

A COMPARISON OF SPIROMETRIC AND PEAK EXPIRATORY FLOW MEASUREMENTS IN MEN WITH AND WITHOUT CHRONIC BRONCHITIS

BY

A. S. FAIRBAIRN, C. M. FLETCHER, C. M. TINKER, AND C. H. WOOD

From the Medical Research Council, the Postgraduate Medical School, and the London School of Hygiene and Tropical Medicine

(RECEIVED FOR PUBLICATION JUNE 28, 1961)

The earliest disturbance of pulmonary function in chronic bronchitis is obstruction to air flow in the bronchial tree. In this paper we compare the merits of several simple tests for such obstruction.

The tests were made in the course of a survey of Post Office employees whose selection has been described by Fairbairn, Wood, and Fletcher (1959). Comparison between subjects with and without chronic bronchitis is confined to samples of the men, since so few of the women had symptoms of the disease (Fletcher, Elmes, Fairbairn, and Wood, 1959). We examined by spirometry 143 postmen, aged 40 to 59 years, using a light spirometer and kymograph with a paper speed of 2 cm. per second. Three successive tracings of vital capacity were made, with the subject breathing out naturally at his own rate, followed by three tracings of forced expiration. The following indices were derived by measurement of the tracings: (1) forced expiratory volume in 1 second (F.E.V._{1.0}); (2) maximum mid-expiratory flow (M.M.F.) (Leuallen and Fowler, 1955); (3) forced vital capacity (F.V.C.); (4) vital capacity performed slowly (V.C.); and (5) the F.E.V._{1.0} expressed as a percentage of F.V.C. or of V.C., whichever of these two measurements was the greater.

We shall refer to the last index as the F.E.V.%, although Gandevia and Hugh-Jones (1957) first used this expression to mean the F.E.V._{1.0} expressed as a percentage of the F.V.C. The term "timed vital capacity" has been widely used in the United States, sometimes for the absolute value of the F.E.V._{1.0} but more often for the F.E.V._{1.0} expressed as a percentage of the V.C. Capel and Smart (1958) call the latter index the forced expiratory ratio (F.E.R.).

Each measure was the mean of the readings from the three tracings. The tracings were made

in a room, the temperature of which at the time of the survey was approximately constant at 68° F. (20° C.), and the figures presented have not been corrected for temperature or barometric pressure.

After spirometry each man was asked a number of questions by one of the three doctors or one of the three nurses about cough, expectoration, recent chest illnesses, and other respiratory symptoms, using a standardized questionnaire which is given in full by Fairbairn *et al.* (1959). The peak expiratory flow (P.E.F.) was then estimated by a Wright expiratory flow meter (Wright and McKerrow, 1959). The instrument was first demonstrated by the interviewer. After one or two trials to ensure that the subject was performing the test correctly, three readings were taken and the mean of these three was recorded.

We interviewed all subjects twice according to a regular scheme whereby an equal number of subjects was allotted to every possible pair from the six interviewers (Fairbairn *et al.*, 1959). The P.E.F. was measured at each interview, so that two mean readings of this test, each one the average of three, were available for every subject. Spirometry was carried out by an independent person only at the second interview.

The peak flow meter was an early model which gave readings in arbitrary units. We later converted these units into litres per minute by comparison with a properly calibrated instrument. To do this, 19 subjects with peak flow rates ranging from 100 to 800 l./min. each blew six times. Each subject used the two instruments alternately, the choice of the first instrument being randomly determined. The mean readings from each instrument were plotted against each other and a linear regression was found; from this regression the arbitrary units were converted into litres per minute.

DISCRIMINATORY POWER OF TESTS

Our chief interest in the relative merits of the tests lies in their ability to discriminate between normal and bronchitic subjects. To compare the discriminatory power of the different tests, two contrasted groups of men were defined from their answers at the second interview when spirometry was done. The groups consisted of (1) 35 normal men, who denied bringing up any phlegm and also denied having had any chest illness during the last three years; and (2) 17 bronchitic men, who said that they brought up phlegm in the winter, both on rising in the morning and throughout the day, and also said that they had had two or more chest illnesses which had caused sickness absence during the past three years. They may be said to have had "chronic bronchitis with recurrent infection," as described by Fletcher (1959).

The remaining 91 men, who were neither definitely normal nor definitely bronchitic, are not considered in relation to discrimination.

A test is, in general, discriminatory if the difference between the mean readings in the two groups of diseased and normal subjects is large in comparison with the standard deviation of the readings within each group. We have therefore used as an index of discrimination for each test Student's *t*, calculated as in testing the significance of the difference between the mean readings in the two groups of 35 normal and 17 bronchitic men. This quantity is obtained by dividing the difference between the two means by the combined estimate of the standard deviation within the two groups, and multiplying the result by a constant factor which depends only on the numbers in each group and which is therefore the same for each test. Its use in this way was suggested by Gilson, Hugh-Jones, Oldham, and Meade (1955). It was calculated for each test which we studied (Table I).

The values of the P.E.F. obtained at the second interview were used for comparison with the spirometric indices. Two tests can of course be compared by this method only if each value of *t* is derived from readings in the same two groups of subjects.

A quantitative test may be used to classify individuals as either diseased or normal, according to whether their test readings lie above or below some critical value. A test of ventilatory function used in this way would be discriminatory if a high proportion of individuals, each independently diagnosed beforehand as normal or bronchitic, were allotted to these same two groups by means of the test. The greater the degree of separation between the distributions of test readings of two such groups, the higher is this proportion of individuals and the more discriminatory is the test.

This proportion of individuals, correctly separated by the test, can be measured directly as follows. A critical value can be determined, within the range of readings where the two distributions overlap, so that the proportion of readings in the bronchitic group which exceed this value is the same as the proportion of readings in the normal group which lie below it. Conversely, the proportion of normal subjects whose readings exceed this critical value is equal to the proportion of bronchitic subjects whose readings lie below it. This second proportion is a measure of the separation of the two distributions, and may therefore be used to compare the discriminatory power of different tests. When, however, there are few subjects, such a proportion can only take a limited number of values and is therefore a less precise test for discrimination than Student's *t*. Nevertheless, it gives a useful indication of the degree of separation of the two distributions of

TABLE I
DISCRIMINATORY POWER OF TESTS OF VENTILATORY CAPACITY
M.M.F. (l.) Vital Capacity (l.)

	P.E.F. (l./min.)	F.E.V. _{1.0} (l.)	M.M.F. (l.)		Vital Capacity (l.)			F.E.V. %
			Original Readings	Logarithm of Readings	Forced Vital Capacity (F.V.C.)	Slow Vital Capacity (V.C.)	Higher of F.V.C. or V.C.	
35 normal men:								
Mean reading	473.4	2.80	2.625	0.39	3.768	3.811	3.890	73.1
Standard deviation	103.0	0.49	1.010	0.18	0.653	0.603	0.593	8.5
17 bronchitic men:								
Mean reading	274.2	1.55	0.994	1.92	2.646	2.802	2.840	54.5
Standard deviation	125.9	0.59	0.649	0.25	0.636	0.579	0.562	13.4
Difference between means	199.2	1.24	1.631	0.46	1.122	1.009	1.051	18.6
Combined estimate of standard deviation	110.8	0.52	0.909	0.20	0.648	0.595	0.583	10.3
<i>t</i>	6.09	8.10	6.06	7.75	5.86	5.73	6.09	6.10
Percentage separation	71	83		77	83	71	75	77

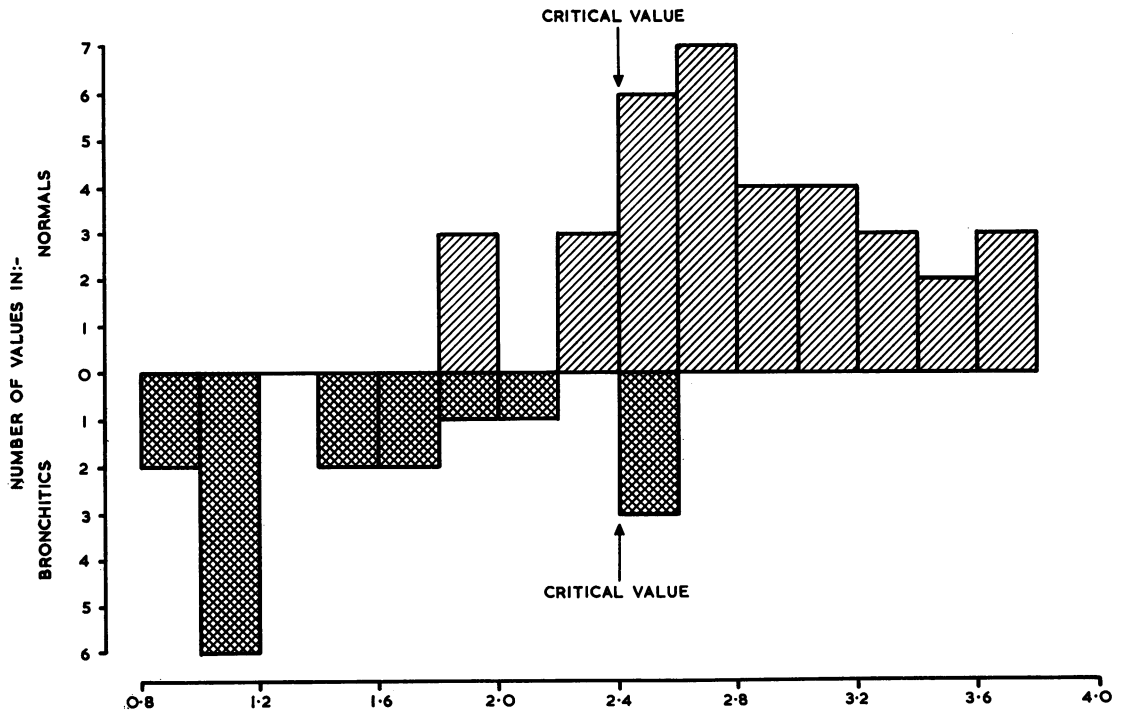


Fig. 1a.—Distribution of values of F.E.V._{1.0} in 35 normal and 17 bronchitic subjects. In normals 6/35=17.1% are below the critical value; in bronchitics 3/17=17.6% are above the critical value.

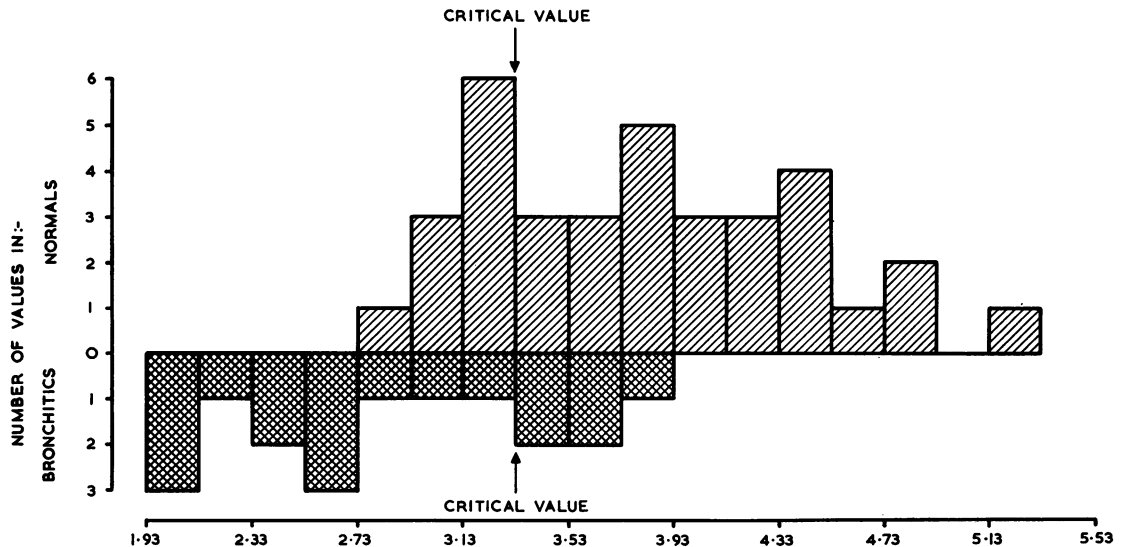


Fig. 1b.—Distribution of values of V.C. in 35 normal and 17 bronchitic subjects. In normals 10/35=28.6% are below the critical value; in bronchitics 5/17=28.4% are above the critical value.

test readings corresponding to different values of *t*, and was calculated as a percentage for each test. This "percentage separation" was greatest (83%) for the F.E.V._{1.0}, which also gave the highest value of *t* (8.1), and smallest (71%) for the V.C., which gave the lowest value of *t* (5.7). Fig. 1 shows, for these two tests, the distribution of readings in the 35 normal and the 17 bronchitic subjects, the extent of the separation between these two distributions, and the critical value used in each case to calculate the percentage separation. The ranking of the other tests by *t* and by the percentage separation showed only rough correspondence.

Greater separation of the two distributions is achieved, and different tests can therefore be more effectively compared, by the choice of two extreme groups such as those which we have defined and by the exclusion of the 91 subjects in the intermediate group. It must be remembered, however, that this choice is deliberately artificial, and that if the whole population of 143 men were categorized as normal or bronchitic by comprehensive definitions, which together included every combination of answers about phlegm and chest illness, the separation of the two distributions of readings for each test would be less than the percentage shown in Table I and Fig. 1.

The percentage separation is determined solely by the ranking of the values and not by their distribution, and is therefore unaffected by simple statistical transformation of these values. If the test values in the population from which the groups are chosen are normally distributed, the percentage separation can be directly estimated, within certain limits due to sampling variation, from the value of *t*, and *t* can in this case be used to compare the discrimination given by different tests. If, on the other hand, the values are not normally distributed, the degree of separation of the distributions cannot be directly estimated from *t*, and it might be misleading to compare the discriminatory value of two tests by using *t*. In this event, normally distributed values can be obtained by a suitable transformation before *t* is calculated.

Two of the tests (M.M.F. and F.E.V.%) showed appreciable departures from normality. An appropriate transformation of the F.E.V.% before calculation of *t* made little difference to its value. The M.M.F. showed a skewed distribution with an excess of high values, and the logarithm yielded a much higher value of *t* than the original value (7.75 instead of 6.05). The discriminatory power of this test would thus be seriously underestimated

if *t* were calculated from the untransformed readings.

The values of *t* found for each test are shown in Table I. The values for F.E.V._{1.0} and for M.M.F. are significantly more discriminatory than each of the remainder. The difference between the P.E.F. and the M.M.F. just falls short of significance at the 5% level.* In particular the F.E.V.% is less discriminatory than the F.E.V. or the M.M.F. Preliminary correction for the small differences in age or sitting height between the groups of subjects made little difference to

TABLE II
VARIABILITY OF PEAK EXPIRATORY FLOW READINGS

A. Variation between Interviews

Subjects		No. of Subjects	Coefficient of Variation (%)
Men seen by:	Any 2 doctors	36	6.4
	,, 2 nurses	36	8.1
Women seen by:	,, 2 doctors	35	7.7
	,, 2 nurses	36	13.3

B. Variation between Successive Readings

Subjects		No. of Interviews	Coefficient of Variation (%)
Men and women seen by	{ Doctors	283	5.4
	{ Nurses	285	6.0

C. Difference between Doctors' and Nurses' Readings in the Same Subjects

Subjects	Mean Excess of Doctors' Readings over Nurses' Readings (l./min.)	Mean Excess Expressed as Percentage of Mean of All Readings in Group at Both Interviews
70 men	43.9	9.6
71 women	27.8	6.0

these results. The value of *t* for the P.E.F. was only slightly greater (6.37) after correction for the observer bias shown in Table II.

VARIABILITY OF MEASUREMENTS OF PEAK EXPIRATORY FLOW

The peak flow meter has recently been used in many surveys, and the variability of the readings merits further discussion. The variation of the mean reading from one interview to another is shown in Table IIA. Each coefficient of variation was derived from a pooled estimate of the standard deviation derived from all pairs of interviews. The coefficients of variation in Table IIB were derived from similar pooled estimates of the standard deviation between triplicate readings at each interview.

*The standard error of the difference between two values of *t*, on the "null hypothesis" that the tests are equally discriminatory, is approximately $\sqrt{2(1-r)}$; *r* is a combined estimate of the correlation coefficient between the paired readings of the two tests in the same individuals within the two groups of normal and bronchitic subjects.

Table IIA shows that the variation between the mean readings obtained by any two doctors at different interviews with the same subjects was less than the variation between the readings obtained by any two nurses: in each case this variation was greater for female subjects than for male subjects. Table IIB shows that the variation between the three readings at one interview was slightly less for doctors than for nurses. Table IIC shows that the readings obtained by doctors were higher than the readings obtained by nurses in those subjects who were seen by both a doctor and a nurse.

These differences between the readings obtained by doctors and nurses were characteristic of all interviewers and were not due to an exceptional performance by one individual. The doctors probably persuaded the subjects to blow harder and rejected more low readings due to faulty technique than the nurses. Their readings would thus correspond more closely to the maximum effort of the subject. This would account for the smaller variability of their readings and their higher average reading in those subjects who had been seen by both a doctor and a nurse. The greater variation found in female than in male subjects accorded with the experience of the observers during the survey. The men were generally competitive and blew into the machine without inhibition. The women, however, needed more persuasion to produce the necessary expiratory effort.

INTERRELATIONSHIPS OF TESTS

The relationships between the different tests are shown in Table III. Correlation coefficients between the readings of each pair of tests have been calculated for all the 72 men who were seen by doctors at the second interview. All these coefficients are statistically highly significant,

TABLE III
RELATIONSHIP BETWEEN TESTS OF VENTILATORY CAPACITY

Correlation coefficients between readings of every pair of tests in 72 men seen by doctors at the second interview

	M.M.F.		F.V.C.		V.C.	V.C.
	F.E.V. _{1.0}	P.E.F.	Original Readings	Logarithm of Readings		
P.E.F.	0.806					
M.M.F.	0.836	0.707				
	Original readings					
	0.895	0.758				
	Logarithm of readings					
F.V.C.	0.835	0.717	0.546	0.639		
V.C.	0.729	0.617	0.488	0.557	0.918	
F.E.V.%	0.774	0.664	0.812	0.895	0.415	0.256

except that between the V.C. and the F.E.V.%, which is only significant at the 5% level. The F.E.V._{1.0} correlates most closely with the other tests, and thus appears to estimate most nearly some characteristic which is common to all of them. Particularly high values of the correlation coefficient are also found between the logarithm of the M.M.F. and the F.E.V.% (0.895) and, as would be expected, between the V.C. and the F.V.C. (0.918). Any departure from normality in the distribution of the readings may affect the magnitude of the correlation coefficient between the tests and so obscure their relationship. The logarithm of the M.M.F., for instance, gave higher values with every other test than the original readings.

The corresponding values of these correlation coefficients for the women were mostly smaller than those for the men. For instance, that between the F.E.V._{1.0} and the P.E.F. was 0.682 compared with 0.806 for the men. This was probably due to the smaller range of disability among the women, so that the errors of measurement were larger in relation to the real differences of ventilatory capacity between the subjects.

The regression of F.E.V._{1.0}(y) on P.E.F.(x) for all subjects seen by doctors was $y=0.00412x+0.622$. This means that the F.E.V._{1.0} in millilitres can be roughly estimated by multiplying the P.E.F. by 4 and adding 600. The observed regression will vary according to the type of population and the age groups chosen for the calculation, and somewhat different figures have been given by other workers (Balgairies, Amoudru, Masure, and Quinot, 1960). The formula which we suggest for the conversion of P.E.F. to F.E.V. appears to work well in hospital practice, but since it depends on a correlation coefficient of only about 0.8 there is a wide scatter about the average F.E.V. predicted in this way.

COMPARISON WITH THE RESULTS OF OTHER WORKERS

Olsen and Gilson (1960) found a significant difference between the ventilatory capacity of men in Bornholm and that of men in two British rural communities, as measured by the F.E.V._{0.75} but not as measured by the P.E.F. Indeed the P.E.F. and the F.E.V._{0.75} ranked the two British populations in a different order of ventilatory efficiency. Our finding that the F.E.V._{1.0} was more discriminatory than the P.E.F. accords with their experience.

Leuallen and Fowler (1955) concluded that the M.M.F. was the most sensitive of a number of

tests of ventilatory capacity in showing abnormality in 100 cases of bronchitis or emphysema. They did not, however, compare this test with the F.E.V. because it is customary in the United States to express the F.E.V. as a percentage of V.C., an index similar to the F.E.V.%. We find, on the other hand, that the discriminatory power of the absolute value of the F.E.V. is at least as great as that of the M.M.F. and exceeds that of the F.E.V.% and that of every other test which we studied.

Our findings on the variability of the P.E.F. and the F.E.V. are compared in Table IV with

TABLE IV

COEFFICIENTS OF VARIATION OF READINGS OF P.E.F. AND F.E.V. OBTAINED BY LOCKHART *ET AL.* (1960) AND IN THE PRESENT STUDY (PER CENT.)

Survey	Lockhart <i>et al.</i>	Present Study	
		London Postmen	Women Sorters
Population	Hospital staff; symptom-free flax workers		
Age group	Not stated; readings standardized for age and sex	40-59	40-59
P.E.F.			
Between subjects	30	27.6	24.0
Between interviews	10	6.4	7.7
Between successive readings	6	5.4	6.0
F.E.V.			
Between subjects	30	26.3	20.5
Between successive readings	11.5	4.3	—

those of Lockhart, Smith, Mair, and Wilson (1960). These workers found a coefficient of variation of 6% between successive readings of the peak flow meter by the same subject compared with our observation of 5.4% in the case of doctors and 6% for nurses. It is remarkable that Lockhart found a coefficient of variation between successive readings of the F.E.V._{0.75} of 11.5%, for this is far greater than our own finding of 4.3% for the F.E.V._{1.0} and that of 1.5 to 3.5% reported by Bovet (1959).

The correlation between the P.E.F. and the F.E.V. for men (Table III) is not quite as high as that reported by Higgins (1957) or by Lockhart *et al.*, but it is of the same order.

DISCUSSION AND CONCLUSIONS

The M.M.F. requires careful measurement of a spiographic tracing of a full expiration and is more laborious than the F.E.V. It is also more sensitive to the degree of completeness of expiration, because this affects the part of the

expiratory curve over which the mean flow is estimated. The F.E.V. is thus the simpler of the two most discriminatory tests of ventilatory capacity which we studied.

The F.E.V. is more repeatable and discriminatory than the P.E.F. because it is less sensitive to the effort which the subject puts into expiration (McKerrow, 1960). Although the exact physiological significance of the minute proportion of expiration which the peak flow meter measures is debatable, and although the peak flow meter is liable to errors of measurement in unskilled hands, its results correlate reasonably well with other indices of ventilatory function (Table III). In large-scale surveys, where convenience and portability are essential, it has been used successfully to demonstrate group differences in the prevalence of respiratory disability (College of General Practitioners, 1961).

SUMMARY

Measurements of peak expiratory flow (P.E.F.) were made by three doctors and three nurses during a survey of respiratory symptoms in London Post Office employees of both sexes. These measurements were repeated after one or two months. On the second occasion spirometric tracings of expirations were made from which measurements of F.E.V._{1.0}, M.M.F., F.V.C., V.C., and F.E.V.% were taken. Significantly high correlations were found between the readings of any two of these tests among the male employees.

The variation between triplicate readings of the P.E.F. was greater at interviews with nurses than at those with doctors, and was greater for female than for male subjects. Doctors obtained higher mean readings of P.E.F. than nurses in those subjects of both sexes who were interviewed by both a doctor and a nurse.

The efficiency of the tests in discriminating between 37 men without and 17 men with symptoms of bronchitis was compared. The F.E.V._{1.0} and the M.M.F. were the two most discriminatory tests.

It is concluded that the F.E.V., which is simpler to measure than the M.M.F., is the most discriminatory test of ventilatory impairment suitable for use in epidemiological studies of chronic bronchitis.

We wish to express our thanks to Sir Walter Chiesman and to Dr. M. C. W. Long, of the Treasury Medical Service, for their goodwill, to the Post Office authorities for giving us permission to carry

out the original survey in 1956, and to the employees and unions concerned for their goodwill. We are especially indebted to Professor Peter Armitage for statistical advice. We would like also to thank many other colleagues for their helpful criticism, and all those who gave computing and secretarial help.

Dr. Fletcher was in receipt of a grant from the Medical Research Council. When the work was carried out Dr. Wood was a member of the scientific staff of the Pneumoconiosis Research Unit of the Medical Research Council.

REFERENCES

- Balgairies, E., Amoudru, C., Masure, P. E., and Quinot, E. (1960). *Rev. méd. minière*, **41**, 24.
- Bovet, J. (1959). *Helv. physiol. pharmacol. Acta*, **17**, 300.
- Capel, L. H., and Smart, J. (1958). *Lancet*, **2**, 771.
- College of General Practitioners (1961). *Brit. med. J.*, **2**, 973.
- Fairbairn, A. S., Wood, C. H., and Fletcher, C. M. (1959). *Brit. J. prev. soc. Med.*, **13**, 175.
- Fletcher, C. M. (1959). *Amer. Rev. resp. Dis.*, **80**, 483.
- Elmes, P. C., Fairbairn, A. S., and Wood, C. H. (1959). *Brit. med. J.*, **2**, 257.
- Gandevia, B., and Hugh-Jones, P. (1957). *Thorax*, **12**, 290.
- Gilson, J. C., Hugh-Jones, P., Oldham, P. D., and Meade, F. (1955). *Spec. Rep. Ser. med. Res. Coun. (Lond.)*, No. 290, p. 41.
- Higgins, I. T. T. (1957). *Brit. med. J.*, **2**, 1198.
- Leuallen, E. C., and Fowler, W. S. (1955). *Amer. Rev. Tuberc.*, **72**, 783.
- Lockhart, W., Smith, D. H., Mair, A., and Wilson, W. A. (1960). *Brit. med. J.*, **1**, 37.
- McKerrow, C. B. (1960). In *Industrial Pulmonary Diseases, A Symposium*, p. 127. Ed. E. J. King and C. M. Fletcher. Churchill, London.
- Olsen, H. C., and Gilson, J. C. (1960). *Brit. med. J.*, **1**, 450.
- Wright, B. M., and McKerrow, C. B. (1959). *Ibid.*, **2**, 1041.