

Quantitative assessment of the value of spirometry

NJ RUSSELL, NJ CRICHTON, PA EMERSON, AD MORGAN

From the Department of Thoracic Medicine, Westminster Hospital, London

ABSTRACT To determine the value of simple spirometric measurements in the diagnostic assessment of breathless patients, doctors requesting such tests were asked to predict the likely ventilatory abnormality, expressing these pretest predictions as probabilities. Comparison of these pretest predictions with the test results allowed an analysis of the doctors' ability to identify lung function abnormalities and an assessment of the diagnostic usefulness of the test. Predictions and spirometric measurements were made in 123 patients. Doctors expressed preference for a particular spirometric category in 112 cases, of which 13 were predicted to have a restrictive defect, 77 were predicted to have an obstructive defect, and 22 were predicted to be normal. Spirometry showed that nine patients had a restrictive defect, 79 had an obstructive defect, and 24 had normal indices. The study showed that 61% of the tests gave a result that doctors predicted as being unlikely. The study also showed that doctors had difficulty in identifying the reversibility of airflow obstruction in patients in whom they correctly predicted obstruction. Spirometry fulfils a useful role in the diagnosis of breathless patients.

Most clinicians consider that measurement of ventilatory capacity by spirometry is useful in the diagnosis of the breathless patient. In many clinical conditions, particularly where more sophisticated measurements of lung function are not available, it may be the only accessible test. The diagnostic "usefulness" of spirometry, however, has never been measured in a quantitative way. To make such a measurement it is necessary to know the requesting doctors' pretest estimations of the likelihood of particular abnormalities and to compare them with the information provided by the test.¹ Thus an unexpected result is "useful" while an expected result has not contributed to the doctor's knowledge about his patient and might not be considered to have provided useful diagnostic information. The purpose of this study was to measure the usefulness, in this sense, of simple spirometry in the diagnostic assessment of breathless patients.

Methods

PATIENTS

The patients studied were inpatients and outpatients sent to the respiratory function laboratory by hospital

Address for reprint requests: Dr PA Emerson, Westminster Hospital, London SW1P 2AP.

Accepted 15 October 1985

doctors of all grades and from all specialties, for whom simple spirometric measurements were requested. The only criterion for selection was that such patients should be complaining of breathlessness.

REQUEST FORM

A new request form was designed on which, in addition to the standard clinical information, doctors were required to predict the likely ventilatory defect, expressing their predictions as probabilities. Three categories of ventilatory impairment can be identified from the use of simple spirometry: restrictive disease, irreversible obstructive disease, and reversible obstructive disease. To allow doctors to predict normal spirometric results for breathless patients, the number of options was enlarged. Thus, breathless

Table 1 *List of the diagnoses offered to doctors on the request form, with an example of a requesting doctor's allocation of probabilities.*

Diagnosis	Probability
1 Restrictive defect	0.00
2 Irreversible obstructive ventilatory defect	0.40
3 Reversible obstructive ventilatory defect present at the time of test	0.30
4 Normal result: reversible obstructive ventilatory defect not present at the time of test	0.15
5 Normal result: psychogenic breathlessness	0.15
6 Normal result: breathlessness due to other causes (eg cardiac, anaemia, etc)	0.00

Table 2 Comparison of requesting doctors' highest pretest probability of the ventilatory defect with the result of spirometry

Test result	Doctors' prediction			Total
	Restrictive	Obstructive	Normal	
Restrictive	5	2	2	9
Obstructive	7	64	8	79
Normal	1	11	12	24
Total	13	77	22	112

patients whom the doctor thought had normal ventilatory capacity could be assigned to outcomes 4, 5, or 6 (table 1), allowing him some flexibility in predicting the cause of the dyspnoea.

SPIROMETRY

Ventilatory capacity was measured with a Vitalograph bellows spirometer and the best of three recordings was used in the analysis.² Reversibility of airflow obstruction was assessed by repeating the measurement 20 minutes after administration of 200 µg of salbutamol via a metered dose aerosol. Predicted normal values are those incorporated into the Vitalograph Spirotrac II System.³⁻⁵

ANALYSIS

The results were returned with a written interpretation of the findings and a post-test probability of the ventilatory defect. The definitions of the categories of ventilatory defect listed above were as follows:

- 1 Restrictive defect if $\frac{FEV_1 \times 100}{FVC} \geq 75\%$ and $FVC \leq 75\%$ of predicted normal.
- 2 Irreversible obstructive defect if $\frac{FEV_1 \times 100}{FVC} < 75\%$ and $< 15\%$ improvement in FEV_1 or FVC after salbutamol.
- 3 Reversible obstructive defect in $\frac{FEV_1 \times 100}{FVC} < 75\%$ and $\geq 15\%$ improvement in FEV_1 or FVC after salbutamol.
- 4 Normal spirometric results if no obstructive or restrictive defect was found.

As these defect categories were defined to be exclusive and exhaustive, the post-test probabilities were in fact certainties and so were always returned as probability 1.0 for the indicated defect category. Pretest probabilities of normal spirometric values in breathless patients, assigned to categories 4, 5, and 6, were considered together in the post-test assessment, since the results of spirometry cannot distinguish between them. From this information, the ability of doctors to predict ventilatory abnormalities in breathless patients undergoing simple spirometry was evaluated.

“USEFULNESS” OF TEST

From a diagnostic point of view, if a test identified a ventilatory defect which was not expected by the requesting doctor, then we considered it to have provided “useful” information. A test could also be considered useful if it confirmed the doctor's most likely prediction, but he expressed uncertainty after this prediction by allocating to it a low probability—that is, less than 0.5. In this analysis we considered a test to have provided useful diagnostic information if the difference between pretest and post-test probabilities was 0.5 or more.

CALIBRATION

The calibration of doctors' predictions for a particular ventilatory category was assessed. In each group of patients falling into consecutive levels of pretest probabilities (<0.2; 0.2–0.39; 0.4–0.59; 0.6–0.79; ≥ 0.8) the patients shown by the test to be in that category were enumerated.

Results

Overall, 134 consecutive cases were entered into the study. Eleven cases were excluded because of an incomplete or inadequate request form or test procedure. Among the remaining 123 cases, there were 64 men and 59 women with an age range of 18–86 (mean 59.2 (SE 1.3)) years.

There were 55 requests for spirometry from 19 junior doctors (housemen and senior house officers), and 68 requests from 16 senior doctors (registrars, senior registrars, and consultants).

The doctors' pretest predictions for the most likely

Table 3 Comparison of doctors' prediction of the reversibility of airflow obstruction with the result of spirometry for the 64 patients in whom airflow obstruction was predicted and demonstrated

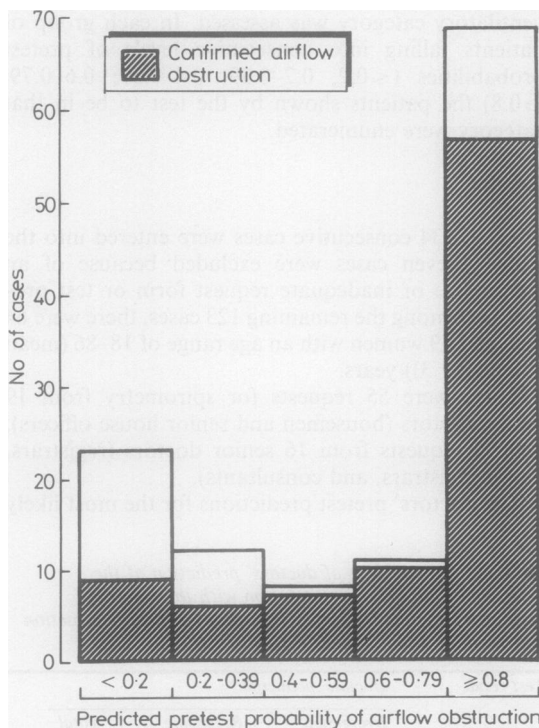
Test result	Doctors' prediction		
	Irreversible	Reversible	Total
Irreversible	22	18	40
Reversible	10	14	24
Total	32	32	64

Table 4 Distribution of the changes between pretest and post-test probabilities for each category of ventilatory impairment

Change in probability*	Defect shown by test				Total
	Restrictive	Irreversible obstructive	Reversible obstructive	Normal	
0.0	2	8	1	5	16
0.01-0.25	1	4	9	5	19
0.26-0.49	2	6	4	1	13
0.5	0	7	0	2	9
0.51-0.75	0	9	4	0	13
0.76-0.99	0	6	10	5	21
1.0	4	16	5	7	32
Total	9	56	33	25	123

*A test was considered "useful" when a change of 0.5 or more occurred.

ventilatory defect categories (that is, the category with the highest pretest probability) were compared with those identified by the test results. Eleven cases were excluded from this part of the analysis because their pretest probabilities were evenly distributed and the doctors' most likely prediction could not be determined. In the remaining 112 cases doctors correctly identified the functional category as either restrictive disease, obstructive disease (irreversible and reversible), or normality in 81 (72%) (table 2).



Distribution of all 123 patients included in the study according to doctors' predicted pretest probability of airflow obstruction.

When doctors predicted an obstructive defect and this was confirmed by the test, reversibility status was correctly identified in only 36 of 64 cases (56%) (table 3). If the categories of ventilatory defect are expanded to four, airflow obstruction being split into reversible and irreversible, then the doctors' accuracy for category prediction drops to 53 of 112 (47%) (tables 2 and 3).

The differences between pretest and post-test probabilities are tabulated for 123 patients (table 4). Overall, 61% of the tests were useful according to our criterion that the change between pretest and post-test probability should be greater than 0.5. Of these useful tests, 46% were in cases where the doctor's most likely prediction was irreversible obstruction and the test demonstrated reversibility, or vice versa.

Considering the prediction of airflow obstruction, we found that when the doctors' pretest probability for airflow obstruction was at least 0.8, then 83% of these patients actually had this defect (figure). This suggests that when the doctor is fairly certain of the presence of airflow obstruction, his probability prediction is reasonably well calibrated. Calibration appears less good, however, when the doctor thinks that airflow obstruction is unlikely.

Discussion

We have attempted to assess doctors' ability to predict the results of simple spirometry in breathless patients and their degree of certainty about the type of abnormality identified. Doctors seem reasonably accurate in their identification of the type of condition as either restrictive or obstructive disease or normality. We accept that the definitions of ventilatory defects used in this study are arbitrary and simplistic; more detailed measurements of ventilatory capacity, however, were beyond the scope of this study. Furthermore, these definitions have been in long term use in this laboratory and are well known to the doctors making the requests.

Prediction of reversibility of airflow obstruction

was less accurate, doctors being correct in only 56% of cases. The level of significant reversibility set in this study is an arbitrary one, albeit widely used in practice.^{6,7} If it is accepted that testing for reversibility of an obstructive pattern of spirometric indices is helpful, then it is reasonable to set a level at which any improvement demonstrated is taken to indicate reversibility. Among the 56 patients classified as having irreversible obstruction on the basis of spirometry, there was less than 10% improvement in both FEV₁ and FVC after salbutamol for 46 of them. Furthermore, of the 33 patients classified as having reversible obstruction, 22 showed an improvement of at least 20% in either FEV₁ or FVC. Thus relatively few of our "obstructive" patients fall in the contentious area of 10–20% improvement in either FEV₁ or FVC (21 of 89; figures relate to the 123 patients entering the study). We are not, of course, suggesting that bronchodilator treatment should be denied to patients who do not fulfil our criterion of reversibility, and indeed some of them may be responsive to corticosteroids, other bronchodilators or a bigger dose of bronchodilator.

Our criterion for determining that a test has been useful requires that the test gives a result which the doctor thought unlikely. It may be argued that changes of less than 0.5 between the pretest and post-test probabilities are useful in that they provide reassurance for the doctor. Of course, the true test of the diagnostic usefulness of an investigation is whether the result influences the patient's management.

Our criterion of usefulness is reasonable only if the doctors' probability predictions are considered to be well calibrated. If, for example, the pretest probability of airflow obstruction in a particular case is 0.9, then we would expect 90% of all similar patients to be shown to have airflow obstruction. Calibrating predictions by the method described indicated reasonable accuracy in the strength of doctors' pretest probabilities, although there was a general tendency to

underestimate the probability of airflow obstruction.

Our results show that simple spirometry has a useful role in the investigation of breathlessness. There is increasing awareness of the need to select tests appropriately.^{1,8} The method of evaluating a test used here could be applied to further studies of the diagnostic value of other everyday tests.

We are grateful to Mrs P Andrews for her help in carrying out the lung function tests, to Miss B Copland for her secretarial assistance, and to all the medical staff of the Westminster Hospital for their cooperation in completing the new request form. This work was carried out with the support of a grant from BUPA.

References

- 1 Griner F, Mayewski RJ, Muslin AI, Greenland P. Selection and interpretation of diagnostic tests and procedures. *Ann Intern Med* 1981;**94**:553–600.
- 2 American Thoracic Society. ATA statement—Snowbird Workshop on Standardisation of Spirometry. *Am Rev Respir Dis* 1979;**119**:831–8.
- 3 Morris JF, Koski A, Johnson LC. Spirometric standards for healthy non-smoking adults. *Am Rev Respir Dis* 1971;**103**:57–67.
- 4 Cherniak RM, Raber MB. Normal standards for ventilatory function using an automated wedge spirometer. *Am Rev Respir Dis* 1972;**106**:38–46.
- 5 Knudson RJ, Slatin RC, Lebowitz MD, Burrows B. The maximal expiratory flow-volume curve. Normal standards, variability and effects of age. *Am Rev Respir Dis* 1976;**113**:587–600.
- 6 Report of the Committee of Emphysema. American College of Chest Physicians. Criteria for the assessment of reversibility in airways obstruction. *Chest* 1974;**65**:552–3.
- 7 Clausen JL. Pulmonary function testing guidelines and controversies. A project of the California Thoracic Society. New York: Grune & Stratton, 1984.
- 8 Griner PF, Glaser RJ. Misuse of laboratory tests and diagnostic procedures. *N Engl J Med* 1982;**307**:1336–9.