The new ATS/ERS guidelines for assessing the spirometric severity of restrictive lung disease differ from previous standards

ASHUTOSH N. AGGARWAL AND RITESH AGARWAL

Department of Pulmonary Medicine, Postgraduate Institute of Medical Education and Research, Chandigarh, India

The new ATS/ERS guidelines for assessing the spirometric severity of restrictive lung disease differ from previous standards

AGGARWAL AN, AGARWAL R. Respirology 2007; 12: 759–762

Background and objectives: The ATS/ERS Task Force on Lung Function Testing recently proposed guidelines for the interpretation of pulmonary function tests and suggested that a reduction in FEV₁ be used for categorizing both obstructive and restrictive abnormalities. This changes the severity stratification algorithm of restrictive patterns diagnosed by spirometry, that are currently categorized based on reduction in VC. This study examined the level of agreement between these two categorization schemes.

Methods: Spirometry records of 2527 adult patients evaluated over 1 year were retrieved; 361 of these patients showed a restrictive pattern. Severity of airway restriction was separately assessed in these patients using the indicative schemes provided in the new ATS/ERS and the earlier ATS guidelines.

Results: There were 212 (58.7%) patients with a restrictive pattern who had identical severity categorization using both guidelines. In most instances of discordance, the severity categorization differed only by a single stratum. Of 149 discordant results, 91 (60.1%) were placed in a better category, and 58 (39.9%) in a worse category, when using the new ATS/ERS recommendations. Overall weighted kappa estimate for agreement between the two schemes of categorization was 0.649.

Conclusion: Based on spirometry results, the level of severity of restriction cannot be described interchangeably between the old and new guidelines for all patients. The new guidelines tend to give lower severity scores for restrictive lung diseases in up to 25% of patients.

Key words: agreement, FEV₁, respiratory function test, restrictive defect, VC.

INTRODUCTION

The ATS/ERS Task Force on Lung Function Testing has recently proposed new guidelines for the interpretation of pulmonary function tests.¹ The algorithm for classifying lung function test results into normal, obstructive, restrictive and mixed patterns depends on four parameters: VC, FEV₁, FEV₁/VC ratio and TLC. At a basic level, these parameters are largely similar to those recommended by the ATS in 1991.² However, an important difference has been suggested for categorizing the severity of abnormal pulmonary function tests. While the earlier ATS guidelines recommended categorization of the severity of obstructive and restrictive defects based upon the degree of reduction in FEV₁ and VC, respectively, the newer ATS/ERS guidelines suggest the use of the degree of reduction in FEV₁, for categorizing all abnormalities.¹,² The routine estimation of static lung volumes in addition to spirometry is not widely practised and in many countries such facilities are available only in a handful of centres. A restrictive pattern of lung disease is therefore inferred and categorized based on observed VC values alone. A shift in the algorithm for severity categorization, from one based on VC to one based on FEV₁, could impact significantly on routine...
clinical practice. Clinicians relying primarily on spirometry results need to be aware of the meaning of a particular level of severity of restriction, and its correlation between the old and new systems of categorization. This study investigated the degree of agreement between these two categorization schemes using data from routine spirometric evaluations.

SUBJECTS AND METHODS

Records of all adult patients referred for spirometry at our institution over a 6-month period after publication of the new ATS/ERS guidelines on interpretative strategies (January to June 2006) were prospectively collected. Reasons for performing spirometry, and other clinical details, were not analysed further. The study was approved by the Hospital Ethics Committee.

All patients performed spirometry on a dry rolling seal spirometer (Spiro RS232; PK Morgan Ltd; Kent, UK), assisted by technicians experienced in pulmonary function testing. Spirometer calibration was frequently checked to ensure consistency. For each patient, the highest measurements of FVC and FEV1 from at least three technically acceptable and reproducible manoeuvres were expressed at body temperature and pressure saturated with water vapour. Age, gender, height and spirometry data were recorded for all patients using a computer software developed earlier. The predicted values for FVC, FEV1, and FEV1/FVC ratio were generated using previously defined prediction equations for north Indian subjects, and observed values expressed as a percentage of their respective predicted values. Lower limits of normal (LLN) for FEV1, FVC and FEV1/FVC were calculated using the lower 95% confidence limits derived from the regression equation being used, and were computed as the difference between the predicted value and 1.645 times the standard error of estimate of the regression equation. Any observed value lower than its corresponding LLN was considered abnormal.

FEV1, FVC and FEV1/FVC were used as the basic parameters to interpret the spirometry data. A spirometry record with a FEV1/FVC less than the LLN for that subject was categorized as having an obstructive pattern. Although a true restrictive defect can be diagnosed only by the demonstration of a reduced TLC, a restrictive pattern was inferred from spirometry results for categorization comparison purposes only. A patient recording an FVC less than its LLN, and an FEV1/FVC ratio above its LLN, was categorized as having a restrictive pattern. Static lung volumes and TLC were not routinely estimated in subjects having restrictive patterns on spirometry.

Spirometry records showing a restrictive pattern were analysed further. The severity of restriction was separately assessed for each patient using the indicative schemes provided in the ATS (1991) and ATS/ERS (2005) guidelines (Table 1). Contingency tables were constructed for the assessment of concordance between the severity categories using the two guide-lines. Overall agreement between the two systems of severity categorization was described using the weighted kappa estimate.

RESULTS

During the study period, 2626 subjects underwent spirometry. After exclusion of the data from 99 subjects aged less than 15 years, records of 2527 subjects (1446 men and 1081 women) were available for analysis. Of these, 1106 (43.8%) subjects had normal spirometry, while 1060 (41.9%) and 361 (14.3%), respectively, had obstructive and restrictive patterns. Of the 361 subjects having a restrictive pattern, 212 (58.7%) had identical severity categorization using the old ATS and the new ATS/ERS interpretative strategies. In most instances of discordance, the severity categorization differed only by a single stratum (e.g. ‘moderately severe’ instead of ‘moderate’) (Fig. 1). Of the 149 discordant results, 91 (60.1%) were placed in a better category, and 58 (39.9%) in a worse category, using the new ATS/ERS recommendation. The overall weighted kappa estimate for agreement between two schemes of categorization was 0.649 (SE 0.017).

DISCUSSION

Although a restrictive defect can only be diagnosed through demonstration of reduced TLC, estimation of static lung volumes is not widely practised for this purpose. A pattern of reduced VC, along with a normal FEV1/VC ratio, is suggestive, though not diagnostic of restriction. Although such a ‘restrictive pattern’ on spirometry will identify a true restrictive defect (based on a reduction in TLC) in about half of such patients, clinicians continue to use spirometry to describe restrictive patterns in settings where measurements of static lung volumes are not available. Based on spirometry results, the severity of restriction has traditionally been assessed by noting the degree of reduction in the observed FVC. Although mathematical models have also been described that help to better predict a reduction in TLC from spirometric data, the accuracy of such approaches in individual patients is not clearly known. Recently published

Table 1 Algorithms for categorization of restrictive patterns on spirometry

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>FEV1% ≥ 70</td>
<td>VC% ≥ 70</td>
</tr>
<tr>
<td>Moderate</td>
<td>FEV1% 60–69</td>
<td>VC% &lt; 70 and ≥ 60</td>
</tr>
<tr>
<td>Moderately severe</td>
<td>FEV1% 50–59</td>
<td>VC% &lt; 60 and ≥ 50</td>
</tr>
<tr>
<td>Severe</td>
<td>FEV1% 35–49</td>
<td>VC% &lt; 50 and ≥ 35</td>
</tr>
<tr>
<td>Very severe</td>
<td>FEV1% &lt; 35</td>
<td>VC% &lt; 35</td>
</tr>
</tbody>
</table>

FEV1% and VC% are the percentages of observed FEV1 and FVC related to their respective predicted values.
Restrictive spirometry defect severity

Figure 1  Severity categorization of 361 restrictive patterns seen on spirometry using the old ATS (1991) and the new ATS/ERS (2005) proposals on spirometry interpretation. Area of each bubble is proportional to the number of patients in that category. See Pellegrino et al. and American Thoracic Society for details.

ATS/ERS guidelines suggest that the severity of any pulmonary function abnormality (both obstructive and restrictive) be assessed using FEV₁. The rationale for such a proposal probably derives from the fact that the FEV₁/VC ratio is preserved (normal or increased) in disorders presenting with restrictive defects, and hence there may be a good correlation between reduction in VC and the corresponding reduction in FEV₁. The new method is also simpler to use since severity of all defects is categorized using a single measurement (FEV₁). Clinicians accustomed to assessing the severity of airway restriction based on the degree of reduction in VC must reorient themselves to doing so using the degree of reduction in FEV₁. It is important to be aware as to how well the previous and the newly proposed algorithms of severity assessment correlate with each other, though such a correlation will be necessarily influenced by the arbitrary numeric cutoff values and the arbitrary number of strata used to describe various categories of severity. This study reports such an assessment using a database of patients undergoing routine spirometry and using largely comparable severity categorization schemes proposed as examples in the guidelines on spirometry interpretation. The overall weighted Kappa estimate, which accounts for the relative hierarchy in each categorization scheme, also suggests a reasonably good agreement between the two guidelines. However, there was more than 40% discordance between the severity strata defined by the two methods of categorization. Among the discordant results, there was a tendency towards underestimation of the severity of airway restriction as compared with the old ATS guidelines. This probably is explained by the fact that in several patients with restrictive defects, the FEV₁/VC ratio is increased, so that the reduction in FVC is relatively greater than the corresponding reduction in FEV₁.

This study has certain limitations. The results are strictly applicable to the group of patients studied and for the categorization systems used in the study. The severity of restriction was based only on spirometric parameters, and therefore the clinical implication of any discordance between severity categorization based on the two algorithms cannot be ascertained. Lastly, a reduction in FVC alone was used to define a restrictive abnormality and static lung volumes were not estimated for all such patients. FVC is likely to represent a true reflection of VC in patients having no airflow limitation. Although it is well recognized that while a normal FVC value can exclude a true restrictive defect (i.e. decreased TLC) with a reasonably high accuracy, a low FVC is associated with a low TLC value in only about 50–60% instances. It is therefore recommended that TLC should be measured to confirm a true restrictive defect if spirometry reveals a low FVC and normal FEV₁/FVC pattern. However, estimation of TLC can be a relatively expensive investigation requiring a higher degree of technical expertise, and as such is not widely available, especially in developing countries. In an office or primary care setting, clinicians still rely on spirometry alone to interpret pulmonary function, and most computerized spirometry systems also interpret results in a similar fashion. This analysis largely addresses the issue of categorising the severity of restrictive patterns seen on spirometry.

These results suggest that the correlation between the severity descriptions from the two guidelines is not perfect. The newer guidelines might underestimate the severity of restriction in some patients. Whether the discordance between severity stratification ultimately has any relevance to clinical decision-making would eventually depend on individual patient profile. Nonetheless, it is important to realize that the level of severity cannot be described interchangeably between the two proposals for several patients.

REFERENCES


