Evaluation of bronchodilator response in patients with airway obstruction

S. A. Quadrelli, A. J. Roncoroni and G. C. Montiel

Instituto de Investigaciones Médicas Alfredo Lanari, Facultad de Medicina, Universidad de Buenos Aires, Argentina

The aim of this study was to define the most useful index of expressing bronchodilator response and to distinguish between asthma and COPD.

A prospective study was carried out of bronchodilator response in 142 asthmatics and 58 COPD patients in a university hospital.

Reversibility was expressed as: 1. absolute change (Δabs); 2. % of initial (Δ%init); 3. % of predicted (Δ%pred) and 4. % of maximum possible response (Δ%max). Dependence on forced expirations volume in 1 sec (FEV₁) as % of predicted and sensitivity and specificity for diagnosis of asthma were established.

A relationship between Δabs and initial FEV₁ was not found in asthma (Δabs vs. % initial FEV₁, r=0.07) or COPD (r=0.02). Δ%pred did not show a correlation in asthma (r=0.10) or COPD (r=0.06). Δ%init was dependent on the baseline value in asthma (r=0.38, P<0.001) but not in COPD (r=0.18, P=n.s.). Δmax was dependent in both. The combination of best sensitivity and specificity to separate asthma and COPD was obtained with Δabs (70.4 or 70.6%). The worst specificity for asthma diagnosis was obtained with Δ%init (50%). The best likelihood ratios were obtained with Δabs and Δ%pred and the worst likelihood ratio with Δ%init.

Δ%init is not recommended as an index for differential diagnosis between asthma and COPD; 2) Δ%init overscores bronchodilator response in patients with low FEV₁. The independence of each bronchodilator response index should be verified in clinical trials for each selected sample.

Introduction

The measurement of spirometric data before and after inhaling a bronchodilator is a commonly ordered pulmonary function test in clinical and research settings. Despite its well-known limitations, it is one of the most useful criteria for asthma diagnosis. Additionally, most clinical trials define 'reversible' vs. 'non-reversible' airway obstruction depending on bronchodilator response, which theoretically determines homogeneous study samples. However, there are no uniformly accepted criteria for defining a 'significant' response (1,2). Questions to answer are: 1. which spirometric indices should be considered to evaluate response [forced expiratory volume in 1 sec (FEV₁), forced vital capacity (FVC), forced expiratory flow between 25 and 75% of FVC (FEF25-75) and FEV₁/FVC]?; 2. which is the best way to express change after inhaling a bronchodilator (absolute change, percentage of initial FEV₁, percentage of predicted FEV₁, percentage of maximal possible response)? and 3. what amount of change should be considered a positive response?

Although a percentage of the initial FEV₁ value is not either a part of the definition of bronchodilator response for international societies nor universally accepted, it is a very common means of expression. An increase in FEV₁ >15% was the most popular definition of a positive bronchodilator response found in a review of the asthma and chronic obstructive pulmonary disease (COPD) literature and some medical settings (3). However, this criterion has recently been challenged due to an inadequate ability to separate diagnostically asthma and COPD patients (4) and mainly because of the generation of a much greater proportion of responsive patients (probably false) among patients with very low initial FEV₁ (5–7). On the other hand, expressing change as an absolute value (for instance, 200 ml) may theoretically require values impossible to achieve for patients with very low FEV₁, thus creating a falsely high proportion of non-responsive tests. Choosing one or the other criterion may result in a different classification of a relevant number of patients (5).

These difficulties have provoked increasing interest in exploring other ways to express bronchodilator response (8,9). The characteristics of an ideal index should be maximal independence of the pre bronchodilator FEV₁ value, greatest power to discriminate asthma and COPD
and the reproducibility of bronchodilator response. We have examined the different ways of expressing bronchodilator response in patients with previously known disease (asthma and COPD) in order to define their dependence on initial FEV₁ and their efficacy to separate patients with asthma from patients with COPD.

Material and Methods

Two hundred patients with airway obstruction were studied during a routine visit in our chest clinic (142 asthmatics, 58 COPD). They were included in the study if they had a previously diagnosed airway obstructive disease and a present baseline spirometry with a FEV₁/FVC relationship 1.6-4 < below the predicted value or lower. All COPD patients showed FEV₁ values lower than 70% of predicted value. To test the power of each index to discriminate between COPD and asthma patients, a comparison was performed studying the subgroup of asthmatics with FEV₁ lower than 55% of predicted value (n = 61) in order to match COPD and asthma patients for baseline lung function. Patients with other chronic respiratory diseases (occupational lung disorders, bronchiectasis, interstitial lung disease, tuberculosis and cancer), previous thoracic surgery, heart failure or any other condition able to interfere with an adequate expiratory manoeuvre or a correct classification of asthma or COPD were excluded. A standardized history of respiratory symptoms was obtained. Patients were classified as asthmatics when they reported attacks of breathlessness and wheeze according ATS criteria and were classified as COPD when they showed breathlessness and wheeze according ATS criteria and were classified as COPD when they showed breathlessness and wheeze according ATS criteria.

Results

Of the 200 patients, 142 were classed as asthmatic (mean age 55.4 ± 19.0 years, 68 women) and 58 as COPD (mean age 67.3 ± 70 years, 12 women). Baseline FEV₁ in asthma patients was 1.57 ± 0.76 L (59.4 ± 19.1% of predicted) and in patients with COPD 1.01 ± 0.34 L (39.7 ± 14.7% of predicted). CVF was 2.59 ± 0.87 L (80.6 ± 20.2% of predicted) in asthmatics and 1.23 ± 0.55 L (60.2 ± 13.7% of predicted) in COPD. Of the COPD patients 23% were current smokers and 77% ex-smokers (68.3 ± 36.5 pack years). Sixty-one asthmatic patients showed an FEV₁ lower than 55% of predicted value vs. COPD.

RELATIONSHIP BETWEEN BASELINE FEV₁ AND BRONCHODILATOR RESPONSE

There was no relationship between dabs and pre-Bd FEV₁ either in asthmatics or in COPD. Changes expressed as
**BRONCHODILATOR RESPONSE IN PATIENTS WITH ASTHMA AND COPD**

Bronchodilator response was significantly different between asthmatics and COPD when expressed as $\Delta$abs (asthma $0.31 \pm 0.22$ vs. COPD $0.16 \pm 0.14$; $P<0.001$) or $\Delta$% pred (asthma $12.0 \pm 7.9$ vs. $6.29 \pm 5.1$%; $P<0.001$). Change expressed as $\Delta$%init or $\Delta$%max was not different between both groups of patients (Table 2, Fig. 4). However, if only patients with reduced FEV$_1$ (lower than 55% of predicted) were considered, all indices reached statistical significance in separating asthma and COPD (Table 2, Fig. 4).

Bronchodilator response did not show a sensitivity higher than 85% to establish asthma diagnosis for any index. The lowest specificity was obtained with $\Delta$%init (50.0%) (Table 3). When predictive values for a positive and negative test were calculated on the basis of an arbitrarily chosen clinical pre-test probability of asthma of 30 and 70%, the best predictive values for a positive test were obtained with $\Delta$abs (50.5% and 84.8%) and $\Delta$% pred (49.2 and 84.1%), while the lowest predictive value for a positive test (39.4 and 78.0%) was obtained with $\Delta$%init. Even with a high predictive value for a positive test,
Bronchodilator response as absolute change (l)

Fig. 3. Relationship between pre-Bd FEV₁ as a percentage of predicted value and Δabs in the whole population (asthma + COPD). Dependence on the degree of baseline airflow obstruction was not found (r = 0.03, P = n.s.; y = 2.9212x + 53.899).

the highest likelihood ratios (2.23 and 2.03); much higher than values for A%init (1.28). If only those asthmatics with lower FEV₁ values were considered (to match the patients for baseline lung function), the results were not notably different.

A subgroup of 8(5.6%) patients were classed as bronchodilator-responsive when applying Δ%pred but as non-responsive when applying Δ%init (Δ% pred >9%, Δ%init <15%). When these patients were compared with the opposite situation (non-responsive on applying Δ%pred but responsive on applying Δ% init: n = 12, i.e. 8.4% of patients), they showed a higher pre-Bd FEV₁ (78.6 ± 14.7 vs. 53.1 ± 8.1 % predicted, P = <0.001). No COPD patient was classed as responsive by only Δ%pred >9%, because all such patients showed a corresponding Δ% init of >15%. However, 10 patients (17.2%) who were classified as non-responsive according to Δ%pred <9% showed Δ% init >15%.

Table 2. Response to bronchodilator: comparison between patients with asthma and COPD

<table>
<thead>
<tr>
<th>Patients with FEV₁ lower than 55% of predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>ΔFEV₁ abs (l)</td>
</tr>
<tr>
<td>Δ% FEV₁ init (%)</td>
</tr>
<tr>
<td>Δ% FEV₁ pred (%)</td>
</tr>
<tr>
<td>Δ% max (%)</td>
</tr>
</tbody>
</table>

Fig. 4. Distribution of bronchodilator response in asthma (■) and COPD (●) expressed by different indices (patients with FEV₁ below 55% of predicted). Bronchodilator response as: (a) absolute change; (b) % of possible maximum; (c) % of baseline FEV₁; and (d) % predicted.
TABLE 3. Response to bronchodilator: sensitivity and specificity of different indices in patients with $FEV_1 <35\%$ (asthma: $n=61$)

<table>
<thead>
<tr>
<th>Index</th>
<th>Cut-off level</th>
<th>True responsive</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$dabs$</td>
<td>200 ml</td>
<td>43</td>
<td>70.4%</td>
<td>70.6%</td>
</tr>
<tr>
<td>$d%^{init}$</td>
<td>15%</td>
<td>52</td>
<td>85.2%</td>
<td>50.0%</td>
</tr>
<tr>
<td>$dabs + d%^{init}$</td>
<td></td>
<td>42</td>
<td>68.8%</td>
<td>70.6%</td>
</tr>
<tr>
<td>$d%^{pred}$</td>
<td>9%</td>
<td>41</td>
<td>67.2%</td>
<td>70.6%</td>
</tr>
<tr>
<td>$A%_{max}$</td>
<td>50%</td>
<td>4</td>
<td>6.5%</td>
<td>98.2%</td>
</tr>
</tbody>
</table>

Bold type indicates significant difference.

TABLE 4. Response to bronchodilator: predictive values for a positive and negative test for patients with $FEV_1$ lower than 55\% (asthma: $n=61$)

<table>
<thead>
<tr>
<th>Pre-test probability of asthma 30%</th>
<th>Pre-test probability of asthma 70%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index</td>
<td>PV for a positive test (%)</td>
</tr>
<tr>
<td>$dabs$</td>
<td>50.5</td>
</tr>
<tr>
<td>$A%^{init}$</td>
<td>39.4</td>
</tr>
<tr>
<td>$dabs + A%^{init}$</td>
<td>48.1</td>
</tr>
<tr>
<td>$A%^{pred}$</td>
<td>49.2</td>
</tr>
<tr>
<td>$A%_{max}$</td>
<td>75.5</td>
</tr>
</tbody>
</table>

The clinical usefulness of each index is highly dependent on clinical pre-test probability of asthma. $A%^{init}$ shows the worst results of predictive values for a positive or negative test in most circumstances. In patients with high clinical pre-test, the presence of a positive response expressed as $A%_{max}$ may be a very specific diagnostic index. Bold type indicates significant difference.

Discussion

This study demonstrates that expressing bronchodilator response in one of the most popular ways (increase in $FEV_1$ as a percentage of baseline $FEV_1$) has important disadvantages: it is strongly dependent on the pre-bronchodilator $FEV_1$, does not reveal significant differences between asthma and COPD patients and shows the weakest power to discriminate between these two conditions.

In order to define reversibility to bronchodilator, three factors should be considered: 1. which will be the applied indices? ($FEV_1$, $FVC$, $FEF_{25-75}$?); 2. which will be the chosen way of expressing this response? and 3. which will be the cut-off limits to define a positive response?

Firstly, $FEV_1$ was better than other commonly used tests for evaluating bronchodilating drugs (5,12). Secondly, the cut-off limits of a positive response may be defined from studies using patient samples to determine confidence intervals for spontaneous variability (13) or placebo response (10,14), or may be calculated from the response to bronchodilator in normal subjects (15). Taking into account that the distribution of bronchodilator response in any of those populations is continuous and unimodal, any definition of the 'cut-off' level for a positive response will be arbitrary. The chosen values for this study were derived from confidence intervals for response to placebo obtained at our pulmonary function laboratory (10), which are similar to currently published values (1,16).

Finally, if the two most commonly used indices for expressing response ($dabs$ and $A%^{init}$) are analysed, it seems obvious that a change of 200 ml may be a small variation in a baseline $FEV_1$ of 2.5 l ($<10\%$), while it may be almost impossible to achieve for a very low baseline $FEV_1$, e.g. 0.4 l (50\%). At the same time, a change of 15\% in initial value may be a very low value in the second example (60 ml), with equivocal clinical significance and even the possibility of under-resolution of the spirometry equipment. The number of patients classed differently according to the different criteria is not small (12\% of asthmatics in a study population) (5) and is important in defining which index is more accurate.

An index with a greater independence of baseline $FEV_1$ increases the comparability between different subjects and also between tests with different baseline $FEV_1$ in the same subject (5,17). According to our results in patients with asthma, only $dabs$ and $A%^{pred}$ were independent of pre-Bd $FEV_1$, while in COPD patients only $A%_{max}$ showed a strong dependence on the baseline $FEV_1$ value. These results in COPD patients contrast with data obtained by Dompeling et al. (17), where $A%^{init}$ was very dependent on
The higher homogeneity of our sample (and heterogeneity of results shows the influence of population pre-Bd FEV1. Differences in patient selection may explain characteristics on diagnosis and severity of airway disease. Perhaps the smaller number, n = 58 vs. 111) may be the reason for the absence of the dependence showed by Dompeling et al. (17). Despite this difference, our data and results from other authors (9, 17) strongly suggest that the most popular index (A%init) is highly dependent on the baseline value in asthma (8) or in a non-selected patient population (9) (r for our population as a whole = 0.33, P = <0.001).

This fact is remarkably relevant when defining bronchodilator response in clinical trials or in patients in different clinical situations. If an inclusion criterion is a pre-determined level of bronchodilator response expressed as A%init (which is very common) (18,19) this will select patients with a more severe pre-bronchodilator airway obstruction. On the other hand, if the sample is composed of more severely affected patients, the response to bronchodilator as a percentage of initial value will be magnified due only to the chosen criterion. Similarly, the response to any bronchodilator in a patient who has experienced great change in his or her baseline FEV1 may be falsely interpreted as different. This potential error would be avoided if response were expressed as Dabs. Finally, in light of the strong dependence of A%init, the prognostic significance of bronchodilator responsiveness in COPD patients should be reconsidered. A worse prognosis for patients with higher reversibility (expressed as A%init) has been reported (20,21). However, is this really the influence of reversibility or solely due to the selection of more severe patients who show a greater response when expressed as A%init? (22). Interestingly, when Postma et al. (23) assessed response to bronchodilator using D% (pred – initial), a higher reversibility was a favourable prognostic criterion.

It is remarkable that several different studies which have analysed the dependence of bronchodilator response on pre-bronchodilator values, while in accord over the strong dependence of A%init, show very different results regarding other methods of expressing reversibility (Table 5). Obviously, the composition of the sample is critical in determining such relationships. This is very important in designing clinical trials and the chosen method of expressing bronchodilator response should be assessed in that particular sample to exclude dependence on baseline FEV1.

Even when advantages of Dabs and A%pred in separating positive and negative responses are clear, the prognostic significance of these indices in evaluating the long-term outcome has not been assessed. Reversibility with a bronchodilator is also a criterion for differential diagnosis between asthma and COPD. Even when inflammation and not bronchodilator response distinguishes between the two disorders, the degree of bronchial responsiveness is a commonly used criterion to separate asthma and COPD in clinical practice and inclusion in clinical trials. In this study, none of the indices resulted in a good differentiation between the two disorders. This is in agreement with previous studies, indicating that up to 23% of asthmatics showed a non-responsive spirometry (5). The most powerful indices found to separate asthma and COPD in this population were Dabs and A%pred, which showed the highest likelihood ratios (Dabs: 2.23 and A%pred: 2.03) and thus the best combination of sensitivity, and specificity, compared to A%init (likelihood ratio 1.28). When analysing the discriminative value of each index solely by considering differences in group means, all indices reached a statistically significant difference if only patients with reduced FEV1 were considered. However, in a more complete analysis, the sensitivity, the specificity and likelihood ratios were not substantially modified by excluding patients with near-normal baseline FEV1. Published studies disagree considerably on this point. While Nicklaus (6), assessing only D%init, showed a strong discriminative power for bronchodilator response, Brand et al. (8) found this method to be a poor diagnostic criterion. This disagreement may again be explained by different population compositions and methods of analysis. Discriminative power may be evaluated using group mean differences (17) or by including sensitivity and specificity calculations (8,24) and, even when the same method is used, different inclusion criteria and doses of bronchodilator may influence the final results. It is remarkable that, in the present study, even the best indices were not sensitive for asthma diagnosis. It is important to remember that clinical background is critically significant in defining predictive values and thus to delimit the increase in diagnostic power using each index. With a low pre-test probability of asthma (30%), even the best indices in this study showed low predictive values for a positive test, however these reached clinical relevance with higher pre-
test probabilities. This emphasizes that, in a clinical setting, the occurrence of a negative test does not exclude the presence of asthma, even when a better diagnosis may be obtained by A%obs or A%pred.

We conclude that: 1. in clinical practice A%init is a poor diagnostic tool to differentiate asthma from COPD; 2. in patients with a previously known diagnosis, A%init overestimates the response to bronchodilator in subjects with very low baseline FEV1; and 3. in clinical trials studying reversibility, the chosen index should be assessed to exclude dependence on the initial FEV1. The expression which correlates best with clinical improvement after long-term treatment has not yet been defined, therefore the inclusion of different indices for expressing bronchodilator response in clinical trials studying the effects of long-term inhalational therapy would be extremely useful.

Acknowledgements

We thank Luis Ortino and Luciana Gambina for their technical assistance and Paola Gambina for valuable help in preparing the manuscript.

References