ASSESSING CHANGE IN THE PNEUMOCONIOSES USING SERIAL RADIOGRAPHS

SOURCES AND QUANTIFICATION OF BIAS[1,2]

R. B. REGER, D. F. BUTCHER AND W. K. C. MORGAN

Reger R. B. (Statistics Branch, ALFORD, P.O. Box 4292, Morgantown, W. Va. 26505), D. F. Butcher and W. K. C. Morgan. Assessing change in the pneumoconioses using serial radiographs: Sources and quantification of bias. *Am J Epidemiol* 98: 243–254, 1973.—The assessment of progression in coalworkers' and other pneumoconioses by comparing serial radiographs is a standard procedure. Despite general acceptance of the principles involved, several sources of bias may be inherent in this practice. This study, designed to identify and quantify the various biases, shows that the assumed chronologic sequence of the paired radiographs, viz. the position, as well as the disease stage existing for a given pair member, can significantly influence the categorization process. It demonstrates that there is marked reluctance to diagnose regression of pneumoconiosis. Although one can not hope to eliminate totally the biases in a side by side reading method, partial remedies are available. Recognition that such biases exist is a start. It is desirable to conduct trials with various readers to ascertain those who are prone to "pull" side by side interpretations towards the middle (i.e. the "concertina" effect) resulting in a lower progression score and those who are oppositely inclined. Side by side readings which differ drastically from those determined by independent interpretation should be viewed with suspicion. Variability between readers will perhaps always be a problem. However, rigid initial training, as well as periodic checks to insure that reading habits do not change, would do much to improve the validity and reliability of the categorization of pneumoconiosis.

coalworkers; pneumoconiosis; radiography, serial films, variability and bias

INTRODUCTION

The diagnosis of coalworkers' pneumoconiosis is dependent on the appearance of

certain fairly distinctive features in the chest roentgenogram in association with a history of coal dust exposure. Furthermore, the severity of the condition is assessed by the extent, profusion, and size of the opacities observed on the film. Accordingly, the chest radiograph is still the only practical method of diagnosing coalworkers' pneumoconiosis in life. Justification for this practice lies in the fact that a definite relationship between the extent and profusion of radiographic opacities and coal dust content of the lungs exists (1).

In Great Britain, the United States, and elsewhere, progression of coalworkers' pneumoconiosis is determined by comparing serial radiographs. In most countries, the readings have a dual purpose. First, it is the responsibility of those performing the work to inform the individual examined of any condition which might affect his health and as such this must be a foremost consideration. Secondly, the readings from the serial films are used for research purposes where inductive logic is used to infer from the part to the whole. The characterization of the relation between radiographic disease and various other factors such as the mine environment, pulmonary function, and symptomatology are examples of such research.

The advantages as well as the drawbacks of reading serial films have been discussed indirectly in several articles in the literature (2–4). The optimal method of determining progression has been a subject of controversy for many years with many of the arguments, both pro and con, remaining unpublished. Nevertheless administrative considerations, usually involving personnel and finance, generally dictate the system to be used. Regardless of the method used, it must be administered in such a fashion that it does not interfere with the prompt delivery of medical care to the participating subject.

Significant bias may be inherent in the determination of progression of pneumoconiosis from the comparison of serial radiographs if the dates on which the films were taken are known to the reader. Whether a particular film occupies the left or right hand position in a pair and hence an implied chronologic sequence, might influence the categorization process. In addition, the category of one film of the pair might exert an influence on the reading of the other film. For example, when a film is paired with an identical film taken on the same day or failing this a copy of itself, is it likely to be given the same reading as when it is paired with another film with an obviously higher or lower category? Further-

more, is the above effect consistent or is it related to the assumed chronologic sequence as inferred from the left or right hand position of the film? Overall effects between readers as well as readers interacting with other variables were also areas of prime concern. Implicit in the above questions is the need to determine if there is reluctance to read regression.

Methods and design

The ideal circumstances for this study would have been the availability of four radiographs on each of several subjects taken over a time interval such that some amount of progression was evident. As such, each subject's films could be arranged in pairs so that every permutation as well as duplicate pairing could be exhibited in a 4 × 4 design. All pairs would then reflect the most extreme progression and regression of the disease as well as all intermediate points.

Unfortunately, this was not possible because our program had not been in existence long enough to have four serial films per subject. In lieu of this, several pairs of radiographs with varying but definite degrees of progression of disease were obtained. Next, an attempt was made to match the pairs in terms of the bony architecture of the chest, lung size, and contours of the other intrathoracic organs so that readers would not suspect they were reading paired films of different subjects. We therefore contrived a situation in which, although we had two pairs of films from two separate subjects, all four were treated as if they were made from the same subject. Eight copies of the four films were obtained and the 32 copies were arranged into all 16 possible pairings of four films. Readers were told, or they assumed from habit, that the earlier film of each pair was always on the left. Identification on all films was hidden from view. The general design is shown in figure 1. The films were arranged in order of severity; 1 showing the least disease and 4 showing the most disease. The numerical identification

ASSESSING CHANGE IN THE PNEUMOCONIOSES 245

in the cells shows which films make up the various pairs in the design. For example, cell 11 contains Film 1, viz. the film of the set of four showing the least disease paired with itself. In cell 12, Film 1 is in the left hand position while Film 2, viz. the film of the set of four showing more radiographic disease than Film 1 but less than Film 3, occupies the right hand position of the pair. The same two films are involved in cell 21; however, their position is reversed. This general procedure in arrangement of pairs of films was used throughout the design. Accordingly, progression of coalworkers' pneumoconiosis is present for those pairs above the principal diagonal while regression is present below the diagonal. No change exists on the diagonal since this is the point where each film is paired with itself. Three distinct sets of four films each

SECOND FILM OF A PAIR

	1	2	3	4
1	Pr 1 / 11	Pr. 2 / 12	Pr 3 / 13	Pr 4 / 14
2	Pr 5 / 21	Pr 6 / 22	Pr. 7 / 23	Pr 8 / 24
3	Pr. 9 / 31	Pr 10 / 32	Pr 11 / 33	Pr. 12 / 34
4	Pr. 13 / 41	Pr. 14 / 42	Pr 15 / 43	Pr. 16 / 44

FIRST FILM OF A PAIR

FIGURE 1. Design of serial film reading trials. This design involves additional dimensions across the separate and distinct film sets as well as across readers.
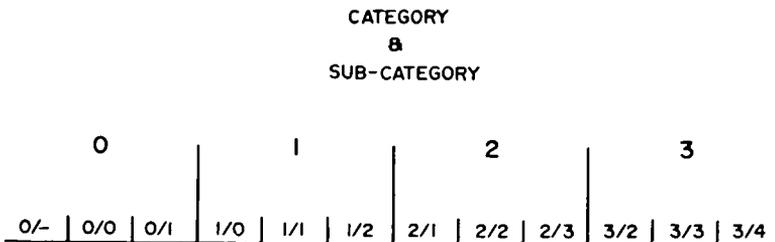
were obtained and each set was then arranged according to the general design. All pairs were read by three physician interpreters. Thus, there were 16 possible pairs of films to be read in each square. A total of 288 readings was obtained; 32 films by three squares by three readers. Two of the physician interpreters were experienced in interpreting films for coalworkers' pneumoconiosis while the third was a neophyte. All paired films were submitted to the readers in a random order. They were read according to the Union Internationale Contre le Cancer (UICC)/Cincinnati Classification System (5) which is based on and closely resembles the National Coal Board elaboration of the 1959 International Labour Organization (ILO) Classification (6) (figure 2). Briefly, the classification system is an expandable four point scheme: Category 0 representing no pneumoconiosis, and Categories 1, 2, and 3 representing stages of simple pneumoconiosis. The expansion of these major classifications, are those designations separated by a slash. The first numeral identifying the subcategory represents the first impression (i.e. major category) assigned by the radiologist while the second numeral represents his judgment of severity within the major category. Thus if a film is interpreted as Category 2/1, the reader eventually decided that the film was Category 2; however, he also seriously considered Category 1. We assumed that the readings represent stages on a continuum which starts with no pulmonary dust deposition and ends with maximal dust deposition and that each step represents an equal

CATEGORY
&
SUB-CATEGORY

0			1			2			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 2. UICC or ILO elaboration of the classification of simple coalworkers' pneumoconiosis.

Downloaded from http://aje.oxfordjournals.org/ at Stephen B. Thacker CDC Library on March 10, 2016

increment of dust deposited. Thus, parametric statistical procedures offer a method of obtaining reasonable estimates of interactions between film categorization and other factors. The base reading (0/−) was given the value 1. The next reading (0/0) was

given 2, and so forth to the ultimate reading (3/4) which was assigned the value 12.

In addition to the criteria mentioned previously, the radiographs were chosen because previous independent readings by several radiologists had revealed unanimity or near unanimity regarding the diagnosis. The independent readings are given in table 1.

Several analytical methods were utilized in testing the various hypotheses, namely, an analysis of variance, the least significant difference (LSD) test, and an F-max (7, 8).

RESULTS

From the analysis of variance (table 2), it becomes obvious that nearly every effect tested was significant. It is quite clear that classification is affected depending on whether a film occupies the left or right hand position. Figure 3 shows the effects of position across all films averaged over readers and squares (quasi replicates). The category assigned to a film in the left hand position of a pair was markedly lower than

TABLE 1

Results of previous independent readings

Film*	Readings		
	Square 1	Square 2	Square 3
1	1/0	0/0	0/0
2	2/1	1/1	1/0
3	2/3	2/1	1/2
4	2/3/3/2†	3/2	2/2

* The film identification (1 through 4) is the same for all squares in the total design. However, there are 12 distinct films involved in the study. Hence, film₁ is not the same from square to square.

† Represents previous readings of 2/3 by two radiologists and 3/2 by two radiologists.

TABLE 2

Analysis of variance

Source of variation*	Degrees of freedom	Sum of squares	Mean squares	F ratio
Squares	2	126.69444444	63.3472	78.78†
Films	3	804.86111111	268.2870	333.65†
Position (film)	4	34.97222222	8.7431	10.87†
Pair (left position)	12	27.16666667	2.2639	2.82†
Pair (right position)	12	102.49999999	8.5417	10.62†
Square X film	6	72.05555556	12.0093	14.94†
Error (a)	56	45.02777777	0.8041	
Reader	2	15.19444444	7.5972	13.15†
Square X reader	4	17.30555556	4.3264	7.49†
Reader X film	6	16.55555556	2.7593	4.77†
Reader X position (film)	8	3.69444444	0.4618	0.80
Reader X pair (position)	48	35.66666667	0.7431	1.29
Reader X square X film	12	13.52777778	1.1273	1.95‡
Error (b)	112	64.72222222	0.5779	
Total	287	1379.94444444		

* Different partitions of the various effects as well as single degree of freedom contrasts are available upon request.

† p < .01.

‡ p < .05.

when it occupied the right hand position. The sum of squares for left versus right hand positions was 31.34 and is not included in table 2. However, it does represent another possible and important partition of the position within film sum of squares. Figure 4 shows that the effect of position was not consistent across different classes of films. It can generally be said that a line of demarcation exists in that those films with the least disease were read much higher when they occupied the right hand position while those films with more positive evidence of disease were read about the same in both positions. For films showing minimal disease, the effect due to position within a pair produced a disparity of around one minor UICC category. On the other hand, for those showing more disease, the disparity relative to position of the film in a pair was only around one-third of a minor category. The four degrees of freedom in table 2 associated with the position (film) effect, when broken down into single contrasts by film class show this quite clearly. For example, the sum of squares for the individual contrasts were: Film 1, 14.22; Film 2, 16.06; Films 3 and 4, 2.35 each.

The effect on categorization of a particular radiograph when it is paired with itself as compared to when it is paired with a higher or lower category film is shown in figure 5. Films paired with themselves are categorized somewhat lower than when paired with other films. The difference, although statistically significant, is of such small magnitude as to be negligible from the viewpoint of clinical management. However, two other inferences are worth consideration. First, which films cause the disparity? Figure 6 shows the dissimilarity by class of film averaged over readers and squares. It is immediately evident that the films showing the least disease account for nearly the entire trend. They are read significantly higher when paired with other films. Virtually no divergence exists for
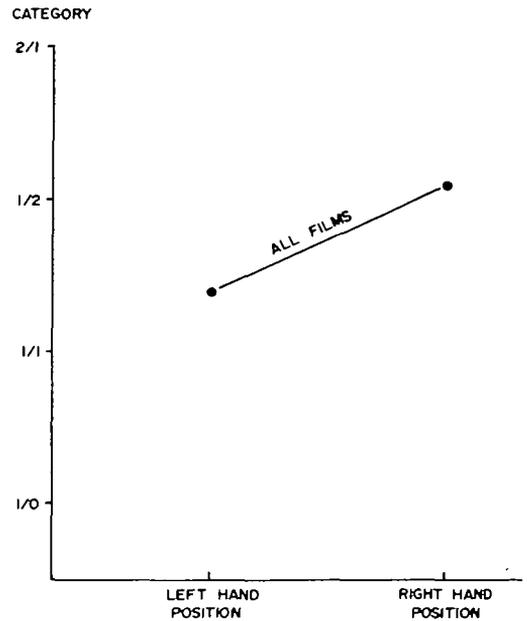

FIGURE 3. Effect due to position (averaged over readers and squares).
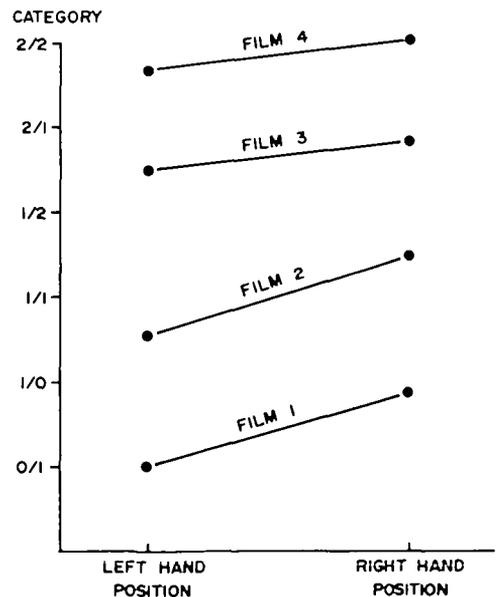

FIGURE 4. Effect due to position by film class (averaged over readers and squares).

those films designated as 2, 3, or 4. The sum of squares for the contrasts portrayed in figures 5 and 6 are not included in table 2. Again, this represents another possible par-
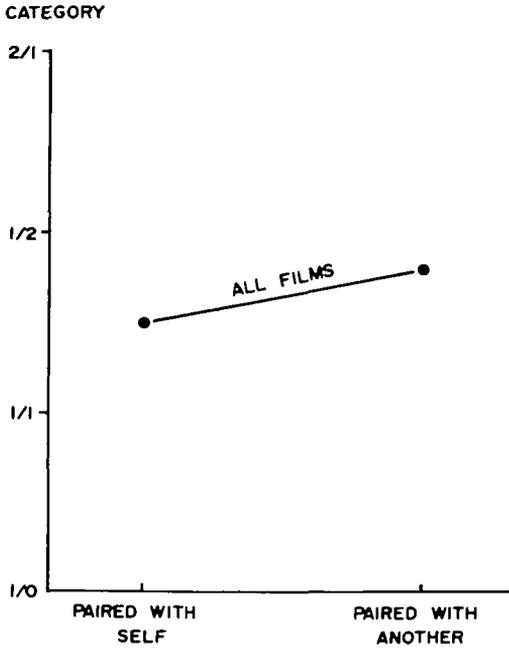
CATEGORY

FIGURE 5. Effect due to pairing (averaged over readers and squares).
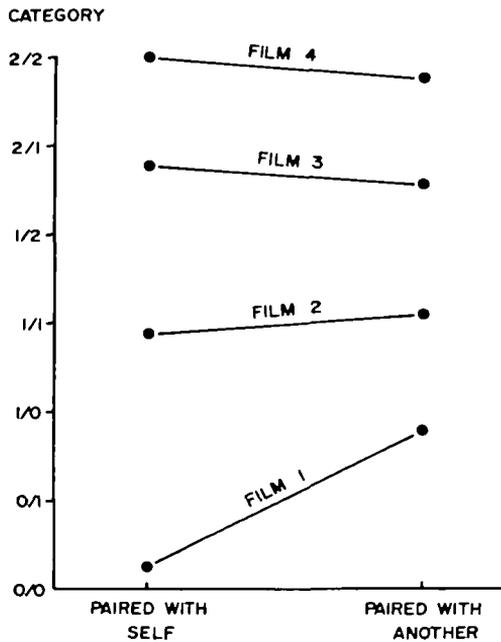
CATEGORY

FIGURE 6. Effect due to pairing by film class (averaged over readers and squares).

tition in the analysis of variance. Secondly, is the disparity consistent regardless of the position of the pair? From the analysis of variance, an obviously different effect occurs depending on whether one is considering the left or right hand member. The effect, although significant for both the left and right hand pair member, is far from consistent. This is evident from the magnitude of the F statistic in the analysis of variance, and also from figure 7. Note the appreciably greater slope of the line for the right hand pair member.

In an attempt to elaborate on the effect of film position on categorization, the subset of films which were paired with themselves were analyzed, thereby holding the disease stage constant. The difference between the left and right hand pair members for this subgroup was only around one-third of a minor UICC category. This is in contrast to a markedly higher discrepancy shown in figure 3 where all films were used and a difference of one minor UICC category was evident.

If each of the trend lines in figure 7 is separated into finer components involving severity of disease, then the trends involv-

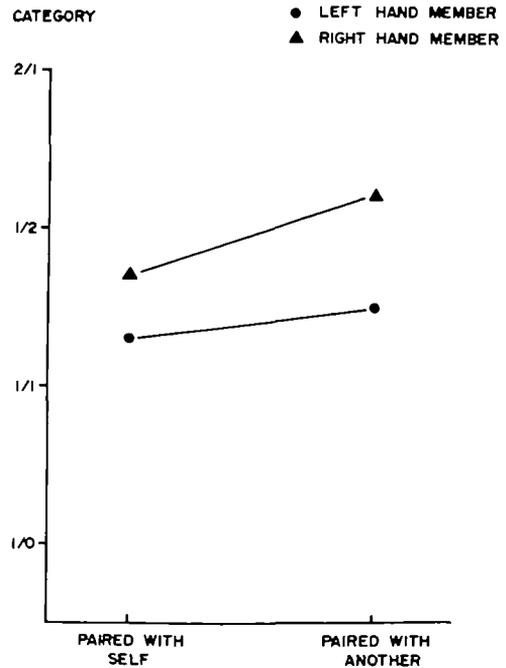CATEGORY   ● LEFT HAND MEMBER   ▲ RIGHT HAND MEMBER

FIGURE 7. Effect due to pairing by position (averaged over readers and squares).

ing more restrictive means come to light. These are shown in figures 8 and 9. The overall trend for the left hand member is almost, if not wholly, influenced by film class 1, i.e. the films with the smallest amount of disease, averaged over readers and squares (figure 8). In contrast, the overall trend for the right hand member (figure 9) is influenced, to some degree, by the stage of the disease in the film. However, films with the least disease seem to have the greatest effect. Moreover, films with higher stages of disease seem to have a slight effect in the opposite direction. Eight of the 24 degrees of freedom shown for the pair (left and right position) in table 2 are portrayed in figures 7, 8 and 9. Table 3 shows the sum of squares for the eight separate contrasts shown in figures 8 and 9.

It is well documented that appreciable interobserver variation exists in the categorization of coalworkers' pneumoconiosis. Surprisingly, the overall reader effect in these trials was somewhat less than has been shown elsewhere. The overall means
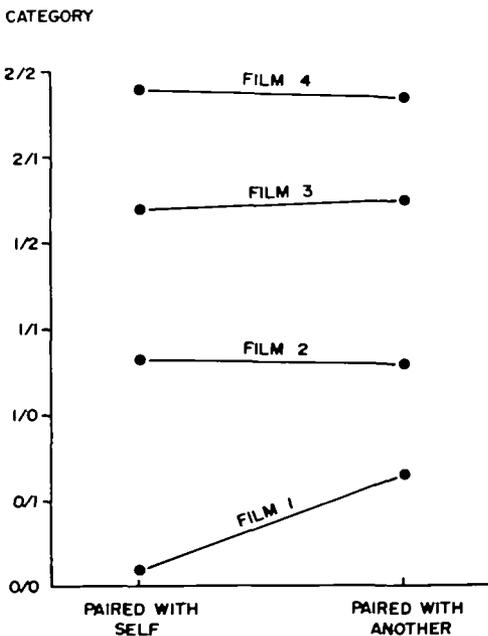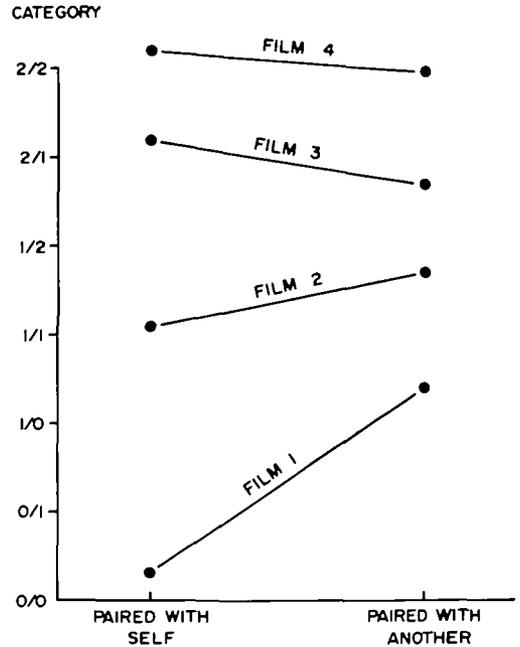

FIGURE 9. Effect due to pairing by film class (averaged over readers and squares). Right hand pair member.

TABLE 3

Sum of squares for single degree of freedom contrasts

	Left hand pair member	Right hand pair member
Film 1	7.26	29.04
Film 2	0.08	2.08
Film 3	0.04	1.56
Film 4	0.08	0.33

for the three readers are shown in figure 10. Although the contrasts may be considered real, it is questionable whether such minor differences have any practical medical importance.

As might be expected, highly significant differences existed between squares as well as between films. Indeed, the study was designed with this in mind, and these effects in and of themselves, are of little concern. However, their interactions with other main effects, e.g. readers, should be considered. Thus, figure 11 shows the relationship between reader and square, square being the quasi replicate of the overall 4 × 4 design.

FIGURE 8. Effect due to pairing by film class (averaged over readers and squares). Left hand pair member.
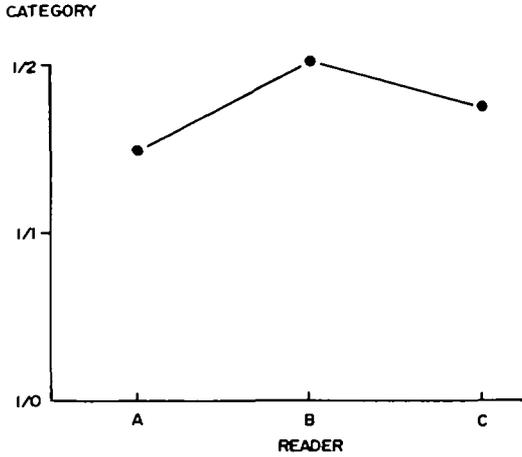
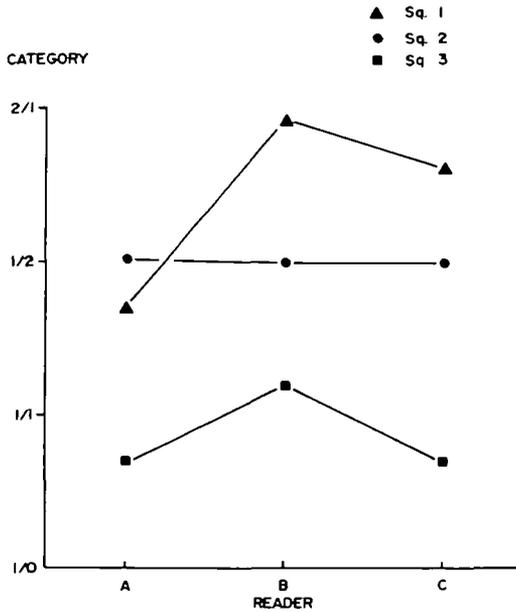FIGURE 10. Effect due to readers (averaged over squares and films).

FIGURE 11. Effect due to reader by square (averaged over films).

Although the general reader effect remains mostly intact, and is reflected again when viewed across squares, some striking anomalies do exist. For square one, the general effect is pronounced and exaggerated, and in square two the reader effect becomes almost non-existent. Figure 12 gives the reader by class of film interaction means. The overall trend appears remarkably con-

sistent, but on closer inspection an important phenomenon is evident. First of all, reader *B* categorizes all classes of films higher than his colleagues. Of the other two readers, *C* categorized those films at the low end of the scale a bit lower than did reader *A*. In contrast, reader *A* awarded a lower category for films at the upper end of the scale. A second order interaction involving readers, squares, and films, though showing some differences, is of no major significance.

The other effects shown in the analysis of variance fall into one of two classes. First, there are the main effects showing highly significant differences which are due to the design of the study, namely, squares and films. In addition, there were insignificant first order interactions; the reader by position (film) and the reader by pair (position) effects.

COROLLARY ANALYSIS

In an attempt to ascertain if reluctance exists towards reading regression of disease, the within pair variances averaged over readers and pooled across squares were de-
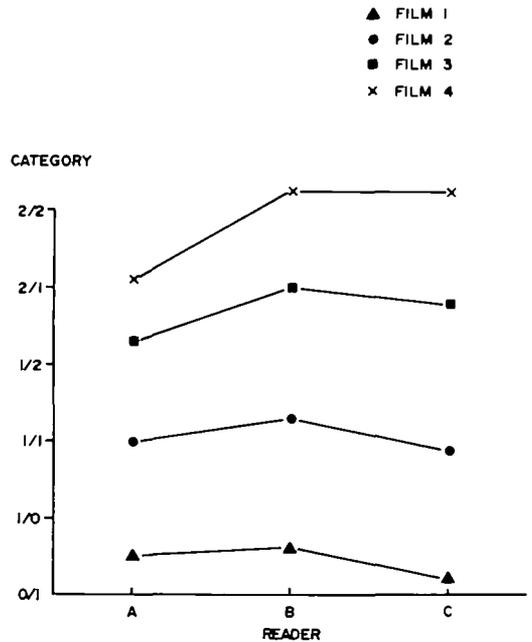

FIGURE 12. Effect due to reader by film class (averaged over squares).

termined for five groups of cells. These five groups are shown in figure 13 and are coded 0, 1, −1, 2, and −2. Theoretically, for cells coded 0, the within pair variability should be non-existent. For code 1 it should be somewhat higher and higher yet for code 2. Since those cells coded as −1 and −2 are the same films that exist above the principal diagonal except that their order is reversed, their within pair variability should resemble that of cells coded 1 and 2, respectively. These within pair variances and two composite variances are given in table 4.

Several striking features are evident from these data. First of all, the within pair variability is significantly greater above than below the principal diagonal. This is compelling evidence to support the belief that

SECOND FILM OF A PAIR

FIGURE 13. Coded components in the general design used in determining 5 different within pair variances.

TABLE 4

Within pair and composite variances for the general design

Code	Description	δ^2
0	Principal diagonal	0.1851
1	One unit above diagonal	1.3867
−1	One unit below diagonal	0.1535
2	Two + units above diagonal	5.5794
−2	Two + units below diagonal	1.8859
1 & 2	All units above diagonal	3.4830
−1 & −2	All units below diagonal	1.0197

TABLE 5

Contrasts involving within pair variances

Contrast	F
1 & 2 vs −1 & −2	3.42*
1 vs −1	9.04*
2 vs −2	2.96† borderline
1 vs 2	4.02†
−1 vs −2	12.29*
0 vs 1	7.49*
0 vs −1	1.21
0 vs 2	30.14*
0 vs −2	10.19*

* $p < .01$.
† $p < .05$.

there is a reluctance on the part of interpreters to read regression. The variance one position above the diagonal is about nine times higher than its counterpart below the diagonal; a highly significant finding. The variances of the extremes, viz. 2 versus −2, does show some disparity but not of this magnitude. Table 5 shows F statistics for these as well as some of the other more important contrasts involving the within pair variances.

DISCUSSION

Two theoretical concepts which are inherent in a consideration of this problem need examination. The first relates to the scale of categorization of coalworkers' pneumoconiosis. The 12-point elaboration of the International Labour Organization Classification presents an unusual problem of how to analyze discrete class data. In reality, there exists a disease continuum which underlies the 12 discrete radiological classifications. Wise and Oldham (9) devised a "Percentile Pairing" technique by fitting a smooth curve to the International Labour Organization readings, thereby transforming the scale. Although the technique may be theoretically valid, it can present a problem in communication. In contrast, Liddell (10) implies that there is no need for such a weighting system and expressed the opinion that the International Labour Organization

readings are primarily related to logarithmic transformations of several variables. The results of his work have often been used to support the proposition that the scale is of equal intervals. Although there may be technical objections to the proposition, lung residue analyses in his work do offer circumstantial confirmatory evidence and furthermore, its use appears justified "faute de mieux" since it offers a practical approach to the problem of determining radiographic change.

The second issue is concerned with the inherent bias which can occur in reading paired films as opposed to an independent randomized system which is a theoretically more pure approach to the problem of assessing progression of coalworkers' pneumoconiosis.

The side by side reading method guarantees that a subjective bias occurs each time a pair of films is categorized. The purpose of reading the films side by side is to insure that categorization of one film of the pair is related to, and modified by, the appearances of the other.

No doubt, some advocates of the side by side system also subscribe to the belief that each film should be separately categorized, without assessing progression directly. This concept, although idealistic, entirely ignores reality. The main attraction of the side by side method is that it allows a comparison of the films and thereby permits a more "accurate" and direct assessment of progression. The thought processes involved in the side by side measurement of progression are probably as follows:

1. The films are given an "overall" assessment.
2. An attempt is made to determine whether there has been any radiographic change.
3. The most easily read film is categorized.
4. Its more difficult partner is then placed in a category in light of the decision on radiographic progression (stage 2 above).

Previous trials in Great Britain (9), have shown that the side by side reading method underestimates the amount of progression of coalworkers' pneumoconiosis as compared to that assessed by the independent randomized method. On the other hand, recent work in the United States (3) has shown an opposite effect. The absence of a consistent directional difference rules out the introduction of a method for a "universal" adjustment between the two methods of determining progression. In both Britain and the USA, the variability associated with the side by side determination of progression has been substantially less and to many this is an appealing argument in favor of the method. However, to bias knowingly an experimental design so that the residual error is less, and to then use the smaller residual as a justification for the inclusion of bias in the design, is statistically invalid and can only be termed self-deceptive sophistry (11).

Aside from these two important issues, the question with which we must concern ourselves relates to the fact that if we use the side by side system, it is vital to know the extent of the bias we are introducing.

Side by side comparison of radiographs is accepted the world over, not only in the diagnosis of disease, but also in the assessment of therapeutic efficacy. In epidemiologic research it is often necessary to temper an obsession with statistical "purism" with practical considerations involving such factors as finance, time, and personnel; all of which influence the choice of a reading system.

We have shown that the position a film occupies in a pair, viz. on the left or right, has a direct influence on the category awarded to it; the left hand film being assumed to have been taken earlier is generally awarded a lower reading than it would had it been placed on the right where it would be assumed that it was the more recent of the pair. This effect is most striking for those films showing minimal disease. The phenomenon was anticipated and

serves to point out that readers are used to reading progression from left to right and are disinclined to read regression. Had they read as much regression as they did progression and had they also read it as frequently, then such an effect would not have been manifest. Furthermore, it was also apparent why the effect of position was more pronounced for the films showing minimal disease. When these films occupied the left hand position, they were paired with films showing equal or more severe disease so that a pair would appear "in-tune" with the interpreter's concept as to what he should see. In contrast, when the position of the films was reversed, so that the film on the left had more disease, there was a distinct tendency to "pull" the right hand film down and the left hand film up rather than read a change in a direction opposite to their thinking. Without question, a general effect due to film position is evident. However, it was demonstrated that this effect is drastically decreased when a subset of films which holds the disease stage constant is analyzed.

We have furthermore shown that films paired with themselves are generally awarded lower readings than when they are paired with films of another disease stage. This phenomenon is not consistent when the position of the film is borne in mind. In the left hand position, films in class 1, i.e. minimal disease, almost entirely accounted for the trend. In the right hand position, films in most all disease stages showed a significant difference regarding whether they were paired with themselves or paired with others.

Some differences existed between the readers and, moreover, they reacted differently to the three squares making up the total design and also to the class of films being read. Reader variability in categorizing single as well as serial films has been reported in numerous articles (13–21). Although total elimination of this source of variability is unlikely, some measure of control is possible.

Owing to the design of the study, the readers should have read about the same magnitude of regression of coalworkers' pneumoconiosis as progression, however, this was not the case. There was a general reluctance to read regression and this phenomenon is no doubt a consequence of training. Most radiologists believe that regression of coalworkers' pneumoconiosis is a rarity if it exists at all, and indeed what is termed regression is artifactual and a consequence of technical factors. Regression might be explained by emphysema and hyperinflation developing concomitantly with pneumoconiosis; the over-distention making the small opacities appear less profuse because there are less of them per unit volume of lung.

Although the pair readings by square and reader are not shown, it should be mentioned that at least one gross anomaly exists in the data. For square one, cell 44, i.e. a film read independently as borderline between Category 2 and 3, when paired with itself was read much lower than expected by two of the physician interpreters. The explanation for this is not clear. An inspection of the original film showed slight over-penetration but the film and the copies made from it were still in the readable range. Prior work (22) has shown a tendency for most readers to award lower categories to over-penetrated films. One can speculate that pairing a slightly over-penetrated film with itself may have caused these unexpected readings.

REFERENCES

1. Rivers D, Wise ME, King EJ, et al: Dust content, radiology, and pathology in simple pneumoconiosis of coalworkers. Br J Ind Med 17:87, 1960
2. McLintock JS: The changing prevalence of coalworkers' pneumoconiosis in Great Britain. International Conference on Coalworkers' Pneumoconiosis. Ann NY Acad Sci 200:278–291, 1971
3. Amandus HE, Reger RB, Pendergrass EP, et al: The pneumoconiosis: Methods of measuring progression. Chest 63:736–743, 1973
4. Liddell FDK, May JD: Assessing the Radio-

logical Progression of Simple Pneumoconiosis. National Coal Board Medical Service, Research Memorandum 4, London, 1966

5. Cooperative study by UICC committee: UICC/Cincinnati classification of radiographic appearances of pneumoconiosis. Chest 58:57–67, 1970

6. Liddell FDK: An experiment in film reading. Br J Ind Med 20:300–312, 1963

7. Steele RGD, Torrie JH: Principles and Procedures of Statistics. New York, McGraw-Hill, 1960

8. Fryer HC: Concepts and Methods of Experimental Statistics. Boston, Allyn and Bacon, 1966

9. Wise ME, Oldham PD: Estimating progression of coalworkers' simple pneumoconiosis from readings of radiological categories. Br J Ind Med 120:124–144, 1963

10. Liddell FDK: Validation of classifications of pneumoconiosis. International Conference on Coalworkers' Pneumoconiosis. Ann NY Acad Sci 200:527–551, 1971

11. Jacobsen M: Choice of a Film Reading Method (unpublished manuscript) and personal communication, 1972

12. Garland LH, Miller ER, Zwerling HB, et al: Studies on value of serial films in estimating progress of pulmonary disease. Radiology 58:161–177, 1952

13. Cochrane AL, Garland LH: Observer error in interpretation of chest films. Lancet 2:505–509, 1952

14. Fletcher CM: Clinical diagnosis of pulmonary emphysema; experimental study. Proc Roy Soc Med 45:577–584, 1952

15. Reger RB, Morgan WKC: On factors influencing consistency in radiologic diagnosis of pneumoconiosis. Am Rev Resp Dis 102:905–915, 1970

16. Garland LH: On reliability of roentgen procedures. Am J Roentgenol Radium Ther 64:32–41, 1950

17. Newell RR, Chamberlain WE, Rigler L: Descriptive classification of pulmonary shadows. Revelation of unreliability in roentgenographic diagnosis of tuberculosis. Am Rev Tuberc 69:566–584, 1954

18. Stradling P, Johnston RN: Reducing observer error in 70 mm chest radiography service for general practitioners. Lancet 1:1247–1250, 1955

19. Yerushalmy J: Reliability of chest radiography in diagnosis of pulmonary lesions. Am J Surg 89:231–240, 1955

20. Garland LH: Studies on the accuracy of diagnostic procedures. Am J Roentgenol Radium Ther 82:25–38, 1959

21. Fletcher CM, Oldham PD: Problem of consistent radiological diagnosis in coal miners' pneumoconiosis. Br J Ind Med 6:168–183, 1949

22. Reger RB, Smith CA, Kibelstis JA, et al: The effect of film quality and other factors on the roentgenographic categorization of coalworkers' pneumoconiosis. Am J Roentgenol Radium Ther 3:462–472, 1972