An elaboration of small opacity types (p, m, and n) in simple pneumoconiosis

D. C. LINDARS

Pneumoconiosis Medical Panel, 92/98 Queen Street, Sheffield S1 1WU

Lindars, D. C. (1971). *Brit. J. industr. Med.*, 28, 131-142. An elaboration of small opacity types (p, m, and n) in simple pneumoconiosis. According to the I.L.O. classification (International Labour Office, 1959), radiographs showing pneumoconiosis may be classified as p, m or n according to the greatest diameter of predominant (small) opacities. Recent work has revealed pathological and physiological differences associated with these appearances.

In response to a request for some refinement of the classification for correlation with pathological data a 'pmn elaboration' has been devised, analogous to the N.C.B. elaboration of the I.L.O. classification (Liddell and Lindars, 1969). It has the following form: o/-, o/o, o/p; p/o, p/p, p/m; m/p, m/m, m/n; n/m, n/n, n/A. Instructions to readers are similar to those for use with the N.C.B. elaboration. The significance of o/- and o/o is identical in the two elaborations; n/A indicates predominance of opacity size close to that of large opacities in the I.L.O. classification.

Two hundred and forty-seven radiographs have been read, on two occasions by four film readers, using both elaborations. Analysis of the results showed that the readers had a slightly greater observer error, in terms of variance, when using the pmn elaboration than with the N.C.B. elaboration, but reading bias was less. Calculation of information transmitted showed a gain in information in the pmn elaboration over conventional p, m, n typing, comparable to the gain in information achieved by the N.C.B. elaboration.

Marginal zones were not used as frequently as in the N.C.B. elaboration. The radiograph series contained too few normal or near-normal radiographs for the lower end of the scale to be adequately studied. Improved results can be expected with increased experience and more careful framing of reading instructions for prior briefing.

It is recommended that the pmn elaboration should be used whenever typing is required for correlation between radiographic appearance and pathological or physiological data.

The I.L.O. classification (International Labour Office, 1959) defines both quantitative and qualitative features in chest radiographs of subjects affected by pneumoconiosis. Considered numerically, the opacities of simple pneumoconiosis present a continuum of increasing abnormality in which the classification arbitrarily defines three points, thus dividing the range of abnormality into four categories, 0, 1, 2, and 3. These categories form a useful index of the stage of progression of the disease, and may be correlated with other parameters, such as pulmonary function,

pathology, and dust content of the lungs. It is a disadvantage, however, that only four points are available for such correlations; even so the error between observers in allocating radiographs to the four categories is appreciable, and if further subdivision of the categories is attempted observer error must be increased.

A method of overcoming this disadvantage was devised by Liddell (1963). He used the doubt in the mind of the observer in classifying 'borderline' cases to give a 12-point scale; the observer was asked to

record the I.L.O. category of his choice, followed (after an oblique stroke) by the alternative considered, if any. Thus, for example, category 2 is subdivided into 2/1, 2/2, and 2/3. Experience with this system, which is now known as the N.C.B. elaboration of the I.L.O. classification, has since shown it to be a practicable procedure, not only in the classification of single radiographs (Liddell and Lindars, 1969), but also in the study of repeated radiographs for estimation of progression (Liddell and May, 1966; Rae, 1969). It has several advantages. It increases the sensitivity of the classification from a 4-point to a 12-point scale without imposing any increase in mental effort on the part of the observer. The I.L.O. definitions are retained, as are the standard films. Reproducibility must obviously be reduced, but as the readings are capable of being grouped back into I.L.O. categories there has been no loss of reproducibility compared with the I.L.O. classification. In addition, prevalence rates can be expressed in terms of I.L.O. categories if this should be required.

Qualitatively the I.L.O. classification defines small opacities as pinhead or punctiform (p) if the greatest diameter does not exceed 1.5 mm, micronodular (m) between 1.5 and 3 mm, and nodular (n) from 3 mm to 10 mm, beyond which they are classed as large opacities. The radiograph is classified as type p, m or n according to the greatest diameter of the predominant opacity. There is, therefore, as with the numerical aspect, a continuum of abnormality, this time of increasing opacity size (or increasing predominance of larger sizes in mixed cases) and within this continuum the classification has arbitrarily defined three points. In its practical application the problem of classifying radiographs as p, m or n is similar to that of the numerical categories; it is not possible to base a decision on actual measurement and the process is essentially a subjective assessment by the observer. It is, therefore, attended by similar difficulties regarding observer error, and the level of error has been reported as being of similar magnitude (Blair *et al.*, 1966).

In contrast with the numerical categories, however, the qualitative types are not so clearly related to progression of the disease, with little apparent movement from one type to another. Evidence is accumulating that they represent a difference in pathological response to the inhaled dust. The nodular type is significantly related to an increase in serum rheumatoid factors compared with the other types (Caplan, Payne and Withey, 1962), and can be seen to progress to typical Caplan's syndrome (Caplan, 1953) in a proportion of cases (Davies and Lindars, 1968). In pinhead cases, on the other hand, a reduced gas transfer factor has been found (Englert and De Coster, 1965; Lyons, Clarke, Hall, and Cotes, 1967), and there is pathological evidence

of a closer association with focal emphysema than in the other types (Ryder, Lyons, Campbell, and Gough, 1970).

During the past decade, on account of the significance of the numerical categories as an index of progression, and the need to find a way of controlling the disease in the face of increasing mechanization in coal mines, British work has mainly concentrated on the quantitative aspect. In view of recent developments, however, the typing of radiographs according to opacity size is assuming greater importance. In particular, pathological and physiological correlations require an expansion of the 4-point scale. The present study was undertaken in response to a request for such an expanded scale to be used in further work on the series of Ryder and his colleagues (1970).

pmn elaboration

Although only three types are defined in the I.L.O. classification, opacity size, as pointed out above, in reality consists of a continuum of increasing dimension from nil to 10 mm diameter. The existing scale is thus 4-point (nil, p, m, and n) and is, therefore, analogous to that of the numerical categories. It was decided to treat it in the same way as the N.C.B. elaboration, producing a 12-point scale, as shown in Figure 1.

This system has the advantage that it is readily understood by those familiar with the N.C.B. elaboration, and instructions for its application are equivalent. Readers enter the type of their choice, followed (after an oblique stroke) by the adjacent type considered, if any; otherwise the same type is repeated. o/- and o/o have the same significance as 0/- and 0/0 in the N.C.B. elaboration; n/A indicates predominance of opacities of such a size that it is difficult to decide whether they should be classed as 'small' or 'large' opacities.

Material and methods

The material used in this study comprised 247 radiographs of coal miners or ex-miners, who had died between July 1965 and May 1967, the last radiograph taken in life being used in each case. The subjects are those used in the study of Ryder and his colleagues (1970), where they are more fully described.

The radiographs were read by four readers, members of the Pneumoconiosis Medical Panels who had previously had some experience of using the N.C.B. elaboration. It so happened that each was based on a different centre, geographically far apart, so that it was necessary for the radiographs to be circulated in batches, and discussion between the readers during the course of the study was not possible. An initial meeting of the readers was held beforehand, however, and details of the reading protocol were agreed.

Because the readings were to be correlated sub-

ILO Category 0			ILO Category 1			ILO Category 2			ILO Category 3		
0/-	0/0	0/1	1/0	1/1	1/2	2/1	2/2	2/3	3/2	3/3	3/4
o/-	o/o	o/p	p/o	p/p	p/m	m/p	m/m	m/n	n/m	n/n	n/A
ILO type p			ILO type m			ILO type n					

FIG. 1. Above: The N.C.B. elaboration of the I.L.O. classification. Below: The pmn elaboration.

sequently with pathological data derived from the left lung in each case, an assessment of radiological features in the left lung field only was included in the reading instructions, in addition to assessment of the radiograph as a whole. The readings for the left lung alone are not considered further here. The instructions to readers are given in the Appendix (p. 141).

It was recognized that the element of doubt in classifying the type of a radiograph concerns two distinct problems. In some cases it arises because the size of the predominant opacity is close to the standard diameter demarcating the I.L.O. types; in others, the relative preponderance of differently sized opacities is the source of difficulty. It was decided to distinguish the latter by encircling the alternative type considered (see Appendix).

Although large opacities were to be considered in some detail they are not dealt with here, except that it was necessary to obtain average readings for large opacities (as category A, B, or C) so that category C radiographs, in which the background of small opacities is frequently difficult to discern, could be separated from the remainder in considering results.

The initial meeting was the only briefing in the method of using the pmn elaboration. The radiographs were subsequently divided into four batches and circulated to the readers. After an interval each batch was randomized from lists of random numbers and recirculated.

Results

The 247 radiographs have been classified, on two separate occasions, by each of the four readers, into both the N.C.B. elaboration and the pmn elaboration. Although for certain radiographs some readers felt unable to give a reading, because of either poor radiographic technique or the presence of very large massive shadows, at least one reading was available for every radiograph; in the majority of cases the full number of eight readings were available.

An average reading was obtained for each radiograph by scoring the sub-categories that had been entered, and dividing the sum of the scores by the number of readings available; the resulting score was taken to the nearest whole number and converted to the corresponding sub-category. The scoring system was as follows:

N.C.B. elaboration	0/-	0/0	0/1	1/0	1/1	1/2	2/1	2/2	2/3	3/2	3/3	3/4
pmn elaboration	o/-	o/o	o/p	p/o	p/p	p/m	m/p	m/m	m/n	n/m	n/n	n/A
Score	1	2	3	4	5	6	7	8	9	10	11	12

The average readings obtained in this way have been used as a reference in the analysis of inter-observer error.

An average reading for large opacities (as I.L.O. category A, B, or C) was obtained in a similar manner. Each radiograph was awarded a quality score by adding together the eight scores given by the readers in the following scheme: film quality satisfactory, score = 3; film quality making film reading difficult, score = 2; film quality making film totally unreliable, score = 1.

In the subsequent analysis the readers' performance in classifying radiographs into the pmn elaboration is compared with their ability to use the N.C.B. elaboration. Results are given for the full series of 247 radiographs; for a reduced number (212), after eliminating radiographs with average category C or quality score 13 or less, referred to as 'acceptable films'; and for 147 'good films', with quality score 21 or over, category C also being excluded.

Use of marginal zones

The distribution of readings in the two elaborations is shown in Table 1 for each of the four readers on his first and second readings. Many of those for which no reading was given were unsatisfactory radiographs, so that the proportions allocated to the various sub-categories were not significantly altered in the 212 acceptable films. These are given (as percentages) for the combined readings of all four readers in Table 2. Table 3 contrasts the use of marginal zones (i.e., sub-categories in which an adjacent alternative was recorded) with that of the central zones for each of the four readers.

In the N.C.B. elaboration use of the marginal zones varied between readers from 33% to 46%.

TABLE 1

USE OF ZONES OF THE ELABORATIONS: NUMBERS OF RADIOGRAPHS ALLOCATED TO EACH ZONE
(ALL FILMS)

(a) N.C.B. elaboration

Reader	Reading	0/-	0/0	0/1	1/0	1/1	1/2	2/1	2/2	2/3	3/2	3/3	3/4	No reading
W	1st	—	35	29	14	49	24	22	50	9	9	6	—	—
	2nd	—	55	26	10	62	10	21	44	10	5	4	—	—
X	1st	—	10	3	5	15	19	22	69	34	28	25	—	17
	2nd	—	7	4	2	15	12	21	84	27	21	24	3	27
Y	1st	—	15	13	13	28	21	20	71	12	5	23	—	26
	2nd	—	20	17	12	45	17	24	68	5	4	15	—	20
Z	1st	—	8	5	9	16	15	21	111	17	17	22	1	5
	2nd	1	10	8	3	16	16	24	122	16	9	17	1	4

(b) pmn elaboration

Reader	Reading	o/-	o/o	o/p	p/o	p/p	p/m	m/p	m/m	m/n	n/m	n/n	n/A	No reading
W	1st	—	35	2	1	29	9	7	85	21	15	40	2	1
	2nd	—	55	—	—	28	1	1	106	11	10	35	—	—
X	1st	—	10	1	1	42	34	27	87	8	9	8	3	17
	2nd	—	6	—	2	73	26	22	67	12	7	5	—	27
Y	1st	—	27	—	—	46	8	3	122	2	4	9	—	26
	2nd	—	18	6	2	58	11	6	105	5	4	9	—	23
Z	1st	—	5	4	3	41	32	25	83	16	7	21	5	5
	2nd	1	4	6	4	46	20	25	87	15	6	24	5	4

They were used much less in the pmn elaboration, from 11% to 38%. Only reader Z made greater use of marginal zones in the pmn elaboration. From Table 1 it can be seen that reader W used marginal zones more in the first reading than he did in the

second; reader Y, who made least use of marginal zones, used them more in the second reading.

The proportion of radiographs placed in marginal zones was almost entirely unaffected by film quality. This held for both the pmn elaboration and the N.C.B. elaboration.

TABLE 2

USE OF ZONES OF THE ELABORATIONS: PERCENTAGE OF RADIOGRAPHS ALLOCATED TO EACH ZONE (ALL READERS, 'ACCEPTABLE' FILMS, 1st AND 2nd READINGS)

N.C.B. elaboration Zone	%	pmn elaboration Zone	%
0/-	0.1	0.1	o/-
0/0	7.9	8.0	o/o
0/1	5.5	1.0	o/p
1/0	3.8	0.8	p/o
1/1	13.4	19.3	p/p
1/2	7.6	7.9	p/m
2/1	9.6	6.6	m/p
2/2	32.5	40.5	m/m
2/3	6.8	4.9	m/n
3/2	5.1	3.3	n/m
3/3	7.4	7.0	n/n
3/4	0.3	0.8	n/A

Inter-observer error

For the measurement of inter-observer error two-way tables were constructed comparing the readings of each reader with those of the average readings as reference. Examples are shown in Figures 2 and 3. Table 4 gives the mean deviation and variance for each of these two-way tables. There was a considerable bias between the readers in the N.C.B. elaboration, but mean deviations were much smaller when the pmn elaboration was used. On the other hand, variance was in all cases increased in the pmn elaboration, but for readers X and Z the increase was very small. The relationships were not altered when based only on 'acceptable' or 'good' radiographs.

Intra-observer error

Intra-observer error has been compared similarly, each reader's second reading being compared with his own first readings. Mean deviations of the two-way distributions (Table 4) show a similar range for

TABLE 3

USE OF MARGINAL AND CENTRAL ZONES: PERCENTAGES OF READINGS IN MARGINAL AND CENTRAL ZONES

Films (1st and 2nd readings)	Reader	N.C.B. elaboration		pmn elaboration	
		Marginal %	Central %	Marginal %	Central %
All films	W	39	61	16	84
	X	45	55	34	66
	Y	36	64	11	89
	Z	34	66	35	65
	Average	38	62	24	76
'Acceptable' films	W	41	59	17	83
	X	46	54	35	65
	Y	37	63	11	89
	Z	33	67	38	62
	Average	39	61	25	75
'Good' films	W	39	61	16	84
	X	45	55	33	67
	Y	36	64	11	89
	Z	33	67	34	66
	Average	38	62	25	75

the N.C.B. elaboration and the pmn elaboration. Variance was again greater in the pmn elaboration, particularly in the case of reader W; this reader had used marginal zones much less in his second reading than he had in the first, so that an appreciable variance was to be expected. Again the results were

not altered when confined to 'acceptable' or 'good' radiographs.

Comparison between the elaborations and conventional I.L.O. classification

The improvement effected by the 12-point scale

FIG. 2. Inter-observer error in the N.C.B. elaboration. Reader Z, 'acceptable' films (1st and 2nd readings combined).

FIG. 3. Inter-observer error in the pmn elaboration. Reader Z, 'acceptable' films (1st and 2nd readings combined).

TABLE 4
OBSERVER VARIATION: MEAN DEVIATION AND VARIANCE FROM TWO-WAY DISTRIBUTIONS OF THE ELABORATIONS
(a) INTER-OBSERVER ERROR AND (b) INTRA-OBSERVER ERROR

Reader	Elaboration	Parameter	Inter-observer error					Intra-observer error		
			All films			'Acceptable' films (1st and 2nd)	'Good' films (1st and 2nd)	All films (1st and 2nd)	'Acceptable' films (1st and 2nd)	'Good' films (1st and 2nd)
			1st reading	2nd reading	1st and 2nd readings					
W	N.C.B.	Mean dev.	-1.18	-1.64	-1.41	-1.32	-1.39	-0.45	-0.49	-0.40
		Variance	1.60	2.14	1.87	1.79	1.96	2.15	2.01	2.12
		Mean dev.	+0.37	-0.12	+0.12	+0.24	+0.21	-0.49	-0.43	-0.38
		Variance	3.25	3.58	3.48	3.32	3.25	5.29	4.48	4.62
X	N.C.B.	Mean dev.	+0.92	+0.98	+0.95	+0.97	+1.04	+0.02	+0.06	+0.23
		Variance	1.45	1.40	1.42	1.41	1.34	2.02	2.04	1.86
		Mean dev.	-0.01	-0.38	-0.19	-0.18	-0.16	-0.40	-0.38	-0.38
		Variance	1.59	1.50	1.58	1.64	1.59	2.13	2.20	2.35
Y	N.C.B.	Mean dev.	-0.16	-0.67	-0.42	-0.43	-0.36	-0.52	-0.54	-0.50
		Variance	1.29	1.44	1.43	1.46	1.48	1.72	1.71	1.64
		Mean dev.	-0.32	-0.42	-0.37	-0.38	-0.29	-0.11	-0.09	-0.09
		Variance	2.68	1.95	2.32	2.38	2.51	2.57	2.68	2.57
Z	N.C.B.	Mean dev.	+0.74	+0.57	+0.65	+0.50	+0.40	-0.17	-0.17	-0.11
		Variance	1.46	1.66	1.56	1.37	1.32	1.34	1.39	1.11
		Mean dev.	+0.28	+0.32	+0.30	+0.23	+0.23	+0.02	-0.02	+0.10
		Variance	1.53	1.89	1.85	1.60	1.54	2.22	2.33	2.06

over the conventional 4-point I.L.O. scale can be quantified by the use of the statistic 'amount of information transmitted'. The method of calculation of this statistic, which is expressed in binary digits, is given in Liddell (1963). Values of $n \cdot \log_2 n$ were kindly supplied by the Safety in Mines Research Establishment.

Table 5 gives the amount of information transmitted from the two-way distributions for inter- and intra-observer error. The data for conventional I.L.O. categories and types were obtained by 'grouping up' the sub-categories. The validity of this procedure has been confirmed for the N.C.B. elaboration (Liddell and Lindars, 1969). The expected improvement was found in the N.C.B. elaboration readings and was apparent for all readers in all situations, both for inter- and intra-observer error. A similar increase in information occurred with the pmn elaboration, and again was found in all applications. Even reader Y, who made least use of the pmn elaboration, gained information from such use as he did make of it. Readers X and Z showed a degree of improvement comparable to that accruing from their use of the N.C.B. elaboration.

Effect of technical quality of radiographs

It has already been noted that the use of marginal

zones was virtually unchanged by elimination of unsatisfactory radiographs (Table 3). Tables 4 and 5 show that observer variation was remarkably little affected also. There was a slight trend towards improvement, however, particularly for reader Z, and apparent for both the N.C.B. elaboration and the pmn elaboration.

Widely discrepant readings

The greater degree of variance in observer error of the pmn elaboration is partly due to some widely discrepant readings which have a marked influence on the results. These have been studied for inter-observer error in the first reading only. A difference of four steps or more of the elaboration between the observed reading and the reference (average) reading has been taken as the definition of a widely discrepant reading.

Ten instances occurred in which the average reading for large opacities was category C; this seems to be a sufficient reason for the appearance of wide discrepancies, and these cases are not considered further.

In Table 6 the remaining cases are listed twice, first in relation to the average reading on the N.C.B. elaboration scale, and secondly in relation to average pmn reading. In 9 cases the discrepant readings were in the N.C.B. elaboration only; in 33 in the

TABLE 5

OBSERVER VARIATION: INFORMATION TRANSMITTED FROM TWO-WAY DISTRIBUTIONS OF ELABORATIONS AND 'GROUPED UP' I.L.O. CATEGORIES AND TYPES: (a) INTER-OBSERVER ERROR AND (b) INTRA-OBSERVER ERROR

Reader	Classification	Inter-observer error					Intra-observer error		
		All films			'Acceptable' films (1st and 2nd)	'Good' films (1st and 2nd)	All films (1st and 2nd)	'Acceptable' films (1st and 2nd)	'Good' films (1st and 2nd)
		1st reading	2nd reading	1st and 2nd readings					
W	N.C.B. elaboration	1.14	1.08	0.96	1.00	1.08	1.04	1.05	1.16
	Conventional I.L.O. cat.	0.66	0.55	0.58	0.58	0.61	0.66	0.63	0.70
	pmn elaboration	0.99	1.04	0.96	0.93	1.05	0.84	0.86	0.91
	Conventional I.L.O. type	0.53	0.61	0.56	0.52	0.58	0.62	0.62	0.54
X	N.C.B. elaboration	0.90	1.00	0.94	0.93	1.05	0.79	0.80	0.95
	Conventional I.L.O. cat.	0.65	0.56	0.57	0.61	0.65	0.50	0.50	0.59
	pmn elaboration	0.93	0.91	0.84	0.79	0.86	0.81	0.85	0.97
	Conventional I.L.O. type	0.49	0.42	0.45	0.41	0.49	0.46	0.44	0.47
Y	N.C.B. elaboration	1.37	1.13	1.15	1.17	1.28	1.13	1.26	1.35
	Conventional I.L.O. cat.	0.88	0.78	0.83	0.84	0.91	0.86	0.88	0.96
	pmn elaboration	0.78	0.92	0.80	0.78	0.87	1.01	0.98	1.06
	Conventional I.L.O. type	0.55	0.61	0.57	0.55	0.63	0.77	0.78	0.79
Z	N.C.B. elaboration	0.96	0.92	0.87	0.94	1.04	1.02	1.09	1.14
	Conventional I.L.O. cat.	0.49	0.49	0.50	0.57	0.62	0.69	0.77	0.82
	pmn elaboration	0.94	0.92	0.92	0.96	1.05	0.95	0.95	1.16
	Conventional I.L.O. type	0.50	0.55	0.53	0.56	0.65	0.50	0.47	0.64

TABLE 6

WIDELY DISCREPANT READINGS: NUMBERS OF READINGS, FOUR OR MORE STEPS FROM THE AVERAGE
READING: ALL READERS, FIRST READINGS ONLY

(a) In relation to average N.C.B. elaboration sub-category

Reading widely discrepant in	0/-	0/0	0/1	1/0	1/1	1/2	2/1	2/2	2/3	3/2	3/3	3/4
N.C.B. elaboration only	—	—	—	—	—	1	5	—	1	2	—	—
pmn elaboration only	—	2	5	10	6	6	1	2	1	—	—	—
N.C.B. and pmn elaborations	—	—	3	1	3	4	2	—	—	—	—	—

(b) In relation to average pmn elaboration sub-type

Reading widely discrepant in	o/-	o/o	o/p	p/o	p/p	p/m	m/p	m/m	m/n	n/m	n/n	n/A
N.C.B. elaboration only	—	—	—	—	3	—	5	1	—	—	—	—
pmn elaboration only	—	—	2	6	7	8	8	—	1	1	—	—
N.C.B. and pmn elaborations	—	—	—	2	1	4	2	1	2	1	—	—

pmn elaboration only; and in 13 they were widely discrepant in both systems.

In relation to average pmn readings these three groups were similarly distributed. Except that there were relatively few at the 'n' end of the scale, their distribution was much the same as that for all readings. A different pattern was apparent in relation to N.C.B. elaboration average readings, however. Wide discrepancy of N.C.B. readings occurred mainly in the upper half of the scale: they are likely to be the result of diagnostic differences, and in fact diagnostic comments mentioned congestive changes, tuberculosis, interstitial fibrosis, and rheumatoid lesions in five of these nine cases. The other two groups fell mainly in the lower half, and although many can be explained on a basis of differing diagnosis, it does seem that pmn typing presents greater difficulty at the level of category 1 or below.

Doubt concerning predominance of mixed opacities

There were 95 readings in which the alternative choice symbol had been encircled to indicate doubt regarding predominance rather than size of opacity. All these readings occurred in the four marginal sub-types, p/m, m/p, m/n, and n/m, in which a total of 409 readings was made. The proportion of encircled entries was therefore 23.2%. Table 7 shows that the proportion varied little between these four sub-types. Individual readers, however, varied considerably in the extent to which they made encircled entries (Table 8). By far the largest contribution was made by reader Z. Reader X, whose use of the elaboration was equivalent to that of reader Z, encircled only 13.8% of such readings, against 36.3% by reader Z. The two readers (W and Y) who made much less use of the elaboration also varied in their proportion of encircled entries, accounting for 10.7% and 32.6%, respectively.

TABLE 7

DOUBT CONCERNING PREDOMINANCE OF MIXED
OPACITIES: PERCENTAGE OF ENCIRCLED ENTRIES IN
EACH MARGINAL ZONE

	p/m %	m/p %	m/n %	n/m %	Total %
All films					
1st reading	28.9	21.0	21.3	17.1	23.3
2nd reading	24.1	22.2	18.6	29.6	23.1
1st and 2nd readings	27.0	21.5	20.0	22.6	23.2
'Acceptable' films					
1st and 2nd readings	28.0	21.8	18.3	23.6	23.5
'Good' films					
1st and 2nd readings	32.6	18.3	16.7	27.8	24.4

There was little consistency between readers in the radiographs for which they made encircled entries. The 95 readings represent 69 radiographs: 45 were encircled once only out of eight available readings, 20 twice, and 3 three times.

TABLE 8

DOUBT CONCERNING PREDOMINANCE OF MIXED
OPACITIES: PERCENTAGE OF ENTRIES IN MARGINAL
ZONES ENCIRCLED BY EACH READER

All films	Reader W %	Reader X %	Reader Y %	Reader Z %
1st reading	11.5	16.7	47.1	32.5
2nd reading	8.7	10.4	23.1	40.1
1st and 2nd readings ..	10.7	13.8	32.6	36.3

Discussion

A valid classification of radiological appearance must not only be reasonably reproducible but must also provide information of positive value. Although it may be introduced solely on the grounds that the appearances are present and are reproducible, it nevertheless assumes that the different classes will be linked in some way to stages in the progression of the disease. Validation of the numerical categories was achieved by the demonstration of a relationship (though not a linear one) between the I.L.O. categories and dust analysis of the lung (Rivers, Wise, King, and Nagelschmidt, 1960). They have also been related to the attack rate of progressive massive fibrosis (Cochrane, 1962). The expansion of the I.L.O. scale to the 12-point N.C.B. elaboration has also now been validated by lung dust analysis (Casswell, Bergman, and Rossiter, 1971), and it has been linked to a time scale of continuing dust exposure (Jacobsen, Rae, Walton, and Rogan, 1971).

The separation of p, m, and n types has not been so clearly validated. However, nodular cases have been shown to be more frequently associated with serum rheumatoid factors (Caplan *et al.*, 1962), and lung dust analysis does not correlate so clearly with category as for the other types (Rossiter *et al.*, 1967). A reduced gas transfer factor has been found to be associated with pinhead opacities (Englert and De Coster, 1965; Lyons *et al.*, 1967), and the latter have also been linked with focal emphysema (Ryder *et al.*, 1970). The pmn elaboration proposed in this study remains to be validated; it is significant, however, that in further work on the relationship between opacity size and pulmonary function values Cotes, Deivanayagam, and Field (1971) have found it necessary to introduce two intermediate stages between p and m, effectively utilizing the central part of the proposed pmn elaboration.

An attempt has been made in this study to assess the reproducibility of the pmn elaboration by comparing the film readers' performance with their use of the N.C.B. elaboration. For a number of reasons the degree of observer error was relatively high in this study. It has been shown that observer variation can be reduced by frequent collaboration between readers (Rae *et al.*, 1963) and by the use of standard radiographs (Fletcher and Oldham, 1951). For any particular study it can be reduced by prior briefing (Liddell and Lindars, 1969). The actual level of error also depends on the distribution of abnormality in the series under review. In the present series there were relatively few completely normal radiographs, so that a high level of error was to be expected. Because of their geographical situation the four readers had had no previous opportunity for collaboration, which might have helped them to achieve more identity of interpretation. In retro-

spect it can be seen that briefing could have been more precise, so that some of the wider discrepancies which arose from difference in interpretation might have been avoided. The radiograph series had not been chosen specifically for this study, but for correlation with lung pathology; and the instruction to readers to assess the left lung before assessing the whole film may well have influenced results.

For the comparison of observer error in the two elaborations two-way distributions have been plotted as if the scales were linear, and the linear average has been used as the reference reading for inter-observer error. The average reading is not the true reading but is probably closer to it than an agreed reading arrived at by discussion between readers (Wise and Oldham, 1963). As the two scales are not linear the measurements of mean deviation and variance derived from them must contain a degree of inaccuracy which affects the real value of making comparisons between them. A linear scale can be produced for the numerical categories, based on the number of readings of each category in large representative groups of miners and assuming a normal distribution. The transformation, due to Wise (in Appendix to Rivers *et al.*, 1960), accords with dust and mineral analysis of the lungs and has been found to fit the observed category distributions in several population surveys (Oldham, 1971). Although originally constructed from I.L.O. categories, the N.C.B. elaboration scale has also been successfully fitted to it (Casswell *et al.*, 1971). No such transformation has been put forward to normalize the scale of p, m, and n, which was treated as if it were linear by Rossiter and his colleagues (1967). Treating both scales as linear the observer bias between readers was found in this study to be less for the pmn elaboration than for the N.C.B. elaboration, but variance in both inter- and intra-observer error was somewhat greater.

In assessing the improvement gained by the use of the elaborated scales over the conventional I.L.O. 4-point system a more valid comparison is available by making use of the statistic of information transmitted. The value of the statistic increases as the number of steps in the classification is increased, but is reduced by observer error; it is not affected by inequality in the scale itself, but is dependent on the distribution of abnormality within the scale, so that its maximum value for a particular number of categories is reached only when the distribution is completely uniform throughout the scale, and when the observer variation is zero. The method of calculation is given in Liddell (1963). When comparing the elaboration with the corresponding I.L.O. scale (by grouping together readings in the sub-categories of the elaboration) we have an identical distribution of abnormality. The success of any proposed additions to the existing classifi-

cation can therefore be simply tested by determining if the gain in information arising from the increased number of steps more than off-sets the fall in information due to increased observer error. This is a better method of assessment than the standard of reproducibility suggested by Ashford and Enterline (in Appendix to Blair *et al.*, 1966), because it can be applied to any set of radiographs showing a good range of abnormality.

Gain in information has been consistently found in the N.C.B. elaboration over the I.L.O. categories (Liddell, 1963; Liddell and Lindars, 1969) and was again apparent in the present study. A similar improvement was found for the pmn elaboration, and this was again apparent in all applications. For the two readers who made greatest use of the marginal zones the gain in information was comparable to that arising from their use of the N.C.B. elaboration.

It is not possible to compare information between the two elaborations, however, because of the influence of distribution of abnormality on the statistic. Although the radiographs are the same the parameters being recorded are different; there is no reason to expect that the distribution of opacity size will match that of numerical profusion.

In this connexion it is of interest to note that, as shown in Table 2, the distribution of abnormality between the two elaborations was in fact remarkably similar. If o/- and o/o, which are by definition identical in the two elaborations, are excluded, Spearman's rank correlation coefficient for the remainder yields $R = 0.94$; the likelihood of this occurring by chance is remote (Student's t test, $P < 0.001$). It probably arises because of a general tendency to under-read (i.e., to read less than the true category) radiographs with predominantly pinhead opacities (postulated by Ryder *et al.*, 1970), and to over-read nodular radiographs, as suggested by the work of Rossiter and his colleagues (1967). It might also be explained by assuming that there is progression by increase in opacity size *pari passu* with increasing profusion of opacities. Or it may be that the three opacity sizes, being pathologically distinct, progress at different rates, the nodular type being the most rapid. Further analysis of a series including multiple radiographs of subjects spread over a prolonged period would be necessary to investigate this aspect further.

A source of particular difficulty is the occurrence in some radiographs of p and n opacities in nearly equal profusion. Such cases are likely to be rheumatoid with a fine dust background (Davies and Lindars, 1968). They are relatively few in number. It is no criticism of the pmn elaboration that they cannot adequately be fitted into it; they are equally troublesome to classify by conventional typing. Readings of 'p/n' and 'n/p' were in fact made

on three occasions and have had to be regarded as 'no reading' for analysis, as also a single entry of 'o/m', which illustrates the difficulty of typing opacities at the lower end of the numerical scale, revealed by the analysis of widely discrepant readings. These anomalies should be anticipated in the pre-reading briefing session.

The instruction to encircle the alternative type in cases of doubt regarding preponderance of mixed opacities was so little used that it appears that this aspect is not as important as has been generally supposed (Blair *et al.*, 1966). However, it may be that readers did not remember to carry out this procedure. Even so reader Z, who encircled the most entries, did so in only one-third of cases where he used marginal zones. Whether there is anything to be gained in making the distinction, from the point of view of correlation with other parameters, remains to be elucidated.

It is concluded from this study that an elaboration of p, m, n typing is a practicable procedure which will give increased information when correlations are being made with physiological or pathological data. The radiograph series contained too few normal or near-normal radiographs for the lowest end of the scale to be adequately examined.

It is not suggested that the pmn elaboration should supplant the conventional I.L.O. typing of individual cases for ordinary clinical purposes. (The same might be said of the N.C.B. elaboration.) The elaborations are essentially research tools for use with radiographic series of defined populations. No doubt a similar elaboration of the s, t, u classification of irregular opacities (U.I.C.C. Committee, 1970) would be useful in appropriate circumstances. For the individual case, however, I.L.O. provides (in the words of Dr. Charles Sutherland in his Hubert Wyers Memorial Lecture) 'convenient short-hand descriptions of films' (Sutherland, 1960). The recent revision of the I.L.O. classification (International Labour Office, 1970) seems to have recognized this concept, retaining the previous system but in addition providing an optional extended classification based on that of the U.I.C.C. Committee.

I wish to thank Dr. J. Watkins-Pitchford, Chief Medical Adviser, Department of Health and Social Security, for permission to publish this article. I am indebted to Dr. J. P. Lyons, Senior Medical Officer, Cardiff Pneumoconiosis Medical Panel, for providing the stimulus and opportunity to carry out the study, and the material on which it was based; also to my colleagues, Drs. D. J. Jones, J. B. Menzies, and R. H. S. Pasqual of the Pneumoconiosis Medical Panels at Cardiff, Newcastle, and Glasgow, respectively, for their willing cooperation in film-reading. I was fortunate to have access to the Safety in Mines Research Establishment, Sheffield; and I am grateful to the Director for allowing me free use of facilities in the Establishment, to Dr. A. Critchlow, who

gave me helpful advice and criticism on many occasions, and to Mr. B. A. Maguire, for statistical advice.

References

- Blair, L. G., Cochrane, A. L., Dick, J. A., McCallum, R. I., Rae, S., Rogan, J., Felson, B., Flinn, R. H., Jacobson, G., Milby, T. H. Jnr., Pendergrass, E. P., Ashford, J. R., and Enterline, P. E. (1966). Radiologic classification of the pneumoconioses. *Arch. environm. Hlth*, **12**, 314-330.
- Caplan, A. (1953). Certain unusual radiological appearances in the chest of coal-miners suffering from rheumatoid arthritis. *Thorax*, **8**, 29-37.
- , Payne, R. B., and Withey, J. L. (1962). A broader concept of Caplan's syndrome related to rheumatoid factors. *Thorax*, **17**, 205-212.
- Casswell, C., Bergman, I., and Rossiter, C. E. (1971). The relation of radiological appearance in simple pneumoconiosis of coal workers to the contents and composition of the lung. *Proceedings of the Third International Symposium, Inhaled Particles, London*. Unwin Brothers. To be published.
- Cochrane, A. L. (1962). The attack rate of progressive massive fibrosis. *Brit. J. industr. Med.*, **19**, 52-64.
- Cotes, J. E., Deivanayagam, C. N., and Field, G. B. (1971). The relation between types of simple pneumoconiosis (p, m, and n) and lung function. *Proceedings of the Third International Symposium, Inhaled Particles, London*. Unwin Brothers. To be published.
- Davies, D., and Lindars, D. C. (1968). Rheumatoid pneumoconiosis, a clinical study. *Amer. Rev. resp. Dis.*, **97**, 617-629.
- Englert, M., and De Coster, A. (1965). La capacité de diffusion pulmonaire dans l'antracosisique micronodulaire. *J. franç. Méd. Chir. thor.*, **19**, 159-173.
- Fletcher, C. M., and Oldham, P. D. (1951). The use of standard films in the radiological diagnosis of coal-workers' pneumoconiosis. *Brit. J. industr. Med.*, **8**, 138-149.
- International Labour Office (1959). Meeting of experts on the international classification of radiographs of the pneumoconioses. *Occup. Safety Hlth*, **9**, No. 2, p. 63. I.L.O. Geneva.
- (1970). *International Classification of Radiographs of Pneumoconioses* (Revised, 1968). Occupational Health Series No. 22. I.L.O., Geneva.
- Jacobsen, M., Rae, S., Walton, W. H., and Rogan, J. M. (1971). The relation between pneumoconiosis and dust exposure in British coal mines. *Proceedings of the Third International Symposium, Inhaled Particles, London*. Unwin Brothers. To be published.
- Liddell, F. D. K. (1963). An experiment in film reading. *Brit. J. industr. Med.*, **20**, 300-312.
- , and Lindars, D. C. (1969). An elaboration of the I.L.O. classification of simple pneumoconiosis. *Brit. J. industr. Med.*, **26**, 89-100.
- , and May, Joan D. (1966). Assessing the radiological progression of simple pneumoconiosis. Medical Research Memorandum 4. National Coal Board, London.
- Lyons, J. P., Clarke, W. G., Hall, A. M., and Cotes, J. E. (1967). Transfer factor (diffusing capacity) for the lung in simple pneumoconiosis of coal workers. *Brit. med. J.*, **4**, 772-774.
- Oldham, P. D. (1971). Numerical scoring of radiological simple pneumoconiosis. *Proceedings of the Third International Symposium, Inhaled Particles, London*. Unwin Brothers. To be published.
- Rae, S. (1969). The reduction of observer variation in categorizing coal workers' pneumoconiosis. *Proceedings of the International Conference on Pneumoconiosis, Johannesburg*. To be published.
- , Ashford, J. R., Morgan, D. C., Pasqual, R. S. H., and Pearson, N. G. (1963). A comparison of some alternative procedures in the classification of chest radiographs for coal-workers' pneumoconiosis. *Brit. J. industr. Med.*, **20**, 293-299.
- Rivers, D., Wise, M. E., King, E. J., and Nagelschmidt, G. (1960). Dust content, radiology and pathology in simple pneumoconiosis of coalworkers. *Brit. J. industr. Med.*, **17**, 87-108.
- Rossiter, C. E., Rivers, D., Bergman, I., Casswell, C., and Nagelschmidt, G. (1967). Dust content, radiology and pathology in simple pneumoconiosis of coalworkers. In *Inhaled Particles and Vapours II*, pp. 419-437. Edited by C. N. Davies. Pergamon, Oxford.
- Ryder, R., Lyons, J. P., Campbell, H., and Gough, J. (1970). Emphysema in coal workers' pneumoconiosis. *Brit. med. J.*, **3**, 481-487.
- Sutherland, C. L. (1960). Pneumoconiosis - its effects and complications. *Trans. Ass. industr. med. Offrs*, **10**, 4-15.
- UICC (International Union against Cancer) Committee (1970). UICC/Cincinnati classification of the radiographic appearances of pneumoconioses. *Chest*, **58**, 57-67.
- Wise, M. E., and Oldham, P. D. (1963). Estimating progression of coal-workers' simple pneumoconiosis from readings of radiological categories. *Brit. J. industr. Med.*, **20**, 124-144.

Appendix

Instructions to film readers

Film reading - Cardiff series

Film no. Column 1

Enter the serial film number.

Small opacities Columns 2 to 5

Obscure the right lung field and assess the left lung as if the condition of the right lung had been identical.

Use 12-point (0/- to 3/4) scale of the N.C.B. elaboration for columns 2 and 3.

Enter 12-point pmn elaboration for columns 4 and 5: o/-, o/o, o/p; p/o, p/p, p/m; m/p, m/m, m/n; n/m, n/n, n/A.

The second-choice symbol is intended to express doubt as to which size of opacity exists. If, however, the doubt arises from the *relative preponderance of mixed opacities* (rather than doubt regarding *size* of the predominant opacity) the second-choice symbol (in column 5) should be encircled, e.g., p/m.

Small opacities Columns 6 to 9

Enter in similar manner to columns 2 to 5 for the whole film.

Large opacities Column 10

Enter '-', A, B, C', as appropriate, for the whole film.

Large opacities Columns 11 and 12

The figure to be entered is the *sum of longest diameters* (in one dimension only for each lesion) of the lesions of progressive massive fibrosis in the left lung only (col. 11) and right lung only (col. 12), in centimetres. Enter the number of lesions measured, in brackets, after each entry (in same column).

Diagnostic comments Column 13

If doubt regarding diagnosis exists, e.g., tuberculosis, interstitial fibrosis, enter the possible alternative.

Diagnostic comments Column 14

Significant additional features should be entered in abbreviated form.

Film quality Column 15

Enter A, B or C.

'A' – film quality satisfactory

'B' – film quality making film reading difficult

'C' – film quality making film totally unreliable

Received for publication November 27, 1970.