Diagnostic value of reversibility of chronic airway obstruction to separate asthma from chronic bronchitis: a statistical approach

N. Meslier*, J.L. Racineux**, P. Six†, A. Lockhart††

ABSTRACT: Our aim was to assess the usefulness of tests of reversibility of airways obstruction (AO) in the differential diagnosis of asthma (A) and chronic bronchitis (CB). We selected on strict clinical criteria 20 CB and 32 A patients with stable AO and measured maximal forced expiratory flows before and after increasing doses of inhaled salbutamol. The highest sensitivity (Se), specificity (Sp) and likelihood ratio (L) were obtained when the improvement in forced expiratory volume in one second (FEV₁) was expressed as a percentage of predicted FEV₁, Se, Sp and L were 0.97, 0.95 and 19.3 (highest possible L: 20) for an increase of 10% of predicted FEV₁, on 0.2 mg of salbutamol, and 1.0, 1.0 and 20, respectively, for an increase of 15% on 1 mg of salbutamol. Conversely, L was only 2.5 when the widely used increase in FEV₁ of 20% over baseline was tested. In a second group of 42 unselected patients, who were referred to the laboratory for routine lung function testing with a presenting diagnosis of A or CB, L was only 5.5 when the 10% increase in predicted FEV₁ was used and even lower for the classical 20% increase over baseline FEV₁. We submit that routine tests of reversibility of AO have little diagnostic usefulness in unselected patients with AO and that the commonly used criteria of 15 or 20% increase in FEV₁ over baseline are of no value in separating A from CB.

The improvement in airways obstruction produced by administration of a bronchodilator is often used in an attempt to differentiate asthma from non-asthmatic obstruction to airflow. In practice a threshold of improvement in a measurement of forced expiratory flow is considered to indicate "reversibility" but the information actually sought is the probability that the patient is asthmatic if the measured change exceeds a given threshold. This amounts to determining the positive predictive value of the test which depends on the sensitivity and the specificity of the threshold chosen and on the prevalence of asthma amongst the population tested [1]. The aim may be to help with the diagnosis and, thus, the treatment of an individual case [2, 3], or to define a population, as for a therapeutic trial [4-8]. There is no consensus on the index to be measured, the mode of expression of the result, and the choice of the threshold of improvement. For example, "reversibility" has been defined by an improvement in forced expiratory volume in one second (FEV₁) of 12% [9], 15% [8] or 20% [4, 5, 7] of its initial value. An improvement in FEV₁ of more than 15% of predicted FEV₁ has been claimed to be characteristic of asthma [10]. PENNOCK et al. [11] proposed a statistical approach based on the reproducibility of the measurements to define a significant variation. Similar discrepancies exist for other indices than FEV₁. As a result the diagnostic value of the test of reversibility for separating asthma from chronic bronchitis is not firmly established.

The objective of our study is to assess the effect of the index, the mode of expression and the dose of inhaled bronchodilator on the results of the reversibility test and, therefore, to better document the value of this test in bronchial obstruction: a) for the selection of well-defined populations of asthmatic patients or chronic bronchitic patients; and b) as a diagnostic criterion in clinical practice.

Patients and methods

The 94 patients studied formed two distinct populations whose characteristics are summarized in table 1.

Group I

All the patients in group I were selected on the basis of strict clinical criteria without any knowledge of their response to a bronchodilator on pulmonary function tests. Thirty two patients with a clinical diagnosis of asthma and twenty patients with a clinical diagnosis of
chronic bronchitis were selected. Patients whose clinical history was not characteristic of asthma or chronic bronchitis were not included in the study. The asthmatic patients had a history of acute dyspneic episodes with wheezing improving with bronchodilator treatment [12] but had no chronic cough and sputum production. Three had been smokers, but their smoking had been interrupted for at least seven years at the time of the study. Chronic bronchitis was defined by the Medical Research Council criteria [13], i.e. as a history of chronic cough and sputum without any acute episode of reversible obstruction. All had been heavy smokers. At the time of inclusion all patients were in a stable state, i.e. they had not suffered from an acute exacerbation and their treatment had been unchanged for at least two months. In all of the patients, inhaled beta-2-agonists were discontinued at least 8 h before the trial. In asthmatic patients, xanthine derivatives were withdrawn 48 h before the trial and the blood theophylline level on the study day was always less than 1 mg·l⁻¹. Five chronic bronchitic patients were not taking xanthine derivatives. The theophylline level in blood measured 10 min before the test was lower than 10 mg·l⁻¹ in the 15 remaining patients (mean±SD=5.36±3.34 mg·l⁻¹). No chronic bronchitic patient was receiving corticosteroid therapy. Eleven asthmatic patients were receiving corticosteroid treatment but the dosage had been unchanged for more than one month.

**Methods**

The maximal expiratory flow rates were measured in the sitting position with a Fleisch No. 3 pneumotachograph. Exhaled volumes were measured by integration of expiratory flows. Recording was carried out simultaneously on X-Y co-ordinates for the flow-volume loop and as a function of time for expiratory flows. We measured four indices of forced expiration: FEV₁, peak expiratory flow rate (PEFR), maximal mid-expiratory flow (MMEF), and forced expiratory flow at 50% of the vital capacity (FEF₅₀). FEV₁ was determined according to Kory et al. [14]. MMEF₂₅-₇₅% before and after use of a bronchodilator was measured between volumes equal to 25 and 75% of the initial slow vital capacity. FEF₅₀ was measured at 50% of the observed slow vital capacity. For each subject, forced expiration was repeated until at least three acceptable flow-volume loops were obtained [15] and the retained value was the mean of the three best measurements [16]. The bronchodilator test was performed by inhaling salbutamol from a metered-dose inhaler. All patients in group I inhaled 0.2 mg salbutamol. Twenty-four of the 32 asthmatic patients and all of the 20 chronic bronchitic patients also inhaled 0.8 mg salbutamol 30 min later. The second dose thus corresponded to a cumulative dose of 1 mg [17]. The forced expiratory manoeuvres were repeated 15 min after each inhaled dose [18]. The improvement was expressed in three different ways: 1) the absolute difference calculated from the raw data (A); 2) the difference expressed as a percentage of the predicted value (A%pred); 3) the difference expressed as a percentage change from the initial value (A%IV). In group II, for routine bronchodilator tests, the patients only inhaled 0.2 mg salbutamol. Forced expiratory manoeuvres were performed with the same equipment and by the same technician in the two groups. The predicted values are those of Amrein et al. [19] for FEV₁, Bass [20] for FEF₅₀, and PEFR, and Prehaut et al. [21] for MMEF₂₅-₇₅%.

The effectiveness of the reversibility test for distinguishing asthma from chronic bronchitis on the basis of the variation in a given index after bronchodilator was defined in terms of its sensitivity (Se), specificity (Sp) [22] and likelihood ratio (L) [23]. The sensitivity, Se=true positives/(T(+))/(T(+)+false negatives(F(-)))

### Table 1. Characteristics of group I and group II

<table>
<thead>
<tr>
<th></th>
<th>Group I</th>
<th></th>
<th>Group II</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>32</td>
<td>20</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td>Age yrs</td>
<td>53.6±12.2</td>
<td>62.0±8.1</td>
<td>53.0±15.9</td>
<td>63.4±7.5</td>
</tr>
<tr>
<td>Initial FEV₁ %pred</td>
<td>43.6±15.7</td>
<td>41.3±16.8</td>
<td>54.0±18.6**</td>
<td>39.7±11.0</td>
</tr>
<tr>
<td>Pred FEV₁ l</td>
<td>2.91±0.55</td>
<td>3.03±0.49</td>
<td>2.90±0.64</td>
<td>2.94±0.65</td>
</tr>
</tbody>
</table>

****: indicates that the difference between the asthmatic (A) and the chronic bronchitic patients (CB) is significant with p<0.01. Values for age and forced expiratory volume in one second (FEV₁) are expressed as mean±SD.
answers the question "if the patient is asthmatic, how likely is he to obtain an improvement greater than a given threshold?". The specificity, $Sp = true negatives (T(-))/[false positives (F(+)) + true negatives (T(-))]$, answers the question "if the patient has a non-asthmatic Airways obstruction, how likely is he to obtain an improvement lower than the threshold?". The positive likelihood ratio, $L = [T(+)/N1]/[F(+)/N2]$, depends on the ratio of the true positives and the false negatives and the size $N1$ and $N2$ of the two populations: $L = [T(+)\times N1]/[F(+)\times N2]$. $L$ maximal value, which is obtained when there are neither false positives nor false negatives, is theoretically infinite. However, in the absence of false positives $F(+) L$ is given the value 1 instead of 0. Thus the maximal possible value of $L$ in our population is 20. The more $L$ approaches its maximal possible value, the better is the diagnostic value of the test. The respective values of true and false positives for different thresholds define the ROC curve (receiver operator characteristic) [23]. On this curve (fig. 1), if the threshold is displaced towards the right, beyond point C, a minimal gain of sensitivity results in a major loss of specificity. Conversely, if the threshold is moved towards the left, beyond point A, a minimal gain of specificity is accompanied by a dramatic decrease in sensitivity. The ideal threshold is located at B, on the convexity, which corresponds to the greatest sensitivity and specificity. We calculated therefore, sensitivity, specificity and likelihood ratio for different thresholds of the three modes of expression ($A$, $A\%pred$ and $A\%IV$) for the four indices under study.

The bronchodilator test used as a diagnostic criterion in evaluating patients with airflow obstruction for clinical practice or for selecting a population for a study must answer the questions: "if the improvement is greater than a given threshold, what is the probability of asthma?" and "if the variation is less than a given threshold, what is the probability of chronic non-asthmatic airflow obstruction?". These questions define the positive and negative predictive values of the test [1]. They are related to the prevalence ($Pr$) of asthma in chronic airflow obstruction, or the prior or pretest probability of asthma. Positive predictive value,

$$PV_+ = Pr\times Se/[Pr\times Se + (1-Pr)\times (1-Se)]$$

and negative predictive value,

$$PV_- = (1-Pr)\times Sp/[Pr\times Sp + Pr\times (1-Se)]$$

We computed predictive values for two likely prevalences of asthma in chronic airflow obstruction: 30 and 50%.

We carried out a discriminant analysis using as a variable the variation of the index expressed as a percentage of the predicted value ($A\%pred$). This analysis was performed for each index and for the association of two, three or four indices.

Student's $t$-test was used for statistical comparisons between populations.

**Results**

**Effects of salbutamol in group I**

*Variation of the forced expiration indices.* The distribution of the raw data changes ($A$) and increase in $FEV_1$ expressed as $A\%pred$ and $A\%IV$ are represented in figure 2. The method of expression of the variation altered the distribution of both A and CB. After inhalation of 0.2 mg salbutamol, a single asthmatic patient had $A\%IV < 20\%$ but seven of the 20 chronic bronchitic patients had an improvement above this threshold. When the variation was expressed as $A\%pred$, 31 of the 32 asthmatic patients achieved an improvement greater than $10\%$ whereas only one chronic bronchitic patient reached this threshold. After inhalation of 1 mg salbutamol, an increase in $FEV_1$ greater than $15\%pred$ or 0.45 $l$ was achieved by asthmatic patients only. For the other indices of forced expiration, whatever the dose and the mode of expression, no threshold completely separated asthmatic from chronic bronchitic patients (figs 3–5).

*Sensitivity, specificity and predictive values as a function of the mode of expression of the results and of the dose of bronchodilator.* Sensitivities, specificities and likelihood ratios of the bronchodilator response in $FEV_1$ for the diagnosis of asthma are presented in tables 2 and 3 for the inhaled dose of 0.2 and 1 mg salbutamol, respectively. Regardless of the mode of expression, the rise in $FEV_1$ after 0.2 mg had a high sensitivity for detecting an obstructive asthma: $Se > 0.9$ for most thresholds. With the expression in $A\%IV$, high thresholds were required to obtain a specificity greater than 0.9 and $Sp$ was only 0.65 with a poor likelihood ratio ($L=2.8$) for a
**Fig. 2.** - Histograms of the variations of forced expiratory volume in one second (FEV₁) after 0.2 (left column) and 1 mg of salbutamol (S) (right column). Ordinate axis: number of asthmatic and chronic bronchitic patients. Increase in FEV₁ is expressed as an increase relative to its initial value (Δ%TV), as a percentage of the predicted FEV₁ (Δ%pred), or as an absolute value in ml (Δ).

**Fig. 3.** - Histograms of changes in peak expiratory flow rate (PEFR). Same format as figure 2.
Fig. 4. – Histograms of changes in mid-expiratory flow rate between 25–75% slow vital capacity (MMEF$_{25-75}$). Same format as figure 2.

Fig. 5. – Histograms of changes in forced expiratory flow at 50% slow vital capacity (FEF$_{50}$). Same format as figure 2.
Table 2. - Sensitivity (Se), specificity (Sp) and likelihood ratio (L) of the response in FEV\(_1\) to 0.2 mg of inhaled salbutamol for diagnosis of asthma

<table>
<thead>
<tr>
<th>(\Delta l)</th>
<th>(\Delta%IV)</th>
<th>(\Delta%pred)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEV(_1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.25</td>
<td>0.3</td>
</tr>
<tr>
<td>Se</td>
<td>1.0</td>
<td>0.97</td>
</tr>
<tr>
<td>Sp</td>
<td>0.55</td>
<td>0.8</td>
</tr>
<tr>
<td>L</td>
<td>2.2</td>
<td>4.8</td>
</tr>
</tbody>
</table>

Se, Sp and L are calculated for three thresholds of each mode of expression of the variation of an index: raw increase (\(\Delta l\)), as a percentage of the initial value (\(\Delta%IV\)) and of the predicted value (\(\Delta%pred\)) in group I. The maximal possible value of L is 20.

Table 3. - Sensitivity (Se), specificity (Sp) and likelihood ratio (L) of the response in FEV\(_1\) to 1 mg of inhaled salbutamol for diagnosis of asthma

<table>
<thead>
<tr>
<th>(\Delta l)</th>
<th>(\Delta%IV)</th>
<th>(\Delta%pred)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEV(_1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>0.45</td>
<td>0.6</td>
</tr>
<tr>
<td>Se</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Sp</td>
<td>0.85</td>
<td>1.0</td>
</tr>
<tr>
<td>L</td>
<td>26.7</td>
<td>20.0</td>
</tr>
</tbody>
</table>

Abbreviations as in table 2. Patients are the 20 CB and 24/32 A of group I.

Table 4. - Positive (PV+) and negative predictive values (PV-) for the diagnosis of asthma of the response in FEV\(_1\) to 0.2 mg of inhaled salbutamol

<table>
<thead>
<tr>
<th>(\Delta l)</th>
<th>(\Delta%IV)</th>
<th>(\Delta%pred)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEV(_1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.25</td>
<td>0.3</td>
</tr>
<tr>
<td>Pr=30%</td>
<td>PV+</td>
<td>PV-</td>
</tr>
<tr>
<td>36</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Pr=50%</td>
<td>PV+</td>
<td>PV-</td>
</tr>
<tr>
<td>80</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

PV+ and PV- are calculated for two arbitrary prevalences (Pr=30% and Pr=50%) of asthma in a population with chronic airflow obstruction, and for three thresholds of each mode of expression of the variation in FEV\(_1\).
was the most powerful index to differentiate asthma from a non-asthmatic airway obstruction (table 5). All of the chronic bronchitic patients were correctly identified whatever the index and the dose. Twenty-six of 32 asthmatic patients were correctly identified by FEV, alone after 0.2 mg salbutamol and 22 of 24 after 1 mg. The association with FEV, of one, two or three other indices added little information. With the association of four indices, 27 of 32 and 22 of 24 asthmatic patients were correctly identified after 0.2 and 1 mg, respectively.

Effect of salbutamol in group II

In patients of group II, sent for a routine bronchodilator test, we tested the threshold of 10%pred increase in FEV, which had shown the highest sensitivity and specificity in group I on 0.2 mg salbutamol. In this group the improvement was lower than this threshold in 18 of 20 patients considered to have a non-asthmatic airways obstruction, but was also lower in 10 of the 22 patients defined as asthmatic. Thus the specificity was 0.9 but the sensitivity was only 0.55 with a likelihood ratio of 5.5.

Discussion

Bronchial obstruction results from several factors: bronchospasm, intraluminal secretions, and oedema and inflammation of the airway wall. The contribution of bronchospasm is greater in asthma than in chronic bronchitis. The response to bronchodilators measured by functional indices depends not only on bronchospasm per se, but also on its relative contribution in relation to the associated factors of obstruction [24].

Our study suggests that the assessment of the difference in reversibility of airway obstruction between asthma and non-asthmatic COLD is affected by the index, its mode of expression and the dose of inhaled bronchodilator. With the discriminant analysis, FEV, appears to be the most effective index of forced expiration to discriminate among the mechanisms of airflow obstruction in asthma and chronic bronchitis. The other commonly used indices of forced expiration may provide information on the site of obstruction, but our results show that they are not as effective as FEV, and add little information when associated with FEV.

The ability of the bronchodilator test to display the difference between the two clinically well-defined populations of group I depends on the mode of expression of the response. The overlap between asthma and chronic bronchitis appears to be greater when the response is expressed as a percentage of the initial value, than when it is expressed as the absolute difference or as a percentage of the predicted value. Ellason et al. [25] also showed that the Δ%IV expression was less effective for separating these clinical entities but they did not investigate the effects of the dose of bronchodilator and of the threshold of improvement. In our study, the discriminant analysis shows that the inhalation of a greater dose of salbutamol (1 mg) enhances the distinction between asthmatic and chronic bronchitic patients.

It could be argued that the difference we found in the degree of reversibility of airway obstruction in asthmatic and chronic bronchitic subjects was related at least in part to the withdrawal of methylxanthines in all of the former but only some of the latter patients. However, the mean improvement after salbutamol was not significantly different in the five bronchitic subjects without and the fifteen with methylxanthines and was 0.22 / in the former and 0.17 / in the latter after 1 mg of salbutamol. These values were small compared to the 0.96 / improvement in the asthmatic patients. Furthermore, if this bias existed, it existed for all the indices and modes of expression we compared.

When the bronchodilator test is used or recommended [26] as a selection criterion of a population, this implicitly means that an improvement of 15% [27, 7], 20% [4, 7, 28] or 25% [5] over the initial FEV, differentiates asthma from non-asthmatic COLD. The same thresholds are often used as a diagnostic criterion in clinical practice. Our results show that a 20% rise in FEV, which is the most widely used criterion, is not specific for asthma (Sp=0.65) and is a poor diagnostic criterion (L=2.8). On the theoretical ROC curve (fig. 1), the threshold with the best diagnostic value is located at the peak of the convexity and has the highest likelihood ratio. For 0.2 mg of salbutamol, in our population, this threshold was 10%pred with Se=0.97 and Sp=0.95, yielding a likelihood ratio of 19.4. As a consequence of the size of our population, a variation of 0.1 in the specificity for the diagnosis of asthma represented a variation in the classification of two chronic bronchitic patients. Therefore, the diagnostic value of high thresholds of the expression in Δ%IV, e.g. 25% (Se=0.97 and Sp=0.85), might be considered close to the diagnostic value of 10%pred. However, such high thresholds are seldom used. Furthermore, the mode of expression in Δ%IV is affected by mathematical bias, i.e. regression towards the mean [29]. It also gives an unjustified advantage to low initial values and implies that the variation of an index after a bronchodilator is related to its initial value which has never been firmly established. Thus, as recommended, this mode of expression should be rejected [30].

The two diagnostic aims of a test of reversibility of airflow obstruction, namely the selection of well-defined

<table>
<thead>
<tr>
<th>Indices*</th>
<th>D-square**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.2 mg</td>
</tr>
<tr>
<td>1</td>
<td>84.83</td>
</tr>
<tr>
<td>1 + 2</td>
<td>91.67</td>
</tr>
<tr>
<td>1 + 2 + 3</td>
<td>94.05</td>
</tr>
<tr>
<td>1 + 2 + 3 + 4</td>
<td>95.40</td>
</tr>
</tbody>
</table>

*: Indices: 1 = FEV; 2 = PEFR; 3 = MMEF; 4 = FEF50.
**: D-square is the generalized Mahalanobis D-square and is calculated from the improvement after 0.2 and 1 mg salbutamol expressed as Δ%pred.
patients and the diagnosis in an individual case, should not be confused and different thresholds of reversibility might be necessary for these two purposes. If the aim is to select asthmatic patients among those presenting a stable airflow obstruction for a trial, e.g. a therapeutic trial, the threshold must yield a high PV+ and therefore a high specificity near to 100% and 1. This will result in a low risk of including in the study non-asthmatic subjects. To avoid the risk of excluding too many true asthmatic patients, these high PV+ and Sp should not be associated with too low PV- and Se. With this objective, the threshold of improvement of 15% of the predicted FEV1 after 0.2 or even better after 1 mg salbutamol is, in our experience, the most effective selection criterion. Our data agree with those of Nicklaus et al. [10] who showed in a population of 50 asthmatic and 50 chronic bronchitic patients that only the former can display an improvement of more than 15%pred in FEV1 after inhalation of isoproterenol. Thus, for a therapeutic trial including patients from different laboratories, the bronchodilator test may be useful for selection of a uniform population and for avoidance of disagreement on the clinical definition of asthma [31, 32].

If the test is used as a diagnostic criterion in clinical practice, simultaneously high PV+ and PV- may avoid both over and underdiagnosis of asthma among patients with chronic airflow obstruction. The results in our group I show that a 10%pred increase in FEV1 after 0.2 mg of salbutamol appears to be the most suitable threshold for this purpose: its positive and negative predictive values for a prevalence or prior probability of asthma of 30% are 89 and 99%, respectively. For a higher prevalence its PV- remains high and therefore the risk of missing the diagnosis of asthma is low (table 4). In the study of Nicklaus et al. [10], the improvement in FEV1 after isoproterenol exceeded 10%pred in only 5 of 50 patients with chronic bronchitis, but was less than this threshold in 10 of 50 patients with asthma. Thus, in their study the sensitivity was 0.64 and the specificity was 0.9. These results are similar to those observed in our group II. In these two populations the threshold of 10%pred remains specific but has a poor sensitivity. As regards the diagnosis of non-asthmatic COLD the result of the test and the clinical findings most often agreed, but as regards the diagnosis of asthma the result of the test and the clinical findings disagreed in more than a third of the cases. These differences in the results between group I and group II may be due in part to a less rigorous withdrawal of the therapy especially with regard to beta-2-agonists, or to the lack of stability of disease in some patients, or to inaccurate use of the terms asthma and non-asthmatic COLD. Furthermore, patients whose clinical symptoms associated acute reversible dyspneic episodes and chronic cough and sputum [33] were excluded from our group I but may have been included in group II with a diagnosis of either asthma or non-asthmatic airway obstruction. In daily routine practice, the reversibility test is therefore a poor diagnostic criterion. It may confirm a clinical diagnosis of asthma, but may not be used to exclude it. A disagreement between a clinical diagnosis of asthma and a small bronchodilator response may suggest that at this moment bronchospasm is not the major component of obstruction and that other treatment is required.

In conclusion: 1) the widely used thresholds of 15 or 20% increase in initial FEV1, have little value for selection of well-defined populations or for the diagnosis of an individual patient; 2) if one wishes to use the test of reversibility as a selection criterion for asthma among chronic airflow obstruction, it would be better to perform the test with a high dose of inhaled bronchodilator, to express the response as a percentage of the predicted FEV1 and to choose a high threshold of variation.

References