Assessment of radiological progression of simple pneumoconiosis in individual miners

F. D. K. LIDDELL

Department of Epidemiology and Health, McGill University, Montreal, Canada

Liddell, F. D. K. (1974). *British Journal of Industrial Medicine*, 31, 185-195. **Assessment of radiological progression of simple pneumoconiosis in individual miners.** The studies reported aimed to determine the best method of assessing radiological progression of simple pneumoconiosis in the individual, so that his progression score could be related to other known information about him. The main concern was with subjects for whom three serial postero-anterior chest radiographs were available at approximately quinquennial intervals.

As in other investigations, the 12-point scale of the National Coal Board elaboration led to markedly lower observer error and variability than did the 4-point International Labour Office classification, without distorting levels of progression.

Side-by-side reading led to substantially lower observer error and variability than did independent reading. Although the levels of progression in side-by-side reading were on average a little lower than in independent reading, the effect varied between readers and sessions, being frequently reversed.

Of the three possible methods of side-by-side assessment, the only one without contraindications was that in which all three films for each subject were viewed together, and there were some specific indications for this approach. Viewing only the first and last films led to some loss of information; viewing all three possible pairs was very expensive of time, both in organization and in actual reading, and was not entirely consistent (additive); and disguise of temporal order of the films proved impractical.

It is concluded that the method of choice for assessing progression in the individual from serial films at roughly quinquennial intervals is to view all films together in known temporal order, recording into the most detailed classification available.

Most early radiographic studies of pneumoconiosis were concerned mainly with reproducibility in the reading of single films for simple pneumoconiosis (small opacities) into the available 4-point (or 5-point) classification. Advances have been made in three dimensions: (1) by consideration of radiological change in serial films of individual subjects, e.g., progression of abnormality; (2) by more specific definition of abnormality into several different types; and (3) through the development of finer scales for recording. Table 1 gives a brief summary of developments, which have of course taken place in different dimensions simultaneously, sometimes within the same investigation.

This paper deals mainly with the assessment of progression of simple pneumoconiosis when three serial postero-anterior radiographs are available for each subject (item 6, Table 1). It is a synthesis of findings from four investigations carried out by the National Coal Board Pneumoconiosis Field Research, aimed to find the best method of assessing radiological progression of simple pneumoconiosis in the individual, to allow correlation with dust exposure and other possible aetiological factors. In order to simplify presentation a standard terminology is used (Appendix I) based on the glossary in Liddell and May (1966). In most reported studies of progression of coalworkers' pneumoconiosis the

TABLE 1
DEVELOPMENT OF RADIOLOGICAL STUDIES OF PNEUMOCONIOSIS

Item no.	Films per subject	Features studied	Classifications and number of scale points	References
1	1 ¹	Simple pneumoconiosis ¹	5 points pre-ILO	Fletcher and Oldham (1949)
1a	1	Large opacities	4 points ILO	Blair <i>et al.</i> (1966)
2	2	Simple pneumoconiosis	4 points ILO	Ashford <i>et al.</i> (1965) (including trial A, this paper)
3	2	Progressive massive fibrosis in relation to simple pneumoconiosis	4 points ILO 4 points ILO	Cochrane (1962)
4	1	Simple pneumoconiosis	12 points NCB	Liddell (1963); Liddell and Lindars (1969)
5	2	Simple pneumoconiosis	12 points NCB	Liddell and May (1966); Jacobsen <i>et al.</i> (1971); Trial B, this paper
6	3	Simple pneumoconiosis	12 points NCB	Trials C and D, this paper
7	2	Progressive massive fibrosis in relation to simple pneumoconiosis	4 points ILO 12 points NCB	McLintock <i>et al.</i> (1971)
8	1	UICC features ²	12 points U/C	UICC Committee (1970)
9	5	UICC features ²	50 points McGill	Eysen and Liddell (in preparation)

¹ Usually

² Small rounded opacities; small irregular opacities; and, with fewer scale points, large opacities; pleural thickening; ill-defined diaphragm; ill-defined cardiac outline; pleural calcification

interval between x-ray surveys has been roughly five years. A pair of films taken at such an interval for one subject is here defined as a *quinquennial diad*. Correspondingly, a set of three, four or five serial radiographs of one subject at intervals of about five years forms, respectively, a *quinquennial triad*, *tetrad* or *pentad*.

Liddell and May (1966, p. 6) have discussed possible approaches to assessing progression. For practical purposes only two exist—side-by-side and independent randomized. In the former, all available films for any one subject are assessed at the same time, taking into account all the available information. Several methods of assembling the films are possible, cf. methods 1, 2a, 2b, 2c, and 3 in trial C, below. In the latter, the films for each subject are separated, pooled with those for all subjects, assembled in a single randomized sequence, and assessed.

The first investigation reported (trial A; Ashford *et al.*, 1965) was of quinquennial diads read into the International Labour Office classification (item 2, Table 1). When its findings were being examined two other developments had to be taken into account. First, Liddell (1963) had published his findings (item 4, Table 1) on the advantages of elaborating the 4-point International Labour Office classification into the 12-point National Coal Board scale (Figure). This had been developed for the study of single films, and further evidence in its favour for this purpose is given by Liddell and Lindars (1969).

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

FIGURE. The National Coal Board elaboration of the International Labour Office classification.

Secondly, the National Coal Board had required a means of assessing radiographic progression for use in its Periodic x-ray Scheme. The material for this purpose also consisted of quinquennial diads, and the aim was to obtain for each colliery an index of progression which could be used as a biological control of dust. Liddell and May (1966) found that the National Coal Board elaboration was preferable to the International Labour Office classification and they gave grounds for recommending a method of obtaining a progression index based on side-by-side reading of the diads (item 5, Table 1). The method was in fact adopted and has since been validated by Liddell (1972) who showed that progression indexes correlated with measures of dust exposure in the quinquennium between films.

Trial B was then carried out using most of the

material from Liddell and May (1966), and it was concluded that the National Coal Board elaboration should be used in Pneumoconiosis Field Research. However, further evidence was still required before the field research could select its own reading method. By this time three collieries in that research had been surveyed radiologically on three occasions, so that quinquennial triads were available for a large number of miners. Trial C was therefore undertaken to examine four methods of assessing progression in triads. The results indicated that two methods could be discarded but it was felt that further investigation was required of the remaining two, and this formed trial D. Trials C and D (item 6, Table 1; see also Rae (1970), Rae and Morgan (1970)), are the main bases of the present paper which puts the findings on reading methods into the context of the pattern of development outlined above, and attempts to give a sound basis for further development, particularly that being carried out currently at McGill University.

Here a team investigating the health effects of asbestos in Quebec (see, for example, Rossiter *et al.*, 1972) co-operated with l'Union internationale contre le Cancer (UICC) to produce a classification in which several different types of abnormality were specified and defined. For small opacities, rounded and irregular, the 12-point scale of Liddell (1963) was adopted (as it had been by the International Labour Office (1970)) and the new classification (UICC Committee, 1970; item 8, Table 1) was accepted as the 'extended classification' by the International Labour Office, after certain minor amendments in 1971 (Jacobson and Lainhart, 1972). The current McGill study is of progression in quinquennial pentads in relation to dust exposure and other factors. As well as recording all the UICC features, most of the scales have been extended, small opacities being recorded on 50-point scales (item 9, Table 1). Eyssen (1973) has shown that the methods are feasible, and a report on the radiological aspects of the main study is in preparation.

Materials and methods

Since 1953 medical surveys have been carried out by the Pneumoconiosis Field Research at selected collieries in the British coalfields. Each colliery has been visited at intervals of roughly five years and many of the original population have been radiographed several times. It was originally hoped that measurement of change in the triad could be determined from assessments of each film on its own shortly after the specific survey. However, a major trial (Ashford *et al.*, 1965; Morgan, 1967) showed that fluctuations in reading standards were so large that this method had to be abandoned. Based on a 'definitive' reading procedure which forced each film into a specific category however much disagreement existed between readers (Rae *et al.*, 1963), it also indicated major differences in levels of progression in diads when assessed side-by-side compared with independent readings of the

films. The evidence for individual readers is given here as the results of trial A, i.e., part of stage 3 of the reading of batches A and B in Ashford *et al.* (1965). For two collieries there were available respectively 1 171 and 1 219 diads which had been divided at random into two equal parts. Reader P assessed one part for each colliery, and reader R the other; also reader Q assessed one part for each colliery and reader S the other. The methods of reading were side-by-side and randomized. The International Labour Office (1959) classification was used.

Trial B

Trial B (Liddell, 1965) was based on 157 of the diads already used by Liddell and May (1966); they included 33 pairs of films taken within a few minutes of each other so that true progression could not have taken place. Readers P, Q, R, and S each assessed all films by each of the three following procedures:

- Side-by-side assessment into the NCB elaboration
- Independent randomized reading into the NCB elaboration
- Independent randomized reading into the ILO classification (unelaborated).

Trial C

For trial C, 200 subjects were selected from all men for whom there were quinquennial triads at three collieries different from those in trial A above. The choice was made at random after stratification for dust exposure over the period covered by the triad, within colliery. The selected subjects were then allocated to two equal groups, i and ii, by a further process of randomization within strata. The methods of assessment, all into the National Coal Board elaboration, were:

1. All three films side-by-side	Groups i and ii
2a. 1st and 3rd films side-by-side	Groups i and ii
2b. 1st and 2nd films side-by-side	Group i
2c. 2nd and 3rd films side-by-side	Group i
3. As 2a, 2b, and 2c but with temporal order disguised	Group i
4. Independent randomized	Groups i and ii

Readers P, Q, R, and S each carried out the above programme on two occasions six months apart. A measure of each subject's dust exposure was made available in terms of the number of particles of airborne respirable dust per millilitre of air at each of his working places multiplied by the time spent in each specific environment (Jacobsen *et al.*, 1971). All tables in the present paper are based on both groups, trial and retriad, except Table 4, group i only, and Tables 9 and 10, trial only.

Trial D

In trial D all the available triads (numbering 1 764) from three more collieries, different from those used in trials A and C, were included. As in trial C the subjects were placed at random into two equal groups, i and ii. The methods of assessment, all into the National Coal Board elaboration, were methods 1, 2a, 2b, and 2c of trial C. Readers P and Q each carried out this programme and subsequently repeated the process each on a different 10% random sample of the 1 764 triads. Dust exposures were available as for trial C. All tables are based on both groups trial, except Table 4, group i only, and Table 8, which uses the material from the 10% retrials.

In all these trials, the only radiographic feature of interest was simple pneumoconiosis. Protocols for the various reading procedures are given in Appendix II. Virtually all the categorization was into the National Coal Board elaboration of the International Labour Office classification (see Figure, which also gives a notional scale of abnormality). Assessment of radiological change in the individual was made in terms of steps of progression, each step being from one sub-category to the next or of one point on the notional scale. The term 'level of progression' is used for the arithmetic mean of steps of progression averaged over the relevant subjects and, where appropriate, readers. Justification for these procedures is provided by Liddell (1972).

Findings

Use of NCB elaboration in assessment of progression

Trial B was the first occasion on which the Pneumoconiosis Field Research readers had used the National Coal Board elaboration. The levels of progression (average steps, Appendix I) they obtained with it and with the unelaborated International Labour Office classification in independent randomized readings of the 157 diads were similar, +0.36 (NCB) and +0.40 (ILO), but the range between readers, i.e., inter-observer variability, was much less with the elaboration, 0.21 (NCB) and 0.61 (ILO). Inter-observer error was also substantially less when the elaboration was used. Disagreements of a reader with the average consisting of a single step occurred in 36% of comparisons, of two steps in 10%, and of three steps (the equivalent of an ILO category) in only 2%. This last figure compares with 19% of at least one category when the ILO classification was used. Finally, in the diads where progression was known not to exist, observer error led to 3% of subjects being assessed as showing change by at least three steps when the elaboration was used, compared with 11% showing at least a category of change in the ILO system.

Comparison of paired and independent reading

Table 2 gives levels of progression in diads when the films were viewed side-by-side and independently in trials A, B, and C for each reader separately and averaged. There was a tendency for less progression to be assessed in side-by-side (paired) than in independent reading in each of the rather different trials. However, the tendency was by no means uniform; it was reversed in four out of the 12 comparisons by reader, at least one in each trial. Inter-observer variability (assessed from the ranges quoted in Table 2) was less in paired reading in two of the three trials, markedly so in trial C (and this is not only due to reader Q). Observer error is compared in Table 3: in each comparison the percentages of disagreements were substantially lower in paired reading than in independent reading. This applied to both inter-observer and intra-observer variation in trial C and inter-observer variation in trial B (where the other form could not be assessed).

A further comparison between paired and independent reading comes from the film pairs where progression is known not to exist. In side-by-side reading only 4% of these pairs were assessed as showing change and in each case by only one step; but in independent reading 38% of the pairs were assessed as showing change, the change being by two or more steps in a substantial proportion.

Additivity in paired assessments

Liddell and May (1966, p. 16) pointed out the desirability of a test of additivity in paired assessments. Thus, the steps of progression assessed from the first and third films of a triad excluding the central film (i.e., method 2a of trial C) should be the same as the total of the steps assessed separately in the first and second diads (methods 2b and 2c). Table 4 gives the results of this test, which was included in trials C and D, for each reader and averaged; true addi-

TABLE 2
COMPARISON OF PAIRED AND INDEPENDENT READING:
LEVELS OF PROGRESSION IN DIADS

	Trial A ¹		Trial B		Trial C ²	
	Paired	Independent	Paired	Independent	Paired	Independent
Reader P	+0.20	+0.24	+0.37	+0.36	+0.38	+0.67
Q	+0.10	+0.19	+0.26	+0.36	+0.20	-0.00
R	+0.14	+0.12	+0.30	+0.26	+0.14	+0.18
S	+0.06	+0.18	+0.18	+0.47	+0.26	+0.64
Average	+0.13	+0.18	+0.28	+0.36	+0.25	+0.37
Range	0.14	0.12	0.18	0.21	0.25	0.67

¹ Nett progression (see Appendix I); both collieries

² In decennial diads

TABLE 3

COMPARISON OF PAIRED AND INDEPENDENT READING: OBSERVER ERROR
(Percentages of disagreements: inter-observer—one reader with mean score; intra-observer—
one reader with himself on retrieval)

Amount of disagreement	Inter-observer error				Intra-observer error	
	Trial B		Trial C ¹		Trial C ¹	
	Paired	Independent	Paired	Independent	Paired	Independent
1 step	19	36	8	20	14	29
2 steps	2	10	$\frac{1}{2}$	8	4	11
3 or more steps	$\frac{1}{2}$	2	$\frac{1}{2}$	3	1	9

¹ In decennial diads

TABLE 4

ADDITIVITY IN PAIRED ASSESSMENTS OF PROGRESSION: LEVELS OF PROGRESSION:
GROUP i IN EACH TRIAL

					1st diad	2nd diad	Total	1st and 3rd films	Test difference
<i>Trial C</i>									
Reader P	..	..	..	..	+0.13	+0.19	+0.32	+0.40	-0.08
Q	..	..	..	..	+0.08	+0.22	+0.30	+0.18	+0.12
R	..	..	..	..	+0.03	+0.08	+0.11	+0.16	-0.05
S	..	..	..	..	+0.04	+0.18	+0.22	+0.21	+0.01
Average	..	..	..	..	+0.07	+0.17	+0.24	+0.24	+0.00
<i>Trial D</i>									
Reader P	..	..	..	..	+0.08	+0.07	+0.15	+0.18	-0.03
Q	..	..	..	..	+0.06	+0.10	+0.16	+0.19	-0.03
Average	..	..	..	..	+0.07	+0.08	+0.16	+0.19	-0.03

tivity would lead to zeros in the final column. In trial C additivity was achieved on average, but there was marked non-additivity for the various readers. In trial D, non-additivity was consistent although fairly small; it was also consistent when the three collieries were considered separately varying from -0.01 up to -0.09.

Disguise of temporal order in paired readings

Trial C incorporated the reading of diads with the temporal order disguised. The order was nevertheless recognized on technical or other grounds in a substantial proportion of cases. Summarizing the results, additivity was achieved, and observer variability and observer error with this method were similar to those in other side-by-side procedures. However, the level of progression assessed by this method was substantially less than by any other method. Adequate disguise was impossible in many instances, and the procedures of disguise and of

randomization consumed a great deal of time. As no major advantages were demonstrated over the methods in which temporal order was known, the technique was not considered further for Pneumoconiosis Field Research nor is it used for comparisons in the following paragraphs.

Effects of the central film in assessing triads

Table 5 shows the levels of progression from first film to third assessed from method 1 (all three films side-by-side) and method 2a (first and third films side-by-side). In trial C, the level of progression was consistently higher when the central film was assessed with the others. In trial D, the levels were the same on average, and neither reader obtained markedly different levels in the two methods. There were no marked nor consistent differences between the methods in observer variability (see ranges in Table 5) or in observer error (Table 6).

One other question of interest is whether the

TABLE 5

EFFECT OF CENTRAL FILM IN TRIAD: LEVELS OF PROGRESSION FROM 1st TO 3rd FILM

		<i>Central film assessed (Method 1)</i>	<i>Central film not used (Method 2a)</i>
<i>Trial C</i>			
Reader P		+0.46	+0.38
Q		+0.22	+0.20
R		+0.24	+0.14
S		+0.34	+0.26
Average		+0.32	+0.25
Range		0.24	0.25
<i>Trial D</i>			
Reader P		+0.21	+0.19
Q		+0.18	+0.21
Average		+0.20	+0.20
Range		0.03	0.02

central film in a triad is useful in method 1 in providing an unequal division of the progression over the whole period. Table 7 shows that in every instance considered the amount of progression in the two diads was assessed as different, in some cases markedly so.

Effects of film quality

Table 8 gives the percentages of films assessed during trials C and D as unsatisfactory (stages a to c of Protocol, Appendix II); because of the different

TABLE 6

EFFECT OF CENTRAL FILM IN TRIAD: OBSERVER ERROR IN PROGRESSION FROM 1st TO 3rd FILM

(Percentages of disagreements: inter-observer—one reader with mean score; intra-observer—one reader with himself on retrieval)

<i>Amount of disagreement</i>	<i>Trial C</i>		<i>Trial D</i>	
	<i>Central film assessed</i>	<i>not used</i>	<i>Central film assessed</i>	<i>not used</i>
<i>Inter-observer error</i>				
1 step	9	8	8	8
2 steps	1	$\frac{1}{2}$	4	4
3 or more steps	$\frac{1}{2}$	$\frac{1}{2}$	1	$1\frac{1}{2}$
<i>Intra-observer error</i>				
1 step	15	14	5	8
2 steps	8	4	2	2
3 or more steps	$1\frac{1}{2}$	1	1	1

TABLE 7

EFFECT OF CENTRAL FILM IN TRIAD: LEVELS OF PROGRESSION IN THE TWO DIADS (METHOD 1)

		<i>1st diad</i>	<i>2nd diad</i>
<i>Trial C</i>			
Reader P	..	+0.15	+0.24
Q	..	+0.02	+0.16
R	..	+0.10	+0.06
S	..	+0.08	+0.22
Average	..	+0.09	+0.17
<i>Trial D</i>			
Reader P	..	+0.09	+0.11
Q	..	+0.05	+0.10
Average	..	+0.07	+0.10
Colliery X	..	+0.04	+0.06
Y	..	+0.15	+0.19
Z	..	+0.06	+0.10

numbers of readers in the two trials, comparisons can be made only within each trial. However, the table shows clear differences in most collieries in the quality of films from the three surveys.

To study how technique might affect the findings so far presented each subject was classified as good, medium or poor according as none, one or two, or all three of his films had been assessed as unsatisfactory in the sense of having inferior quality, or showing abnormality interfering with the assessment of simple pneumoconiosis. Not unexpectedly, there was a tendency for good subjects to have less progression recorded and for bad subjects to show more progression; this was slight in trial C but marked in trial D. Again, both inter-observer and intra-observer variability was least for good subjects and greatest for bad subjects in both trials C and D, while observer error was much lower in good subjects and much higher in bad subjects.

TABLE 8

RADIOGRAPHIC TECHNIQUE: PERCENTAGES OF FILMS ASSESSED AS UNSATISFACTORY

<i>Survey</i>	<i>Trial C</i>			<i>Trial D</i>		
	<i>Colliery</i>			<i>Colliery</i>		
	<i>U</i>	<i>V</i>	<i>W</i>	<i>X</i>	<i>Y</i>	<i>Z</i>
1st	18.0	18.3	6.8	19.4	15.1	15.9
2nd	17.4	16.5	11.5	15.1	13.4	11.2
3rd	5.3	8.2	17.5	8.3	11.7	5.2

However, there was no sign that film quality affected the relative findings. In other words, no evidence was obtained from this portion of the study for or against any one method of progression assessment compared with another.

Comparison of methods in identification of progressors

For trials B and C we calculated the coefficients of correlation between the assessments of progression in the individual (mean scores) obtained from each method. In trial B, the three coefficients all lay between +0.76 and +0.78. For trial C, the coefficient for the two side-by-side methods was +0.76, but the coefficients for these two methods each related to independent reading were a little lower at +0.70 (methods 1 and 4) and +0.65 (methods 2a and 4).

Progression in relation to dust exposure and other factors

In trial B each individual's dust exposure was related to his progression as assessed by the various methods of reading his diad. They were all similar, between +0.31 and +0.39, and larger than in previous studies of these methods. However, they were neither large enough for satisfactory prediction nor sufficiently different to allow a choice of method. Those for the triads of trials C and D are given in Table 9. They are seen to be even smaller and to give no assistance in the choice of method.

It will be noted that coefficients of partial correlation which could be calculated for trial D were generally lower than the crude coefficients, i.e., than those calculated without regard to other possibly interrelated factors. Table 10 gives the correlation matrix of pooled (crude) correlations for trial D.

TABLE 9

DUST EXPOSURE AND PROGRESSION IN TRIADS
Coefficients of correlation between (a) exposure to dust¹ in the period covered by the triads and (b) mean score of progression in the individual subject

Method of assessing progression	Trial C (pooled)	Trial D					
		Crude coefficients			Partial coefficients		
		X	Colliery Y	Z	X	Colliery Y	Z
1 All 3 films side-by-side	+0.07	+0.03	+0.29	+0.09	+0.07	+0.18	+0.07
2a 1st and 3rd films side-by-side ..	+0.18	-0.00	+0.25	+0.10	+0.09	+0.11	+0.09
2b, c Adding paired assessments of 1st and 2nd diads	+0.06						
4 Independent randomized	+0.10						

¹ In terms of the number of particles of airborne respirable dust per millilitre of air at each of his working places multiplied by the time spent in each specific environment (Jacobsen *et al.*, 1971)

TABLE 10

INTERRELATIONSHIPS OF VARIABLES
Matrix of correlations (pooled over the 3 collieries of trial D) between age, x-ray status, previous history, exposure to dust, and progression

Variable	Variable				
	1	2	3	4	5
1 Age at time of 1st survey	1	+0.55	+0.27	-0.27	+0.14
2 Years worked at coalface before 1st survey ..		1	+0.39	+0.23	+0.23
3 X-ray status at time of 1st survey			1	+0.02	+0.39
4 Dust exposure ¹ in period covered by the triad ..				1	+0.11
5 Progression (mean score of method 1)					1

¹ See note to Table 9

This shows, not surprisingly, that on average older men had had longer at the coal face (+0.55) and had a higher prevalence of abnormality at first survey (+0.27) but had been less exposed in the period between surveys (-0.27) than younger men. It also shows the expected positive correlation between x-ray status at the time of the first survey and years at the coal face until that time (+0.39). The variable which had the largest correlation with progression was x-ray status at the time of the first radiograph.

Discussion

The trials reported here were searching for a method of radiological assessment of progression of simple pneumoconiosis in the individual so that his progression score could be related to other known information about him. The more important criteria for evaluating reading methods became observer error (lack of reproducibility of progression scores for the individual), the extent to which radiological change was assessed where pathological change was known not to exist, and relationships with dust exposure of individuals. Less important were levels of progression assessed by the various methods, and observer variability (variation inter-observer and intra-observer in levels of progression). The order of importance of these criteria is almost exactly the reverse of those summarized by Liddell and May (1966, p. 18), and indeed their most important criterion is not listed above. This is not a paradox, for the aims of the earlier investigation were different—to choose the most appropriate method of assessing the average amount of progression in a group of subjects such as a colliery population. However, the main findings of the two investigations were very similar.

The evidence presented by Liddell and May (1966, p. 36) strongly favoured side-by-side reading (with temporal order known) into the National Coal Board elaboration; so does the evidence in this paper, which also suggests that all three films of a triad should be assessed together.

Disguise of temporal order was found to be impracticable. Despite great care in the masking of identification for method 3 of trial B, all readers were able to distinguish the earlier and later films in a substantial proportion of pairs. As there was no suggestion that any of the more important criteria were met better when the temporal order had been disguised, this approach cannot be recommended.

The use of the National Coal Board elaboration in the assessment of single films (Liddell and Lindars, 1969) and of diads (Liddell and May, 1966) has not been challenged. From trial A the 4-point International Labour Office (1959) classification was again seen to be too coarse and it was not thereafter used

unelaborated in Pneumoconiosis Field Research.

It has often been taken for granted that progression scores obtained from independent randomized reading are inherently without bias and that such lack of bias is an essential requirement of any measure of progression. Neither assumption is entirely valid. As to the former, should there be consistent differences in quality between earlier and later films of a set (and there were such differences in trials C and D; see Table 8) and should a reader's allowances for technique be inadequate in any consistent way, important biases could arise (Liddell and May, 1966, p. 16); again, imperfect adjustment for the subjects' ages could introduce bias in independent reading. As to the latter assumption, while the same authors stated that they would view with suspicion any method of reading which produced levels of progression differing markedly from those obtained in independent randomized reading, they went on to explain that such differences had to be considered, for their purposes, in relation to observer variability.

In the present paper the requirement is for a score of progression for each miner which has the best balance of sensitivity and consistency without loss of validity, so that it can be related to past dust exposure and other possible aetiological factors. Thus, here too, progression does not have to be unbiased in relation to differences in prevalence. However, it must be emphasized that a statement of **prevalence** at one particular survey cannot necessarily be obtained satisfactorily from assessments made in side-by-side reading.

Levels of progression were lower in paired reading than in independent reading on average in trials A, B, and C, but only reader S was consistent in this regard over the three trials (Table 2). Liddell and May (1966, pp. 24 and 32) showed a similar pattern in their main investigation but a reversal in their preliminary enquiry. Consistent reversal has recently been reported by Amandus *et al.* (1973). Further, the relationship between mean scores of progression obtained for a number of subjects from side-by-side and from independent reading appears reasonably simple. This was illustrated by Liddell and May (1966, p. 86) and is confirmed by the correlations in the trials reported here, which ranged between +0.65 and +0.78 despite dilution by observer variation. Thus scores of progression from paired reading seem to be, in the light of the comments in the previous paragraph, just as good as those obtained from independent reading. However, independent reading when compared with paired reading has poorer observer variability (Table 2) and markedly worse observer error (Table 3). It also indicates much more difference in film pairs when pathological change is known not to exist, and it does not lead to substantially better correlations with dust ex-

posure (Table 9). It is thus less suitable and indeed was abandoned after trial C.

Two methods remained for assessing triads—all three films viewed side-by-side or in separate pairs. The first is conceptually more attractive, being more akin to clinical practice and considerably easier to organize. Thus the use of three separate pairs would be acceptable only if the more important criteria are met more closely by it, and the necessary condition of additivity would also have to be met.

The evidence on additivity is slightly equivocal (Table 4). Levels of progression may be a little less when the central film is not assessed (Table 5); there was nothing to choose between the two methods in terms of observer variability (see ranges in Table 5) and of observer error (Table 6), and correlations with dust exposure favoured the two methods differently in trials C and D (Table 9). Thus there were no indications for reading in pairs and there do not appear to be any contra-indications to assessment of the complete triad side-by-side. The use of only the first and last films for each subject would be a simplification but would involve discarding possibly significant information contained in the central films (Table 7).

The degree of correlation between progression scores in the individual, as obtained by all methods examined, was satisfactory. It was not greatly diluted by observer error when independent reading was one of the methods being compared. This helps to confirm the comparative unimportance of the level of progression in the present context.

The choices of methods were not affected by considerations of radiographic technique. In particular, independent randomized reading was discarded because it did not meet the basic criteria as well as did the other methods, and disguise of temporal order was unworkable.

A final comment is required on the correlations of progression scores with dust exposure. These have been too similar and too low to be useful for comparing reading methods. It has been suggested that higher values might have been obtained if gravimetric measurements of dust concentrations had been available instead of particle counts. However, for all 10 collieries for which exposures have been calculated for these purposes the coefficients of variation between individuals have been very large. Part of the explanation for low correlations may be because exposure throughout the period covered by the triad is not a relevant correlate with progression. Radiological change may arise from exposure which occurred many years before the change appears on the radiograph. Again, correlations will be heavily diluted by differences in such personal factors as ventilation during exertion, initial penetration of dust into the lungs, its elimination, and biological reaction to retained dust.

Certainly the crude correlations (which are all that could be calculated except for trial D) are unsatisfactory and other interrelated factors have to be taken into account. The correlation matrix of Table 10 might suggest that progression arose mainly in those whose earliest films in the triad already showed abnormality, itself due to exposure during coal-face work in the previous years. Correspondingly, Liddell (1966) and McLintock, Rae, and Jacobsen (1971) have reported more progression in diads where the earlier film was abnormal than in those where the earlier film was normal (item 7 in Table 1) and Jacobsen, Rae, Walton, and Rogan (1971) have reported a similar finding in paired reading of decennial diads. These phenomena require careful study and light should be cast on them as a result of the work presently in hand at McGill University. Here, progression has been assessed in quinquennial pentads and is to be related to exposure over various intervals from each subject's entry into the occupational environment.

The work reported here was carried out while I was Head, Medical Statistics Branch, National Coal Board, and all the statistical analyses were carried out within the Branch by Mrs. Janet Gray and her staff to whom I am most grateful. The trials were designed on a co-operative basis and the film-reading was done by the medical officers of the Board's Pneumoconiosis Field Research, Drs. D. C. Morgan, R. S. H. Pasqual, N. G. Pearson, and S. Rae. Dust exposures were calculated from occupational history data obtained in the same research and kindly supplied by Mr. W. H. Walton and his staff. The opportunity has been taken to correct certain inaccuracies and repair some important omissions in previous partial presentations. I have benefited greatly from discussion with many colleagues from the National Coal Board Medical Service and at McGill, and I thank them all sincerely. However, the views expressed here are my own and do not necessarily represent those of any of the above-mentioned persons or organizations; in particular, the generalization in the last sentence of the abstract is a reflection of my own opinion.

References

- Amandus, H. E., Reger, R. B., Pendergrass, E. P., Dennis, J. M., and Morgan, W. K. C. (1973). The pneumoconioses: methods of measuring progression. *Chest*, 63, 736-743.
- Ashford, J. R., Rae, S., Pearson, N. G., Morgan, D. C., and Pasqual, R. S. H. (1965). The measurement of change in the prevalence of simple pneumoconiosis on repeated surveys of coalworkers: a comparison of three methods of reading. *Report PFR/P/263 (revised)*. National Coal Board Medical Service, London.
- 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. Jr., Pendergrass, E. P., Ashford, J. R., and Enterline, P. E. (1966). Radiologic classification of the pneumoconioses. *Archives of Environmental Health*, 12, 314-330.

- Cochrane, A. L. (1962). The attack rate of progressive massive fibrosis. *British Journal of Industrial Medicine*, **19**, 52-64.
- Eyssen, G. E. (1973). *McGill Side-by-Side Reading: Report on the Feasibility Study*. (Privately distributed.)
- Fletcher, C. M. and Oldham, P. D. (1949). The problem of consistent radiological diagnosis in coalminers' pneumoconiosis: an experimental study. *British Journal of Industrial Medicine*, **6**, 168-183.
- International Labour Office (1959). Meeting of experts on the international classification of radiographs of the pneumoconioses, 1958. *Occupational Safety and Health*, **9**, 63-69. ILO, Geneva.
- (1970). *International Classification of Radiographs of Pneumoconioses* (Revised, 1968). Occupational Safety and Health Series No. 22. ILO, Geneva.
- Jacobsen, M., Rae, S., Walton, W. H., and Rogan, J. M. (1971). The relation between pneumoconiosis and dust-exposure in British coal mines. In *Inhaled Particles III*: Proc. int. symposium organized by the British Occupational Hygiene Society, London, 1970, edited by W. H. Walton, pp. 903-917. Unwin, Old Woking, Surrey.
- Jacobson, G. and Lainhart, W. S. (Eds.) (1972). ILO U/C 1971 International classification of radiographs of the pneumoconioses. *Medical Radiography and Photography*, **48**, 67-110.
- Liddell, F. D. K. (1963). An experiment in film reading. *British Journal of Industrial Medicine*, **20**, 300-312.
- (1965) Trial of methods of film reading (Trial 65x1). *Report PFR/P/288* National Coal Board Medical Service, London.
- (1966). Epidemiological assessment of progression: 2nd and 3rd interim reports on methods. *Reports RSC/P(66) 9 and 15*. National Coal Board Medical Service, London.
- (1972). Validation of classifications of pneumoconiosis. *Annals of the New York Academy of Science*, **200**, 527-551.
- and Lindars, D. C. (1969). An elaboration of the ILO classification of simple pneumoconiosis. *British Journal of Industrial Medicine*, **26**, 89-100.
- and May, J. D. (1966). *Assessing the Radiological Progression of Simple Pneumoconiosis*. Medical Research Memorandum No. 4. National Coal Board, London.
- McIntock, J. S., Rae, S., and Jacobsen, M. (1971). The attack rate of progressive massive fibrosis in British coalminers. In *Inhaled Particles III*: Proc. int. symposium organized by the British Occupational Hygiene Society, London, 1970, edited by W. H. Walton, pp. 933-952. Unwin, Old Woking, Surrey.
- Morgan, D. C. (1967). Fortschritte der Staublungenforschung. *Internationale Staublungentagung, Munster*, p. 357.
- Rae, S. (1970). The reduction of observer variation in categorizing coal workers' pneumoconiosis. In *Pneumoconiosis*, edited by H. A. Shapiro, pp. 268-272. Oxford University Press, Cape Town.
- , 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. *British Journal of Industrial Medicine*, **20**, 293-299.
- and Morgan, D. C. (1970). The selection of a film reading technique for the assessment of progression in coal-workers' pneumoconiosis. *Report ISM/P(70) 24*, National Coal Board, Edinburgh.
- Rossiter, C. E., Bristol, L. J., Cartier, P. H., Gilson, J. G., Grainger, T. R., Sluis-Cremer, G. K., and McDonald, J. C. (1972). Radiographic changes in chrysotile asbestos mine and mill workers of Quebec. *Archives of Environmental Health*, **24**, 388-400.
- UICC Committee (1970). UICC/Cincinnati classification of the radiographic appearances of pneumoconioses. *Chest*, **58**, 57-67.

Received for publication 28 November 1973

Accepted for publication 20 February 1974

Appendix I: Definitions of terms used in this paper

Diad (Triad, Tetrad, Pentad): a set of 2 (3, 4, 5) serial films for one subject; the term *quinquennial* is used to indicate the average interval between adjacent films in the set.

ILO classification: the International Labour Office (1959) classification of simple pneumoconiosis which was at that time into four categories (0, 1, 2, and 3).

NCB elaboration of the ILO classification: a 12-point classification of simple pneumoconiosis in which the reader records not only the ILO category but also if he seriously considered an adjacent category (Liddell, 1963; Liddell and Lindars, 1969).

Step: the difference between any two adjacent points on the notional scale of abnormality illustrated in the Figure; a positive step indicates an increase in abnormality and *vice versa*.

Progression score: a measure of the degree of radiological progression assessed in an individual subject, expressed as *steps* (q.v.).

Mean score: for an individual subject, progression scores obtained by a number of readers, averaged over the readers.

Level of progression: the average for a group of subjects of the progression scores obtained by one reader, expressed as *average steps*.

Average level of progression: the average for several readers of their levels of progression for a group of subjects.

Nett progression: the level of progression for a group of subjects, when the ILO classification was used unelaborated and no account was taken of the magnitude of progression or regression; thus each individual is taken as showing +3, 0 or -3 steps of progression.

Observer variability: the variation between readers in levels of progression.

Observer error: the variation between readers in progression scores.

Appendix II: Protocols

There follows the full protocol for the assessment of all three films of a triad side-by-side (method 1) adopted for trial D. That for assessing pairs was closely similar but simpler at stages (d), (e), and (f). These protocols had developed out of those used in earlier trials but the changes, particularly between trials C and D, were small.

Stage (a)

General scrutiny of the three films to ensure that they relate to the same man and to take account of differences in breathing and other technical or pathological factors. It should be appreciated that these may affect the classification of the film and/or the assessment of radiological change; if they do, enter T for each film concerned.

Stage (b)

Consider whether any opacities present in each film are of simple pneumoconiosis or not, and record where diagnosis of pneumoconiosis is doubtful (D) on any film.

Stage (c)

Record other abnormalities (including progressive massive fibrosis) present in each film according to the agreed schedule and record whether or not these interfere with the diagnosis (D) or categorization (C) of simple pneumoconiosis or both (B).

Stage (d)

Recognize whether or not there is radiological change of simple pneumoconiosis (progression, regression or no change) in any pair of films without making any attempt to quantify such change. Recognition will involve scrutiny of (1) first and second films; (2) second and third films; and (3) first and third films. It will lead to one of four possible situations, as follows:

- A Change recognized in none of the three pairs
- B Change recognized in one only of the three pairs
- C Change recognized in two only of the three pairs
- D Change recognized in all of the three pairs

Stage (e)

Decide which film to classify first, as follows:

- Situation A Select the easiest film to classify
 - Situation B Select the easier film to classify from the pair showing change
 - Situation C Select the easiest film to classify from a pair showing change
 - Situation D Select the easiest film to classify
- Place the selected film into the NCB elaboration of the ILO classification.

Stage (f)

Place the two remaining films into the NCB elaboration of the ILO classification as follows:

- Situation A Record for both films the same (NCB) class as for the first film
- Situation B Classify second the other film from the pair showing change. Record 'M' for the third film
- Situation C Classify second a film recognized at stage (d) as showing change from that already classified in stage (e). Thereafter classify the third film; its reading will be the same as one of those already recorded
- Situation D Classify second the film next easiest to classify. Thereafter classify the third film

Classification must *in each case* be in the light of the decisions made at stage (d) (except for the third film in situation B). Thus, if at stage (d) no change (or progression or regression) was recognized in any particular pair of films, the readings of these two films recorded at stages (e) and (f) should be the same (or should be higher on the newer film; or lower on the newer film; respectively).

Stage (g)

Record for each film its classification in terms of diameter of majority opacity, using the notation:

- PP to correspond as closely as possible with 'p' of Blair *et al.* (1966).
- PQ for a film where classification 'p' is seriously considered.
- QQ for a film where neither 'p' nor 'r' is seriously considered.
- QR for a film where classification 'r' is seriously considered.
- RR to correspond as closely as possible with 'r' of Blair *et al.* (1966).