted)</td>
<td>92.4 ± 24.1</td>
</tr>
<tr>
<td>FVC (% predicted)</td>
<td>102.4 ± 21.5</td>
</tr>
<tr>
<td>FEV₁ (%)</td>
<td>73.4 ± 10.3</td>
</tr>
<tr>
<td>Maximum exercise (max)</td>
<td></td>
</tr>
<tr>
<td>VO₂max (ml/min)</td>
<td>1452.10 ± 518.07</td>
</tr>
<tr>
<td>VO₂max (% predicted)</td>
<td>70.2 ± 20.0</td>
</tr>
<tr>
<td>VO₂max (ml min⁻¹ kg⁻¹)</td>
<td>20.3 ± 7.0</td>
</tr>
<tr>
<td>VO₂AT (%VO₂max predicted)</td>
<td>44.3 ± 11.0</td>
</tr>
<tr>
<td>HRmax (% predicted)</td>
<td>80.7 ± 11.8</td>
</tr>
<tr>
<td>O₂-Pulse max (ml/b)</td>
<td>10.6 ± 3.6</td>
</tr>
<tr>
<td>VEmax/MVV</td>
<td>0.49 ± 0.15</td>
</tr>
<tr>
<td>SaO₂ max (%)</td>
<td>95.2 ± 1.6</td>
</tr>
</tbody>
</table>
Results

Seventy-five subjects (67 males and 8 females) participated. Mean age was 52.2 years and many subjects had a history of smoking. Fifty-eight subjects (55 males and 3 females) were classified as having small opacities and 17 (12 males and 5 females) as having large opacities by roentgenographic analysis (10). Clinical, spirometric and exercise findings are presented in Table 2.

In the new classification, the lower limit of normality (70% of predicted VO₂max) was the best possible combination between sensitivity and specificity changing the VO₂max % predicted values against the gold standard (VO₂AT). Thus, the point closest to the upper left corner of Figure 1 was selected as the best cutoff for normality for the new method.

In a global comparison between the two schemes, we found that of 75 studied subjects, only 19 (25.3%) were classified as normal by the AMA/ATS criteria (VO₂max >25 ml min⁻¹ kg⁻¹). On the other hand, 40 of 75 subjects (53.3%) had a VO₂max >70% predicted (no loss of aerobic capacity), including 21 of 56 subjects with VO₂max <25 ml min⁻¹ kg⁻¹. No subject with VO₂max >25 ml min⁻¹ kg⁻¹ had a VO₂max <70% predicted (Table 3).

We anticipated that age might have an impact on the different schemes due to the generally lower predicted VO₂max in older subjects (above 50 years). As shown in Table 3, there was a significant difference regarding the gradation of aerobic dysfunction just in the older group (k = 0.03, P = 0.29). Among patients older than 50 years (group II), of 22 patients with VO₂max >70% predicted, only 5 (22.7%) had normal remaining aerobic ability. Furthermore, only 4 of these older subjects were considered to have severe functional loss (VO₂max <40% predicted), compared to 17 classified as severe on the basis of VO₂max <15 ml min⁻¹ kg⁻¹. Therefore, in patients >50 years old, VO₂max expressed as ml min⁻¹ kg⁻¹ frequently suggested a lower remaining ability compared to the milder degree of loss of capacity found by looking at VO₂max % predicted. In summary, the interpretation of the two VO₂max analyses was concordant for 61% of subjects younger than 50 years as opposed to only 23% of 44 older subjects (Table 3).

We also hypothesized that overweight individuals would be more likely to show a difference between the two schemes. Those subjects who showed concordance for VO₂max (normal or abnormal for both meth-
ods of analysis) were as a group closer to normal body weight than those who were discordant, and this was true for both younger and older subjects (Figure 2). Using a logistic regression analysis that considered disagreement between the classifications as dependent event and demographic and anthropometric variables as independent variables, only age and actual weight contributed to explaining the discordance (age, P = 0.0006; actual weight, P = 0.04; sex, P = 0.15; height, P = 0.84).

We also compared the normality suggested by the two classifications with that indicated by a clinical (dyspnea score) and respiratory function (FEV₁ above lower 95% confidence limit) indicator. In older subjects (group II), a significant association was found between the absence of functional loss (VO₂max >70% predicted) and a dyspnea score <1 and a normal FEV₁. A similar relationship between these variables and VO₂max (ml min⁻¹ kg⁻¹) was not found (Figure 3).

Finally, a direct comparison between ROC curves and the areas under the curves demonstrated a better performance of the new proposal (VO₂max, % predicted) compared to current recommendations (VO₂max, ml min⁻¹ kg⁻¹) (Figure 1). In addition, contrasting the lower limits of normality according to the two classifications showed a significant gain in specificity using the 70% value of VO₂max predicted. Thus, while sensitivity remained unaltered, the new proposal allowed an improvement in specificity from 32.6% to 70% with a likelihood ratio increasing from 1.26 to 4.06 (Figure 1).

Discussion

In this comparative study of two ways of expressing aerobic dysfunction (VO₂max, ml min⁻¹ kg⁻¹, and VO₂max, % predicted) in a group of silicotic claimants, we found important disagreement in classifying normal subjects and in estimating the degree of abnormality. In particular, for subjects older than 50 years and for overweight subjects (actual weight above the ideal weight for height), using VO₂max (ml min⁻¹ kg⁻¹) would frequently lead to the conclusion of a greater aerobic dysfunction than when using the VO₂max % predicted expression. Our findings suggest that the present recommendations for rating aerobic dysfunction in pulmonary patients could cause considerable misinterpretation about the presence and degree of impairment.

Work-related respiratory limitation of exercise has long been quantified in classifications of aerobic dysfunction, usually by correcting the VO₂max values for body weight (1-5,18,19). The basic premise for the use of VO₂max (ml min⁻¹ kg⁻¹) to assess impairment is that it allows an immediate comparison with the metabolic demands of the job, therefore providing an answer to the presumed main question: what is the residual ability of the worker in performing his task with comfort and safety? This traditional approach based only on remaining
ability is appropriate for job placement or “fitness-for-duty evaluations”, but ignores the potential physiological importance of loss of aerobic capacity.

Our results indicated that recommendations based on VO\textsubscript{2max} (ml min\textsuperscript{-1} kg\textsuperscript{-1}) can be significantly different from a classification based on loss of aerobic capacity (Table 3). The disagreement between the classifications is particularly noticeable in older and heavier individuals. Thus, in those older than 50 years, with low absolute VO\textsubscript{2max} values (l/min), the actual weight correction in overweight individuals led to lower remaining ability than functional loss. In these patients, the loss of aerobic capacity was proportionally small in view of their low predicted values. These findings are particularly important considering that this demographic and anthropometric profile is the most prevalent in claimants with respiratory diseases (20).

VO\textsubscript{2max} normalization for actual weight\textsuperscript{-1}, a common procedure in sports medicine, should be justifiable only in activities where the total dynamic displacement of the body is the rule (fast walking, running) (21) and not the exception, as in most industrial plants and extractive sites (22). In activities without active body movement, such as short distance walking, cycling and weight lifting, the normalization of actual weight for height (ideal weight) would prevent the overestimation of the predicted values of VO\textsubscript{2max} in overweight subjects (16).

The lower limit of normality established by the new proposal (70% of predicted VO\textsubscript{2max}) was a best decision threshold obtained from a ROC curve constructed with VO\textsubscript{2AT} as the criterion for aerobic dysfunction (Figure 1). VO\textsubscript{2AT} is defined as the level of exercise VO\textsubscript{2} above which oxidative-based energy production is overlapped by anaerobic mechanisms. Typically, this is associated with increases in blood lactate, ventilatory and cardiac responses, catecholamine release, and a reduction of endurance time (13). Furthermore, the effort-independent characteristic of this parameter is especially interesting for the aerobic evaluation of claimants (23). The intermediate impairment cutoffs follow those usually recommended for pulmonary function tests (11). The limit of severe aerobic impairment was established at 40% of predicted VO\textsubscript{2max} considering that it is the lower limit of normality for VO\textsubscript{2AT} (16). Thus, if a claimant shows a VO\textsubscript{2max} below this value, obviously his VO\textsubscript{2AT} will be very low and he is likely to utilize predominantly unsustainable anaerobic sources of energy in his activities.

Usually, the validation of a classification of aerobic dysfunction for occupational purposes is done in comparison with ergonomic assumptions. In order to use a less controversial criterion, we compared the normality detected by the two tested schemes with a clinical score of dyspnea (9), and a measure of ventilatory limitation at rest (FEV\textsubscript{1}) (11). The significantly stronger association between normality detected by the new proposal (VO\textsubscript{2max} >70\% predicted) and that suggested by such parameters (Figure 3) indicated that a large number of patients over 50 years and overweight were incorrectly considered impaired on the basis of the present recommendations.

The utilization of VO\textsubscript{2AT} as a gold standard of aerobic dysfunction in a ROC analysis permitted a comparison of the two classifications along the entire spectrum of possible decision thresholds. The visual inspection of the curves and the direct comparison between the cutoff points of normality suggested by the two classifications showed a significant gain in specificity in the new proposal with the sensitivity of the AMA/ATS scheme being maintained. When we analyzed the points on the ROC curves above 60\% sensitivity and specificity we observed a striking superiority of VO\textsubscript{2max} expressed as % of predicted values (Figure 1). These findings are consistent with the tendency of excessive false-positive rates for aerobic ab-
normality obtained when using the AMA/ATS classification (ml min⁻¹ kg⁻¹).

In conclusion, our data indicate that in the assessment of aerobic dysfunction in pulmonary patients, loss of aerobic capacity expressed as VO₂max % predicted should be an important parameter to be considered in the evaluation of potential impairment. Thus, instead of the comparison with the metabolic requirements of the work (remaining aerobic ability, VO₂max in ml min⁻¹ kg⁻¹ or l min⁻¹), loss of capacity should be considered to optimize the maximum VO₂ analysis. However, additional research is necessary to evaluate whether or not these findings are also applicable to normal aged overweight subjects.

Acknowledgments

The authors thank Profs. Neil F. Novo and Adauto Castelo Filho, Universidade Federal de São Paulo, Escola Paulista de Medicina, for their help with the statistical analysis.

References