Eur Respir J, 1993, 6, 983–990
Printed in UK - all rights reserved


European Respiratory Journal
ISSN 0903 - 1936

# The contribution of respiratory function tests to clinical diagnosis

G. Laszlo*, G.N. Lance**, G.T.R. Lewis+, A.O. Hughes++

*The contribution of respiratory function tests to clinical diagnosis. G. Laszlo, G.N. Lance, G.T.R. Lewis, A.O. Hughes. ©ERS Journals Ltd 1993.*

**ABSTRACT: The purpose of this study was to investigate the amount of diagnostic information contained in a set of routine lung function studies and to attempt to determine which tests could be omitted without significant loss of discrimination.**

**Cluster analysis was performed on a set of physiological and questionnaire data, collected prospectively in 1,542 male patients, referred consecutively for measurement of forced expired volumes, static lung volumes and measurements of the transfer factor for carbon monoxide. A respiratory questionnaire was completed for each patient. A physician assigned the patients to a rigorously defined diagnostic category, based on supporting clinical information, as well as pulmonary function, apart from 241 patients with unusual diagnoses and those in whom the criteria did not apply satisfactorily. This diagnosis was never included as a classification variable.**

**Basing the classification on three independent measurements, total lung capacity, the ratio of forced expiratory volume in one second ($FEV_1$) to vital capacity, and the transfer factor for carbon monoxide, the computer generated six groups: one normal, one showing an isolated gas exchange defect, and four with varying degrees of restriction and obstruction. This classification performed well in separating the patients with the clinical diagnoses of chronic airflow obstruction, bronchial asthma and interstitial lung disease from those with ischaemic and valvular heart disease and other miscellaneous disorders. Omitting total lung capacity resulted in some loss of specificity, but valid information was still obtained. The inclusion of all the static and dynamic lung volumes and of carbon monoxide transfer coefficient made little difference. Inclusion of information from the respiratory questionnaire about smoking, sputum and breathlessness was unhelpful, as the resulting groups were less well recognizable physiologically.**

**We conclude that in addition to confirming the validity of historic classifications, disorders of ventilation and gas exchange, these results show that standard respiratory function tests can contribute useful diagnostic information even when considered in isolation, and may be allowed an important role in the diagnostic process.**
*Eur Respir J., 1993, 6, 983–990.*

* Respiratory and + Medical Physics Departments, Bristol Royal Infirmary, Bristol, UK. ** Dept of Computer Science and ++ Dept of Epidemiology, University of Bristol, Bristol, UK.

Correspondence: G. Laszlo
Respiratory Department
Bristol Royal Infirmary
Bristol BS2 8HW
UK

Keywords: Classification
cluster analysis
diagnosis
lung volumes
respiratory function tests
transfer factor (diffusing capacity) for carbon monoxide

Received: March 10 1992
Accepted after revision February 26 1993

Pulmonary function tests can be used to identify different patterns of respiratory impairment. According to time-honoured convention, reductions of maximum breathing capacity are classified into "obstructive" and "restrictive" defects [1]. Abnormalities of gas exchange are generally found to accompany these disturbances, but can occur in isolation [2]. Much, but not all, of the available information is contained in the derivatives of the forced expired volume-time curve.

Statistical methods, aided by computers, are now able to generate classifications which are as free as possible from bias, and to identify those variables which contribute most information. Over a period of 3 yrs we accumulated, prospectively, a set of standard lung function tests (spirometry, lung volumes, single breath carbon monoxide transfer), with additional clinical information, which we have found helpful in the composition of lung function reports (clinical diagnosis, Medical Research Council disability grade, regular expectoration of sputum and smoking habit). We have now attempted to classify our patients on the basis of this information, using a form of cluster analysis which has been studied extensively [3–6]. The classification procedure used is "divisive", in that we start with the whole population and divide it until the optimal number of groups are formed. The available variables are split into two types: 1) intrinsic, those used to form the groups; and 2) extrinsic, those which are relevant but which must not be allowed to influence the classification, such as diagnosis.

The diagnosis of chronic respiratory disorders and other conditions leading to breathlessness depends on clinical, radiological, pathological and physiological information.

In this study, we set out to determine whether the groups generated by an unbiased cluster analysis of lung function tests were identifiable physiologically and clinically and, if so, to explore ways of selecting the tests most likely to discriminate patients in different, rigorously defined, diagnostic categories. We report the results of four computations. In the first, we classified the subjects using eight commonly reported measurements (forced expiratory volume in one second ($FEV_1$), slow vital capacity (EVC), $FEV_1$/EVC (forced expiratory ratio (FER)), total lung capacity (TLC), functional residual capacity (FRC), residual lung volume (RV), transfer factor of the lungs for carbon monoxide (TLCO) and carbon monoxide transfer coefficient (KCO)). Secondly, we studied a "biased" classification, based only on FER, TLC and TLCO, which is theoretically the smallest amount of information needed to generate groups with and without obstruction, restriction and gas exchange defects, respectively. Thirdly, we examined the performance of the three most easily performed tests, FER, EVC and TLCO. Finally, we included clinical information based on a respiratory questionnaire obtained at the time of the lung function test.

## Methods

Data were collected from 1,542 consecutively referred male patients, at their first visit to the respiratory laboratory for measurement of lung volumes, spirometric tests and transfer factor for carbon monoxide. By means of a structured interview, information was obtained about present and past smoking habit, current Medical Research Council dyspnoea grade [7], and whether sputum was expectorated for three months in each year. This information was incorporated into the database. Additional information available to the reporter included current medication, history of episodic wheezing or nocturnal dyspnoea, exertional chest pain, oedema and occupational or environmental exposure to biological or mineral dust. Peak expiratory flow (PEF) was measured on a Wright meter (Clement Clark). $FEV_1$ and EVC were measured on a dry bellows wedge spirometer (Vitalograph). These were measured again (PEF-PB, $FEV_1$-PB and EVC-PB) after the inhalation of 200 μg of salbutamol when airflow obstruction was present, as determined by a ratio of $FEV_1 \times 100$/EVC (FER) of less than 70%, or when requested by the referring physician. The patients were asked to withhold bronchodilators for 4 h before the test, but were not excluded if this request was not followed. Lung volumes were measured by helium dilution and subdivided as FRC, RV and TLC. TLC was obtained from RV plus inspired vital capacity. Carbon monoxide uptake was measured using automated apparatus (Morgan transfer test), using the single breath alveolar volume (VA) for calculation of TLCO. Single breath residual volume (RVSB) was calculated by subtracting inspired volume from VA. KCO was calculated as TLCO/VA (body temperature, pressure and saturation (BTPS)). Carbon dioxide tension ($PCO_2$) was measured by the rebreathing method when $FEV_1 \leq 1$ *l*, or when requested [8]. Values were expressed as the percentage of the predicted normal derived from European Coal and Steel Community (ECCS) standard formula [9]. KCO was not transformed.

Diagnostic classification was performed, within four weeks of reporting, by one physician with access to clinical information. Ischaemic heart disease (361 patients) and the other cardiac disorders (421 patients) were classified on the basis of cardiac catheter data. Shunt fraction was entered as an independent variable in cases of septal defect, and was calculated from ventricular and arterial oxygen saturations. The interstitial lung diseases included separate lists of 200 patients with sarcoidosis, fibrosing alveolitis, scleroderma, occupational lung diseases and a group of patients with miscellaneous fibrosing disorders. The clinical diagnosis of sarcoidosis was accepted without histological confirmation where there was bilateral hilar adenopathy or a characteristic chest radiograph in association with uveitis. Not all patients so classified had demonstrable pulmonary involvement. Cryptogenic fibrosing alveolitis was diagnosed histologically, or by the presence of clubbing and characteristic crackles, in the absence of asbestos ex-posure or any other known autoimmune cause of alveolar disease. Lung fibrosis otherwise appears as "other interstitial lung disease". The 85 patients defined as having bronchial asthma had a history of variable wheezing and dyspnoea prior to the onset of chronic symptoms. The diagnosis of chronic airflow obstruction (234 patients), therefore, includes a heterogeneous group of patients with low FER. It also includes some with chronic bronchial asthma, who did not fit with certainty into the above definition. The patients in whom a diagnosis of emphysema was made confidently on radiographic grounds could be identified separately. The combination of chronic bronchitis with cardiac disease was identified separately but assigned to the cardiac group. When two separate diagnoses were present, a judgement was made as to which was dominant. There is a miscellaneous group of 241 patients, including those in whom the diagnosis was not known.

Diagnosis was not included as a variable in the computations.

### *Computer analysis*

*The raw data.* We analysed the results of 1,542 male patients, aged 18 yrs and over (mean age 52 yrs).

For each patient, many variables were recorded, but some of the less usually performed tests readings were not always obtained. All of the programs used for the analyses can handle "missing data", although they do to some extent modify the results. The variables collected are shown in table 1.

*The transformations applied to the raw data.* The following variables were expressed as a percentage of expected normal [8]: PEF, $FEV_1$, EVC, PEF-PB, FEV-PB, EVC-PB, FRC, RV, TLC, TLCO. No other variables were transformed.

Table 1. – The lung function variables, their types, and whether or not they have been transformed

| Variable | Type of variable* | Data available for patients | Transformation YES/NO$^{\emptyset}$ |
|---|---|---|---|
| Identifier | - | - | - |
| Age yrs | N | All | No |
| Sex | Q | All | No |
| Height m | N | All | No |
| Weight kg | N | All | No |
| PEF $l \cdot min^{-1}$ | N | 1507 | Yes |
| $FEV_1$ $l$ | N | All | Yes |
| EVC $l$ | N | All | Yes |
| PEF-PB $l \cdot min^{-1}$ | N | 332 | Yes |
| FEV-PB $l$ | N | 342 | Yes |
| EVC-PB $l$ | N | 342 | Yes |
| FRC $l$ | N | 1409 | Yes |
| RV $l$ | N | 1412 | Yes |
| TLC $l$ | N | 1412 | Yes |
| RVSB $l$ | N | 1281 | No |
| $T_{LCO}$ $mmol \cdot min^{-1} \cdot kPa^{-1}$ | N | 1419 | Yes |
| $K_{CO}$ $mmol \cdot min^{-1} \cdot kPa^{-1} \cdot l^{-1}$ | N | 1416 | No |
| $Pco_2$ kPa | N | 81 | No |
| Dyspnoea | M(5) | 1457 | No |
| Smoking | M(7) | All | No |
| Sputum | Q | 1453 | No |
| Shunt % | M(9) | 69 | No |
| Disease classification† | M(35) | All | No |
| FER % | N | All | - |

*: the variable types are Q: binary, two; N: numerical, eighteen; or M: multistate, four; and the number of states are shown in parenthesis. †: the program could not handle the 35 state variable "diagnosis" so this was compressed into six diagnostic multistate variables. This had no effect on the classifications, since diagnosis was not actually used as a classificatory variable. $^{\emptyset}$: lung volumes and $T_{LCO}$ were expressed as percentages of standard reference values calculated from age and height. The regression formulæ used were taken from QUANJER [9]. PEF: peak expiratory flow; $FEV_1$: forced expiratory volume in one second; EVC: slow vital capacity; PB: postbronchodilator; FRC: functional residual capacity; RV: residual volume; TLC: total lung capacity; RVSB: single breath residual volume; $T_{LCO}$: transfer factor of the lungs for carbon monoxide; $K_{CO}$: carbon monoxide transfer coefficient; $Pco_2$: carbon dioxide tension; Shunt: shunt fraction; FER: forced expiratory ratio, *i.e.* $FEV_1 \times 100/EVC$.

*Preliminary data surveys.* Before any classifications were attempted, the distribution of the variables was examined. Physiologically impossible results attributable to recording errors were removed.

*The classifications [3, 4].* The first program determined which variables should be used for splitting the population into two initial groups in such a way as to optimize the inhomogeneity. If the variable selected happened to be missing from any individuals, these were arbitrarily placed in one of the two groups. Next, all individuals were examined to see if they were "more like" the group in which they do not reside; if so, they were moved to the other group. This procedure was repeated until convergence was obtained. The effect of this "reallocation" was to further improve the homogeneity of the two groups, and it corrected any misplacements caused by the random placement of individuals from which the division variable was missing.

The two groups were now examined and the least homogeneous was the next candidate for division. The chosen group was now divided and the reallocation procedure was applied to the two derived groups, and this process was repeated until the optimum number of groups had been obtained, yielding the greatest homogeneity within the groups and the greatest differences between them.

A second program used these results and gave all individuals the opportunity of reallocating to any of the final groups. This tidied the groups to again improve their overall homogeneity. The revised membership of the groups was printed out for study if required.

A third computation [4] was used to calculate the Cramer value for each variable, a statistical measure which indicates the importance of the variable to defining the classification [5]. A Cramer value can be obtained for both continuous and discrete variables. It lies between 0 and 1: a zero value indicates that the variable is randomly distributed across groups, and so does not contribute to the classification, whereas a value of one

indicates a variable which is a perfect discriminator. A scaled Cramer value can be used to determine the optimum number of groups in a classification [6]. When applied to the data for the 1,542 patients, six groups were indicated as being optimum, and it is these results which are described below. When large numbers of variables are studied, those with Cramer values exceeding the average value are considered to be the most useful discriminators. Variables with low Cramer values are not discussed, because they contribute little to distinguishing between groups.

For the purpose of classification, the variables are separated into two sets: 1) intrinsic, *i.e.* those which are used in the classification; and 2) extrinsic, *i.e.* those which are masked out. The Cramer values for the extrinsic variables indicate whether they too are good discriminators, even though they played no part in the determination of the groups.

We classified the subjects using four different sets of intrinsic variables. The mean values for each test were listed to provide a physiological profile of the most typical members of each group. We computed:

1. The classification based on the eight variables usually reported ($FEV_1$, EVC, FER, RV, FRC, TLC, TLCO, KCO).
2. A classification based on FER, TLC and TLCO.
3. A classification based on FER, EVC and TLCO.
4. The classification based on all the information collected, including questionnaire data.

## Results

The correlations between the lung function variables are given in table 2. In general, static and dynamic lung volumes and their derivative FER were highly correlated. TLCO, FER and TLC appeared to be relatively independent of each other.

### *Eight static and dynamic lung volume variables with CO transfer*

Data from 1,542 adult male patients were sorted by the computer into six groups. Table 3 lists the eight variables which were included, the first four having Cramer values greater than the mean. FER, $FEV_1$ and EVC contributed most to the classification, TLCO and KCO the least.

Group 1: This group was characterized by somewhat high FER and low lung volumes. Mean TLCO was 92% of predicted. Broadly, this may be considered as "restriction".

Group 2: These results were within normal limits. Mean EVC was 117% and mean TLCO was 98% predicted, which are similar to the results in a group of healthy middle-aged men studied in our laboratory [10].

Group 3: This group showed a virtually normal FER at 94% predicted, but low static lung volumes and TLCO, and normal KCO.

Table 2. – Table of correlations between subdivisions of lung volume, the derivatives of the forced expired volume curve and the indices of carbon monoxide transfer

| | $FEV_1$ | FER | RV | EVC | FRC | TLC | TLCO | KCO |
|---|---|---|---|---|---|---|---|---|
| $FEV_1$ | 1.000 | 0.769 | -0.525 | 0.829 | -0.306 | 0.283 | 0.467 | 0.136 |
| FER | | 1.000 | -0.647 | 0.306 | -0.593 | -0.261 | 0.337 | 0.299 |
| RV | | | 1.000 | -0.266 | 0.845 | 0.556 | -0.201 | -0.253 |
| EVC | | | | 1.000 | 0.017 | 0.632 | 0.408 | -0.048 |
| FRC | | | | | 1.000 | 0.692 | -0.166 | -0.342 |
| TLC | | | | | | 1.000 | 0.183 | -0.224 |
| TLCO | | | | | | | 1.000 | 0.742 |
| KCO | | | | | | | | 1.000 |

For abbreviations see legend to table 1.

Table 3. – Six groups derived from 1,542 adult males, using eight variables only

| Variable | Population mean | Group 1 n=395 | Group 2 n=355 | Group 3 n=199 | Group 4 n=225 | Group 5 n=191 | Group 6 n=177 | Cramer value |
|---|---|---|---|---|---|---|---|---|
| FER | 91 | 103 | 100 | 94 | 91 | 84 | 51 | 0.881 |
| $FEV_1$ | 85 | 98 | 114 | 63 | 88 | 67 | 35 | 0.799 |
| EVC | 94 | 98 | 117 | 69 | 100 | 84 | 73 | 0.798 |
| RV | 93 | 64 | 89 | 69 | 104 | 113 | 161 | 0.788 |
| FRC | 93 | 74 | 97 | 70 | 103 | 99 | 133 | 0.777 |
| TLC | 91 | 84 | 104 | 67 | 97 | 89 | 100 | 0.765 |
| TLCO | 84 | 92 | 98 | 67 | 67 | 98 | 55 | 0.692 |
| KCO | 1.59 | 1.77 | 1.64 | 1.62 | 1.16 | 1.93 | 1.13 | 0.604 |

The variables are listed in descending order of their Cramer value. The group number is assigned in decreasing value of FER. The results are expressed as mean % predicted for the group except for KCO, which is untransformed. For abbreviations see legend to table 1.

Group 4: This group was characterized by a slightly lower FER at 91% of predicted, normal lung volumes, and low Kco.

Group 5: This group had moderate airflow obstruction, with well-preserved TLco, and high Kco.

Group 6: These showed severe airflow obstruction, normal TLC, high FRC and RV, and low Kco.

Groups 4, 5 and 6 were predominant in patients with chronic airflow obstruction, and Groups 1–3 in those with interstitial lung diseases.

*FER, TLC and TLco*

The data were again classified into six groups. Table 4 describes the characteristics of the groups again arranged in descending order of FER. TLco and FER dominated this classification. The six groups could be interpreted broadly as:

Group 1: Normal, with a tendency to show slightly lower values for FRC and RV than published normal figures.

Group 2: Normal FER and $FEV_1$, with marginally reduced values for TLco and Kco.

Group 3: A severely restricted group, with low static lung volumes and TLco.

Group 4: Minimal airflow obstruction, with normal static lung volumes and TLco.

Group 5: Moderate airflow obstruction, reduced TLco, and modestly increased FRC and RV.

Group 6: Severe airflow obstruction, very low TLco, and high FRC and RV.

*FER, EVC and TLco*

The subjects were again classified into six groups. Two groups were essentially normal, two had restrictive disease of varying severity, and two had, respectively, mild and severe airflow obstruction. TLco and FER were the most important discriminators in this classification. These results suggested that, of the subdivisions of static lung volume, residual volume was the most sensitive indicator of obstruction, whilst TLC appeared the most useful at separating the normal and obstructive groups from those with restrictive ventilatory defects.

*Lung function tests supplemented by respiratory questionnaires*

The results of the lung function tests in the same 1,542 patients were classified again, together with answers from the respiratory questionnaires and additional information (sputum, dyspnoea, smoking, $P_{CO_2}$ cardiac catheter data, age, height and weight). The presence or absence of sputum had a major effect on the classification. Patients with severe airflow obstruction were divided into two otherwise similar small groups on the basis of this, whilst over 1,000 patients with a variety of disorders were distributed into two large groups, 1 and 2.

*Relationship of classification to diagnosis*

The ability of the classification to separate the diagnostic groups for those subjects diagnosed as having chronic airflow obstruction, asthma, ischaemic heart disease, valvular heart disease and interstitial lung disease is shown in table 5. The unbiased classification employing the eight standard lung function variables performed well ($\chi^2$=733 with 25 DF). Classification using three related independent variables, FER, TLC and TLco, performed similarly ($\chi^2$=806 with 25 DF). The use of the spirometric tests FER and EVC with TLco also yielded a valid separation, $\chi^2$=713 with 25 DF.

Another way to compare the classifications with the true diagnoses is to permute the columns of the square matrices in table 5 and to determine which permutation results in the largest number of cases lying in the diagonals. If there were perfect separation, such that all patients with heart disease were in Group 5, all with obstructive disease in Group 4, and so on, then the diagonals would add up to the total number of patients, in this case 1,542. The diagonal sums for the permuted tables 5a and b and for the case of the three variables FER, EVC and TLco are 498, 514 and 514, respectively. This confirms the $\chi^2$ test that the use of FER, TLC and TLco is marginally better than the use of all eight variables. Fourteen percent of patients with the diagnosis of chronic airflow obstruction were classified into the normal or restrictive groups, whilst 19–20% of those with interstitial lung disorders appeared in the obstructive group. The proportions were similar in both classifications.

Table 4. – Six groups, derived from 1,542 adult males using only variables FER, TLC and TLco (marked by *). Groups are arranged in descending order of FER

| Variable | Group 1 n=329 | Group 2 n=520 | Group 3 n=185 | Group 4 n=231 | Group 5 n=155 | Group 6 n=122 | Cramer value |
|---|---|---|---|---|---|---|---|
| FER* | 105 | 101 | 99 | 84 | 65 | 49 | 0.878 |
| $FEV_1$ | 106 | 96 | 72 | 85 | 51 | 37 | 0.757 |
| EVC | 105 | 98 | 76 | 104 | 81 | 77 | 0.534 |
| TLC* | 93 | 89 | 72 | 101 | 91 | 102 | 0.542 |
| RV | 80 | 81 | 73 | 105 | 120 | 157 | 0.628 |
| FRC | 86 | 85 | 77 | 101 | 107 | 134 | 0.605 |
| TLco* | 109 | 81 | 55 | 100 | 74 | 42 | 0.878 |
| Kco | 1.95 | 1.54 | 1.31 | 1.74 | 1.48 | 0.87 | 0.632 |

For abbreviations see legend to table 1.

We also compared the performance of the eight variable classification and classification employing only FER, TLC and TLCO in separating three small groups of patients with lung diseases which have distinct physiological characteristics:

1. A subgroup of chronic airflow obstruction patients in whom the diagnosis of emphysema could be made with confidence on radiographic grounds. These differed from the whole group of patients with airflow obstruction.
2. Bronchial asthma.
3. Cryptogenic fibrosing alveolitis and related disorders, excluding rheumatoid lung, sarcoid and extrinsic alveolitis.

The two classifications performed comparably. The use of additional lung volume information in the eight variable classification appeared to improve the separation between asthma and the interstitial lung disorders ($\chi^2$=102 and 75, respectively), but otherwise the $\chi^2$ values for the comparisons were marginally higher when only three variables were included. Nine percent of patients with "chronic bronchitis and emphysema" were placed in the "restrictive" group in the eight variable classifications, the corresponding figure being 17% in the three variable run. For fibrosing alveolitis, the proportions placed outside the "restrictive" group were 13 and 14%.

Table 5. – Relationship of computer classification to clinical diagnosis

a) Eight variables ($FEV_1$, EVC, FER, RV, FRC, TLC, TLCO, KCO)

| | Group no. | | | | | | |
|---|---|---|---|---|---|---|---|
| Diagnosis | 1 | 2 | 3 | 4 | 5 | 6 | Total |
| 1 | 5 | 8 | 20 | 38 | 40 | 123 | 234 |
| 2 | 12 | 15 | 4 | 3 | 33 | 18 | 85 |
| 3 | 67 | 45 | 49 | 23 | 11 | 5 | 200 |
| 4 | 146 | 116 | 27 | 53 | 17 | 2 | 361 |
| 5 | 117 | 94 | 78 | 70 | 54 | 8 | 421 |
| 6 | 48 | 77 | 21 | 38 | 36 | 21 | 241 |

b) Three variables (FER, TLC, TLCO)

| | Group no. | | | | | | |
|---|---|---|---|---|---|---|---|
| Diagnosis | 1 | 2 | 3 | 4 | 5 | 6 | Total |
| 1 | 4 | 15 | 14 | 39 | 70 | 92 | 234 |
| 2 | 8 | 9 | 1 | 45 | 15 | 7 | 85 |
| 3 | 45 | 63 | 55 | 23 | 7 | 7 | 200 |
| 4 | 107 | 182 | 29 | 36 | 6 | 1 | 361 |
| 5 | 111 | 160 | 62 | 55 | 28 | 5 | 421 |
| 6 | 54 | 91 | 24 | 33 | 29 | 10 | 241 |

Diagnoses: 1: chronic airflow obstruction; 2: asthma; 3: cryptogenic fibrosing alveolitis and other similar disorders; 4: ischaemic heart disease; 5: valvular heart disease; 6: other. The groups refer to the computer-generated classification described in tables 3 and 4. Values are patient numbers. For description of the variables used, see text. The use of three selected variables gave better separation between the diagnoses ($\chi^2$=806 for 24 DF) than did 8 variables ($\chi^2$=773). For abbreviations see legend to table 1.

## Discussion

Our classification allocates the subjects into discrete groups according to all available information. In this study, the final "optimum" number was determined mathematically from the scaled Cramer value [6], yielding the best available homogeneity within the group, and the greatest heterogeneity between the groups. In this type of analysis, information which reinforces the main classifier is given considerable weight. By contrast, principal component analysis [11], canonical correlation analysis [12], and other multivariate analyses aim to give less weight to the reinforcing information and to look for factors which provide independent information to identify how many components can significantly be described.

During this study, great care was taken to ensure that the diagnostic categories were as accurate as possible. One physician assigned the diagnosis on the basis of the clinical information provided by the referring specialist, the respiratory questionnaire, the lung function results, and the results of detailed cardiological investigations, supplemented by a personal knowledge of the patient or perusal of the notes, using radiographs in many instances. The miscellaneous group of 241 unclassified patients includes those subjects to whom the rigorously defined diagnostic criteria did not apply, as well as those with rarer conditions. These are included in the analysis (table 5). The diagnoses are, therefore, not independent of physiological information of various kinds, including previous measurements of $FEV_1$, but they are probably the best that can be achieved in such large numbers under these circumstances, especially as the answers to the standard respiratory questions were available to the reporter in each case. This analysis enables us to explore the amount of diagnostic information that is contained in the data that can be collected in the laboratory without any subjective clinical input, using only various combinations of physiological tests and answers to standard questions in a structured interview. We have confined our exploration of the diagnostic significance of the groups to those categories which were rigorously defined.

The distributions of the groups among the various diagnostic categories were influenced by the referral pattern. At this time, the laboratory received referrals from the cardiac unit for routine assessment at the time of cardiac catheterization, and from the rheumatology department for the assessment of systemic sclerosis; no renal unit existed in the hospital and many patients with chronic airflow obstruction were not referred for measurements other than spirometry. Moreover, some disorders in which the pulmonary function laboratory plays an important part in management, such as cryptogenic fibrosing alveolitis, are relatively rare. Patients with this diagnosis accounted for only 26 of our reports, while other interstitial lung diseases including sarcoid and scleroderma bring the total in this category to only 200.

In theory, classifications of this type purport to be able to handle missing information. Whilst this is true, it is essential that there should be no systematic reason why information is excluded if bias is not to be introduced. For instance, in this study, $P_{CO_2}$ was measured only when

$FEV_1$ was below 1.1 *l*, that is for only 81 patients, and mean $P_{CO_2}$ was 8 kPa (60 mmHg).

With these constraints in mind, we were able to test the ability of this type of divisive classification to generate recognizable clusters, and to separate the patients into groups using only the information that could be obtained by the technician.

This study shows the considerable extent to which the pattern of abnormality of lung function is a useful pointer to a clinical diagnosis. In clinical practice, reports of standard lung function tests concentrate initially on the presence or absence of airflow obstruction and on the presence of defective gas exchange. Ideally, an estimate of lung volume should be obtained to identify restrictive ventilatory defects, which are rigorously defined as a reduction of total lung capacity [1, 13, 14]. Only three variables (FER, TLC and $T_{LCO}$) are needed to generate this classification. As would be expected, the computer used these variables to generate a normal group, and five recognizable types of physiological disturbance. The constitution of the groups was significantly different from the clusters obtained using eight variables, and including more derivatives of lung volume.

We tested the extent to which patients with different diagnoses were separated by these classifications, selecting first five clearly definable categories: chronic airflow obstruction, asthma, interstitial lung disease, ischaemic heart disease and valvular heart disease. As these categories contain mildly and severely affected cases, one would expect considerable overlap in the mildly abnormal groups, and good separation in the severely abnormal groups. This did indeed occur, with each disorder behaving as expected. Patients with ischaemic heart disease were classified mainly as normal, mildly restricted or mildly obstructed. Those with valvular heart disease had a wide variety of physiological defects. The overall distinction between the five diagnostic categories was better using these three variables than using eight, though only marginally, although the additional lung volume information and transfer co-efficient helped to separate the patients with chronic bronchitis and emphysema from the normal subjects. Most of the patients with cryptogenic fibrosing alveolitis were concentrated into the restrictive groups; separation between patients with interstitial lung disease from those with chronic bronchitis and emphysema was not improved by the inclusion of more than three variables

There is an increasing tendency to rely on FER and VC to diagnose restrictive lung diseases and to omit lung volume measurements [14]. This leads to some loss of information, and has the theoretical objection that some patients with airflow obstruction may be misdiagnosed. Nevertheless, a classification which includes $T_{LCO}$, VC and FER but omits TLC appears to work fairly well, partly because $T_{LCO}$ falls *pari passu* with FER in obstructive diseases, but is inversely proportional to it in patients with alveolitis.

There is at present no well-established, computerized, expert system for reporting the results of lung function tests, although several are being developed. These go well beyond the crude groupings that we have generated, as the handling of individual results requires a more subtle approach, often involving exploration of the relationship between different tests. For example, central airflow obstruction is sometimes diagnosed incorrectly as bronchial asthma, and the few cases studied would be included with it in our classification. To avoid errors of this type, it is necessary to request extra tests, some of which yield useful information only in a very small number of cases. In this instance, examination of the relationship of PEF to $FEV_1$ provides a clue to the presence of central, as opposed to diffuse, airflow obstruction. This examination is easily performed by a program in which the results are examined logically in a systematic order, but it requires minute examination of the results rather than a process which lumps the information.

The value of our approach is that it takes a global view of the information generated about the population under study. Our results are reassuring in that they confirm that an unbiased examination of static and dynamic lung volumes and $T_{LCO}$ in a large series of patients yields a classification which is recognizable physiologically, goes part of the way towards distinguishing patients with different conditions, and alerts the reporter to possible errors in clinical diagnosis. The data should be complete for each patient. A sequential approach, defining subgroups according to additional information, is likely to be helpful. For diagnostic purposes, selecting a small number of tests is as useful as collecting a large number of highly correlated observations.

***Acknowledgements***: The authors thank D. Smith and J. Dennison for technical assistance and B. Ellison for typing the many versions of this paper.

## References

1. Baldwin E de F, Cournand A, Richards DW. – Pulmonary insufficiency. I. Physiological classification, clinical methods of analysis, standard values in normal subjects. *Medicine* 1948; 27: 243–278.
2. Austrian R, McClement JH, Renzetti AD, Donald KW, Riley RL, Cournand A. – Clinical and physiological features of some types of pulmonary disease with impairment of pulmonary diffusion: the syndrome of alveolar capillary block. *Am J Med* 1951; 11: 667–685.
3. Lance GN, Williams WT. – REMUL: a new divisive polythetic classificatory program. *Aust Comput J* 1975; 7: 109–112.
4. Lance GN, Williams WT. – Attribute contributions to a classification. *Aust Comput J* 1977; 9: 128–129.
5. Cramer H. – *In*: Mathematical Methods of Statistics. Princeton University Press, 1946; pp. 443.
6. Ratkowsky DA, Lance GN. – A criterion for determining the number of groups in a classification. *Aust Comput J* 1978; 10: 115–117.
7. Warley ARH, Finnegan OC, Nicholson EM, Laszlo G. – Grading of dyspnoea and walking speed in cardiac disorders and in chronic airflow obstruction. *Br J Dis Chest* 1987; 81: 349–355.
8. Campbell EJM, Howell JBL. – Simple rapid methods of estimating arterial and mixed venous $P_{CO_2}$. *Br Med J* 1960; i: 458–462.
9. Quanjer PhH (Ed). – Standardized lung function testing. *Bull Eur Physiopathol Respir* 1983; 19 (Suppl. 5): 1–95.

10. Nicholson EM. – The diagnosis of emphysema in life. Thesis, University of Bristol, 1982.
11. Cowie H, Lloyd MH, Soutar CA. – Study of lung function data by principal components analysis. *Thorax* 1985; 40: 438–443.
12. Van Pelt W, Quanjer PhH, Borsboom GJJM, Van der Lende R. – Respiratory symptoms and the maximum expiratory flow-volume curve; a multivariate approach. *Eur Respir J* 1988; 1: 122–132.
13. Gibson GJ. – *In*: Clinical Tests of Respiratory Function. London, McMillan Medical 1984.
14. American Thoracic Society. – Lung function testing: selection of reference values and interpretative strategies. *Am Rev Respir Dis* 1991; 144: 1202–1218.