Determination of mortality risk in obstructive pulmonary syndromes: physical examination compared with measurement of forced expired volume in one second

NE PARDEE, RH WINTERBAUER, EH MORGAN, JD ALLEN

From the Section of Pulmonary Diseases, Department of Internal Medicine, The Mason Clinic, Seattle, Washington, USA

ABSTRACT A standardised physical examination based on four clinical signs was compared with the FEV\textsubscript{1} with regard to ability to determine five-year mortality risk in subjects with varying degrees of obstructive airway disease. Both evaluation methods identified low, intermediate, and high risk groups within the population studied. Individuals with no positive physical signs, or with an FEV\textsubscript{1} of 70\% of predicted or more had mortality not significantly greater than that predicted on the basis of standard mortality tables. There was no difference between the FEV\textsubscript{1} and the physical examination in ability to predict mortality.

In an earlier paper we explored the relationships of four physical signs to the forced expired volume in one second (FEV\textsubscript{1}) and evaluated the 11 possible combinations of two or more of the signs for their ability to estimate severity of airways obstruction.\textsuperscript{1} The four signs were intensity of breath sounds heard over peripheral lungs during rapid, deep inspiration, palpable recruitment of scalene muscles in inspiration during quiet breathing at rest, expiratory wheeze, and occurrence of a wet, crackling sound in the expiratory air-flow phase of voluntary cough. Presence of each sign proved to have a statistically significant association with reduction in the FEV\textsubscript{1}. This was true for the study population as a whole and also for each of the participating examiners individually. Scalene recruitment differed from the other three signs in that with this observation there were frequent false positives; when this sign was present alone, FEV\textsubscript{1} was always normal. At the same time, it was clear that neither scalene tensing nor any of the other signs was redundant. Analysis of the accuracy of FEV\textsubscript{1} estimation when combinations of signs were evaluated indicated that discrimination of normal subjects from those with airway obstruction and judgments of the severity of obstructive disease were better when all four variables were taken into account than with any combination of two or three of the observations. Clinical estimates were improved further by use of two procedures for weighting the data. Performance of the different examiners varied somewhat; the pattern of differences indicated observer error as the probable cause of this variation. Differences between examiners were not large, however, and agreement between the FEV\textsubscript{1} and clinical judgments of the degree of airways obstruction was close enough to suggest that physical examination might share with the FEV\textsubscript{1} the ability to determine mortality risk.\textsuperscript{2} The present report describes follow-up of the original population, begun with the purpose of learning whether or not clinical staging of severity of airway obstruction would, in fact, have a significant relationship to mortality.

Methods

There were 183 subjects in the original study population, 109 men and 74 women. Ages ranged from 19 to 84 years, with a mean of 54. Subjects were not pre-selected; the only criterion for including a patient in the study was his having been referred to the pulmonary function laboratory for spirometry.

Address for reprint requests: Dr NE Pardee, Section of Pulmonary Diseases, Department of Internal Medicine, The Mason Clinic, 1100 Ninth Avenue, Seattle, Washington 98101, USA.
The FEV₁ was the spirometric standard for the study. Predicted values used were those of Kory and associates. Seventy-three patients had FEV₁ values of 80\% of predicted or more. With 68 patients, the FEV₁ was between 45 and 79\% of predicted. Forty-two subjects had FEV₁ determinations less than 45\% of predicted.

Each patient was examined by one of four examiners. Examinations were “blind”—that is, the results of FEV₁ measurements were not known to the examiners and data from an examination was discarded if the examiner had any previous knowledge of that patient. Subjects were seated for the examinations. Scalene muscle recruitment in inspiration was judged with the patients breathing quietly at rest, not talking or coughing. Scalene tensing was considered present when bilateral firming of scalene muscles was palpable with each breath. Cough quality was classified as either dry (when the sound heard with voluntary cough consisted only of a brief explosive pop) or moist and crackling (when there was a rattling or crackling sound in the period of expiratory air flow just after opening of the glottis). Cough quality was judged as heard by the examiner standing near the subject, not during auscultation. The examiners used diaphragm stethoscopes for auscultation. Expiratory wheeze was recorded present when it was conspicuous with rapid deep breaths or with any less vigorous ventilatory effort. For evaluation of breath sound intensity, auscultation was performed at six sites: bilaterally over the upper anterior chest, in the axillae, and at the posterior bases. Examiners coached the subjects to take the fastest and deepest breaths possible, with mouth open, so that sounds produced were those related to maximum inspiratory air-flow. Inspiratory breath sound loudness was rated from zero to four, as follows: zero, silent; one, barely audible; two, easily heard but definitely less loud than normal; three, normal; and four, louder than the usual normal. Level three was the major reference point for the scale. Examiners based their judgments of expected normal loudness on sounds heard over peripheral lung fields during rapid deep inspiration in previous examinations of individuals known to have normal ventilatory function. Ratings from the six areas were added to give a total breath sound score (range of possible values, 0-24).

The method of tabulating the physical examination data which gave the best correlation between clinical judgments of air flow obstruction and the FEV₁ in the original study was used for the current investigation. A point was given for each positive sign. The sum of the points of abnormality awarded for an examination produced an index of clinical abnormality for that subject. Two weighting procedures were used: one took advantage of the semi-quantitative nature of breath sound intensity scoring; the second eliminated scalene tensing from consideration if it occurred as a single positive sign.

Table 1 summarises the staging system. There were seven index categories, 0 to 6. Category 0 represented a normal examination, with normal breath sound intensity and no wheeze, no scalene tensing, and no abnormality of cough quality. Category 1 included subjects having borderline breath sound intensity (breath sound scores 13-15), or expiratory wheeze, or crackling cough. Category 2 was composed of individuals with any two points of abnormality occurring together: moderate reduction in breath sound intensity (score 7-12), or borderline breath sounds and wheeze, or wheeze along with scalene tensing, and so forth. Category 6 indicated maximal abnormality, with marked reduction in breath sound loudness (three points) and with crackling cough, expiratory wheeze, and scalene recruitment (one point each) all present.

Mortality data came from review of hospital records, from State Health Department vital statistics records, and from a survey of referring physicians. Predictions of mortality for subgroups of the study population, adjusted for age and sex, were based on the 1965-70 Ultimate Basic Tables, Graduated Mortality Rates per 1000, obtained from the Society of Actuaries. Five-year mortality rates for the different groups within the study population were determined by the life-table method. Judgment of statistical significance depended on modified t tests, based on calculation of standard errors of differences between percentages.

Results

It was possible to obtain at least one year follow-up for 166 subjects. Data were complete for a full five years for 108 individuals, and for six years for 75. There were 40 deaths, 37 of which occurred in the first five years. Table 2 presents the deaths by cause,
Table 2 Causes of death in the study population

<table>
<thead>
<tr>
<th>Cause</th>
<th>Number</th>
<th>Percentage of total deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respiratory failure and/or cor pulmonale</td>
<td>13</td>
<td>32.5</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>12</td>
<td>30.0</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>5</td>
<td>12.5</td>
</tr>
<tr>
<td>Non-respiratory cancer</td>
<td>5</td>
<td>12.5</td>
</tr>
<tr>
<td>Renal failure</td>
<td>2</td>
<td>5.0</td>
</tr>
<tr>
<td>Trauma</td>
<td>2</td>
<td>5.0</td>
</tr>
<tr>
<td>Unknown</td>
<td>1</td>
<td>2.5</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 3 Five-year mortality data, subjects categorised by FEV₁

<table>
<thead>
<tr>
<th>Percent of predicted FEV₁</th>
<th>Number of subjects</th>
<th>Median age (yr)</th>
<th>Expected mortality (%)</th>
<th>Observed mortality (%)</th>
<th>Significance level of excess mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>70 or more</td>
<td>92</td>
<td>49</td>
<td>6</td>
<td>9</td>
<td>&gt; 0.30</td>
</tr>
<tr>
<td>50-69</td>
<td>46</td>
<td>58</td>
<td>9</td>
<td>26</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>Less than 50</td>
<td>45</td>
<td>60</td>
<td>11</td>
<td>49</td>
<td>&lt; 0.01</td>
</tr>
</tbody>
</table>

Table 4 Five-year mortality data, subjects categorised by index of clinical abnormality

<table>
<thead>
<tr>
<th>Index category</th>
<th>Number of subjects</th>
<th>Median age (yr)</th>
<th>Expected mortality (%)</th>
<th>Observed mortality (%)</th>
<th>Significance level of excess mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>91</td>
<td>50</td>
<td>6</td>
<td>9</td>
<td>&gt; 0.30</td>
</tr>
<tr>
<td>1-3</td>
<td>70</td>
<td>53</td>
<td>10</td>
<td>32</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>4-6</td>
<td>22</td>
<td>59</td>
<td>11</td>
<td>69</td>
<td>&lt; 0.01</td>
</tr>
</tbody>
</table>

as recorded on the death certificates. Respiratory failure, with or without cor pulmonale, was the most common cause of death, followed by cardiovascular disease and cancer, in that order.

When the population was sorted by FEV₁, five-year mortality was low at all levels of FEV₁ above 69% of predicted. A trend for increased mortality appeared below 70% of the predicted FEV₁. There was uniform high mortality when the FEV₁ was less than 50% of predicted. The population thus arranged itself into three categories, one with low risk, one with intermediate risk, and one with high risk (see table 3). None of the deaths in patients whose initial FEV₁ values were 50% of predicted or more were caused by respiratory failure or cor pulmonale. In the low risk group, mortality was slightly above that predicted, but the difference was not statistically significant with the number of subjects involved. Mortality in patients whose FEV₁ measurements were between 50 and 69% of predicted, on the other hand, was significantly increased, even though none of the death was caused by respiratory failure. When the FEV₁ was less than 50% of predicted, two-thirds of the deaths were caused by respiratory failure or cor pulmonale. Grouping the patients by physical examination gave a mortality distribution very similar to that found when the population was subdivided on the basis of the FEV₁, as shown in table 4. Five-year mortality in subjects with no positive signs (category 0) was the same as when the FEV₁ was 70% of predicted or more. There were no respiratory deaths in categories 0, 1, or 2. Overall mortality was significantly greater than expected for both categories 1 and 2, however (p = < 0.05). Mortality increased further in category 3, and here five of the seven deaths (71%) were the result of respiratory failure. Categories 4 to 6 had the greatest mortality and formed a homogeneous high-risk group. Five-year mortality for these patients was more than six times that predicted on the basis of age and sex (p = < 0.01). Eight of the 11 deaths (73%) were caused by respiratory failure. There was no significant difference between the mortality figures obtained when the study population was categorised by FEV₁ and those which resulted from grouping the subjects by degree of clinical abnormality.

Discussion

The plan of our original investigation aimed to minimise bias in selection of the subjects and to ensure that examiners' observations would be distorted as little as possible by previous knowledge of the patients' diagnoses and symptoms. The "blind" nature of the study makes it necessary to consider several possibilities when interpreting the data. It is likely that there were individuals with asthma, bronchitis, or mild emphysema among the 73 subjects with normal FEV₁ values, considering the nature of the population being studied. There are at least two selection factors which act to produce differences between the study group and a random sample of the population at large. These selection factors operate both with the individuals whose FEV₁ measurements are normal and with the patients with abnormal spirometry. In the first place, each patient has some symptoms or concern which has brought him to the medical centre for evaluation. Secondly, each subject, after his or her initial medical evaluation, has been referred to the pulmonary function laboratory for testing. Thus, even the patients with normal spirometry would be expected to have more respiratory symptoms than would be the case in the sample of individuals of the same ages drawn from the general community. The slight excess mortality in the groups with FEV₁ values of 70% of predicted or more, or with no positive
physical signs, and the greater excess mortality in subjects with moderate abnormality on physical examination or whose FEV₁ determinations are between 50 and 69% of predicted both support this notion. The low risk observed in the patients without abnormalities on physical examination or with FEV₁ measurements of 70% of predicted or more occurs in spite of the selection bias. Normal spirometry and negative clinical signs—at least when a group of several signs is taken into account—gain credibility as indicators of minimal risk. The relative risk is clearly increased when there are positive physical signs or when spirometry is abnormal, but the magnitude of increase in mortality is probably overestimated by our data, since all the subjects perceived themselves as being in need of medical examination, and were perceived by their physicians as having, or possibly having, a respiratory abnormality.

The choice of the four signs included in this investigation was arbitrary. Day-to-day clinical experience had suggested that these observations were of value, but had not made it clear how best to use them in patient evaluation. We did not entirely anticipate either how useful the individual signs would be or how much the use of combinations of the signs would add to the information obtained from the observations taken singly. The results of this study do not imply that these four variables are the only bits of evidence which should be included in clinical evaluation of patients suspected of having airway obstruction. Other observations or other combinations of observations may be as good as, or better than, those reported on here. Barrel chest, reduced chest expansion, impaired cardiac dullness, and absent cardiac apical impulse have all been shown to be significantly associated with presence of obstructive airway disease. Nath and Capel have presented convincing evidence that cracks heard early in inspiration favour a diagnosis of severe airways obstruction. Our own earlier finding that accuracy of FEV₁ estimation increases progressively as more signs are included in an evaluation format argues that clinical judgments of the severity of obstructive disease and of mortality risk can be better than those described in the present study, if additional observations are taken into account.

Two features of clinical evidence, one illustrated in the initial analysis of our data and one at this point speculative, may help explain the observed advantage of using several physical signs in combination to evaluate severity of airway obstruction. Each of the four signs has relatively high specificity and low sensitivity. Each one alone fails to identify individuals with airway obstruction more often than it incorrectly indicates that obstructive disease is present. Use of several signs together would be expected to reduce the effects of false-negative errors. Further, if observer error occurs, the chance that an observer will be wrong in making every one of a series of observations should in theory be less than the probability that he will be wrong only part of the time. If all the signs in a set bear on a single clinical question—for example, whether or not airway obstruction is present—use of several observations together should lead to more accurate diagnoses than consideration of the individual signs taken into account one by one.

References

4. Obtained April 1977 through the Society of Actuaries, 208 S LaSalle St, Chicago, IL 60604 (from the chairman, committee on mortality under ordinary insurances and annuities).
Determination of mortality risk in obstructive pulmonary syndromes: physical examination compared with measurement of forced expired volume in one second.

N E Pardee, R H Winterbauer, E H Morgan and J D Allen

Thorax 1981 36: 928-931
doi: 10.1136/thx.36.12.928

Updated information and services can be found at: http://thorax.bmj.com/content/36/12/928

Email alerting service

These include:

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/