Thorax (1975), **30**, 548.

Variability of routine pulmonary function tests[1]

JOHN HRUBY and JOHN BUTLER

Division of Respiratory Diseases, Department of Medicine, University of Washington, Seattle, Washington 98195, USA

Hruby, J. and Butler, J. (1975). *Thorax*, **30**, 548–553. **Variability of routine pulmonary function tests.** Pulmonary function tests sometimes indicate a progressive deterioration and at other times a 'stepwise' worsening which may be followed by improvement. Interpretation depends on the extent of random or diurnal variations in function. Routine pulmonary function tests (VC, FEV_1, FRC, and airway resistance (Raw)) were repeatedly measured in normal subjects, patients with stable irreversible airways obstruction, and patients with stable restrictive disease. In all groups there was a significant ($P<0.001$) diurnal variation in Raw, with high values in the morning, low values at noon, and rising values in the evening. The midday Raw values were about 80% of the highest daily values. The considerable random and diurnal variability seen in all tests is reflected in the range of high and low values (% of mean individual response) in individuals. The largest variation in an individual between measurements taken at two different times was 81% in Raw (range: 40% above to 41% below the mean). There was less variation in FEV_1 (29%), FRC (62%), and VC (30%). Thus the finding of a stepwise change in function could reflect its natural variability. When repeated studies are done to assess progress or the effects of therapy on disease, there are many factors, including the time of day at which the tests are performed, which should be standardized as far as possible.

The course and rate of deterioration in diseases causing airways obstruction are currently the subject of controversy as a result of studies using repeated tests of pulmonary function. Reports have indicated that patients may show a linear deterioration of function (Fletcher, 1968) or an exponential decline of function (Emirgil and Sobol, 1971), or that they may show stepwise worsening which is followed by a return of pre-existing or even improved values in some cases (Bates, 1973). These conflicting observations could be better understood if the extent of spontaneous variation in function tests were known.

Reports that there is spontaneous day-to-day and diurnal variation of pulmonary function have appeared since the time of John Hutchinson (1846). The practical question, then, is how reliable are values obtained on one occasion in relation to assessment of severity of disease, response to therapy, and changes with time. This study attempts to answer this question using an experimental protocol which approximates the conditions in a hospital pulmonary function laboratory.

[1]Supported by NHLI Training Grant HL 05819 and NHLI Pulmonary SCOR HL 14152

MATERIAL AND METHODS

PATIENTS Slow vital capacity (VC), one-second forced expiratory volume (FEV_1), functional residual capacity (FRC), and airway resistance (Raw) were measured repeatedly on each of two days in 15 individuals of both sexes aged between 45 and 61. Five patients with known stable airways obstruction and four with known stable restrictive disease were chosen with regard only to their representing the types of patient referred for pulmonary function testing and their willingness to take part in the study. No asthmatics were included. Six normal subjects in the same age range, one of whom was a cigarette smoker, served as controls (Table).

APPARATUS Vital capacity and FEV_1 were measured with a 13·5 litre water-sealed spirometer. Airway resistance and FRC were obtained using an isopressure body plethysmograph (DuBois *et al.*, 1956; Mead, 1960). The plenthys-

TABLE
SUBJECTS

	Sex	Age (yr)	Ht. (in)	Wt. (lb)	History of Smoking	Diagnosis
Normal	F	47	62	189	S	No symptoms
	F	52	64	130	NS	No symptoms
	F	54	62	115	NS	No symptoms
	F	52	64	125	NS	No symptoms
	M	46	64	127	NS (PS)	No symptoms
	M	57	67	155	NS	No symptoms
Obstructive disease	F	56	61	107	S	Chronic bronchitis Moderate obstructive disease
	M	61	68	149	S	Chronic bronchitis Mild to moderate obstructive disease
	M	50	72	138	S	Chronic bronchitis Old tuberculosis Severe obstructive disease
	M	45	68	143	S	Chronic bronchitis Old tuberculosis Moderate obstructive disease
	M	58	68	145	S	Chronic bronchitis Old tuberculosis Moderate obstructive disease
Restrictive disease	F	51	62	117	S	Fibrosis—unknown aetiology Moderate restrictive disease
	F	58	65	102	NS	Old tuberculosis Pneumonectomy, thoracoplasty Restrictive disease Scoliosis
	F	51	57	85	NS	Old tuberculosis right apex Left-sided fibrothorax Restrictive disease
	M	55	67	170	S	Left pneumonectomy for tuberculosis Restrictive disease

S = smoker; NS = non-smoker; PS = past smoker.

mograph traces were displayed and measured at flow rates between 0 and 0·5 l/sec using a rotatable protractor and a precalibrated scale on the oscilloscope face (Briscoe and DuBois, 1958). The body plethysmograph was calibrated each morning. Experimental error for plethysmographic measurement of FRC has been found to vary with a standard deviation of ±0·17 l. (DuBois *et al.*, 1956). All measurements of volume are at BTPS.

PROCEDURE Measurements were made at five fixed times between 8 am and 6 pm. Each test period lasted about 20 minutes. Every patient or subject was tested during all five periods of the day on two, not necessarily consecutive, days. All patients and subjects were asked to get up at 6.45 am and only to eat before the 8 am study and after the 12.30 pm session. They were not permitted to smoke during the 30 minutes before test periods. Medications were unchanged during the period of the study. On the day of the test no bronchodilator drugs were given.

All measurements were made by one observer to avoid inter-observer variability (Rosner, Abraham, and Caceras, 1965). They were always made in the order Raw, FRC, FEV_1, and VC in case the VC and FEV_1 manoeuvres should influence Raw. Observer-subject interaction during testing was standardized by using the same routine and instructions. Five measurements of Raw and FRC were made at each session and the average was recorded. Since Raw varies with lung volume we endeavoured to correct for this by multiplying Raw by thoracic gas volume (Lloyd and Wright, 1963). This product is termed specific airway resistance (SRaw). Three measurements of VC and FEV_1 were made in the sitting position, the highest being recorded. To avoid observer bias, Raw and FRC values were recorded only as a slope quotient, and VC and FEV_1 were calculated after both test days were completed.

ANALYSIS Since the object was to see the extent of change which might be found in a patient with stable disease studied on different occasions and to find if there exists a diurnal trend in function, the data were analysed in the following manner:

For each individual, the mean, standard deviation, and range of values were calculated for the five studies done during each day and for all 10 studies done over the two days of testing of each individual. For each of the groups, an analysis of variance of the mean values for each time period and the daily means was made to assess the significance of any diurnal or day to day variability. Individual FEV_1, VC, and FRC values were expressed as a per cent of predicted normal to standardize the effects of sex, age, height, and

weight. Nomograms derived by Kory *et al.* (1961) and Bates, Macklem, and Christie (1971) were used to obtain the predicted values. Daily air pollution indices were obtained from the local Environmental Protection Agency to ascertain whether environmental pollution might be related to large individual variations in function.

RESULTS

AIRWAY RESISTANCE

(1) The average Raw in controls was 1·8±0·5 $cmH_2O\ l^{-1}\ s$; both the obstructive patients (Raw 3·4±0·6 $cmH_2O\ l^{-1}\ s$) and those with restrictive disease (Raw=3·3±1·9 $cmH_2O\ l^{-1}\ s$) had significantly higher resistance. When values were standardized for lung volume (Lloyd and Wright, 1963), the restrictive group (SRaw=6·6) did not differ from the controls (6·7) although the obstructive group, as expected, had higher values (15·7).

(2) There was a significant ($P<0·01$) diurnal variation for Raw and SRaw for the control, the obstructive, and the restrictive groups. To facilitate comparison of the diurnal trends seen in individuals within each group, the Raw values were calculated as a percentage of the highest daily recorded value. In all groups there was a similar decrease in resistance during the morning followed by an increase toward late afternoon and evening (Fig. 1).

FLOW RATES The average FEV_1 in controls (2·9±0·4 l.) was 114±4·5% of the predicted value, in the obstructive group (1·9±0·5 l.) it was 64±12·9% of predicted, and in the restrictive group (1·1±0·5 l.) it was 46±11·4% of predicted. Average FEV_1/VC in controls was 75±3%, in the obstructive group 53±6%, and in the restrictive group 76±7%.

There was a diurnal variation in FEV_1, the highest flows being recorded at midday, as expected from the Raw variation, but this was not statistically significant.

VOLUMES

Functional residual capacity The average FRC in controls (2·3±0·5 l.) was 81±18·9% predicted, in the obstructive group (3·7±0·9 l.) it was 105±11·1%, and in the restrictive group (1·6±0·4 l.) it was 57±10·7%.

There was no statistically significant diurnal variation in FRC.

Vital capacity Normals were 116±12·1% of predicted, obstructive 88±15·4%, and restrictive 46±11·6%.

Diurnal variability was slight and not statistically significant but showed a rise in VC during the morning followed by a fall in the afternoon.

VARIABILITY

Airway resistance The variability of Raw and SRaw between readings in any individual was surprising. The average standard deviation for individuals in the control group was ±0·41 $cmH_2O\ l^{-1}\ s$, in the restrictive group ±0·41 $cmH_2O\ l^{-1}\ s$, and in the obstructive group ±0·60 $cmH_2O\ l^{-1}\ s$. However, the SD does not highlight the variation

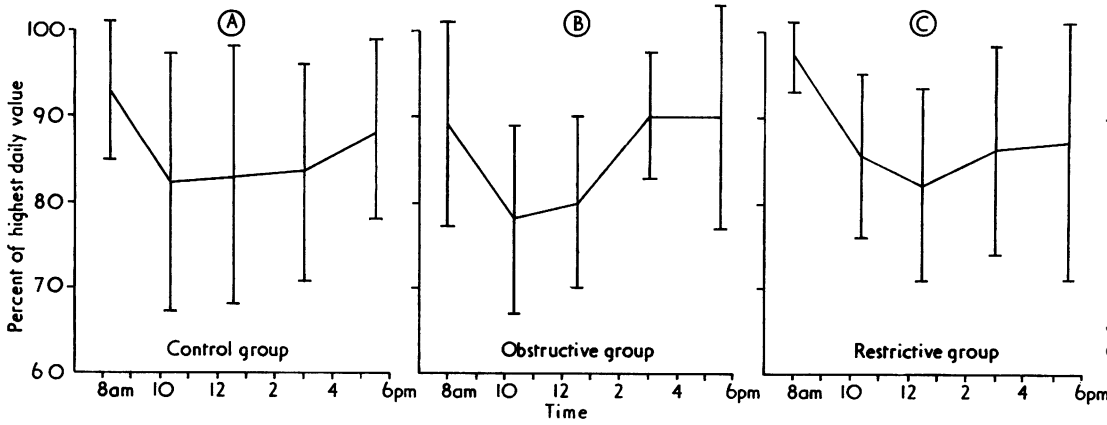

FIG. 1. *Diurnal variation of Raw. Raw for each individual for each time period was calculated as a percentage of the individual's highest daily value. The plotted points are the mean and standard deviation for all individuals of a group.*

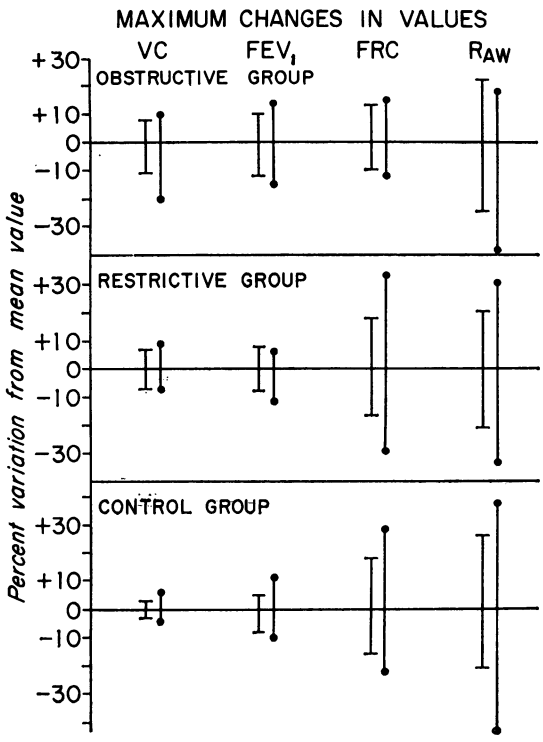

FIG. 2. *Range of recorded values plotted as a per cent change from the mean of all 10 values over the two days: ●—●=the range in the one individual with the largest variability; |—|=the average range of all individuals of a group.*

found on different occasions. This can be appreciated from the range of recorded values in individuals (Fig. 2). Among the controls the individual with the largest variability in function had a range as great as 81% (40% above and 41% below the mean), among the obstructive group 57% (19% above and 38% below the mean), and among the restrictive group 60% (30% above and 30% below the mean). Airway resistance appeared to be the most variable function of those studied. Pollution was not a factor.

Flow rate The range (Fig. 2) was not so impressive but was still considerable. The largest individual range observed in FEV_1 was: in the control group 21% (11% above and 10% below the mean), in the obstructive group 29% (14% above and 15% below), and in the restrictive group 17% (7% above and 10% below).

Functional residual capacity The largest individual range observed was almost as great as in the Raw values: control 50% (27% above and 23% below the mean), obstructive 27% (15% above and 12% below), and restrictive 62% (32% above and 30% below).

Vital capacity This was the most stable function. The maximum individual range observed was: control 10% (6% above and 4% below the mean), obstructive 30% (10% above and 20% below), and restrictive 15% (8% above and 7% below).

DISCUSSION

A highly individualized 24-hour rhythm for the VC of patients with tuberculosis was observed by Dissmann (1950), who found that there was usually a fall in VC during the evening and night and a rise during the morning. This general pattern, which was seen but was not significant in our data, was also noted for both VC and FEV_1 in normals and patients with airways obstruction by Lewinsohn, Capel, and Smart (1960) and in similar patients by de Vries *et al.* (1962). Aepli (1968) reported that, in asthmatics, airway resistance diminishes during the course of the morning with a subsequent increase during the afternoon. This was the pattern we found. However, his normals showed no periodic variation. Nor did those studied by Graham, Heim, and Constantine (1967), whose study included only a few normals, or those studied by Guyatt *et al.* (1967), who were cigarette smokers.

Large spontaneous day-to-day variations have been reported for certain of the pulmonary function tests. Rahn, Fenn, and Otis (1949) observed an average standard deviation of 36% for measurements of VC on 17 subjects taken over a two-month period, while Mills (1949) similarly reported a considerable day-to-day variation in the VC of 17 healthy male subjects which exceeded the variation expected between repeated measurements on one occasion. These results have more recently been supplemented by the observations of Goldberg and Cherniack (1965), who measured VC, TLC, and Raw, and by Spicer and Kerr (1966) who noted day-to-day variations of 1 l. in RV, FRC, and TLC and twofold variation in airway conductance in patients with airways obstruction and in normals.

The variability that we have described here is caused by three factors—the variability in the patient, in the observer, and in the measurement itself. We did not try to identify and separate

these three factors since they are all normally operative when pulmonary function testing is performed. Pulmonary testing laboratories make use of predictive tables derived from studies by Ferris and Smith (1953), Needham, Rogan, and McDonald (1954), Lyons, Tanner, and Picco (1960) (24), Kory *et al.* (1961) (20); and more recently Morris, Koski, and Johnson (1971). The experimental designs of these studies did not consider the possibility of diurnal and spontaneous variability in their individual normal subjects.

Do these data reflect the sort of values expected in a routine pulmonary function laboratory? The number of subjects and patients was limited, and this may have been the reason why the diurnal variation of FEV_1 described by Lewinsohn *et al.* (1960) was not found to be statistically significant in our study, although changes in airway resistance were. The methods of our study differed from those normally used in that

(1) the standard water-sealed spirometer used by us has been superseded in many laboratories by more sophisticated wedge and flow integrating apparatus which may be inaccurate (Fitzgerald, Smith, and Gaensler, 1973);

(2) it is not usual for calibration of the plethysmograph to be performed before and after each test as was done in this study;

(3) these tests were carried out by a single observer and this is not always the case when repeated tests are done on a patient;

(4) rigorous avoidance by the patients of drugs and smoking was insisted upon in this study, but is unusual before routine testing.

Thus in our series the possibility of instrumental or observer error as a cause of variation might be expected to have been reduced and it is likely that the variation in data that we found is a minimal estimate of that which may usually be expected.

Several conclusions may be drawn from this study:

(1) Data from a single set of routine function tests must be interpreted with considerable caution and slight abnormalities or changes from previous values disregarded.

(2) Repeated testing is necessary before an exact delineation of the abnormality can be assured.

(3) The finding of a 'stepwise change' may be just a reflection of the natural variability of function rather than significant of a deterioration or improvement.

(4) When repeated studies are done to asses the results of therapy on the course of a disease, it is important not only to try to standardize the patient's condition (anxiety, smoking, bronchodilator drugs) but also the time of day at which the tests are performed.

One might speculate that the variability we found in our routine tests, which are said to be a rather coarse measure of functional changes, may be evident in recently designed tests to measure minor small airway or gas distribution abnormalities. Recognition of this diurnal or spontaneous variability may explain some of the conflicting reports as to their 'sensitivity' and specificity.

REFERENCES

Aepli, R. (1968). Le rhythme journalier de la résistance bronchique chez l'asthmatique. *Respiration*, **25**, 405.

Bates, D. V. (1973). The fate of the chronic bronchitic. *American Review of Respiratory Disease*, **108**, 1043.

——, Macklem, P. T., and Christie, R. V. (1971). *Respiratory Function in Disease*, 2nd edition, p. 94. Saunders, Philadelphia.

Briscoe, W. A. and DuBois, A. B. (1958). The relationship between airway resistance, airway conductance and lung volume in subjects of different age and body size. *Journal of Clinical Investigation*, **37**, 1279.

de Vries, K., Goei, J. T., Boody-Noord, H., and Orie, N. G. M. (1962). Changes during 24 hours in the lung function and histamine hyperreactivity of the bronchial tree in asthmatic and bronchitic patients. *International Archives of Allergy*, **20**, 93.

Dissmann, E. (1950). Zur Frage von Eigenrhythmus und Grundrhythmus in den Tagesschwangungen der Vitalkapazität. *Acta Medica Scandinavica*, **137**, 441.

DuBois, A. B., Botelho, S. Y., Bedell, G. N., Marshall, R., and Comroe, J. H. (1956). A rapid plethysmographic method for measuring thoracic gas volume: a comparison with a nitrogen washout method for measuring functional residual capacity in normal subjects. *Journal of Clinical Investigation*, **35**, 322.

Emirgil, C. and Sobol, B. J. (1971). Long-term course of chronic obstructive pulmonary disease. *American Journal of Medicine*, **51**, 504.

Ferris, B. G. and Smith C. W. (1953). Maximum breathing capacity and vital capacity in female children and adolescents. *Pediatrics*, **12**, 341.

Fitzgerald, M. X., Smith, A. A., and Gaensler, E. A. (1973). Evaluation of 'electronic' spirometers. *New England Journal of Medicine*, **289**, 1283.

Fletcher, C. M. (1968). Bronchial infection and reactivity in chronic bronchitis. *Journal of the Royal College of Physicians*, **2**, 183.

Goldberg, I. and Cherniack, R. M. (1965). The effects of nebulized bronchodilator delivered with and without IPPB on ventilatory function in chronic

obstructive emphysema. *American Review of Respiratory Disease*, **91**, 13.

Graham, W. G. B., Heim, E., and Constantine, H. P. (1967). Measurement of airway variation and bronchial reactivity in normal and asthmatic subjects. *American Review of Respiratory Disease*, **96**, 266.

Guyatt, A. R., Alpers, J. H., Hill, I. D., and Bramley, A. C. (1967). Variability of plethysmographic measurements of airways resistance in man. *Journal of Applied Physiology*, **22**(2), 383.

Hutchinson, J. (1846). On the capacity of the lungs and on the respiratory functions with a view of establishing a precise and easy method of detecting disease by the spirometer. *Medico-Chirurgical Transactions (London)*, **29**, 137.

Kory, R. C., Callahan, R., Hollis H. G., and Syner, J. C. (1961). Clinical spirometry in normal men. *American Journal of Medicine*, **30**, 243.

Lewinsohn, H. C., Capel, L. H., and Smart, J. (1960). Changes in forced expiratory volumes throughout the day. *British Medical Journal*, **1**, 462.

Lloyd, T. C. and Wright, G. W. (1963). Evaluation of methods used in detecting changes of airway resistance in man. *American Review of Respiratory Disease*, **87**, 529.

Lyons, H. A., Tanner, R. W., and Picco, T. (1960). Pulmonary function studies in children. *American Journal of Diseases of Childhood*, **100**, 196.

Mead, J. (1960). Volume displacement body plethysmograph for respiratory measurements in human subjects. *Journal of Applied Physiology*, **15**, 736.

Mills, J. N. (1949). Variability of the vital capacity of the normal human subject. *Journal of Physiology*, **110**, 76.

Morris, J. F., Koski, A., and Johnson L. C. (1971). Spirometric standards for healthy nonsmoking adults. *American Review of Respiratory Disease*, **103**, 57.

Needham, C. D., Rogan, M. C., and McDonald, I. (1954). Normal standards for lung volumes, intrapulmonary gas-mixing, and maximum breathing capacity. *Thorax*, **9**, 313.

Rahn, H., Fenn, W. O., and Otis, A. B. (1949). Daily variations of vital capacity, residual air, and expiratory reserve including a study of the residual air method. *Journal of Applied Physiology*, **1**, 725.

Rosner, S. W., Abraham, S., and Caceras, C. A. (1965). Observer variation in spirometry. *Diseases of the Chest*, **48**, 265.

Spicer, W. S. and Kerr, D. H. (1966). Variation in respiratory function. *Archives of Environmental Health*, **12**, 217.

Requests for reprints to: Dr. J. Butler, Division of Respiratory Diseases, Department of Medicine, University of Washington, Seattle, Washington, 98195, USA.