Assessment of ventilatory abnormality in the asymptomatic subject: an exercise in futility[1]

BRUCE J. SOBOL with the assistance of BEATRICE WEINHEIMER

From the Division of Medicine and the Cardiopulmonary Laboratory, Grasslands Hospital, Valhalla, New York 10595

The rising morbidity and mortality from chronic pulmonary disease has produced a great deal of interest in simple ventilatory testing among various population groups, as well as in the use of spirometry as part of annual physical examinations in business and industry: this growing concern in a serious health problem can only be commended. The use of these tests to compare pulmonary function in two differing population groups is a proper application of both statistics and ventilatory testing. However, the temptation and the practice of diagnosing individuals without symptoms as suffering from chronic pulmonary disease because of ventilatory abnormalities alone are to be deplored, and in this paper we shall try to show that this is not justified.

METHODS

The methods used in this study have been described previously (Sobol and Emirgil, 1964). All our subjects were from adult working and college populations chosen at random, and had ventilatory function tests and routine chest radiographs taken during tuberculosis screening by the Westchester Tuberculosis and Public Health Association. Two forced vital capacities (F.V.C.) were performed properly in the estimation of the technician, and the two F.V.C.s differed from each other by no more than 10%. All the subjects had a normal chest radiograph, had no history of respiratory allergies or cardiopulmonary-chest wall disease, and had never used tobacco in any form (except where indicated).

RESULTS

Because the data used in this study were obtained in a field survey, some assurance was first necessary that the results were no more erratic than those obtained under the more ideal conditions of the laboratory, where the subjects usually co-operate well and time can be spent to ensure a proper performance. We compared our data for the F.V.C. with that of the Veterans Administration-Army Cooperative Study, and ours with that of Leuallen and Fowler (1955) for a measurement of flow rate, the maximal mid-expiratory flow (M.M.F.). In this study the correlation coefficient (r) and the standard error of the estimate (S.E.E.) for the F.V.C. of 168 healthy, non-smoking men between 20 and 59 years were 0·633 and 0·567 respectively. In the Veterans Administration-Army Cooperative Study the comparable values were 0·64 and 0·58, which were almost identical. The comparative values for the M.M.F. are given in the Table. We think this good agreement between the findings in the laboratory and in the field justified the use of the data we obtained.

Figure 1 demonstrates the wide scatter which the statistical values of r and S.E.E. represent. Since the F.V.C. is a function of age as well as height, in order to present the data on a two-dimensional graph the factor of age was eliminated by graphing only one decade, 20–29. In order to increase the number of points on the graph healthy smokers were also plotted. In this age group there is no statistical difference between smokers and non-smokers in terms of their F.V.C. The extreme variability seen in measurements of flow rate is shown in Figure 2. This is a plot of the M.M.F. against the F.V.C. in men 20 to 59 years of age, all of whom were thought to have normal lungs and were non-smokers. Age was not included in the regression formula since it explains only 1·6% of the unexplained variation and the P value is 0·3–0·4. Note the
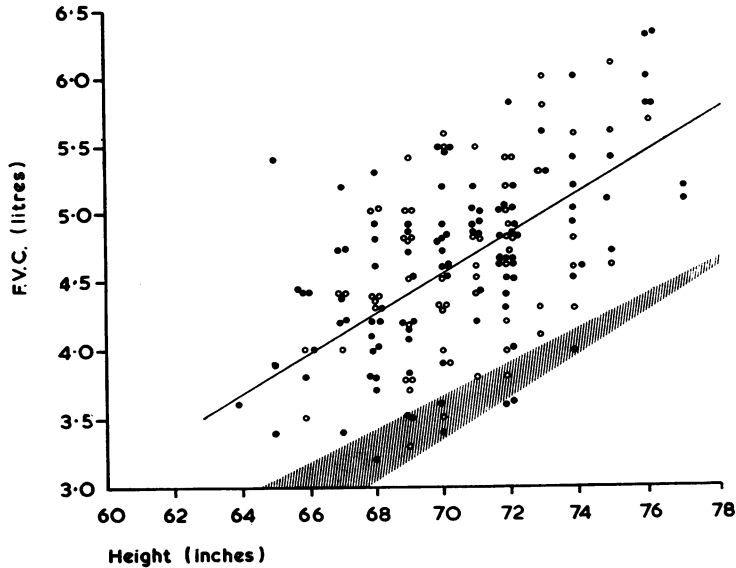
FIG. 1. *Regression of the F.V.C. on height. Open circles: non-smokers, N=64; closed circles: smokers, N=107. The shading encloses an area demarcated above by values of −20% of predicted (the regression line at the centre of the graph) and below by −2 S.E.E. Note that by using the −20% criterion, 11 subjects would be considered abnormal who are normal by the criterion of −2 S.E.E. Note also the enormous scatter of the data (see text).*
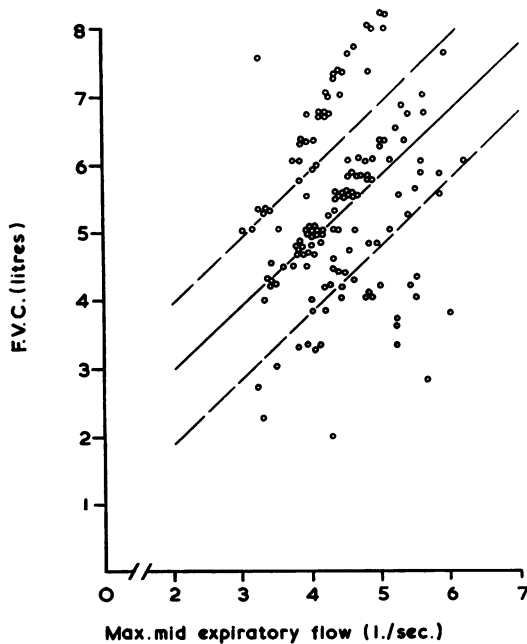

FIG. 2. *Plot of the M.M.F. on the F.V.C. in 155 healthy non-smoking men. The oblique line is a plot of the regression formula: M.M.F.=0·89+F.V.C. (1·004); r=0·572. p=<0·001, S.E.E.=1·028. Broken lines represent 1 S.E.E.*

enormous scatter of the data. Using the predicted F.V.C. rather than the observed F.V.C. Leuallen and Fowler (1955) got 0·44 for r, again suggesting that the scatter obtained in the field is no worse than in the laboratory. Figure 3 is a plot of the first of the two expiratory efforts against the second for three measurements of flow rate—the peak flow rate (P.F.R.), the flow rate at the point when 50% of the F.V.C. had been expelled (50%), and the flow rate at the point when 75% of the F.V.C. had been expelled (75%). Although the maximum difference between the F.V.C.s was 10%, the average difference between each of the flow rate measurements was 20%. The per cent. variation between the first and second efforts for these three flow points is similar (Sobol and Emirgil, 1964), and further work has shown that the variation between two efforts for the M.M.F. is no less than for the P.F.R. (Sobol, 1965).

DISCUSSION

Figure 1 shows the likelihood of an enormous overlap between the normal and the abnormal. Despite this, many asymptomatic individuals are described as normal or abnormal on the basis of ventilatory findings alone.

Bates and Christie (1964) wrote 'a useful general rule is that a deviation of 20% from the predicted normal value probably is significant'. (Since the current discussion is limited to the F.V.C. and measurements of flow rate this would refer to −20% only.) This is probably a fair statement of
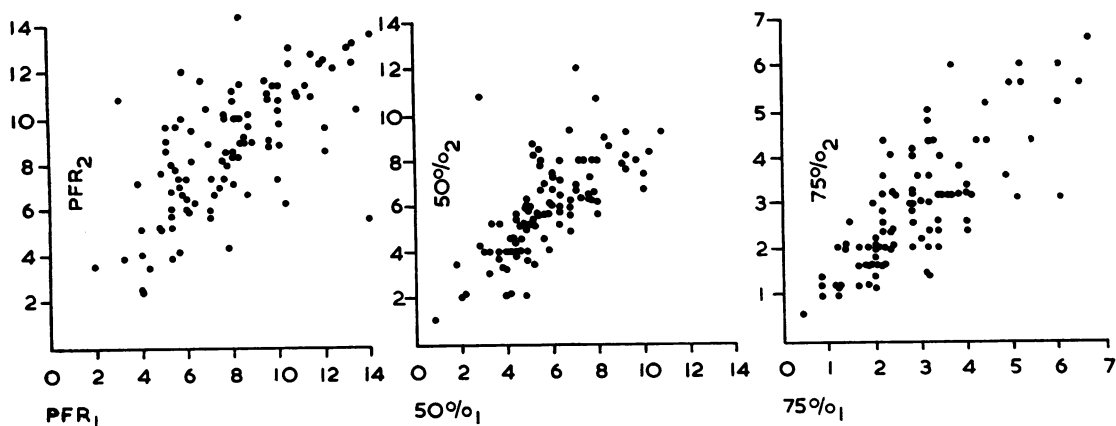
FIG. 3. *Plot of three flow points on an initial expiratory effort against the same flow point on a second effort. The three flow points plotted are described in the text. All values are in litres/second. Note the change of scale in the third graph. When the difference in flow rate between the first and second expiratory efforts is calculated as a per cent. change the difference between the first and second efforts for all three flow points is the same. Note the considerable variation between the first and second efforts.*

the generally used criteria for the assessment of ventilatory function, but there is very little experimental evidence for the validity of such a judgment, which may be fallacious on two counts. First, ventilatory data tend to have a scatter which is homoscedastic rather than heteroscedastic, *i.e.*, the variance of the data around the regression line tends to be uniform. This can be noted in Figures 1 to 3. The scatter is just as wide at the smaller values as it is at the larger values. Obviously a single percentage used as a prediction of normal for all individuals is invalid, since it makes the limits of normal too wide at large values and too small at small values (Fig. 1).

Since the use of a regression formula for any data has as a prerequisite the homoscedasticity of the data (Downie and Heath, 1959), and since the determination of the range of normal for ventilatory data is usually taken from such a formula, the use of a fixed percentage to determine the normal is probably frequently in error.

Despite this statistical prerequisite, the variance of the data around this regression line is not usually discussed by pulmonary physiologists. However, Needham, Rogan, and McDonald (1954) did test for homoscedasticity by first plotting their data. They stated that in adults all the parameters under study (including the vital capacity and maximum voluntary ventilation) complied with this requirement. When scattergrams are given it seems obvious that homoscedasticity is present, not only in measurements of expiratory volume but in

expiratory flow rate as well (Needham, Rogan, and McDonald, 1954; Birath, Kjellmer, and Sandquist, 1963; Ritchie, 1962; Tinker, 1961; Bouhuys, 1963; Hepper, Fowler, and Helmholz, 1960), though the authors do not usually mention it. Oddly enough, this seems to be true for other aspects of respiratory function such as lung volumes (Needham *et al.*, 1954; Hepper *et al.*, 1960) and the relation between oxygen consumption and body weight (White and Alexander, 1965). This is contrary to what we might expect, which is that the per cent. variation rather than the absolute variation would remain constant. Although the observation is clear no explanation is forthcoming.

There is a second reason why the use of −20% as the limit of normal for ventilatory function is probably invalid. In deciding the criteria for the normal we should not exclude any sizable percentage of normal subjects. The rule of −20% should therefore conform to the statistical delineation of normal, which is generally considered to fall within two standard deviations (S.D.) of the mean or within 2 S.E.E. The Table gives in part the data of Leuallen and Fowler (1955) for the M.M.F. as well as the findings of this study. Two S.D. is considerably in excess of 20% of the mean, so that if the 20% criterion were used a large number of presumably normal individuals would be considered abnormal.

There is a further problem in the prediction of the normal range of ventilatory function for a

448   Bruce J. Sobol and Beatrice Weinheimer

TABLE

COMPARISON OF THE M.M.F. IN 66 NORMAL MEN FROM LEUALLEN AND FOWLER (1955) AND IN 155 SUBJECTS IN THIS STUDY

Decade Mean Age	No. of Subjects	M.M.F. Mean (±1 S.D.) Coefficient of Variation		$\dfrac{-2\ \text{S.D.}}{\text{Mean M.M.F.}} \times 100$	
3 27·2 24·3	10 64	4·68 (1·10) 24%	5·50 (1·42) 26%	−47%	−52%
4 34·2 34·3	23 38	4·38 (1·01) 23%	5·27 (1·02) 19%	−46%	−38%
5 44·7 44·4	19 33	4·37 (1·10) 25%	5·21 (1·07) 21%	−50%	−41%
6 54·3 53·5	14 20	3·71 (1·00) 27%	4·80 (1·05) 22%	−54%	−44%

In each column the first set of figures is from Leuallen and Fowler; the second set is from this study. The fact that the scatter of the data from this field study is no greater than that from the laboratory study of Leuallen and Fowler is demonstrated by the fact that the co-efficients of variation are quite similar and, if anything, tend to be lower in this study.

The last column demonstrates how much more than −20% 2 S.D. actually represents. The use of −20% for the delineation of normal for the M.M.F. would include a large number of normal subjects (see text).

particular individual, which is the impressive variability of ventilatory tests on repeated efforts. In this study, although the F.V.C.s on two quick successive efforts differed at maximum by 10%, the average difference between flow rate measurements was 20%. For the M.M.F. the differences varied from 0 to 286%. Forty-one per cent. of the total group had differences between the two M.M.F.s of more than 20%, and only 16% had no difference between the first and the second effort. Leuallen and Fowler (1955) found a standard deviation for the difference between paired efforts in 100 subjects with pulmonary disease to be 0·17 1./min. for a mean value of the corresponding M.M.F.s of 0·89 1./min.—again a rather impressive difference between paired efforts. We can agree that at any one time a single result might fall well below the expected value, whereas later it might well be above that value. These striking differences occur although to the observer a maximal effort was made on both occasions by the subject, with nothing in the record to suggest that one effort was submaximal.

A final problem is the tendency in the laboratory or in the field to use normal values established by others. This is unsatisfactory because these values are affected by variations in technique, equipment, and environment as well as by the selection of 'normal' records. The multiplicity of varying regression formulae for the same ventilatory test

shows the danger of using the values of other workers. In addition, no matter how rigidly the investigator chooses his normal subjects, he will then elect to have the ventilatory performance of some of them fall outside the range of normal. Statistically he usually defines the range of normal as 2 S.D.s or 2 S.E.E.s. In such circumstances approximately 5% of his normal group will be considered abnormal (2·5% if only those below the 2 deviations are considered abnormal), and anyone who uses his criteria must expect to find approximately 2·5% of his normal population with ventilatory abnormalities.

We have shown that the normal individual can deviate below the expected for his group: he can also deviate considerably above it. When such a patient develops pulmonary impairment it may be that as ventilatory function worsens during the course of his disease he will produce apparently normal pulmonary function tests. If the individual has an F.V.C. which falls from +2 S.E.E. to −2 S.E.E., using the S.E.E. of Kory, Callahan, Boren, and Syner (1961), it is conceivable that he might fall from an F.V.C. of 5·2 1. to an F.V.C. of 2·8 1. (his predicted being 4·0), a loss of 46% in his F.V.C. all the while remaining within the bounds of normal. Such a patient would discredit ventilatory function testing by swelling the numbers of false negative tests.

The use of statistics in the study of ventilatory function has its primary place in relating some aspect of ventilation to something which can be measured, such as the establishment of the regression of F.V.C. on age and height. Another proper use is the comparison of ventilatory performance between one group and another. However, the attempt to predict any given subject's performance on a statistical basis is a highly uncertain practice. Correlations which demonstrate statistical relations between physiological phenomena and measurable events are too inexact for prediction in the individual case.

Unfortunately, there is nothing in the armamentarium of routine ventilatory function testing which will compare with the proteinuria of the renal physiologist. Perhaps some test may be developed which is suitable to both field and laboratory use. At the present time assessment of the individual patient is hampered by the necessity for a high degree of co-operation and maximum effort by the subject. Until such time as ventilatory tests have been improved, abnormal findings in a subject without symptoms can be due to either poor performance by him or the innate variability of normal values from subject to subject.

SUMMARY

Simple ventilatory function testing yields results which are too scattered to be useful in the detection of pulmonary abnormality in the individual asymptomatic subject. The parameters evaluated were the forced vital capacity, as a measure of volume, and the maximal mid-expiratory flow. We think this simple ventilatory function testing is too inexact to permit identification of pulmonary abnormality in the asymptomatic subject.

REFERENCES

Bates, D. V., and Christie, R. V. (1964). *Respiratory Function in Disease*, p. 91. Saunders, Philadelphia and London.

Birath, G., Kjellmer, I., and Sandquist, L. (1963). Spirometric studies in normal subjects. II. Ventilatory capacity tests in adults. *Acta med. scand.*, **173**, 193.

Bouhuys, A. (1963). The forced expiratory volume ($FEV_{0.75}$) in healthy males and in textile workers. *Amer. Rev. resp. Dis.*, **87**, 63.

Downie, N. M., and Heath, R. W. (1959). *Basic Statistical Methods*, p. 82. Harper, New York.

Hepper, N. G. G., Fowler, W. S., and Helmholz, H. F. (1960). Relationship of height to lung volume in healthy men. *Dis. Chest*, **37**, 314.

Kory, R. C., Callahan, R., Boren, H. G., and Syner, J. C. (1961). The Veterans Administration-Army Cooperative Study of Pulmonary Function. I. Clinical spirometry in normal men. *Amer. J. Med.*, **30**, 243.

Leuallen, E. C., and Fowler, W. S. (1955). Maximal midexpiratory flow. *Amer. Rev. Tuberc.*, **72**, 783.

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.

Ritchie, B. (1962). A comparison of forced expiratory volume and peak flow in clinical practice. *Lancet*, **2**, 271.

Sobol, B. J. (1965). The maximal midexpiratory flow: a re-examination. *Amer. Rev. resp. Dis.*, **92**, 914.

—— and Emirgil, C. (1964). Subject effort and the expiratory flow rate. *Ibid.*, **89**, 402.

Tinker, C. M. (1961). Peak expiratory flow measured in the Wright peak flow meter: distribution of values in men aged 30–59 who denied respiratory symptoms. *Brit. med. J.*, **1**, 1365.

White, R. I., Jr., and Alexander, J. K. (1965). Body oxygen consumption and pulmonary ventilation in obese subjects. *J. appl. Physiol.*, **20**, 197.