Coalminers' pneumoconiosis and lung function, and exposure to dust of variable quartz content.

Brian G. Miller

Doctor of Philosophy
University of Edinburgh
1988

CONTENTS

DECLARATION 1

ACKNOWLEDGEMENTS 2

ABSTRACT 4

CHAPTER 1 INTRODUCTION 6

1.1 Background 6

1.2 The Pneumoconiosis Field Research of the National Coal Board 8

1.3 Quartz and pneumoconiosis at one Scottish colliery 10

CHAPTER 2 OBJECTIVES OF THE PROJECT 12

CHAPTER 3 THE PFR RESEARCH PROGRAMME AND ITS LEGACY OF DATA 13

3.1 Chest radiographs 13

3.2 Additional data from medical surveys 14

3.3 Data on exposure to respirable coalmine dust 15

CHAPTER 4 RE-READING THE CHEST RADIOGRAPHS 18

4.1 The ILO (1980) system for classifying pneomoconioses 18

4.1.1 Technical quality 19

4.1.2 Parenchymal abnormalities 21

4.1.3 Pleural and other abnormalities 25

4.1.4 Comments 26

4.2 The IOM film-reading panel 26

4.3 Organisation of the independent randomised film-reading exercises 29

4.3.1 Availability of films 29

4.3.2 Design of batches of films 29

4.3.3 An additional reading exercise to include the 1980 NCB Medical Service survey films 33

4.4 Reading films serially for progression 35

4.4.1 Selection of a subset of the men 37

4.4.2 Design of batches of films 37

CHAPTER 5 ORGANISATION AND VALIDATION OF DATA 45

5.1 Personal data 45

5.2 Data from medical surveys 46

5.3 Data on occupational exposures 47

5.4 Organising data from film-reading exercises 48

5.5 Summary descriptions of data 49

5.5.1 Distributions of age and smoking habits 49

5.5.2 Distributions of height and weight 52

5.5.3 Distributions of lung function variables 53

5.5.4 Distributions of dust exposure variables 55

CHAPTER 6 STATISTICAL METHODS 60

6.1 The analysis of film-reading data 60

6.1.1 Logistic regression for binary variables 61

6.1.2 The Rasch model 63

6.1.3 Estimation in a model with dependent observations .. 67

6.1.4 A multivariate logistic model 72

6.1.5 Analysis of ordered categorical data 76

6.2 The analysis of lung function data 78

CHAPTER 7 RESULTS OF ANALYSES OF FILM-READING DATA 79

7.1 Data from readings by the IOM panel 79

7.1.1 Film quality 80

7.1.2 Profusion of small opacities 82

7.1.3 Between-reader differences in assessments of profusion 85

7.1.4 Within-reader variations in assessing profusion 88

7.1.5 Shape and size of small opacities 92

7.1.6 Profusion of large opacities 93

7.2 Logistic regression analyses of small opacities profusion 94

7.2.1 Summarising small opacities assessments by medians . 94

7.2.2 The form and presentation of the logistic regression analyses 96

7.2.3 Logistic regression analyses of 4th survey films ... 100

7.2.4 Logistic regression analyses of 5th survey films ... 103

7.2.5 Logistic regression analyses of 6th survey films ... 107

7.2.6 Logistic regression analyses of 1980 films 111

7.2.7 The effects of smoking on profusion of small opacities 115

7.2.8 Adequacy of the median as a summary of readings 122

7.2.9 Additional checks on adequacy of logistic models ... 125

7.2.10 Investigations of alternative models 128

7.3 Data from serial readings by medical readers 138

7.3.1 Film quality 138

7.3.2 Progression of small opacities 141

7.3.3 Agreement between the medical readers 145

7.3.4 Progression of large opacities 151

7.4 Logistic regression analyses of serial readings 153

7.4.1 Choice of variable for analysis 153

7.4.2 The form and presentation of the analyses of progression 155

7.4.3 Regression analyses of progression 4th to 5th survey 157

7.4.4 Regression analyses of progression 5th to 6th survey 161

7.4.5 Regression analyses of progression 6th to 1980 survey 165

7.4.6 The effects of smoking on progression of small opacities 170

7.5 Summary and overview of findings 173

7.5.1 Findings from analyses of panel's readings 173

7.5.2 Findings from medical readers' assesments of change 178

7.5.3 Concluding remarks 182

CHAPTER 8 RESULTS OF ANALYSES OF LUNG FUNCTION 183

8.1 Regression analyses of lung function data at each survey 184

8.1.1 The form and presentation of the regression analyses 184

8.1.2 Regression analyses of 4th survey lung function data 186

8.1.3 Regression analyses of 5th survey lung function data 189

8.1.4 Regression analyses of 6th survey lung function data 193

8.1.5 Checks on the adequacy of the regression models 197

8.2 Regression analyses of change in lung function between surveys 201

8.2.1 The form and presentation of the analyses 201

8.2.2 Changes in lung function from 4th to 5th survey 202

8.2.3 Changes in lung function from 5th to 6th survey 204

8.2.4 Checks on the adequacy of the regression models 209

8.3 Small opacities and lung function 214

8.3.1 Small opacities and lung function at individual surveys 215

8.3.2 Small opacities and changes in lung function 217

8.4 Summary and overview of findings on lung function data .. 219

8.4.1 Findings from cross-sectional analyses 220

8.4.2 Findings from analyses of change in lung function .. 221

CHAPTER 9 DISCUSSION 223

REFERENCES 239

APPENDIX A Protocol for film readings by IOM panel 246

APPENDIX B Protocol for medical readers' film readings 251

APPENDIX C Database schemas 257

APPENDIX D Copy of: Heederik D and Miller BG. (in press) "Weak associations in occupational epidemiology: adjustment for exposure estimation error" 270

Men working at the coal face, Colliery P, Midlothian, c. 1959.
(Photograph by permission of The Scotsman Publications Ltd.)

DECLARATION

This thesis is submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy. None of the material herein has been submitted for any other degree. Except where due acknowledgement has been given, the work described and the composition and preparation of the manuscript were performed by the author.

ACKNOWLEDGEMENTS

I am indebted to a number of people for help in the course of the work described here, most of which was carried out as part of a research project by the Institute of Occupational Medicine (IOM), Edinburgh, jointly funded by the National Coal Board (now British Coal Corporation) and the European Coal and Steel Community, under research contract 7260-04/025/08. Professor Anthony Seaton (Director, IOM) gave permission to register for postgraduate study in connection with the project, and Dr Michael Jacobsen (Deputy Director) and Mr Fintan Hurley (Head, Statistics Group) gave encouragement, support and advice. Dr Robin Prescott (Faculty of Medicine, University of Edinburgh) supervised the preparation of this thesis, and his comments at all stages led to innumerable improvements in content and presentation. The Dutch epidemiologist Dick Heederik (Wageningen University) spent part of 1987 as a visting researcher at the IOM, and we had a number of stimulating discussions on the analysis and interpretation of lung function variables.

Examination and classification of radiographs were performed by Drs Peter Pern and Colin Soutar, and by the members of the IOM film reading panel; Mrs Betty Henderson, Mrs Kay Duncan, Mrs Valencia White, Mrs Catherine Scott, Ms Jose Jeffryes, and (in the early stages) Mrs Gillian Chalmers. Mrs Isobel McIntosh and Mrs Helene Murray recorded the panel's classifications, and they and Ms Margaret Burnett and Ms Mags Parker entered them to computer data files. Ms Ann Kinnear's programming skills were invaluable in extracting and collating existing data from various computer files, and in constructing sections of the project database in which to store them. Mr Alex Cowie's help was useful in interpreting some of the historical data on environmental measurements. Mrs Brenda McGovern and Ms Fiona Small assisted in extracting and preparing bibliographic references.

My appreciation and gratitude go to all of these people, and to all of my colleagues past and present whose efforts over more than thirty years have made the National Coal Board's Pneumoconiosis Field Research a landmark in occupational epidemiology.

I composed and prepared this thesis using facilities for text preparation available on the IOM's mainframe and personal computers, and take full and sole responsibility for errors, omissions and inaccuracies.

Throughout this period, my somewhat preoccupied manner and my irregular and curtailed appearances at home have been treated by my immediate family with great understanding, without which the completion of this thesis would have been considerably more difficult. Every word and figure herein is dedicated, with the deepest gratitude for their patience and forbearance, to Jan, Kirsteen and Siobhán.

ABSTRACT

In the course of the British National Coal Board's Pneumoconiosis Field Research programme, medical officers examining chest radiographs taken in 1978 from workers at a colliery in Midlothian considered that a small number showed unusually rapid progression of pneumoconiotic abnormalities. A case–control study based on these radiographs suggested an association with workplace exposure to dusts containing higher proportions of quartz than had previously been seen in the research, and further investigations were initiated.

This thesis describes the design and execution of a study in which existing radiographs for men at this colliery were subjected to intensive re–examination, with the objective of relating any evidence of radiographic abnormalities to data already held on the individual men's exposures to respirable airborne dust in the coalmine, and on lung function and smoking habits.

All available radiographs from over 1400 men employed at that colliery who had attended medical surveys in 1970, 1974 and 1978 were collected, along with some taken by the NCB's Medical Service in 1980; these were classified for pneumoconiotic abnormalities according to the ILO (1980) scheme, both by an experienced panel of non–medical readers reading each radiograph independently and in randomised order, and by two medical readers who examined series of films for progression of disease over time.

Analyses by logistic regression, supplemented by other statistical modelling techniques, confirmed that appearances of small pneumoconiotic shadows of profusion at least 1/0 on the ILO (1980) scale were associated most strongly with the estimates of individuals' exposures to respirable coalmine dusts from before the

1970 to after the 1974 surveys, and particularly with estimates of exposures to the quartz components of these dusts. There was no evidence that the men's smoking habits were an important modifying factor in this association.

Analyses by linear regression methods of the lung function variables measured at the surveys suggested that a slightly higher rate of loss of lung function between 1974 and 1978 was experienced by men with higher exposures to the non-quartz fraction of the dust, but this effect was not observed between 1970 and 1974. Cross-sectional analyses comparing different men's lung function at the same survey suggested an association between higher exposures and higher lung function. This was interpreted as probably due to a selection effect in the population studied. There was no evidence of a relationship between apparently dust-related lung function effects and the radiographic abnormalities ascribed to quartz exposures in the same men.

Suggestions are made for further work on suitable methods of statistical analysis for data from the subjective assessment of radiographs. The need is identified for a follow-up study to investigate progression and any new incidence of pneumoconiosis in these men, taking account of whether they continued in coalmining employment after this colliery closed in 1982.

1 INTRODUCTION

1.1 Background

Coalworkers' pneumoconiosis (CWP) is a disease of the lungs, which develops in some colliery workers as a response to the inhalation and retention in the lung of airborne coalmine dust. Its risk increases with increasing exposure to respirable dust, that is the fraction of the aerosol which can pass into the finest airspaces of the lung. The severity of the disease varies greatly. In its simple or early form, which is characterised radiographically by the appearance of small opaque shadows, it is usually accompanied by no detectable impairment of lung function. In many cases, the disease never becomes more serious. In the small proportion of more advanced cases, the shadows are larger, and signal the presence of fibrotic lesions whose progression is usually accompanied by respiratory disability, and may lead ultimately to the death of the worker.

As with so many disease processes, individual susceptibility appears to be a large factor in determining whether an individual worker exposed to quantities of respirable coalmine dust will develop a pneumoconiotic response; there are no known factors which allow accurate predictions of such response in individuals, and the picture is further clouded by large variations between the disease rates in different regions of the British coalfields.

That the statements in the two paragraphs above may almost be taken for granted is due to the great amount of research which has taken place in the middle part of this century on lung disease in coalworkers. However, the recognition of silicosis as a lung disease of workers exposed to dusts containing high proportions of quartz predated any recognition of pneumoconiosis as a separate disease entity, and indeed when symptoms similar to those of silicosis were first documented in coalworkers, it was presumed that these were due to quartz in the coal mining environment, an impression which appeared to be confirmed by the similarities displayed by radiography when this technique became available (Seaton, 1983).

The principal evidence for a distinction between the two diseases was obtained by a series of studies carried out by the Medical Research Council (MRC) in mining communities in South Wales, stimulated by concern over an apparent epidemic of lung disease in that region's coalminers. These studies demonstrated that the disease was found in men exposed to dust containing little or no quartz (Medical Research Council, 1942, 1943), and Gough (1940) described the disease in coal trimmers at docks in South Wales whose exposure was to virtually pure coal dust. Coal miners' pneumoconiosis was prescribed as an occupational disease eligible for compensation in 1943.

The MRC's surveys continued. Following nationalisation of the coal industry in 1947, the first British dust standards for coalmines were introduced by the National Coal Board (1949), but these were based on little hard quantitative evidence. To obtain such evidence, the NCB began a longitudinal study of coalminers in collieries selected to cover the range of types of coal and mining conditions in the British coalfields. The stated aim (Fay, 1957) was "to establish what environmental conditions should be maintained if mineworkers are not to be disabled by the dust they breathe". The research was initially organised by staff of the NCB's Medical Service, but a separate unit, known as The Institute of Occupational Medicine, was established in Edinburgh by the NCB in 1969, specifically to carry on this research.

1.2 The Pneumoconiosis Field Research of the National Coal Board

The first phase of the PFR centred on three rounds of medical surveys, about five years apart, which were carried out at each of 13 collieries in England, 7 in Wales and 4 in Scotland, with the intention of examining the entire industrial workforce at each. At the first round, the examination was limited to a full-sized postero-anterior chest radiograph and a questionnaire to establish the working history of each worker. At later surveys, trained staff measured anthropometric and lung function variables, and administered a questionnaire on respiratory symptoms and smoking habits. Teams of investigators were stationed full-time at the collieries, and carried out a statistically designed programme of samples to quantify dust concentrations in locations typical of the range of colliery occupations. Information on times worked in these various occupations was extracted routinely from colliery payroll records for each man. New recruits were automatically included in this system, and at their first survey appearance a history was obtained of their employment prior to engagement at that colliery.

Interim results from these surveys were published as they became available, culminating in the publication of the results from a major interim analysis, by Jacobsen *et al* (1971), which showed a clear association between radiographic evidence of pneumoconiotic changes and estimates of exposures to concentrations of respirable dusts expressed in mass units (Dodgson *et al*, 1971). These results were pivotal in the determination of new gravimetric dust standards for British coalmines (National Coal Board, 1969), which were introduced in 1970. Further publications (Rae *et al*, 1971; Rogan *et al*, 1973) reported analyses which showed associations of dust exposure with increased prevalence of bronchitic symptoms, and with lower average levels of the lung function variable Forced Expiratory Volume (FEV).

Meanwhile the research continued, but because of various factors including the closure of some of the original 24 collieries, only 10 collieries continued into the second phase. Two further complete rounds of surveys took place at these collieries, and two of the collieries received a sixth survey visit, all at intervals of around four years. At the same time, the research diversified. A cohort for the study of mortality was defined restrospectively as a 56% sample of those attending the first round of surveys, and was later extended to all of those attending (Miller and Jacobsen, 1985); the 56% sample was also used as the basis for a follow-up study to include survivors who had left the industry (Soutar *et al*, 1986); additional laboratory studies were instituted to investigate the mechanisms of coalminers' lung disease.

Hurley *et al* (1982) reported on the principal findings on pneumoconiosis and dust exposure, from data from surveys up to the fifth round. These confirmed unambiguously the correlation between quantitative estimates of coalmine dust exposure and simple pneumoconiosis, and refined the risk estimates. Important unexplained differences in risks remained between some of the collieries, and were not explicable in terms of differences between the collieries in the levels of quartz in the coalmine dusts; these levels were generally low (average 5%). However, a case-control study of men with unusually rapid progression of pneumoconiosis (Jacobsen and MacLaren, 1982) suggested that higher quartz exposures were associated with a more aggressive form of the disease, and perhaps influenced the risks of developing PMF. Further important quantification of pneumoconiotic risks, particularly for the development of PMF, was reported by MacLaren and Soutar (1985).

1.3 Quartz and pneumoconiosis at one Scottish colliery

Colliery P was the one Scottish colliery at which PFR medical surveys were continued into the second phase of the research. It received a total of six PFR surveys between 1954 and 1978. Ex-miners from the cohort defined by attendance at the first survey received a follow-up examination at the same time as the PFR survey team were conducting the 5th routine survey.

The prevalence of pneumoconiosis at colliery P since the research started was low, and this reflected the Scottish experience generally. However, the medical officer who read and classified the radiographs taken at the 6th PFR survey in 1978 considered that 21 radiographs from the 623 taken showed progression of simple pneumoconiosis when compared with radiographs taken four years previously from the same men at the 5th PFR survey.

This observation spurred a case-control study (Seaton *et al*, 1981), which took as cases the 21 men from whom these radiographs had been taken, and 21 age-matched controls selected from the men who attended the 5th and 6th surveys at the same colliery, and whose radiographs had been classified as showing no pneumoconiosis. A second reader examined the radiographs without knowledge of their status as cases or controls, and classified progression in 18 of the 21 cases and in none of the controls, confirming the pattern of the original readings. Comparison of the estimates of exposure for cases and controls in the period between the 5th and 6th surveys showed significantly higher values on average among the cases for the respirable mixed-dust exposure, quartz exposure, and highest for percentage quartz in dust, which averaged 13.0% for the cases against 8.4% for the controls. Similar patterns of difference were apparent in the estimates of the men's lifetime exposures to dust and quartz.

The authors reported that quartz levels in this period were higher than had previously been typical, because adverse geological conditions in one seam had led to excursions of the powerful coal-getting machinery into silicaceous rock in the roof and floor of that seam, leading to the creation of free silica dust amongst the coal. They concluded that this unusually high release of quartz had been responsible for the unusually rapid pneumoconiotic response of miners exposed to dust levels which, in normal circumstances, would have been expected to carry a low risk for the development of pneumoconiosis. They also noted that these findings had implications for the setting of control limits for mixed dusts containing relatively high proportions of quartz.

It was considered desirable to investigate further, and accordingly a project was designed in which it was proposed to make an extensive study of all the radiographs taken from men working at this colliery in the 1970's, and to collate these classifications with the detailed data already available on the environmental conditions in which these men had worked, along with the questionnaire and measured data from the medical surveys. This report describes the planning, execution and analysis of that project, and the conclusions which were drawn.

2 OBJECTIVES OF THE PROJECT

The principal objectives to which the design and execution of this project were addressed were as follows:

1. To make an intensive study of recent chest radiographs from men working at one coalmine in Scotland, classifying systematically any pneumoconiotic abnormalities observed;

2. To examine the relationships between estimates of exposures to respirable coalmine dust and its constituent components, particularly quartz, and the presence of radiograph abnormalities;

3. To examine relationships between estimated exposures and the progression of abnormalities over time;

4. To examine the relationships between estimates of exposures and spirometric measurements of lung function taken at the same surveys as the radiographs;

5. To examine relationships between the radiographic appearances and the measurements of lung function.

3 THE PFR RESEARCH PROGRAMME AND ITS LEGACY OF DATA

This chapter contains a brief description of the procedures by which the PFR research programme was carried out, and provides some background on the nature and extent of the data which it collected. At each of the collieries studied, surveys took place at intervals of four or five years, with the visits lasting typically between one and two months. The aim was to examine on each occasion all industrial workers employed at that time. Attendance was not compulsory, but non-response was generally lower than is often the case with industrial surveys. In particular, the earlier surveys often achieved coverage of well over 90% of the workforce. At colliery P, there were a total of six surveys, which took place around February 1954, April 1958, February 1964, December 1970, October 1974 and November 1978.

3.1 Chest radiographs

At each survey, a full-sized postero-anterior chest radiograph of each man was taken. Over the period of the research, there were small changes in radiographic technique, but the principal techniques remained the same. At first attendance, each man was given an identifying "x-ray number", by which all subsequent film and other data were indexed. A manually maintained card system provided a link between these numbers and the men's names and National Insurance numbers.

Soon after they were taken, all films from each survey were assessed by medically qualified personnel for any signs of clinically significant disease, and information on any abnormal features noticed was passed to the individual's GP. During these clinical readings, any earlier films from the same individual were available for comparison.

During the first phase of the PFR research, which embraced all investigations from the 1st to the 3rd survey, at least two of the four experienced Medical Officers on the research team read each film to provide an agreed "definitive" classification (Fay and Ashford, 1960). The classifications were restricted to the 4 major categories of simple pneumoconisis 0, 1, 2, and 3, and complicated pneumoconiosis (progressive massive fibrosis, PMF) classified as category A, B, C or PMF undefined (ILO, 1959). Films taken at 4th or later PFR surveys were classified clinically by one of five or six experienced NCB medical officers (Jacobsen *et al*, 1977), to provide assessments of simple and complicated pneumoconiosis on scales comparable to the earlier readings. Research directed to specific topics has often required more detailed classifications of specially chosen subsets of available films, by panels of readers.

3.2 Additional data from medical surveys

From the second survey onwards, additional data were collected from the participants. Weight and standing and sitting heights were measured, and trained personnel administered a questionnaire on smoking habits and respiratory symptoms, whose design (Rae *et al*, 1971) was based on a questionnaire devised for similar purposes by the Medical Research Council. Lung function measurements were taken using a modified Gaensler wet spirometer; forced expired volume in one second (FEV) and forced vital capacity (FVC) were measured from spirometer traces, the subject making three forced expirations after a practice blow (Rogan *et al*, 1973).

The questionnaire collected data on whether an individual was a smoker, and if not whether he ever had been, and on the amounts of cigarettes, hand-rolling tobacco, cigars and pipe tobacco smoked on weekdays and weekends. In addition, respondents were questioned on persistent cough, persistent phlegm, shortness of breath on light exertion, and recent chest illnesses.

3.3 Data on exposure to respirable coalmine dust

The aim of the PFR research was to quantify the health-related risks presented to coalworkers by the airborne coal dusts to which they were exposed in their work, and to assess how any risks might be minimised. It was clear from the outset that, if this aim were to be met, one of the principal tasks of the research programme would be to derive estimates of individuals' exposures which usefully differentiated between the degrees of dustiness inherent in different jobs within the colliery. Such estimates would also have to take account of changes in workers' conditions induced by, for instance, transfers to other duties, or by the natural trend for older coalface workers to be replaced by younger men.

This was achieved by the development of a system based on the concept of an occupational group; that is, a reasonably homogeneous set of specific activities taking place at a particular place within a colliery, where individuals working within the same occupational group might be expected to be exposed to similar amounts of dust. Two distinct data gathering exercises came into being, the first to measure the conditions of dustiness within each occupational group, and the the second to record the amount of time spent by each worker in each occupational group.

Initially, occupational groups were defined on the basis of local knowledge of mining conditions, and series of exploratory measurements of dust concentrations. Dust concentrations in the occupational groups were then monitored regularly, using sampling instruments positioned close to the men throughout selected working shifts. Initially, the concentrations were based on particle counts obtained by using the standard thermal precipitator (Watson, 1936), whose dust sampling characteristics were related during a series of side-by-side instrument comparisons (Dodgson *et al*, 1971) to those of the MRE gravimetric sampler (Dunmore *et al*, 1964) which was used subsequently. As a result, the earlier particle counts were converted to equivalent gravimetric units (grammes of dust per cubic metre of sampled air, $g.m^{-3}$), which have been used in all subsequent work.

The proportions of the dusts collected by the sampling instruments which were made up of combustible and volatile components were later determined by high temperature ashing. The proportions in the ash residue of quartz, and of kaolin and mica, were measured variously by interference microscopy and by infra-red spectroscopy (Dodgson, 1963; Dodgson *et al*, 1971; Dodgson and Whittaker, 1973); from time to time, and less systematically, measurements were made of the proportions of other mineral components such as carbonates.

The measurements of the respirable dust concentrations were routinely monitored for stability, and the occupational groups were revised as necessary to reflect changes in the typical concentrations, or changes in mining methods and conditions.

Information on the times spent by each man in each occupational group was extracted from the colliery systems which recorded attendance at work for the purpose of payroll calculations. In the early days of the research, this extraction was performed by clerks at the collieries. Later, when the payroll systems were computerised, it became possible to extract attendance information direct from the computer system, on magnetic tape.

Individuals' exposures were then characterised by multiplying the times worked over a period in the different occupational groups by mean dust concentrations appropriate to each group. The cumulative total of the products, over all the groups, was used as a measure of exposure in the period. Similarly, products of time with concentrations of the components of the dust, which in turn had been obtained as products of dust concentrations and compositional proportions, were used to estimate exposures to the components of the dust.

Obviously, no information on working patterns or on dust concentrations prevailing before the start of the research, or at other collieries, could be obtained in this way. At each man's first appearance in the research, a working history questionnaire was applied, and the times recorded were classified into six broad "combined groups". Mean concentrations for these combined groups were obtained by combining data from individul occupational groups sampled during the first ten years of the research. As before, estimates of exposure were produced by summing the products of times and concentrations over all the combined groups.

Further details of the processes by which exposure estimates were calculated are given in reports by Hurley *et al* (1979; 1982).

4 RE-READING THE CHEST RADIOGRAPHS

4.1 The ILO (1980) system for classifying pneumoconioses

A chest radiograph, even in the absence of any sign of pathological abnormality, is a complex visual image, composed of a variety of shapes and patterns. Anatomical features can be difficult to observe or distinguish, and abnormal features may be even more difficult to distinguish from the normal. All features observed are likely to be difficult to describe succinctly and accurately.

On inspection of a film for the presence of abnormalities, implicit issues are:

- IDENTIFICATION ; is any (or a particular) abnormality present?
- DISCRIMINATION ; could it not be something else of similar appearance?
- QUANTIFICATION ; how widespread or serious does it appear to be?

There is no abnormal feature which may be observed on a radiograph which in itself implicates respirable dust as its cause; pathological changes due to a wide variety of causes can have similar or identical appearances. However, there are patterns which are widely regarded as "typically pneumoconiotic", in that they have been repeatedly observed in men with exposure to respirable dust. The appearance of such a pattern on the radiograph of an individual who is working or has worked in dusty conditions is likely to be considered indicative of the presence in the lung of pneumoconiotic lesions. Simple pneumoconiosis, which is usually not associated with respiratory dysfunction or disability, cannot be diagnosed during life except by examination of the chest radiograph.

For epidemiological studies of pneumoconiosis and its determinants, therefore, it has been recognised that there exists a need to standardise the description, quantification and documentation of radiographic features seen in all types of pneumoconiosis (not just that of coalworkers). To this end, the International Labour Office has over the last 50 years promoted discussion and published a series of guidelines on ways of classifying features of chest radiographs of persons with pneumoconiosis. The International Classification of Radiographs of Pneumoconioses (ILO, 1980) is the most recent revision of the resulting scheme, which exists to facilitate standardisation of description of several typical patterns of abnormality. The scheme covers the presence of such patterns, how much of the lung is affected, and how badly.

This section describes the use of the Classification for description of abnormalities, with reference to the data recording form shown as exhibit 4.1.1. This form, currently in use at the IOM, is based on the sample form given in the ILO (1980) booklet, but differs in the way it is laid out. It also incorporates some small departures from the Classification which have been found convenient in everyday use.

4.1.1 Technical Quality

The reader first assesses the technical quality of the film and classifies it under one of four grades; 1=good, 2=acceptable, 3=poor, 4=unacceptable. If the film is classified as grade 4, the implication is that the quality is so bad that the possibility of detecting abnormalities is seriously impaired or excluded, and no classification of abnormalities is attempted. Grades 2 and 3 both imply some defect, the distinction being that in grade 2 the defect is not likely to impair classification of the radiograph for pneumoconiosis, while in grade 3 the film is considered impaired in some way but still acceptable for classification. If the quality grade is assessed as other than 1, a comment should be made about the defect(s)

Exhibit 4.1.1. Example of form used for recording readers' assessments of radiographs in the independent randomised reading exercises. The identities of the films were preprinted on the forms by computer program, from the batch lists.

INSTITUTE OF OCCUPATIONAL MEDICINE - I.O.M. - FILM READING PANEL PAGE: 1

READING ID: XRLV BATCH ID: 01 READER ID: 101 [BATCH NUMBER:] DATE: / /

FILM ID	SEQ	QUAL 12 34	1- 9	CPA OFF RLB	N X	PROF 012 3-+	ZONES X R L	TYPE P.Q.R S.T.U	LGE OPS O.A B.C	PLAQUES FACE Y,N R L	PLAQUES WID ABC R L	PLAQUES EXT 123 R L	DIFFUSE FACE Y,N R L	DIFFUSE WID ABC R L	DIFFUSE EXT 123 R L	DIA RLB	CPA RLB	PLEURAL CALC CHEST WALL X R L	PLEURAL CALC DIA X R L	PLEURAL CALC OTHER X R L	PLEURAL CALC EXT 123 R L	OTHER ABN	C X
PFRP0097-786	001					/		/															
PFRP4565-704	002					/		/															
PFRP1872-786	003					/		/															
PFRP1118-786	004					/		/															
PFRP1331-786	005					/		/															
PFRP0865-786	006					/		/															
PFRP1837-786	007					/		/															
PFRP3015-786	008					/		/															
PFRP2136-786	009					/		/															
PFRP0328-745	010					/		/															

present in the film. In the IOM, commonly occurring defects have been allotted codes, which are used to record these defects when they are observed (see exhibit 4.1.2 for codes); a written comment is necessary only if a defect is not on the standard list.

In one further local addition to the Classification, readers at the IOM record a separate technical defect which can occur. Poor positioning of the subject relative to the radiographic equipment can result in the loss off the lower edge of the film of the image of the costophrenic angles, i.e. the angular points at the lower and outer extremes of the lungs (see exhibit 4.1.3). Such loss at either or both sides is recorded when it occurs, since its occurrence will make the assessment of costophrenic angle obliteration (see below) impossible on the side affected.

4.1.2 Parenchymal abnormalities

4.1.2.1 Small Opacities - profusion

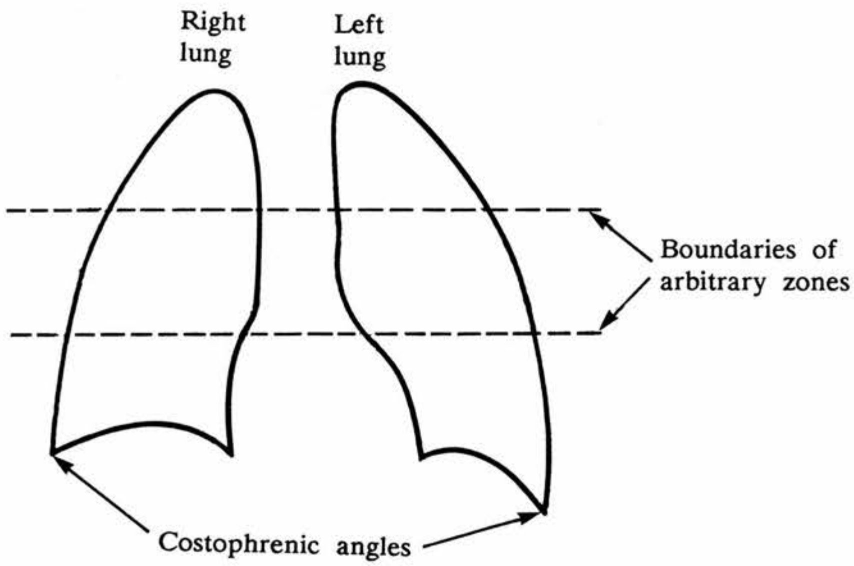
The abnormality considered most important in the classification of pneumoconiosis is the occurrence within the lung parenchyma of small fibrotic lesions, which show on the radiograph as small (10mm or less in diameter or width) opaque markings. These can be round in appearance or more irregular in shape; it is believed that pneumoconiosis in coalworkers is typified by the rounder shadows.

Small opacities when they occur usually do so in sufficient numbers to make counting an inconvenient system of quantification. Instead, the profusion of the small opacities is assessed by the reader's subjective judgment in comparison with standard films. These films are distributed by the ILO as an integral part of the Classification system, and the categories of profusion are defined in terms of the

Exhibit 4.1.2. List of codes for specific defects of film quality used in all film readings. (IOM local adaptation of ILO (1980) scheme.)

Numeric code	Defect
1	Dark film
2	Pale film
3	Too much contrast
4	Lack of contrast
5	Patient movement
6	Poor inspiration
7	Marks on film
8	Ill-defined lung markings
9	Other; see comment

Exhibit 4.1.3. Diagram of lung image on typical postero-anterior radiograph, showing position of the costophrenic angles.



films themselves.

Classification of profusion begins with the four categories 0=no opacities, and 1 to 3 showing increasing amounts of opacities. There are sets of standard films illustrating each category, and up to 1971 a film was allocated to that category whose standard film it most resembled in profusion. Since then, an extended version of the scale of categories has been introduced, to allow finer distinctions between profusions lying between those of the standard films. This extension involves the use of a secondary digit, as follows. A film which would definitely have been allocated to category 1, say, under the pre-1971 scheme is recorded as 1/1; similarly 0/0, 2/2, 3/3. However, a film which is allocated to category 1, but for which the reader seriously considered category 0, is recorded as 1/0. Allocation of the extreme categories 0/- in a conspicuously normal or healthy-looking case, and 3/+ for a film where profusion is markedly higher than on the standard category 3 film, completes the 12-point extended ILO scale of profusion:

0/-, 0/0, 0/1, 1/0, 1/1, 1/2, 2/1, 2/2, 2/3, 3/2, 3/3, 3/+.

This scale gives an ordered set of 12 categories of profusion, which is by definition an elaboration strictly within the categories of the original 4-point scale. Ignoring the second digit effectively collapses the 12-point back into the 4-point scale.

4.1.2.2 Small opacities - zones

Opacities, when present, are often not evenly distributed through the radiographic image of the lung. Provision is made for the description of the distribution of opacities by the definition of three arbitrary zones in each lung, demarcated by horizontal lines drawn (or, more realistically, imagined) at one-third and two-thirds of the vertical distance between the lung apices and the domes of the diaphragm. The zones in which small opacitites are observed are indicated diagrammatically on the recording form.

The category of profusion is determined by considering the observed profusion over only the affected zones of the lung, when making comparisons with the standard films. Where the zones show markedly different profusion, then the zone(s) showing the lesser profusion are ignored for the purpose of classifying the profusion.

4.1.2.3 Small opacities - shape and size

Provision is made for the classification of the shape and size of observed small opacities, using alphabetic codes. Two kinds of shape are recognised, rounded and irregular. Within each kind, three sizes are differentiated, and are illustrated by standard radiographs. The sizes correspond approximately to ranges of diameters or widths (for rounded and irregular opacities respectively) up to 1.5 mm, exceeding 1.5 and up to 3 mm, and exceeding 3 and up to 10 mm. These ranges are denoted by the symbolic codes p, q, r for the rounded and s, t, u for the irregular shapes.

Two letters are used to record shape and size, separated by an oblique. If all the opacities are considered to be the same shape and size, the same letter is used twice; otherwise, the first letter describes the predominant type and the second describes the next most common.

Since there are six letters, each of which can appear in either position, there are 36 possible combinations which can be recorded.

4.1.2.4 Large opacities

Opacities greater than 10 mm in diameter are described as large opacities, are distinguished from the small opacities discussed above, and are recorded separately. Three sizes of large opacities are defined as follows:

Category A - diameter exceeding 10 mm but not greater than 50 mm. Where more than one such opacity is present, the sum of the diameters is not greater than 50 mm.

Category B - exceeding category A, but combined area not exceeding the equivalent of the right upper zone of the lung.

Category C - combined area exceeding the equivalent of the right upper zone of the lung.

4.1.3 Pleural and other abnormalities

The aim of the 1980 ILO Classifiation Scheme is to provide descriptive tools for all radiographic abnormalities associated with the different types of pneumoconiosis. Traditionally, simple pneumoconiosis in coalworkers has been characterised by small, mostly rounded, parenchymal opacities; the more severe, progressive form of the disease corresponds in general to the appearance of large opacities on the radiograph.

Exposure to other kinds of mineral dusts and fibres has been associated with radiographic abnormalities of the chest wall, pleura and diaphragm, as well as with parenchymal abnormalities. The Classification provides for description of these as well as of some more or less common abnormalities observable on some radiographs of the chest. For further details, reference should be made to the appropriate sections of the Guidelines to the Classification (ILO, 1980). Because these abnormalities are unlikely to be related to exposure to coalmine dust, and more specifically because the readers of the IOM film reading panel have little experience in classifying films from industries other than coal mining, they are not considered here. Presentation and analysis of data from the assessments of radiographic abnormalities in this study concentrates on the small and large pneumoconiotic opacities of the parenchyma.

4.1.4 Comments

Provision is made for the recording of pertinent comments, "particularly if some other cause is thought to be responsible for a shadow which could be thought by others to have been due to pneumoconiosis" (and which, by implication, is not otherwise represented by facets of the Classification). Typically, a comment is written in a book kept for the purpose, and the fact is recorded on the form.

The other use for Comments suggested by the Classification is to note where technical quality may have affected the reading materially. In a variation from the letter of the scheme adopted locally, codes are used for specific technical defects, as described above in section 4.1.1, and written comments are thus less frequently employed for this purpose.

4.2 The IOM film-reading panel

Classification of appearances of pneumoconiosis and related abnormalities on chest radiographs is time-consuming and therefore, if performed by medically qualified persons, it is also expensive. In a clinical context, where the aim is diagnostic, it is essential that the inspection of radiographs be performed by medical staff. For epidemiological purposes, what is required is more descriptive than diagnostic, and the ILO classification scheme reflects this emphasis. The current guidelines (ILO, 1980) state explicitly, "The scheme is designed for classifying the appearances of postero-anterior chest radiographs.", "The Classification does not define pathological entities", "The object of the Classification is to codify the radiographic abnormalities of pneumoconiosis in a simple reproducible manner.".

It follows that a full knowledge of the physiology and pathology of the lung is not a necessary prerequisite for the successful application of the Classification, and that it is possible that persons without medical training can be trained to apply the Classification. This was demonstrated in practice by Peters *et al* (1973), who showed results from two non-medical readers that agreed more closely with those of their medically qualified trainer than did classifications of the same radiographs by two other experienced radiologists. This study thus also demonstrated how easily trainee readers can be influenced by the particular reading habits and interpretations of their trainer, in an area where even the most experienced and apparently expert radiologists can differ widely in their classifications (for an extreme example, see Jacobsen *et al*, 1984).

Copland *et al* (1981) reported on a study which was designed to establish whether radiological novices could teach themselves to use the Classification without expert guidance. Eight people without medical qualifications were selected at random from responses to newspaper advertisements, and were invited to train themselves, under controlled conditions, to use the 1971 revision of the ILO Classification scheme (ILO, 1972). The trainees were given: an introduction to the scheme, the standard radiographs and the booklet of guidelines; a short demonstration of the radiographic anatomy of the chest; and a glossary of medical terms used in the ILO booklet. No further help was given, and there were no meetings or discussions between readers during the self-training period.

The readers trained themselves by examining radiographs in conveniently sized batches, and classifying them by comparison with the standard radiographs, referring only to the ILO text for assistance. Amongst the radiographs presented was a set selected to represent a range of abnormality, and these were read by the readers on repeated occasions (although the readers were not aware of the identities of the films they classified). It was thus possible to monitor each reader's learning phase and the eventual stabilisation of their reading habits, as their confidence and speed of reading increased.

Copland *et al* (1981) judged their experiment a success, the most important evidence being the ability of the readers to produce classifications of films from the PFR research which related to estimates of exposure to airborne dust, in a way that was consistent with results from the earlier "definitive" readings of PFR films (Jacobsen *et al*, 1971); they inferred that readers trained in this way "...may generate valid data for epidemiological studies.". After the training period, therefore, the experiment was continued with the six readers who chose to remain, and these six became the first readers of the IOM film reading panel, as it is now known. They continued to read batches of films, with or without clerical help; each of the five readers who now remain currently assesses a batch of around 170 radiographs at a single session, according to the current revision of the Classification (ILO, 1980), which was adopted shortly after its publication.

The assessments made by the panel have proved useful in several large-scale epidemiological studies. Examples of the results they have obtained are given by

Soutar *et al*, 1981; Maclaren, 1985; Maclaren and Soutar, 1985; Soutar *et al*, 1986; Seaton *et al*, 1986; Louw *et al*, 1986. The work of the panel continues to be regarded as a cost-effective system for the measurement and differentiation of levels of pneumoconiotic abnormality on chest radiographs.

4.3 Organisation of the independent randomised film-reading exercises

4.3.1 Availability of films

It was intended to study all the chest radiographs taken since 1970 for the miners who had worked at colliery P during the 1970's. Of the 1416 men, the computer files holding earlier results of readings of survey radiographs indicated that 401 had attended all three of the 1970, 74 and 78 surveys and therefore had three films available; 392 had two films available; and 623 had only one film taken during that time. Exhibit 4.3.1 shows the details of recorded survey attendance for all of these men. Some of the 1974 films were taken during a follow-up study of men no longer working in the industry.

4.3.2 Design of batches of films

In total, the films to be read by the members of the IOM panel numbered 2610. At each reading session, a panel member routinely assessed a batch of (around) 200 films; if the batches are much larger than this, fatigue can be observed in the readers. It was therefore planned that each reader should complete her readings over thirteen weeks at one such batch each week. Three batches of 200 and ten of 201 films would accommodate the required numbers.

Exhibit 4.3.1. Table showing numbers of films available from three surveys, for 1416 men broken down by combinations of surveys attended, according to existing computer records. (In brackets are shown two films not recorded on computer but found in envelopes.)

No of films per man	Year of survey 1970	1974	1978	Total
One	433	(1)		433 (1)
		102		102
			88	88
Two	259	259		518
	18	(1)	18	36 (1)
		115	115	230
Three	401	401	401	1203
Total	1111	877 (2)	622	2610 (2)

Exhibit 4.3.2. Table showing compositions of batches of films for independent randomised reading exercise, in terms of numbers of men for whom different numbers of films were available.

Number of films per man 3	2	1	Batch size (films)	No. of batches	Total films
31	30	47	200	3	600
31	30	48	201	8	1608
30	31	49	201	1	201
31	31	48	203	1	203
			Total	13	2612

Each film was to be read and classified independently, taking the films one at a time. The most important consideration in the allocation of films to batches was considered to be the investigation of radiological change as evidenced by the readings for successive films from the same man. If such films were read on different occasions, it was apparent that differences between the readers' levels of performance would confound the effects of interest, and that such confounding would be minimised by allocation of all films for any man to the same batch.

The men were therefore stratified according to the number of films available for each, and the allocation to each batch made in proportion from each stratum. A little trial and error yielded the numbers in exhibit 4.3.2, which were taken as targets for the allocation.

Films are stored in large envelopes, each of which contains the whole series of films for an individual man. The envelopes required for this study were collected from storage and collated while the allocation of films to batches was being devised. It was discovered that in a few cases the contents of an envelope were not wholly consistent with the computer record of a man's appearances at the different surveys, with, most often, a film or films missing from the envelope. In three cases where a film was missing there was another film present which had been taken within a few months of the date of the relevant survey. It can sometimes happen that a survey film is of poor quality, and a replacement may then be taken as soon as convenient; this may also be done if a film is lost. Since the time span between the survey and the substitute film did not exceed a year, it was decided to proceed as if these particular films had been taken at the surveys, and to include them in the present reading exercise. For other men, single films or whole envelopes could not be found. Some were known to be on loan to the Pneumoconiosis Medical Panel, but most were simply misplaced, at least temporarily. The total number of films missing was 47. Men with missing films were allocated to batch 13, which was to be read last, so that additional searches could be made for them while the earlier batches were being read. In two cases, envelopes contained films from the period of interest whose existence had not been recorded on the computer file. The total number of films for which readings were desired was therefore updated to 2612.

The remaining men and their films were allocated to the batches according to the plan in exhibit 4.3.2, randomising within strata by an algorithm for fixed-size sampling based on one by Bebbington (1975).

Each film was then allotted its place in the order within the batch by simple randomisation. Previous film reading exercises had shown that, if two films from the same subject were read consecutively, the reader might notice the similarity. To avoid this, the batches were checked for adjacent films from subjects and where necessary these were separated arbitrarily.

Batches of films were made up according to the completed randomised batch lists. The lists were also used for the production of forms for recording the results of the film reading sessions; separate sets of forms were produced for each reader for each batch, each containing the preprinted film identities in their randomised order. An example of a prepared blank form is shown in exhibit 4.1.1. This is designed for recording assessments of films made according to the ILO (1980) classification scheme, as described in 4.1. The form departs in a few details from the letter of the ILO (1980) booklet. A detailed protocol for its completion is given in Appendix A; this is the protocol which applies to all reading sessions by the IOM panel.

The thirteen batches were read twice, in the same order, and without alteration of the positions of the films within each batch. The first reading exercise took place between January and May 1984, and the second exercise between May and September 1984. As noted above, it was not possible to obtain all the required films at the time of the batching, but all the missing films had been allocated to batch 13; by the time this batch was read for the first time, many of these had been located, and only 18 of the intended films were not available for reading then. Three of these were from one miner and were on loan to the Pneumoconiosis Medical Panel, and they were returned in time to be included in the second reading. The whereabouts of the remaining 15 remains unknown. In the time intervening between the two reading exercises, some of the films read in the first reading were unavoidably removed for an entirely different project, and were not returned for the second reading; these numbered 69. Taken with the three returned from the Medical Panel, this meant that 72 films were read only once, while the rest, numbering 2525, were each read twice.

4.3.3 An additional reading exercise to include films from the 1980 survey

All the films read as described in the previous section derived from the surveys which took place as part of the Pneumoconiosis Field Research programme in 1970, 74 and 78. No more PFR surveys were carried out at this colliery, but the colliery was visited in 1980 by the NCB Medical Service as part of their Periodic X-ray (PXR) scheme, and men attending then had a film taken. At the time the readings were being planned, however, the Institute had no knowledge of which men were both seen in 1980 and belonged in the study group.

It was obviously desirable to include films taken in 1980 for men in the study, for radiological assessment of any changes over the additional two years. Medical

Service do not use the same system of so-called 'X-ray' numbers, for the storage and organisation of their films, as that used by the IOM for PFR surveys. Instead, each man's envelope is updated with a new Film Serial Number (FSN) whenever a man is surveyed, and the envelope is stored under the most recent FSN. This meant that individuals could not be matched on the key identity number used on all the IOM's computer data files. From Medical Service, a listing was obtained of all men who attended for X-ray at the 1980 survey at pit P. It was then necessary to trace these nearly 600 men, by name and date of birth, in the IOM's manual card filing system. The list of X-ray numbers produced was then matched on the computer to identify those in the study population.

This matching identified 462 members of the study population who had attended for X-ray in 1980. These films then had to be located and withdrawn from the storage system of Medical Service, a task made more difficult by two factors. Firstly, Medical Service had recently moved premises and the reorganisation of the filing of films was not quite complete. Secondly, many of the men had continued in employment with the NCB at other sites after the closure of pit P, and some of these had attended for later X-ray examination at the routine surveys at those sites; the required films would then be stored under a different FSN from that on the supplied listing.

The card filing system kept by Medical Service provided a key, alphabetic by surname, to the most recent FSN in all but one case, who could not be traced. Envelopes were found for 448 of the men, although in one case the required 1980 film was missing from the envelope. In all, 15 of the films sought could not be found.

For 447 men who were in the study population, then, films from the 1980 Medical Service survey were discovered and brought to the IOM. It was decided that while the IOM panel were reading these they should also read films, mostly from the 1978 survey, which as described above were read only once during the main duplicate reading exercise. In order to build in some comparability with readers' performances at earlier readings, all films were included for each of the 65 men for whom a film had only been read once, even though some of these films had been read twice. They were amalgamated with the 1980 films to give a total of 590 films, and arranged at random into five batches with the constraint that all films for one man appear in the same batch. Three earlier films from two men could not be found for this exercise. The films from 1980 found to be missing were allocated to Batch 5 in the hope that some of them might be discovered before the batch was read. (In the event, none was.) The readings for this supplementary exercise took place between July and September 1986.

4.4 Reading films serially for progression

As has been described above, the desire to assess radiological change over the periods between surveys motivated the placement of different films from the same individual within one batch. Within the span of a single reading session, however, random fluctuations can and do occur in the reading patterns of even the most experienced readers; and these fluctuations may well be of a magnitude which could obscure relatively small progressions of profusion. While such progressions might not have immediate clinical significance, they could contain useful information on the magnitude of any dose-response relationship, and it would be of interest to quantify such progressions as precisely as possible.

The question of how best to describe and quantify radiological progression has been controversial, but the general consensus is that fluctuations in a reader's perceptions of the ILO standards, and the influence of factors such as the technical quality of individual films, will be minimised if a series of two or more films from one individual is assessed together; that "side-by-side" viewing of films from the same man, with their temporal order disclosed to the reader, will allow (in theory at least) the assessments to take account of film quality; and that the differences in observations of pneumoconiotic abnormalities across such a series will provide a more coherent description of the progression of disease within an individual.

The panel of readers who participated in the independent classification of the films had little or no experience in the side-by-side assessment of series of films. We therefore enlisted the cooperation of three readers who had the necesssary experience; each was a chest physician with many years of familiarity with occupational respiratory disease in general, and pneumoconiosis in particular. All three were also busy men, and it was apparent that the ideal of fullest information, which would be got by having all the films read in triplicate, was not a feasible objective. At the same time, it was anticipated that there might be differences in the reading habits of even such experienced readers, so it was not considered wise simply to split the reading burden into three parts; although this would minimise the reading requirement of each reader, comparisons between readers would be wholly confounded by such a scheme. A compromise between these extremes was reached by the application of techniques of experimental design, while the reading load was further limited by omitting chosen subsets of the study population from this exercise.

4.4.1 Selection of a subset of the men

Exhibit 4.4.1 displays the information available at this stage on the numbers of films available from the men's attendance at different combinations of the four surveys. This table differs in structure from exhibit 4.3.1 in that it includes information on the availability or otherwise of a film from the 1980 Medical Service survey. But collapsing over 1980 film availability would not reproduce exhibit 4.3.1, because exhibit 4.4.1 incorporates updated information on availability of earlier films which came to light during the batching for the earlier exercises. (It also includes an extra 1980 film which had been discovered for one of the study population. This film had not been identified in the matching of PFR and Medical Service records, because the man's surname had been spelt Miller on one computer file and Millar on the other.)

Because the focus of this reading exercise was on change within an individual between surveys, and because the limiting resource was the readers' time, all singleton films were eliminated. One further economy was effected by omitting the doubletons from the 1970 and 1974 surveys, since they predated the period of the main peak in quartz exposures.

Design of this exercise thus involved allocation of 632 series of two or more films, totalling 2035 films, to conveniently-sized batches; and the allocation of these batches to the three readers.

4.4.2 Design of batches of films

Exhibit 4.4.2 shows the structure of the design chosen to allocate batches to readers. It required 18 batches; of these, 12 would be read by only one of the

Exhibit 4.4.1. Table showing numbers of films available from four surveys, for 1414 men broken down by combinations of surveys from which films were found (two men had no films available).

No of films per man	1970	1974	1978	1980	Total
	427				427
		80			80
One			37		37
				2	2
Two	236	236			472
Total unused	663	316	37	2	1018
	8		8		16
	6			6	12
Two		51	51		102
		23		23	46
			50	50	100
	119	119	119		357
Three	21	21		21	63
	11		11	11	33
		66	66	66	198
Four	277	277	277	277	1108
Total used	442	557	582	454	2035
Total films	1105	873	619	456	3053

Exhibit 4.4.2. Intended and actual allocation of batches to medical readers for serial reading exercise.

Batch no.	Reader code 010	011	013
1	x		i
2	x		i
3	x		
4	x		
5	x		
6	x		
7	x	x	
8	x	x	
9		x	
10		x	
11		x	
12		x	
13	s	x	i
14	s	x	i
15	s		i
16	s		i
17	s		i
18	s		i

Key to entries:

x - Reading executed as intended

i - Reading intended but not executed

s - Substitution for intended reading

Exhibit 4.4.3. Table showing compositions of batches of films for serial reading exercise, in terms of numbers of men for whom different numbers of films were available.

Batch type	Number of films per man 4	3	2	Batch size (films)	No. of batches	Total films
A	16	12	7	114	7	798
B	15	12	9	114	2	228
C	15	12	8	112	8	896
D	15	13	7	113	1	113
				Total	18	2035

readers (four batches each), and the remaining six would be read by pairs of readers, allocated as in a simple balanced incomplete blocks design (Cochran and Cox, 1957). This design would give a manageable batch size; 17 batches of 113 films and one of 114. Each reader would assess eight batches.
For the allocation of series of films to construct batches of these sizes, it would be necessary to stratify them according to the number of films in each series, as was done for the earlier exercises, and to arrange them by randomisation in a non-systematic order. There was a further practical constraint which conditioned the allocation, however. The recording forms used for routine accumulation of film-reading data from independent randomised readings of single films were designed to accommodate ten films each (see exhibit 4.3.3), and were printed by a Fortran 77 computer program which used lists of identities of films allocated to a batch. It was desired to use an adaptation of this program, rather than to write a new one from scratch, to produce recording forms for this exercise. It was also desired to have films from one series recorded in temporal order on adjacent lines of the forms, and to separate different series on the same page by a blank line.

The design of this allocation took place in stages. Firstly, possible allocations of men with two, three and four films in similar proportions were examined, and the combination shown in exhibit 4.4.3 was adopted as suitable; this was convenient, but one of many similar possibilities. Secondly, the different ways of arranging series within the ten lines of a form were identified, and four of them were taken as candidates; (4,4), (4,3), (3,3,2), and (3,2,2), where the numbers represent numbers of films in a series. Within each of the four types of batch structure, the numbers of each of these had to satisfy certain linear constraints in the numbers of films and numbers of series. Suitable solutions to these constraints were obtained, as shown in exhibit 4.4.4. As before, these were not unique solutions, but they gave a convenient basis for an allocation scheme with the desired

Exhibit 4.4.4. Allocation of series of films to batches, and pages of forms within batches, for serial reading exercise. Tabulated are the numbers of pages containing each of four combinations of lengths of film series, as allocated to each of the four batch types designated in exhibit 4.4.3.

Page type	Page structure	Batch type (see exhibit 4.4.3) A	B	C	D	
a	4 - 4 -	6	6	5	5	
b	- 4 - 3 -	4	3	5	5	
c	3 - 2 - 3	3	3	2	3	
d	2 - 3 - 2 -	2	3	3	2	
	No. of batches	7	2	8	1	Total
	Total no. of men	245	72	280	35	632
	Total no. of films	798	228	896	113	2035

properties, to arrange different patterns or combinations of film numbers on the printed page.

These "page patterns" were then randomised within each batch type, to give a skeleton allocation scheme for series to pages within batches, in terms of the size but, not yet, the identity of each series. Selection of individual series to complete the pattern was then performed from randomised lists, stratified by number of films in a series. This was achieved by a mixture of simple programs written in FORTAN and in GENSTAT (Alvey *et al*, 1983).

The batch lists were then printed for the construction of the batches, and the forms printed; a sample page is shown as exhibit 4.4.5. It will be noticed that the layout of the form differs from that in exhibit 4.3.3, since it was decided to use a short form of the Classification, concentrating on small and large pneumoconiotic opacities, and ignoring pleural and other abnormalities. A copy of the protocol supplied to the readers, and making explicit the conventions for recording their assessments, is given in appendix B.

It was intended that each of the three readers should make appointments, as convenient to them, to read batches of films. However, one of the three readers was prevented, by pressure of other work, from being able to read at around the same period of time as the other two; and it became apparent that completion of this exercise was likely to be delayed by several months if we waited for this reader to have the necessary amount of time available to read these films. It was therefore decided to ask the other readers if they could help, and reader 010 kindly agreed to complete the missing readings, excluding the two batches he had already assessed. This last-minute pragmatic arrangement meant that the readings did not conform to the extended balanced incomplete blocks design; the actual

Exhibit 4.4.5. Example of form used for recording readers' assessments of radiographs in the serial reading exercise. The identities of the films were preprinted on the forms by computer program, from the batch lists.

```
I.O.M SERIAL FILM READING EXERCISE  PAGE:   3

READING ID: XRES   BATCH ID: 00

READER ID: 000

¦-------------------------------------------------¦
¦ FILM  ID  ¦SEQ¦        ¦ S M A L L  O P S¦   ¦  ¦
¦-------------------------------------------------¦
¦               ¦ QUAL¦ N¦PROF ¦ZONES¦TYPE ¦LQE¦ C¦
¦               ¦  ¦  ¦  ¦     ¦     ¦     ¦OPS¦  ¦
¦               ¦12¦1-¦ X¦ 012 ¦  X  ¦P.Q.R¦O.A¦ X¦
¦               ¦34¦9 ¦  ¦ 3-+ ¦ R L ¦S.T.U¦B.C¦  ¦
¦-------------------------------------------------¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦PFRP1714-70 017¦  ¦  ¦  ¦  /  ¦  .  ¦  /  ¦   ¦  ¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦-------------------------------------------------¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦PFRP1714-74 018¦  ¦  ¦  ¦  /  ¦  .  ¦  /  ¦   ¦  ¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦-------------------------------------------------¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦PFRP1714-78 019¦  ¦  ¦  ¦  /  ¦  .  ¦  /  ¦   ¦  ¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦-------------------------------------------------¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦PFRP1714-80 020¦  ¦  ¦  ¦  /  ¦  .  ¦  /  ¦   ¦  ¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦-------------------------------------------------¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦---------------¦  ¦  ¦  ¦  /  ¦  .  ¦  /  ¦   ¦  ¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦-------------------------------------------------¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦PFRP4171-70 021¦  ¦  ¦  ¦  /  ¦  .  ¦  /  ¦   ¦  ¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦-------------------------------------------------¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦PFRP4171-74 022¦  ¦  ¦  ¦  /  ¦  .  ¦  /  ¦   ¦  ¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦-------------------------------------------------¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦PFRP4171-78 023¦  ¦  ¦  ¦  /  ¦  .  ¦  /  ¦   ¦  ¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦-------------------------------------------------¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦PFRP4171-80 024¦  ¦  ¦  ¦  /  ¦  .  ¦  /  ¦   ¦  ¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦-------------------------------------------------¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦---------------¦  ¦  ¦  ¦  /  ¦  .  ¦  /  ¦   ¦  ¦
¦               ¦  ¦  ¦  ¦     ¦  .  ¦     ¦   ¦  ¦
¦-------------------------------------------------¦
```

allocation of batches read by the two readers is shown in exhibit 4.4.2, where it can be seen that these readers still had an overlap of four common batches, on which comparisons of their reading patterns could be made.

5 ORGANISATION AND VALIDATION OF DATA

The data required for this project comprised existing data from the PFR research programme, and new data obtained from the film reading exercises whose design was described in chapter 4. In this chapter, sections 5.1 to 5.4 describe how they were brought together and organised into a computer database designed specifically for the project, using the facilities of the database management system SIR (Robinson *et al*, 1980). The structure of a SIR database is defined by "schema definitions"; Appendix C reproduces a selection of the schemas defining the current database, and the record type numbers quoted in this chapter refer to those schemas.

Descriptions of the distributions of the existing data extracted and brought together for this project are presented in 5.5. Descriptions of the results of the various film-reading exercises will be found in 7.1 and 7.3.

5.1 Personal data

Throughout the PFR research, each participant coalworker was identified by a code which combined a letter identifying the colliery with a four-digit number allocated sequentially within each colliery. (This identification was quite distinct both from any national reference such as National Insurance number and from any industry identifier such as that used in the payroll systems.) All data on the men in the study, as they were loaded into the project database, were identified and indexed by their PFR code number.

Name, National Insurance number and date of birth were extracted from existing files for each man. Dates of birth recorded in different files were often recorded on different occasions, and multiple versions of these were cross-checked for consistency. In the few cases of discrepancy, the date which appeared on the majority of records was taken. Personal data were held in record type 16 (Appendix C).

5.2 Data from medical surveys

The principal features of the PFR medical surveys have been described in 1.2. Questionnaire data on smoking habits, and data on lung function and anthropometric measurements, were extracted for the men in the study from the records of the 4th, 5th and 6th PFR surveys, and loaded into record type 15 within the database (Appendix C). Extensive work on all questionnaire and lung function data, carried out in connection with another recent PFR study, had given assurance on the quality of these data as currently stored.

There was, however, one feature of the 5th survey lung function data from this particular colliery which was already known to be unusual. Normally, all lung function measurements were carried out by one experienced technician, but at the fifth survey at this colliery a new, inexperienced technician was employed. Initial tests suggested that her results were comparable with those of existing technicians, but to give additional information on the situation in the field a more experienced technician completed about 120 of the measurements at this survey. Subsequent examination of the data suggested strongly that the inexperienced technician was recording results at a slightly lower level than the experienced one. In order that this could be allowed for if desired, an indicator was placed on file with the 5th survey lung function data to show which technician had recorded them. No other survey was affected.

5.3 Data on occupational exposures

The defined study population was all men in employment at PFR colliery P who attended the 4th, 5th or 6th survey. It was decided to retain individually the exposure data from inter-survey periods ISP 3, ISP 4 and ISP 5, which covered the periods immediately before the 4th and up to the 6th surveys, and to form a cumulative total of any exposure from ISP 0, ISP 1 and ISP 2, that is up to the 3rd survey. The latter data were extracted from existing files and the measured and unmeasured exposures added together, and the cumulative totals were formed of estimated exposures to respirable dust, ash, and quartz. The data from ISP3 and ISP 4 were treated similarly, except that the measured and unmeasured components of each exposure were kept and stored separately. All of these data were loaded into record type 10 in the database (Appendix C).

Measured exposure data from ISP 5 were available on computer in more detail than from the other inter-survey periods. From around the 5th survey, times worked in occupational groups in each quarter had been retained on file, whereas times from earlier periods had been cumulated over whole inter-survey periods before storage. Occupational group concentration measurements and corresponding compositional data had been stored by quarter since after the 3rd survey, so that it was possible in principle to associate these directly with measured times by quarter. The individual times and the individual environmental measurements were loaded to record types 11 and 12 respectively in the database (Appendix C). Unmeasured exposure data for this period had not been processed onto computer, and the small amount of such data relevant to colliery exposure was extracted from manuscript records. Although negligible in magnitude, it was also loaded to the database for completeness, in record type 14. Record type 13 was used to hold some few records of measured time corresponding to occupational groups for which no environmental data were available.

5.4 Organising data from film reading exercises

So that the database could stand as a self-contained record of the structure of the study population and its data, the information on survey attendance and consequent availability of films which was used to design the film reading exercises, updated in the light of the ability actually to find those films, was loaded to record type 1 (Appendix C). These data formed the basis of the case structure of individuals included in the database. Information on the allocation of films to batches, and the dates of the reading sessions, were assigned to database record types 2, 3 and 4 and loaded into the database.

Data from the three film reading exercises by the IOM panel, described in 4.3, were punched directly into computer files. They were then loaded into record types 5, 6 and 7, with almost identical schemas; only record type 5 is shown in Appendix C. The shorter readings by the medical reders, described in 4.4, were also punched, and loaded into the simpler record type 8.

The alphabetic and numeric codes which are used to record the various features covered by the ILO (1980) classification were rigidly defined by local protocols, based on the principles laid down in the ILO booklet, and reproduced in Appendix B. As the data were loaded to the database, any which did not conform to the specification, either by invalid codes or because of logically inconsistent combinations of codes, were detected by validation instructions inserted in the schema definitions (Robinson *et al*, 1980). Invalid records thus identified were denied entry to the database, and listed. By reference to the original coding forms, queries were resolved and the corrected data entered, the process being repeated as necessary until all the data passed all the validity and consistency checks.

5.5 Summary descriptions of data

This section provides brief summary descriptions of the distributions of the data which were already available for the men in the study population, and which were extracted from the existing computer records of the PFR research programme, and brought together for the purposes of this project. In 5.5.1 to 5.5.4 the distributions of age, smoking habits, height, weight, lung function and occupational exposure to dust and quartz are tabulated and described, with comments.

5.5.1 Distributions of age and smoking habits

Exhibit 5.5.1 tabulates the men who appeared and completed questionnaires at each of the 4th, 5th and 6th PFR surveys, by age and smoking habits. Because the dates on which each individual man was surveyed were not readily available, the convention has been adopted to calculate the ages on the first day of the months in which most of each survey took place; these were December 1970, October 1974, and November 1978. Adoption of such conventions is common and is not considered likely to introduce any detectable distortion into any of the analyses.

The data on smoking habits utilised in this table are those collected at the survey in question; thus a man who appeared at different surveys and replied differently to the questions on smoking, for example if he had stopped smoking since the previous survey, will contribute to different smoking groups at the different surveys. Within each survey, however, the breakdown is exhaustive and mutually exclusive.

Exhibit 5.5.1. Table showing breakdown of numbers of men completing questionnaires at each survey by smoking habits and by age (at start of month of survey) in 10-year groups.

Survey	Smoking habit	Age at survey 15-24	25-34	35-44	45-54	55-64	65-	Total
1970	Nonsmoker	36	34	22	31	27	0	150
	Exsmoker	5	15	24	49	41	0	134
	Smoker	77	105	161	272	211	0	826
	Unknown	1	0	0	0	0	0	1
	Combined	119	154	207	352	279	0	1111
1974	Nonsmoker	33	33	36	21	18	4	145
	Exsmoker	10	17	21	48	43	1	140
	Smoker	64	58	110	248	101	9	590
	Unknown	0	0	2	2	0	0	4
	Combined	107	108	169	319	162	14	879
1978	Nonsmoker	28	19	20	20	18	0	105
	Exsmoker	8	12	13	38	32	0	103
	Smoker	54	35	80	123	76	0	368
	Unknown	0	1	14	20	11	0	46
	Combined	90	67	127	201	137	0	622

Exhibit 5.5.2. Table summarising data from each survey on height and weight for men completing tests.

Variable	Survey	No. of Men	Mean	Standard deviation
Height (cm)	1970	1110	170.2	6.4
	1974	876	170.4	6.8
	1978	578	171.1	6.7
Weight (kg)	1970	1110	73.2	11.5
	1974	876	73.6	11.7
	1978	578	75.1	11.6

From both the 4th and 5th surveys, there were a very few men for whom medical and other questionnaire data were not found, even though the computer records indicated that they had attended the relevant survey. For the 6th survey, however, there were 46 men without questionnaire data. These corresponded to a batch of questionnaires which are believed to have been lost due to clerical error before the data they contained were entered to the computer; there seems no way that these lost data can now be regained. It is notable that the age distribution of the men with lost questionnaires looks very unlike a random sample of those attending the survey, and this apparently systematic loss of data has consequences for analyses which refer to smoking data from the 6th survey.

The data in exhibit 5.5.1 display several trends worthy of mention. The first and most obvious is the contraction of the workforce over the period; the number at the 6th survey is little more than half that at the 4th. It should be noted, however, that although the attendance at PFR surveys was generally high, it never reached 100%; that it was higher in the earlier years of the research than in the seventies, when industrial unrest was more prevalent; and that as the closure of this colliery was already foreseen, disaffection amongst the workforce would be expected to be greater than in times of more optimistic outlook. Assuming, though, that the surveyed population was not entirely untypical of the population employed at the time of each survey, it is noteworthy that the reduction in numbers over the period was relatively greater in the older men, as would be expected in a time of contraction.

The 4th and 6th surveys included nobody over the age of 65, which was the normal retirement age throughout this period. However, 14 men over 65 attended at the 5th survey; this survey combined routine PFR examinations of current employees with follow-up surveys of men who had left the industry, and we included data from follow-up examinations in this study. All the men for whom follow-up was attempted had attended the very first survey, which took place at this colliery in February 1954. The present study population included only those with routine examinations at one of the three surveys in the 1970s, and so no worker was included solely on the basis of an attendance for follow-up. Most of these men had left the colliery after attending the 4th survey. The oldest ex-worker attending the 5th survey was 68.

The distribution of smoking habits across the age groups is also of interest. As is often observed in industrial populations, there was a preponderance of current smokers, who constituted nearly 75% of those at the 4th survey. The corresponding figure at the 6th survey was 64%; this change is consistent with the increased proportion of young non-smokers generally, but in this data set appears to be more clearly associated with a preferential loss, over the intervening years, amongst the oldest smokers. At all surveys, the proportion of non-smokers was considerably lower among the younger men than among the older.

5.5.2 Distributions of height and weight

Data on the heights and weights of those attending the surveys are summarised in exhibit 5.5.2. These data are generally unremarkable, the only trend being towards a slight increase with time in average height. This is consistent with the lower proportions at the later survey of older workers, whose physiological development predated the National Health Service and modern nutrition.

5.5.3 Distributions of lung function variables

Data available on the lung function measurements taken at the surveys are summarised in exhibit 5.5.3. Each man performed one practice blow, and then three blows for the purpose of measurement; on occasions, a man may not have been able to complete three such blows, and some men failed to complete even one. In a few cases, the complete expiration necessary for the assessment of FVC was judged unsatisfactory from the spirometer trace, even though the initial part of the trace, from the first second of expiration, was suitable for the assessment of FEV. All men with at least one satisfactory measurement are included in the tabulation of each variable from each survey. The mean of between one and (usually) three values was calculated for each man, and the tabulated values are unweighted means of these means.

At each survey, the highest values for both FEV and FVC are amongst nonsmokers, with little difference between the ex- and current smokers. This observation is consistent both with the higher proportion of younger workers among the non-smokers, and with the known capacity of exposure to tobacco smoking to induce loss of lung function. The relative importance of these two factors cannot be discerned from this table; detailed statistical analysis would be necessary, to estimate the effects simultaneously.

Another notable feature of the data in exhibit 5.5.3 is the apparent transitory decline in the mean values at the 5th survey. This is most evident in the mean FVCs, and is believed to be an artifact of measurement due to the inexperience of the technician who performed most of the 5th lung function measurements; see 3.2 above. Such an artifact has a potentially serious effect on inferences from observed patterns of change in lung function between surveys.

Exhibit 5.5.3. Table summarising data from each survey on lung function measurements (mean of up to three satisfactory blows) for men completing tests, broken down by smoking habits.

Variable	Survey	Smoking habit	No. of Men	Mean	Standard deviation
FEV (1)	1970	Nonsmoker	150	3.70	0.83
		Exsmoker	131	3.30	0.84
		Smoker	824	3.24	0.81
		Combined	1105	3.31	0.83
	1974	Nonsmoker	144	3.70	0.84
		Exsmoker	139	3.20	0.82
		Smoker	588	3.17	0.81
		Combined	871	3.27	0.84
	1978	Nonsmoker	105	3.71	0.82
		Exsmoker	102	3.24	0.80
		Smoker	366	3.23	0.82
		Combined	573	3.32	0.84
FVC (1)	1970	Nonsmoker	147	4.66	0.86
		Exsmoker	129	4.37	0.85
		Smoker	815	4.33	0.84
		Combined	1091	4.38	0.85
	1974	Nonsmoker	138	4.36	0.88
		Exsmoker	138	4.03	0.85
		Smoker	580	4.01	0.83
		Combined	856	4.07	0.85
	1978	Nonsmoker	105	4.68	0.90
		Exsmoker	102	4.27	0.86
		Smoker	364	4.27	0.86
		Combined	571	4.34	0.88

5.5.4 Distributions of dust exposure variables

Exhibit 5.5.4 summarises the data on the total working times recorded for each man in the periods up to the 3rd survey in 1964, and between the 3rd, 4th, 5th and 6th surveys thereafter; also summarised are the estimates derived from these times of individuals' exposures to respirable coalmine dust, and to the proportion of that dust measured as quartz.

The values for time are derived from the numbers of recorded shifts, converted at a standard conventional rate of 7.25 hours per working shift, and expressed here in units of one thousand (1000) hours. These were totalled across the different occupational groups in which time was recorded for each individual, and the figures shown are the mean and the maximum of these totals. In ISPs 3 and 4, data were kept separately concerning times recorded directly from the colliery payroll systems (designated as "measured") and from later interviews with the men, at which working history questionnaires were applied ("unmeasured"); the totals of measured and unmeasured times for each man are also summarised. By definition of the study population, most of the data from 1970 onwards were collected directly and are designated "measured". There were a few men who attended the follow-up survey at 1974, and these would have had a questionnaire applied. There were larger numbers of men with some unmeasured time in ISP3, between 1964 and 1970; during this period, transfers from the closures of other collieries were taking place, and these data are likely to concern time at other collieries, or time spent at this colliery prior to a man's inclusion in the PFR research population, or time for which data had been lost and was later recovered by interview. During ISP 5, when the colliery popultion was declining, almost all the men's time was accounted for direct from the payroll, and the number of men with any computer record of unmeasured time was very small; these data were

Exhibit 5.5.4. Summary of data on exposure to the non-quartz and quartz fractions of respirable dust for all 1416 men in study population. Tabulated are the mean and maximum values, and the number of non-zero values, for each variable. Data derived from questionnaire on working history (unmeasured) are separated from routine monitoring data (measured) after 3rd survey.

Variable	Period	Mean	Max	No. >0.0
	Up to 3rd survey	26.62	80.65	1118
	ISP 3* measured	7.10	34.83	1220
	ISP 3* unmeasured	1.38	11.66	294
	ISP 3* total	8.48	34.83	1256
Time (1000 hrs)	ISP 4 measured	3.86	10.46	1318
	ISP 4 unmeasured	0.11	6.96	47
	ISP 4 total	3.98	10.46	1327
	ISP 5 measured	3.49	11.23	940
	Up to 3rd survey	34.48	144.41	1118
	ISP 3* measured	10.22	48.01	1220
Respirable	ISP 3* unmeasured	1.74	33.89	294
non-quartz	ISP 3* total	11.96	48.01	1256
dust				
exposures	ISP 4 measured	5.17	27.80	1318
(g.hr.m-3)	ISP 4 unmeasured	0.16	21.33	47
	ISP 4 total	5.32	27.80	1327
	ISP 5 measured	3.66	21.74	940
	Up to 3rd survey	1.756	7.560	1118
	ISP 3* measured	1.040	5.835	1220
	ISP 3* unmeasured	0.123	3.022	294
Respirable	ISP 3* total	1.163	5.835	1256
quartz				
exposures	ISP 4 measured	0.656	7.893	1317
(g.hr.m-3)	ISP 4 unmeasured	0.021	3.829	47
	ISP 4 total	0.678	7.893	1326
	ISP 5 measured	0.505	5.100	940

* ISP 3 values include only 1415 men

omitted from consideration.

The last column displays the number of men for whom a non-zero time was recorded in the measured and unmeasured categories in the different periods. The great majority of men had recorded exposures in ISPs 3 and 4, and fewer had times in ISP 5, as expected from knowledge of the workforce contractions. Consideration of the numbers of non-zero values for the measured, unmeasured and total times shows that while 1220 men had non-zero measured times in ISP 3, inclusion of unmeasured times produced non-zero values for only another 36 men. In ISP 4 only 9 more men had non-zero values on inclusion of the unmeasured time. In general, then, the unmeasured times for ISPs 3 and 4 served to fill in sporadic periods of unrecorded times.

The second and third sections of exhibit 5.5.4 contain similar summaries of the estimated exposures to the quartz and non-quartz fractions of the respirable dust. Since these exposures are calculated from the times recorded in the various occupational groups, they show the same numbers of non-zero values, with the exception of ISP 4 measured (and total) quartz exposures which are one less; here, one man had time allotted to an occupational group with a low dust concentration and a zero recorded proportion of quartz.

The footnote to exhibit 5.5.4 shows that the values for ISP 3 are based on only 1415 men. Initial data summaries showed that the whole (quartz + non-quartz) dust exposure for one man was around 110 $g.hr.m^{-3}$, over twice the next highest value. Investigation of the individual recorded times from which this value had been cumulated revealed an impossibly large total number of recorded shifts over all the occupational groups, and it was concluded that the ISP 3 figures were unreliable for this man, and they were excluded from analyses using the ISP 3 exposures.

It is not possible to discern from exhibit 5.5.4 whether those who had the highest exposures to dust also had the highest quartz exposures. In fact, the compositional data showed a considerable degree of variation in the proportions of quartz in the dust exposures of different men at different periods. These are summarised in exhibit 5.5.5. For each man in each period, and retaining the distinction between measured and unmeasured time in ISPs 3 and 4, the cumulated quartz exposure was expressed as a percentage of the cumulated whole dust exposure; and the table shows the minimum, arithmetic mean, and maximum of these percentages. (Since the concept of a percentage is only defined for a positive denominator, values corresponding to zero exposures are undefined, and the means in exhibit 5.5.5 were calculated only over the defined values, whose numbers are shown in the table. This contrasts with the tabulations in exhibit 5.5.4, where zero exposures were included in the calculation of means.)

Exhibit 5.5.5. Summary of data on proportion of quartz in non- zero exposures to respirable dust. Tabulated are the minimum, mean and maximum values of that proportion, expressed as a percentage, and the number of values over which they were calculated.

Variable	Period		No of Values	Min	Mean	Max
	Up to 3rd survey		1118	0.36	4.78	7.58
	ISP 3	measured	1220	2.80	7.46	17.54
Fraction	ISP 3	unmeasured	294	1.82	5.13	9.12
of quartz	ISP 3	total	1256	1.82	7.23	17.54
in respir-						
able dust	ISP 4	measured	1318	0.00	7.78	29.36
(%)	ISP 4	unmeasured	47	2.40	7.60	15.22
	ISP 4	total	1327	0.00	7.77	29.36
	ISP 5	measured	940	1.44	8.75	31.79

6 STATISTICAL METHODS

6.1 The analysis of film-reading data

Data collected by the application of the ILO (1980) classification scheme are, by their nature, not recorded as interval scaled variables, but as nominal variables. Those which attempt to quantify the apparent severity of any observed pneumoconiotic abnormalities, such as the profusions of small and large opacities, and the extents of pleural plaques, diffuse pleural thickening, and pleural calcification, are scored as ordered categorical variables; the increasing categories imply greater degrees of abnormality, even if it is not clear how these categories relate to some underlying continuum of severity of pathological change. Reconciliation and summary of assessments of the same films by different readers, often with demonstrably different patterns of reading habits, complicates the problem.

The analysis of ordered categorical responses has long presented problems for rigorous statistical analysis, particularly where it is desired to relate the response to explanatory variables in an analogue of the regression methods used for the same purpose for interval-scaled responses. Jacobsen (1975) reviewed the methods available at that time, for the analysis of categories of radiological profusion. One simple possibility covered there is the assigning of arbitrary scores, for example based on the natural integers, to the categories, so that they may be treated as if they were measured on an interval scale. Another system of assigning scores transformed the observed distribution of responses to correspond to the frequencies of a Gaussian probability distribution; this, however, required the not always reasonable assumption that the Gaussian described the distribution of abnormality in the study population (rather than of the unexplained residual variation, which is assumed Gaussian in linear regression models). The latter method has not proved popular in the intervening years, no doubt in reflection of its unreasonable assumptions. The assignation of whole integers has often been used, implicitly or explicitly, in presentations of results, or in the calculation of simple arithmetic indices such as the "progression index" used by the then NCB Medical Service in its annual reports, for monitoring and comparing the progression of simple pneumoconiosis in British coalminers (Liddell, 1974). However, the assumption implicit in its use, that the categories represent some underlying categories which are in some sense equal in size or equidistant on some continuum, has likewise militated against the development of statistical models based on such assumptions. More recent advances have adopted an approach based on parametric modelling of the random processes underlying the development of abnormalities and their classification by readers. The most successful have all been in some sense or another extensions or generalisations of logistic regression.

6.1.1 Logistic regression for binary variables

Logistic regression models (Cox, 1970) have developed for the analysis of binary response variables, by expressing the logistic transform (or logit) of the expected value of the response as a linear function of explanatory variables. Specifically, if there are n random variables Y_i with $\Pr(Y_i = 1) = \pi_i$, we seek to express π_i as a function of a linear combination $\underline{\beta}^T \underline{x}_i$ of the known explanatory variables $\underline{x}_i$, where $\underline{\beta}$ is a vector of regression coefficients which have to be estimated. The logistic function is commonly chosen as the linking function, because of its attractive computational properties, although other functions have also been used. The logistic may be written as

$$E[Y_i] = \Pr(Y_i=1) = \pi_i = \frac{\exp(\underline{\beta}^T \underline{x}_i)}{1 + \exp(\underline{\beta}^T \underline{x}_i)} \qquad (6.1)$$

Assuming the random variables Y_i are independent, the log-likelihood function is

$$L(\pi_1 \ldots \pi_n) = \sum_{i=1}^{n} \{y_i \log(\pi_i) + (1-y_i)\log(1-\pi_i)\} \qquad (6.2)$$

into which the parameterisation (6.1) may be substituted directly. Estimation of the parameters $\underline{\beta}$ proceeds by maximising this log-likelihood with respect to the parameters.

The logistic regression model is one special case of a generalized linear model (McCullagh and Nelder, 1983); This family of models is characterised as an extension of linear regression models to permit error distributions other than the Gaussian, so long as the distribution is a member of an exponential family, and some differentiable monotonic function is assumed to act as a link between the expected values μ_i and the linear predictor $\underline{\beta}^T\underline{x}_i$. Maximum likelihood estimates can be obtained by iterative application of least-squares estimation techniques, as detailed in the next section.

When a single reader classifies each of several films without reference to any other film, it is reasonable to assume statistical independence of the classifications (although it would be prudent to test the assumption whenever possible). If the classification is or can be dichotomised into a binary response, the logistic model provides a useful framework within which research questions about the response and its relationship to explanatory variables may be addressed. The principal objection to its use in the analysis of small opacities profusion, or indeed any ordered response of more than two categories, is the loss of information inherent in dichotomising at a single point on the ordered response scale. Some insight into the seriousness of this loss can be gained by the pragmatic approach of applying the same analysis with different cut-points, and comparing the results.

Recent years have seen the development of extensions and generalisations of these methods, some of which are described in 6.3 and 6.4.

6.1.1.1 Estimation by the method of iteratively reweighted least squares

Following Nelder and Wedderburn (1972), McCullagh and Nelder (1983) show how the usual least-squares methods for the estimation of linear regression coefficients may be adapted to estimate the coefficients of generalized linear models. The method can be derived by a Taylor series linear approximation to the log-likelihood function. From a provisional estimate $\underline{\beta}_0$ of the coefficients, estimates of the linear predictor $\eta_0=\underline{\beta}_0^T\underline{x}$, and of the fitted value μ_0 can be derived. Other quantities required are the derivative of the link function, $d\eta/d\mu$, and the variance function $V(\mu)$ of y, both evaluated at the current provisional estimate. Then an updated estimate of the coefficients is obtained by regressing the adjusted variate

$$z_0 = \eta_0 + (y-\mu_0)(\frac{d\eta}{d\mu})_0 \qquad (6.3)$$

on the explanatory variables $\underline{x}_i$ with weights

$$w_0 = (\frac{d\mu}{d\eta})_0^2 V^{-1} \qquad (6.4)$$

The process is repeated as necessary until convergence to the required precision is achieved. Computer programs which have inbuilt facilities for such iterative estimation include BMDP (Dixon, 1983), GLIM (Baker and Nelder, 1978) and GENSTAT (Alvey *et al*, 1983).

6.1.2 The Rasch model

The binary logistic model extends readily to the situation where more than one observer independently classifies the same binary response. The Rasch model

(Andersen, 1980) was developed in the field of psychometrics, to model the probabilities of subjects completing a series of tests; but by substituting for example radiographs for subjects, and readers' opinions for tests, a direct transfer of applications is achieved.

The Rasch model, applied here, would assign to each reader a parameter α_j, so that the differences in these reflect the differing propensities of the readers to classify films as showing an abnormal feature; the differences in "reading levels" so often observed when human observers are employed as measuring instruments. The "amount" of a particular abnormality present on each film is parameterised by θ_i. The Rasch logistic model expresses the expected values

$$\Pr(Y_{ij}=1) = \frac{\exp(\theta_i+\alpha_j)}{1 + \exp(\theta_i+\alpha_j)} \qquad (6.5)$$

Assuming the classifications are independent, the log likelihood can be written as

$$L = \{\sum_i \theta_i y_{i.} + \sum_j \alpha_j y_{.j}\} - \sum_i \sum_j \log\{1+\exp(\theta_i+\alpha_j)\} \qquad (6.6)$$

This model defines an exponential family with canonical parameters θ_i and α_j. Some linear constraint on the parameters is necessary to eliminate a joint indeterminacy. However, straightforward maximum likelihood estimation of the parameters is problematical, in that the estimates of the reader differences α_j are inconsistent as the number of films $n \to \infty$.

The canonical exponential form of (6.6) shows that $T_i=Y_{i.}$ is sufficient for θ_i. Therefore, by conditioning on $T_i=t_i$, the observed totals of positive responses for each film, Andersen (1980) shows that the conditional log likelihood for the reader parameters can be derived as

$$L_c = \sum_j \alpha_j y_{.j} - \sum_k n_k \log\{ \sum_{y_{i.}=t_i} \exp(\sum_j \alpha_j y_{ij})\}$$

where n_k is the number of films for which $t_i=k$. For any film, the number of readers recording abnormality can lie between 0 and r, the number of readers. The n_k films with any particular value k form a set within which the information about the reader parameters α_j is contained in the data on which k readers made the positive assessments. The conditional likelihood equations have a finite solution set which is unique if $0 < y_{.j} < n$ for all j, that is if no reader classifies all the films as either normal or abnormal; these cases produce unbounded estimates for the reader involved, but since such a reader provides no information about the differences between films, his assessments could be ignored without loss.

A likelihood-ratio test can be derived (Andersen, 1980) for the adequacy of the assumption of constant reader differences α_j over different values of θ_i. The data set is indexed by, and can be arranged in order of, the values of t_i=k. Partitioning the conditional likelihood for different values of k yields separate estimates of the α_j within each partition (combinations of adjacent partitions can be formed to avoid sparseness of data). The difference between the likelihoods from the overall estimates and the partitioned estimates yields a likelihood-ratio statistic, with degrees of freedom equal to the difference in the numbers of parameters fitted.

The Rasch model gives a satisfactory framework for the description of reader differences, and conditional maximum likelihood estimation provides a method of obtaining consistent estimates of these differences. Estimation of the θ_i may be achieved by substitution of the estimates of the α_j in (6.6), but each film in the set defined by a particular t_i=k will have the same estimated θ_i; the estimates are a non-linear but monotonic function of the number of positive assessments. However, primary interest usually lies not in attempting to label each film with an "abnormality" parameter, but in relating the risk of a film being assessed abnormal to explanatory variables such as exposures to dust. An intuitively appealling extension is therefore to replace the film parameters θ_i with a regression function $\underline{\beta}^T\underline{x}_i$, to yield the link function

$$\pi_{ij} = Pr(Y_{ij}=1) = \frac{\exp(\alpha_j+\underline{\beta}^T\underline{x}_i)}{1 + \exp(\alpha_j+\underline{\beta}^T\underline{x}_i)} \qquad (6.7)$$

With only one reader, this is the logistic regression model (6.1), with the regression constant explicitly stated. With more than one reader, the model implies parallel regressions with different intercepts for the different readers, and on the assumption of independence estimates can be obtained from any logistic regression program.

However, the assumption of independence here is not appropriate, because the contributions to the likelihood from the model (6.7) for two readers reading the same film are identical to those for the same two readers examining two different films which have the same values of the explanatory variables. It is more reasonable to assume that identical stimuli, as represented by the explanatory variables, will produce different degrees of pathological response in individuals, and that these differences betwen individuals will be visible on their radiographs. Different readers' opinions of the same film will thus be expected to show correlation which is not modelled by (6.7) with an independence assumption.

This could be modelled by including a term for the deviation, say ε_i, of an individual's response to stimulus from the average $\underline{\beta}^T\underline{x}_i$. Then (6.7) becomes

$$\pi_{ij} = \frac{\exp(\underline{\beta}^T\underline{x}_i+\varepsilon_i+\alpha_j)}{1 + \exp(\underline{\beta}^T\underline{x}_i+\varepsilon_i+\alpha_j)} \qquad (6.8)$$

Since, as in (6.5), the number of parameters increases with the number of films, unconditional likelihood estimation is not satisfactory.

6.1.3 Estimation in a model with dependent observations

One possible approach to the problem is to model the variation in the ε_i rather than to attempt to estimate each value. This approach, as detailed below, adapts a technique used by Williams (1982) for dealing with over-dispersion in logistic regression, and appeals again to the iteratively reweighted least squares (IRLS) method, as generalised by Green (1984).

We assume that interest is not in explicit solutions for the ε_i, but rather that it is desired to allow for the correlation which they induce in the readers' assessments. Assume that ε_i is a random variable sampled from a distribution characterised satisfactorily by a mean 0 and a variance σ^2. Conditional on a known ε_i, y_{ij} has mean π_{ij} and variance $\pi_{ij}(1-\pi_{ij})$, and all the y_{ij} are uncorrelated. Unconditionally, this is not the case, but we may derive approximations to the unconditional moments up to the second order. To simplify notation, let K_{ij} be a random variable with mean $\alpha_j+\underline{\beta}^T\underline{x}_i$ and variance σ^2. Then the conditional moments given k_{ij} are

$$E[Y_{ij}|k_{ij}] = \pi_{ij} \simeq \frac{\exp(k_{ij})}{1+\exp(k_{ij})}$$

$$Var[Y_{ij}|k_{ij}] = \pi_{ij}(1-\pi_{ij})$$

$$Cov[Y_{ij},Y_{il}|k_{ij},k_{il}] = 0 \qquad j\neq l$$

Using established results for conditional moments,

$$E_Y[Y_{ij}] = E_K[E_Y[Y_{ij}|k]] = \pi_{ij}$$

$$\begin{aligned}
Var_Y[Y_{ij}] &= Var_K[E_Y[Y_{ij}|k]] + E_K[Var_Y[Y_{ij}|k]] \\
&= Var[\pi_{ij}] + E[\pi_{ij}(1-\pi_{ij})] \\
&= E[\pi_{ij}^2] - E^2[\pi_{ij}] + E[\pi_{ij}] - E[\pi_{ij}^2] \\
&= E[\pi_{ij}](1-E[\pi_{ij}]) \\
&\simeq \exp(\underline{\beta}^T\underline{x}_i+\alpha_j)/\{1+\exp(\underline{\beta}^T\underline{x}_i+\alpha_j)\}^2
\end{aligned}$$

$$\mathrm{Cov}_Y[Y_{ij},Y_{il}] = \mathrm{Cov}[E[Y_{ij}|k],E[Y_{il}|k]] + E[\mathrm{Cov}[Y_{ij},Y_{il}|k]]$$

$$= \mathrm{Cov}[\pi_{ij},\pi_{il}], \text{ since } Y_{ij} \text{ and } Y_{il} \text{ are conditionally independent.}$$

Now $\mathrm{Cov}[K_{ij},K_{il}] = \mathrm{Cov}[(\underline{\beta}^T\underline{x}_i+\alpha_j+\varepsilon_i),(\underline{\beta}^T\underline{x}_i+\alpha_l+\varepsilon_i)]$

Since all but the ε terms are constant, this reduces to

$$\mathrm{Cov}[K_{ij},K_{il}] = \mathrm{Cov}[\varepsilon_i,\varepsilon_i] = \mathrm{Var}[\varepsilon_i] = \sigma^2$$

By Taylor series approximation,

$$\mathrm{Cov}[\pi_{ij},\pi_{il}] \simeq \frac{\partial\pi_{ij}}{\partial k_{ij}}\frac{\partial\pi_{il}}{\partial k_{il}}\mathrm{Cov}[K_{ij},K_{il}]$$

$$= \pi_{ij}(1-\pi_{ij})\pi_{il}(1-\pi_{il})\sigma^2$$

Unconditionally, therefore, we have the same expectation for Y_{ij} as conditionally, and we have an expression for the off-diagonal elements of the covariance matrix V=Var(Y). If we order the data set by reader within film, then Var(Y) is a block diagonal matrix with r*r symmetric blocks at the diagonal and zeroes elsewhere.

Given a known value of σ^2, we can set up the regressions as above, but the matrix V is not of the correct form to permit use of its inverse as a column vector of weights. We can, however, proceed by transforming the problem orthogonally in such a way that the weight matrix is diagonal.

Unless σ^2 is too large (when the model anyway becomes unrealistic) V=Var(Y) is positive definite and we can thus find eigenvectors in a matrix P and eigenvalues in a diagonal matrix Λ such that

$$P^TVP = \Lambda \quad ; \quad P\Lambda P^T = V \quad ; \quad PP^T = P^TP = I$$

In Green's (1984) formulation, the IRLS step consists of regressing the adjusted response $(y-\eta)+D\underline{\beta}$ on columns of D using weights V^-, where D is the matrix of derivatives $\partial\eta/\partial\underline{\beta}$ and V^- is a (possibly generalised) inverse of V. This is equivalent to (6.3) and (6.4). It is also equivalent to regressing $P^T(y-\eta)+P^TD\underline{\beta}$ on columns of P^TD, using weight matrix $(P^TVP)^- = \Lambda^-$, which is diagonal and can be arranged as a column vector of length n*p.

Thus, for a given value of σ^2, there is a computational scheme which can be expressed as the following algorithm:

1. Obtain a first estimate of $\underline{\beta}$
2. Read y_{ij} and $\underline{x}_i$ for the first film (i=1)
3. Calculate η, D and V as functions of $\underline{\beta}$
4. Obtain P^T and Λ from V, calculate Λ^-, P^TD
5. Calculate $y'=P^T(y-\eta)+P^TD\underline{\beta}$
6. Copy y', Λ^- and P^TD into the next available P rows of a response vector, weight vector and explanatory variable matrix.
7. Get data for next film and return to 3; if all data have been read, goto 8.
8. Calculate new $\underline{\beta}$ by standard regression calculations. If convergence is satisfactory, stop, otherwise go to 2 with new $\underline{\beta}$.

This algorithm is straightforward to program, for example in GENSTAT (Alvey *et al*, 1983); the computation time involved is large, which is due to GENSTAT's orientation towards data structures, rather than individual values. An implementation in a high-level language such as Fortran or Pascal would doubtless run faster, but GENSTAT has the advantage of ready availability of regression calculations and square-root decomposition of matrices. Choice of how to implement the algorithm on other computing facilities would depend upon the tradeoff between execution speed and cost of custom-building a special program, and the latter would depend very much upon the availability of subroutines to perform important parts of the calculations.

The above algorithm assumes that V can be calculated from the current fitted values, and this requires that σ^2 is known. In applications, it is likely that estimation of σ^2 will be necessary. Because the transformation of the regression problem is orthogonal, the residual sum of squares is unchanged, and is given by

$$RSS = (y-\hat{\eta})^T V^- (y-\hat{\eta})$$

Its expected value is

$$E[\,(y-\hat{\eta})^T V^- (y-\hat{\eta})\,]$$

$$= E[\,y-\hat{\eta}\,]^T V^- E[\,y-\hat{\eta}\,] + \mathrm{Trace}[\,V^-\Sigma\,]$$

$$\text{where } \Sigma = \mathrm{Var}[\,y-\hat{\eta}\,]$$

From a standard result of regressing y on, say, X,

$$\Sigma = \mathrm{Var}[\,y-\hat{y}\,] \simeq V - X(X^T V^- X)^- X^T \text{ , and}$$

$$\mathrm{Trace}[\,V^-\Sigma\,] = \mathrm{Trace}[\,V^-V\,] - \mathrm{Trace}[\,V^-X(X^TV^-X)^-X^T\,]$$

$$= \mathrm{Trace}[\,V^-V\,] - \mathrm{Trace}[\,X^TV^-X(X^TV^-X)^-\,]$$

$$= \mathrm{Trace}[\,V^-V\,] - \mathrm{Trace}[\,I_p\,]$$

$$= \mathrm{Trace}[\,V^-V\,] - p$$

If V^- is assumed known, that is as a function of current estimates of $\underline{\beta}$ and σ^2, the multiplication V^-V can be expressed in terms linear in known functions of $\underline{\beta}$ and the unknown σ^2, to give the possibility of solving for an updated estimate of σ^2. From current values of the parameters, V^-_{r*r} is calculated, with elements v^-_{jk}. This can be done separately for each film, since V^- is block diagonal. Then, summing over all the films,

$$\mathrm{Trace}[\,V^-V\,] = \sum_{i=1}^{n}[\,\sum_{j=1}^{r} v^-_{jj}\eta_j(1-\eta_j) + 2\sigma^2 \sum_{j<k} v^-_{jk}\eta_j\eta_k(1-\eta_j)(1-\eta_k)\,]$$

Thus we may obtain an updated estimate as

$$\hat{\sigma}^2 = \frac{RSS + p - \sum_i \sum_j v^-_{jj}\eta_j(1-\eta_j)}{2\sum_i[\,\sum_{j<k} v^-_{jk}\eta_j\eta_k(1-\eta_j)(1-\eta_k)\,]}$$

The calculations to estimate σ^2 would be best organised in a loop exterior to that in which iterations are performed to estimate $\underline{\beta}$, otherwise computations would become extremely heavy. Space is not a problem, since only the two summation terms need to be accumulated over what is already required for the regressions, and the residual sum of squares is already available as a by-product of those calculations.

Because, even for known σ^2, V is not expressible in the form $\sigma^2 V(\eta)$, this model is not a case of quasi-likelihood as redefined by McCullagh (1983), and it is therefore not possible to appeal directly to existing results on convergence in quasi-likelihood models. Nor is it clear whether the approximations used in the derivations above are adequate, or whether the inclusion of further terms in the approximating expansions might improve the convergence properties at the undoubted cost of more calculation.

In fact, the actual value of σ^2 will not be of primary importance, and interest will be limited to the effect that non-zero values of σ^2 have on estimates of the coefficients α_j and $\underline{\beta}$. It will probably be adequate to examine this by performing the estimation of these latter parameters assuming initially that $\sigma^2=0$, and then for a series of pre-chosen values of σ^2. Such a scheme will avoid the iterative estimation of σ^2.

The above model was developed as one possible approach to the problem of modelling dependence between binary observations. Another possible approach, based on a recently published paper, is discussed in the next section.

6.1.4 A multivariate logistic model

In 6.1.3 above, the attempt to model the joint dependences amongst a set of binary response variates rested upon linear approximations to transformations of the data to a scale on which the dependences could be expressed as a covariance matrix. The computational load which this method incurs is heavy, and the model is clumsy to specify. A method which described a set of binary responses by a more direct modelling procedure, and which could model such dependences, would seem preferable.

Cox (1972) discussed the analysis of multivariate binary data. One logistic model mentioned in that paper used the device of transforming the binary variables Y_i to $Z_i = 2Y_i-1$, so that the Z's take the values ±1, and then expressing the joint probability distribution of the Z's as

$$\log(\Pr(Z_1=z_1,\ldots,Z_r=z_r)) = \alpha_1 z_1+\ldots\alpha_r z_r+\alpha_{12}z_1 z_2+\ldots$$
$$+\alpha_{r-1,r}z_{r-1}z_r+\ldots-\Lambda \qquad (6.9)$$

where Λ is a normalising constant. The full model has 2^r-1 parameters if all terms up to $z_1\ldots z_r$ are included, and is thus equivalent to fitting a full multinomial distribution. Fitting models which only included the lower order terms would decrease the required number of parameters, although the model with only first and second order terms still requires r(r+1)/2, which is almost as many as for the normal-theory model for continuous variables.

This class of models has not gained widespread popularity since Cox's (1972) paper was published, and this may be due partly to the fact that the marginal probabilities do not have a simple expression in terms of the α's. However, the model has some attractive properties. Firstly, maximum likelihood estimation based on (6.9) is computationally tractable. Secondly, it is straightforward to add regression terms in explanatory variables to the right-hand side of equation (6.9). Thirdly, a fundamental result on conditional probabilities has far-reaching practical implications.

Bonney (1987) notes that if we condition each Y_k on the observed values of all the preceding variables $y_{k-1},\ldots,y_1$, we can express the joint probability of $\mathbf{Y}$ given some explanatory variables $\underline{x}$ as a product of conditional probabilities as follows;

$$\Pr(Y_1,\ldots,Y_r|\underline{x}) = \Pr(Y_1|\underline{x})\Pr(Y_2|y_1,\underline{x})\ldots \Pr(Y_r|y_1,\ldots,y_{r-1},\underline{x}) \qquad (6.10)$$

Then (6.9) can be reparameterised by expressing it as a product as in (6.10), and defining a conditional logit of each of the factors as a linear function of the z's;

$$\theta_1 = \log\frac{\Pr(Y_1=1|\underline{x})}{\Pr(Y_1=0|\underline{x})} = \gamma_1+\underline{\beta}^T\underline{x} \;;$$

$$\theta_i = \log\frac{\Pr(Y_i=1|y_1,\ldots,y_{i-1},\underline{x})}{\Pr(Y_i=0|y_1,\ldots,y_{i-1},\underline{x})} \qquad (i=2\ldots r)$$

$$= \gamma_i+\gamma_{i1}z_1+\ldots+\gamma_{i,i-1}z_{i-1}+\gamma_{i12}z_1z_2+\ldots$$
$$+\gamma_{12\ldots i-2,i-1}z_1z_2\ldots z_{i-2}z_{i-1}+\underline{\beta}^T\underline{x} \qquad (6.11)$$

This model is called by Bonney (1987) the "regressive logistic" model, because of the manner in which each Y-variable is made to depend on the preceding Y's. In situations where there is some underlying order to the i's, for example where i indexes the temporal order of the data, the model allows the expression of dependence of a variable on those preceding it in time. Where the variables are dependent but not obviously ordered, this formulation still allows the parameters γ with more than one subscript to represent the dependence structure, even if the individual γ values do not then have quite so obvious an interpretation.

Because the factors of (6.10) are independent, this regressive logistic model for r dependent observations can be fitted in the same way as the logistic regression model for r independent observations, and by the same computer programs, requiring only that the data matrix be augmented by the z's calculated from the y's. However, it will be noted that the equations for the Y_i given in (6.11) are of differing lengths, increasing with i. The data set is made rectangular in shape by the simple device of redefining the Z's in the form

$$Z_{ij} = \begin{cases} Z_j = 2Y_j - 1, & \text{if } j < i, \\ 0, & \text{if } j > i, \end{cases}$$

so that the conditional logits are expressed as

$$\begin{aligned} \theta_i = \ & \gamma_i + \gamma_{i1} z_{i1} + \gamma_{i2} z_{i2} + \ldots + \gamma_{ir} z_{ir} + \gamma_{i12} z_{i1} z_{i2} + \ldots \\ & + \gamma_{r-1,r} z_{i,r-1} z_{ir} + \ldots + \gamma_{12\ldots r-1,r} z_{i1} z_{i2} \ldots z_{ir-1} z_{ir} \\ & + \underline{\beta}^T \underline{x} \qquad\qquad (6.12) \end{aligned}$$

The model (6.12) will not be useful if it is necessary, for a particular data set, to fit all cross-product terms of every order. Some (probably considerable) reduction in dimensionality will be sought, as in most model-fitting situations, so as to arrive at a suitable but parsimonious description of the data. Three special cases are of particular interest. Firstly, the cross-product terms can be limited to those of second order, γ_{ij}, with all higher terms $\gamma_{ijk}=\ldots=0$. Secondly, an analogue of the "equal correlation" case of the normal-theory model can be set up by further restricting the parameters so that $\gamma_{ij}=\gamma$ for all i,j. Thirdly, we can set $\gamma_{ij}=0$. This last submodel is equivalent to assuming complete independence amongst the variables, so that independence is seen to be a nested submodel of the second case, which in turn is nested within the first. Thus likelihood ratio tests are directly available for testing the assumption of independence against the alternatives of equal or unequal second-order dependence parameters. Similarly, by including other cross-product terms of successively higher orders, it is straightforward to test at what stage such further inclusions do not benefit the model's fit to the data.

The model represented by (6.12) requires some care in setting up the data matrix, particularly in the construction of the Z's from the Y's, although this can be automated by some programming in a suitable package or high-level language. However, once the Z's are placed in the data matrix, the model has the considerable advantage over most of its competitors that the estimation can be performed by any computing facility for logistic regression models, which includes most of the major statistical analysis packages. Further, these packages generally contain facilities for arithmetic transformations on the data and the construction of new variables, so the calculation of the cross-products of the Z's can also, if desired, be performed within the package, which may be a practical advantage if it avoids the necessity of constructing and storing fairly sparse data matrices with large numbers of cross-product variables. It is to be stressed, however, that the calculation of the Z's is not iterative, and that the only iteration required by this model is that for fitting the logistic model; this may be expected to consume computing resources at a rate similar to any logistic regression problem with the same number n*r of data points, and the same number of fitted parameters. It seems likely that, although this model has been in relative neglect since Cox's (1972) paper, Bonney's (1987) demonstration of the transformation to logistic form, and of the consequent simplicity in estimation procedures, will result in a greater variety of researchers investigating the usefulness of the multivariate logistic model, and testing its applicability to their data sets.

One less desirable feature of this model is that not all of its submodels are invariant to permutation of the order of the variables. Where, as with film-reading data, the order is arbitrary, it may be necessary to rerun analyses on one or more permutation of the responses, for reassurance that the results are not a specific result of assuming any particular order. The practical aspects of performing such reanalyses might require reruns of programs in the chosen package, with a permutation of the data during or after reading; with packages in which the data all remain in memory, this will be easily programmed without the necessity of rereading extensive data sets.

6.1.5 Analysis of ordered categorical data

All the models mentioned above are for the analysis of binary data. Any ordered variable can always be dichotomised at a chosen category boundary and thus converted to a binary variable, but there is an arbitrary element to the choice of boundary, and it is seldom clear how much information, if any, is lost by ignoring the given ordering.

Models for the analysis of ordered categorical variables have received increasing attention in recent years. The two approaches which have found widest acceptance are those of McCullagh (1980) and Anderson (1984). Each has theoretical advantages and disadvantages, but they can be expected to give similar results in practice. Here Anderson's approach has been adopted; it is a special case of a wider class of models, and is relatively straightforward to implement in existing computer packages.

Anderson (1984) defines the stereotype ordered regression model. This predicts the probability P_s that Y will take category y_s, given regressor variables $\underline{x}$, by the logistic equation

$$P_s(\underline{x}) = \Pr(Y=y_s|\underline{x}) =$$

$$\exp(\beta_{0s} + \varphi_s\underline{\beta}^T\underline{x})/\sum_{t=1}^{k} \exp(\beta_{0t} + \varphi_t\underline{\beta}^T\underline{x}) \quad (s=1...k) \qquad (6.13)$$

The model is over-parameterised, and it is necessary to assign arbitrary values to some parameters before the remainder are estimable. Here we adopt the convention of referring all the relative odds to the first category, which requires $\varphi_1 = \beta_{01} = 0$, and the remaining indeterminacy of scale is eliminated by setting $\varphi_k = 1$. If the categories are truly ordered with respect to the explanatory variable(s) $\underline{x}$, the φs should obey the order restrictions

$$0 = \varphi_1 < \varphi_2 < \ldots < \varphi_k = 1.$$

(Note that the sign within the bracketed expression, and the order of the φs, are reversed from Anderson's conventions, because we have chosen to refer relative odds to the first rather than to the last category.) The regression coefficients $\varphi_k\underline{\beta}$ are thus proportional across the categories in this model, and if the estimates are indeed ordered, the model will predict that units with higher values of any x for which the estimate of β is positive should arise from higher categories, other factors being equal.

Estimation is by maximum likelihood. Because the model is not log-linear in the parameters, the usual asymptotic distributional arguments do not necessarily hold, but it is unlikely that their use will prove seriously misleading if the iteration converges to a finite solution. Categories which are not distinguishable by regression on $\underline{x}$ will have $\varphi_s=\varphi_t$, which hypothesis can be tested by a likelihood ratio test. Estimates of φ are not however constrained to obey the order restriction, and a seriously non-monotonic set of estimates would suggest that the ordered model is not appropriate. With categories of increasing abnormality on radiographs, the expectation is that the data will show good ordering, although it may not be possible to distinguish all the categories of the 12-point ILO (1980) scale.

The model (6.13) shows that the contribution made to the likelihood by each observation depends only on the parameters and the observation itself, and is relatively straightforward to calculate. The model can therefore be fitted by standard statistical packages, and for the applications here this was done in the package BMDP (Dixon, 1983), using the derivative-free non-linear regression program BMDPAR with a user-supplied subroutine to define the likelihood contribution for each observation.

6.2 <u>The analysis of lung function data</u>

All the regression analyses of lung function were performed using standard linear regression methods (Draper and Smith, 1981). All the computations were performed in the package GENSTAT (Alvey *et al*, 1983).

7 RESULTS OF ANALYSES OF FILM-READING DATA

7.1 Data from readings by the IOM Panel

In this section are summarised the assessments made by the IOM film reading panel. All the readings were of individual films assessed independently in random order, and classified according to the ILO (1980) classification scheme.

The IOM panel always read and classify the full range of radiological abnormalities covered by the ILO (1980) scheme (with the exception of some of the codes for "other diseases", since they lack the medical training necessary to make differential diagnoses). For the purposes of this project, the assessments of pleural abnormalities were considered relatively unimportant; they were observed infrequently, as might be expected, since they are not usually strongly associated with exposure to respirable coalmine dust. The assessments which *a priori* were considered the most relevant were those concerning the profusions of small and large opacities, and, to a lesser extent, the shape and size of the small opacities. No detailed work was done on the small numbers of positive assessments recorded for pleural abnormalities, and we do not present any results here regarding those assessments.

Data presented in this section are from the three independent randomised film reading exercises decribed in 4.3. These were the large readings of all available films from surveys 4, 5 and 6, (coded XRLV and XRVL) and the third, smaller, exercise (coded XREX) which included films missed from either of these, plus the films from the 1980 PXR survey of the NCB Medical Service.

During the second and third reading exercises, the IOM panel consisted of five readers. Prior to that, it also included a sixth member, who performed assessments for the entire first exercise before resigning from the work of the panel. That member was noted for reading profusion of small opacities at a considerably higher level than the other panel members, and this pattern was apparent also in her readings in the first exercise for this project. Because she did not read in the other exercises, the inclusion of data from her assessments would have distorted the pattern of assessed abnormality in the first reading, reducing the comparability of the different exercises. Therefore, all data from that reader's assessments from the first exercise have been excluded from further consideration, and they are not presented here.

The remaining five readers are identified throughout by the code numbers 101, 104, 105, 106, 112, by which their identities are routinely coded and distinguished on the IOM computer.

7.1.1 Film quality

Exhibit 7.1.1 shows a tabulation of the frequencies with which each of the readers assigned film quality scores, aggregated over all three film reading exercises. Individual films are therefore often represented more than once within the values for each reader, since there were duplicate readings. The vast majority of films were classified by all readers as being of quality 1=good or 2=acceptable. A small proportion were classified as 3=poor, and on a very small proportion of occasions the quality was scored 4=unreadable. Data (not presented here) on agreement between the readers on the quality of individual films showed, in general, reasonable agreement. Comparing readers pairwise, exact agreement on quality score was reached in around 80% of cases.

Exhibit 7.1.1. Table of numbers of films assigned to four grades of film quality, by five readers; totals over all independent randomised film-reading exercises.

Reader	Film Quality: Good	Acceptable	Poor	Unreadable	Total
101	4669	1024	0	0	5693
104	4656	1022	1	8	5687
105	4462	1128	97	0	5687
106	4738	948	0	0	5686
112	4147	1510	31	0	5688
Total	22672	5632	129	8	28441

7.1.2 Profusion of small opacities

Exhibit 7.1.2 shows the frequencies with which each of the readers assigned the 12 levels of profusion of small opacities to the films they considered readable (that is, almost all of those presented to them). The three reading exercises are shown separately, so that each film makes at most a single contribution to each reader's assessments for each exercise, though the exercises had many films in common.

Several features are readily observed in all three of these tables. Firstly, none of the readers uses the code 0/−, which some readers like to apply to films which look even "healthier" than the ILO standard films defining category 0. The panel readers are thus operating on a maximum 11−point rather than a 12−point scale of profusion. Additionally, it is noticeable that only reader 106 has made any allocations to the highest profusion category 3/+. However, because few films attracted profusion scores around category 3, it is not possible to deduce that the remaining readers never use category 3/+; we may be simply observing zero occurrences, in finite samples, of events with low but nonzero probabilities.

The overall distributions of profusion are typical of those seen throughout the PFR and much other research in cross−sectional studies of employed populations. The majority of films show little or no signs of abnormality, and the numbers showing progressively higher profusion categories decrease progressively; the distributions are heavily skewed, with long right tails. The distributions show another feature, often called the "middling tendency"; that is, the tendency for assessments of profusion with the same category on the 4−point scale to concentrate in the corresponding central category on the 12− (or 11−) point scale. For example, in exercise XRLV, the five readers made a total of 3556 allocations of films to category 1, but these were not split even approximately equally to the three categories 1/0, 1/1, 1/2.

Exhibit 7.1.2. Table of numbers of assessments by five readers of the profusions of small opacities on films read in the three independent randomised reading exercises.

Exercise : XRLV

Reader	Profusion of small opacities												
	0/-	0/0	0/1	1/0	1/1	1/2	2/1	2/2	2/3	3/2	3/3	3/+	Total
101	0	674	779	0	758	170	0	213	0	0	0	0	2594
104	0	1379	427	277	329	115	1	57	4	0	5	0	2594
105	0	1204	185	484	512	81	41	84	2	1	0	0	2594
106	0	2461	17	20	28	19	11	15	9	7	6	1	2594
112	0	1084	612	299	294	170	63	37	25	9	1	0	2594
Total	0	6802	2020	1080	1921	555	116	406	40	17	12	1	12970

Exercise : XRVL

Reader	Profusion of small opacities												
	0/-	0/0	0/1	1/0	1/1	1/2	2/1	2/2	2/3	3/2	3/3	3/+	Total
101	0	333	931	1	977	133	2	148	2	0	0	0	2527
104	0	1391	386	279	301	114	0	45	5	0	3	0	2524
105	0	1117	147	616	411	109	4	115	8	0	0	0	2527
106	0	2369	6	18	13	34	19	28	19	10	6	4	2526
112	0	1063	369	193	353	280	104	110	44	8	4	0	2528
Total	0	6273	1839	1107	2055	670	129	446	78	18	13	4	12632

Exercise : XREX

Reader	Profusion of small opacities												
	0/-	0/0	0/1	1/0	1/1	1/2	2/1	2/2	2/3	3/2	3/3	3/+	Total
101	0	49	136	11	211	83	24	52	6	0	0	0	572
104	0	314	107	45	65	22	2	12	0	0	0	0	567
105	0	243	96	91	79	41	2	16	4	0	0	0	572
106	0	472	9	15	18	19	9	15	4	7	3	1	572
112	0	235	113	49	83	58	16	15	3	0	0	0	572
Total	0	1313	461	211	456	223	53	110	17	7	3	1	2855

Category 1/1 accounted for 1921 of these, well over a half. The same phenomenon can be observed elsewhere in the tables, nor is it concentrated in the assessments of any one reader.

This tendency is interpreted as a consequence of the definition of the 12 categories as an elaboration of the older 4-point scale. The instructions for arriving at a classification on the elaborated scale may be paraphrased as follows: "by comparison with the standard films for categories 0, 1, 2, 3, decide which category the film most resembles; if the decision is made to allocate to a category without seriously considering an adjacent category, make the first and second digit of the classification correspond to the category of the standard film; if a category is chosen after an adjacent category is seriously considered, make the first digit of the classification the chosen category, and the second that considered and rejected". It is clear that without any assistance or advice on how the reader may quantify her degree of certainty, there can be considerable variation between readers in how often they feel certain enough to use a central category. Both personality and experience can be expected to play important parts in individuals' interpretations of these instructions.

One further observation is the obvious differences between the readers in the relative frequencies with which they assign films to the categories representing higher degrees of abnormalities. For example, in exercise XRLV, reader 106 made 2594 assessments, and 2478 of these, that is over 95%, were in categories 0/0 and 0/1; of the same number, reader 105 allocated 1389, or 54%, to these lowest categories, and reader 101 allocated 1453 or 56%. Similar differences in reading patterns are seen in the other two exercises.

7.1.3 Between-reader differences in assessments of profusion

The tables in exhibit 7.1.2 are marginal totals for each reader, and they cannot show how often readers agreed on the classification of profusion of individual films. A complete documentation of the agreement would require a 5-dimensional table, which would occupy a great amount of space, and which would anyway not be easily interpreted without intensive examination and complicated statistical analysis to extract the principal features of its structure. A first approximation to such an examination is provided by examining the two-way tables of marginal totals obtained by comparing the assessments of pairs of readers. With five readers, there are ${}^5C_2$ = 10 such pairwise combinations.

Exhibit 7.1.3 shows all 10 tables summarising pairwise reader agreement on the assessments of exercise XRLV, as an example. Although assessments were made on the 12-point scale of profusion, for these tables they have been collapsed to the 4-point scale, to reduce the effect on agreement of different patterns of usage of the side categories relative to the central categories.

These tables show both the extent and the nature of the agreement and disagreement regarding the major categories of profusion assigned to individual films. Numbers lying on the diagonals of these square tables are counts of films where the two readers concerned both assigned the same major category, and counts off the diagonals are of cases where the readers assigned different categories. In an obvious sense, the extent of any disagreement is indicated by distance from the diagonal.

In some instances, for example in the table comparing readers 101 and 106, the differences in levels of assessed profusion are clearly apparent. In others, that is

Exhibit 7.1.3. Tables showing agreement between pairs of readers on profusion of small opacities on the 4-point scale, for all films seen in exercise XRLV.

Reader 101 Profusions

Rdr 104 Profns	0	1	2	3	Total
0	1262	495	49	0	1806
1	190	399	132	0	721
2	1	33	28	0	62
3	0	1	4	0	5
Total	1453	928	213	0	2594

Reader 101 Profusions

Rdr 105 Profns	0	1	2	3	Total
0	1078	292	19	0	1389
1	368	573	136	0	1077
2	7	63	57	0	127
3	0	0	1	0	1
Total	1453	928	213	0	2594

Reader 104 Profusions

Rdr 105 Profns	0	1	2	3	Total
0	1285	104	0	0	1389
1	519	537	121	0	1077
2	2	80	40	5	127
3	0	0	1	0	1
Total	1806	721	262	0	2594

Reader 101 Profusions

Rdr 106 Profns	0	1	2	3	Total
0	1445	870	163	0	2478
1	7	36	124	0	67
2	1	19	15	0	35
3	0	3	11	0	14
Total	1453	928	213	0	2594

Reader 104 Profusions

Rdr 106 Profns	0	1	2	3	Total
0	1803	660	15	0	2478
1	2	48	17	0	67
2	1	12	20	2	35
3	0	1	62	3	14
Total	1806	721	62	5	2594

Reader 105 Profusions

Rdr 106 Profns	0	1	2	3	Total
0	1386	1030	62	0	2478
1	2	41	24	0	67
2	1	6	28	0	35
3	0	0	13	1	14
Total	1389	1077	127	1	2594

Exhibit 7.1.3 continued.

Reader 101 Profusions

Rdr 112	0	1	2	3	Total
P 0	1142	489	65	0	1696
r 1	296	369	98	0	763
f 2	15	66	44	0	125
n 3	0	4	6	0	10
s					
Total	1453	928	213	0	2594

Reader 104 Profusions

Rdr 112	0	1	2	3	Total
P 0	1479	216	1	0	1696
r 1	318	425	20	0	763
f 2	9	78	34	4	125
n 3	0	2	7	1	10
s					
Total	1806	721	62	5	2594

Reader 105 Profusions

Rdr 112	0	1	2	3	Total
P 0	1220	472	4	0	1696
r 1	165	539	59	0	763
f 2	4	64	57	0	125
n 3	0	2	7	1	10
s					
Total	1389	1077	127	1	2594

Reader 106 Profusions

Rdr 112	0	1	2	3	Total
P 0	1690	4	2	0	1696
r 1	710	44	7	2	763
f 2	77	17	22	9	125
n 3	1	2	4	3	10
s					
Total	2478	67	35	14	2594

Exhibit 7.1.4. Table of pairwise consistency coefficients summarising agreement between pairs of readers on major category of small opacities profusion. Tabulated is the proportion (%) of films seen in exercise XRLV which both readers classified to the same category of the 4-point scale.

Reader	101	104	105	106
104	65.1			
105	65.8	71.8		
106	57.7	72.2	56.1	
112	59.9	74.7	70.0	67.8

amongst the readers with similar marginal frequencies for the levels of profusion, it is clearer that there is considerable evidence of association between the readers' assessments. For example, in the table comparing 101 and 105, there is a clear concentration of the values towards the diagonal, with complete agreement on the major category of 1078 + 573 + 57 = 1708 films, representing 66% of the total.

This figure, the percentage of films to which the readers assigned the same category, is sometimes known as a "consistency coefficient". These can be calculated for each of the 10 two-way tables, and they are tabulated together in exhibit 7.1.4, ranging from 56% to 75%. The corresponding values from the second exercise XRVL showed very similar patterns in the 10 two-way tables and in their summary consistency coefficients, and they are not shown here.

The consistency coefficient measures only one aspect of the extent of agreement between readers; consistency can be low when one reader assesses profusions at a higher level than another, even if their overall ranking of the amounts of abnormality on each film is the same, because then the counts in the two-way marginal tables will tend to lie just off the diagonal. It is possible to construct other arithmetical coefficients to represent other aspects of agreement, but this is not pursued here; later analyses which require allowance for differences and similarities in readers' patterns of reading habits appeal to more general statistical modelling techniques, in which such features are represented by explicit parameterisations.

7.1.4 Within-reader variations in assessing profusion

The independent randomised film-reading exercise XRVL was designed to be a duplicate rereading of the first exercise XRLV. Data from a very few films were lost by their omission from the first exercise, and a few more from the second, as described in 4.3.2. Apart from these, every film was read by each of the readers on two separate occasions many months apart. The data from these readings therefore provide important information on the repeatability of each reader's classifications.

Exhibit 7.1.5 shows, for each reader, a table cross-classifying the major category of profusion of small opacities, on the ILO 4-point scale, assigned to each film at each of the two exercises. For all the readers, it is clear that the data values cluster towards the diagonals, showing agreement between the categories assigned at the two occasions; the agreement is more readily apparent for those readers who recorded appreciable numbers of films in categories other than 0. Reader 112 tended to assign higher categories to the same films in the second reading exercise than in the first. None of the other readers showed quite such obvious evidence of a shift in reading level over the intervening time period.

Consistency coefficients were calculated from these tables, to summarise the proportion of films for which the classifications assigned to each film were the same, on the 4-point scale, at the two occasions. Exhibit 7.1.6 displays these. These confirm the general impression of a reasonable level of agreement; for the different readers, however, they are obviously affected in part by the different frequencies with which those readers recorded the presence of opacities. Most obviously, the tendency of reader 106 to read few films as showing abnormalities has resulted in a high self-consistency compared with those of the other readers, who all recorded more abnormalities.

Exhibit 7.1.5. Tables showing agreement within each reader between assessments of profusion on the 4-point scale on the same films seen on two different occasions.

Reader : 101

Exercise XRVL Profusions

XRLV	0	1	2	3	Total
P 0	1019	390	13	0	1422
r 1	230	584	85	0	899
f 2	12	137	54	0	203
n 3	0	0	0	0	0
s					
Total	1261	1111	152	0	2524

Reader : 104

Exercise XRVL Profusions

XRLV	0	1	2	3	Total
P 0	1573	187	1	0	1761
r 1	200	479	15	0	694
f 2	0	27	30	1	58
n 3	0	0	3	2	5
s					
Total	1773	693	49	3	2518

Reader : 105

Exercise XRVL Profusions

XRLV	0	1	2	3	Total
P 0	1137	227	0	0	1364
r 1	123	869	46	0	1038
f 2	0	39	79	0	118
n 3	0	0	1	0	1
s					
Total	1260	1135	126	0	2521

Reader : 106

Exercise XRVL Profusions

XRLV	0	1	2	3	Total
P 0	2341	45	24	1	2411
r 1	27	16	17	2	62
f 2	2	4	21	7	34
n 3	0	0	3	10	13
s					
Total	2370	65	65	20	2520

Reader : 112

Exercise XRVL Profusions

XRLV	0	1	2	3	Total
P 0	1296	343	23	0	1662
r 1	131	452	147	2	732
f 2	2	29	82	6	119
n 3	0	0	5	4	9
s					
Total	1429	824	257	12	2522

Exhibit 7.1.6. Consistency coefficients summarising the proportion (%) of films classified by each reader in the same category of profusion on the 4-point scale, on two separate occasions.

	Reader				
	101	104	105	106	112
Consistency (%)	66	83	83	95	73

Exhibit 7.1.7. Distribution of films from reading exercise XRLV by predominant and secondary classification of shape (rounded = p,q,r; irregular = s,t,u) of all opacities classified as >= 0/1 in profusion, by each reader.

Shape of Opacities	Reader 101	104	105	106	112
Round	1620	598	171	72	9
Round/Irreg	0	106	387	30	250
Irreg/Round	0	0	0	7	138
Irregular	300	511	832	24	1113
Any opacities	1920	1215	1390	133	1510
None (0/0)	674	1379	1204	2461	1084
Total	2594	2594	2594	2594	2594

Exhibit 7.1.8 Table showing distribution of readers' assessments of profusions of large opacities on the films seen in the first reading exercise XRLV.

	Large opacities profusion				
Reader	0	A	B	C	Total
101	2593	0	1	0	2594
104	2574	17	3	0	2594
105	2585	9	0	0	2594
106	2582	9	3	0	2594
112	2578	12	4	0	2594
Total	12912	47	11	0	12970

7.1.5 Shape and size of small opacities

Exhibit 7.1.7 summarises, for the films in reading exercise XRLV, how each of the five readers used the codes in the ILO (1980) classification scheme for the description of the types of small opacities seen. A two-letter code is used to record the reader's opinion; each letter can be one of the six p,q,r (representing increasing sizes of rounded opacities), s,t,u (representing increasing sizes of irregular opacities). It is compulsory under the ILO (1980) scheme, whenever small opacities are recorded, that is when the recording of profusion of small opacities is 0/1 or higher, to describe by these alphabetic codes the type of opacities most commonly observed, and to add a second different letter representing the next most common type (or, if there is only one type, the same letter repeated). The 36 possible combinations have been condensed in this table by grouping together those letters describing the same general shape of opacities, regardless of their size. Thus, the four possible groupings are those where: both first and second letter were size variants of the same rounded shape; the first showed predominantly rounded, and the second that irregular opacities were also common; the first showed predominantly irregular, and the second that rounded were also common; and both first and second showed irregular.

The table shows, for each reader, the breakdown to these four condensed categories of shape, of all the films which the reader classified as having any small opacities. It is clear that, even allowing for the differences across readers observed above in the proportions of films classified as showing small opacities, the proportional allocations of the films to the four shape categories differ enormously. Of 1920 films classified by reader 101 as showing small opacities, 1620 (84%) were classified as being all round; whereas of 1510 films classified by reader 112 as showing small opacities, only 9 (0.6%) were classified as all round.

Two-way tables similar to those of exhibit 7.1.5 merely highlighted these differences in the readers' opinions about the shapes of the opacities they observed, and are not reproduced here. Results from the panel's readings in the other exercises, XRVL and XREX, showed very similar patterns.

7.1.6 Profusion of large opacities

Exhibit 7.1.8 shows the distribution of the assessments made by each reader of the profusion of large opacities present on each film, during reading exercise XRLV. Each reader except 101 made a few classifications of opacities of profusion A (approximate (sum of) maximum diameter(s) between 10 mm and 50 mm), and each reader except 105 made at least one classification of profusion B (...between 50 mm and the area of the right upper zone of the lung). There were no classifications of the highest grade C of profusion of large opacities.

The fraction of the readings in which large opacities were recorded as present was small overall; but the presence of large opacities could be a sign of the severe abnormality PMF (progressive massive fibrosis), and it was of interest to examine the extent to which the readers agreed on the presence of this sign. Of the 58 cases in which the presence of large opacities (of profusion A or B) was recorded on a film, 33 were from films for which only one reader made such a recording, and 5 films were classified as showing large opacities by two readers. Classifications of large opacities were made by three, four and all five readers on 2, 1 and 1 films respectively. Consensus on large opacities in this exercise is thus seen to have been low, although, because of the high proportion of negatives on this assessment, calculated consistency coefficients would have high values.

Again, similar patterns were apparent in these readers' results from the duplicate reading exercise XRVL. Consensus between readers' opinions on the presence of large opacities on the same films was equally poor in this exercise.

7.2 Logistic regression analyses of small opacities profusion

This section reports the results of analyses in which the film-reading panel's assessments of small opacities profusions were related to data on the individuals' exposures to respirable coalmine dust and quartz, with investigation into the necessity of allowing in the analyses for nuisance variables such as age and smoking habits. The principal analytical tool for these analyses was the application of logistic regression modelling, with estimates and tests of statistical significance obtained by maximum likelihood methods. The response variable analysed was a median of the readers' classifications, intended to summarise the consensus of those classifications; as discussed above, comparison of readers' opinions of the profusions of small opacities on the same films showed reasonable agreement in the ranking of those films by profusion.

Because of the poor consensus between readers on assessments of shape of small opacities and profusion of large opacities, it was judged that regression analyses of these variables were unlikely to prove informative; they are not considered further here.

7.2.1 Summarising small opacities assessments by medians

Because of the complexities inherent in working with several readers' different assessments of the same film, it was decided to base much of the initial statistical investigation of the data on a summary measure of the profusion of small opacities assessed for each film. The measure chosen was the median of the profusions assigned by the five readers. It was felt that the arithmetic mean was less suitable, requiring the adoption of arbitrary scores for the categories, and being heavily influenced by isolated assessments of high profusion.

Most of the films from the 4th, 5th and 6th PFR surveys had been read twice by all five readers, and for these films the median was calculated from all ten profusion assessments. The 1980 PXR survey films had been read only once, in the third exercise, so medians for these films were necessarily based on only five values. This exercise also provided readings for the films omitted from the earlier exercises, particularly the second, and these values were also incorporated into the calculation of the medians. No median was calculated in any case where less than five assessments of profusion were available (for example in the very rare cases where films were declared to be of unreadable quality). Medians were always calculated from the data as originally recorded on the 12–point scale of profusion, on the principle that they could be grouped more coarsely by subsequent calculation as deemed necessary.

It was not possible to use the same number of assessments in the calculation of the median for each film; most importantly, the medians for the 1980 films could be based only on the five assessments from the last reading exercise, and this represented a systematic difference in the way these films were treated, compared with the other films. However, it is believed that any effects of this systematic difference will be relatively unimportant, for two reasons. Firstly, any systematic shift of a reader's level over time, as suggested by exhibit 7.1.5, was small compared to the substantial differences between readers; and secondly, while comparisons between the 1980 and the other surveys would undoubtedly be distorted by any such shift, the results given in the next sections are specific to individual surveys, and are therefore not subject to distortion from this source.

Exhibit 7.2.1 describes in tabular form the distribution of the calculated median categories of profusion for all films from each survey for which such calculation was possible. This table shows that no film had a median profusion greater than 2/3; also that the "middling tendency" shown in exhibit 7.1.2 and discussed in 7.1.2 was much less evident in the summarised data.

Comparisons between the columns of exhibit 7.2.1 are not easy to make, because the values differ greatly between rows and because the column totals are different. In Exhibit 7.2.2, the numbers are re-expressed as percentages of the column totals. This table presents the impression of the amount of pneumoconiotic abnormalities in the population decreasing over the first three surveys, which runs against intuition. However, this comparison is between different sub-populations which are not necessary directly comparable in structure; complex selection effects, due to the contraction of the workforce over the period, require that the study of change should be based on observation of differences in the classifications of films from the same individuals. The patterns in exhibit 7.2.2 are consistent with preferential loss of older workers, which is in turn consistent with the descriptions of the age distributions observed at the three PFR surveys, as shown in exhibit 5.4.1.

7.2.2 The form and presentation of the logistic regression analyses

The form of the analyses which were carried out on the data from the film-reading exercises was moulded by the desire to discover whether the assessed profusions of small opacities bore any relation to the variables describing the individuals' occupational exposures; whether any such relationships differed for the quartz and non-quartz components of the dust, or according to the time elapsed

Exhibit 7.2.1. Distribution of numbers of films by calculated values of median profusion of small opacities as assessed by panel readers, by year of survey.

Median profusion	Year of survey 1970	1974	1978	1980	Total
0/-	0	0	0	0	0
0/0	442	501	400	243	1586
0/1	232	164	98	82	576
1/0	148	96	59	32	335
1/1	192	81	39	48	360
1/2	50	15	7	25	97
2/1	17	4	5	7	33
2/2	22	9	6	7	44
2/3	2	1	4	1	8
3/2	0	0	0	0	0
3/3	0	0	0	0	0
3/+	0	0	0	0	0
Total	1105	871	618	445	3039

Exhibit 7.2.2. Distribution of proportions of films at each survey by calculated values of median profusion of small opacities as assessed by panel readers; numbers tabulated are percentages, to nearest integer.

Median profusion	Year of survey 1970	1974	1978	1980
0/-	0	0	0	0
0/0	40	58	65	55
0/1	21	19	16	18
1/0	13	11	10	7
1/1	17	9	6	11
1/2	5	2	1	6
2/1	2	0	1	2
2/2	2	1	1	2
2/3	0	0	1	0
3/2	0	0	0	0
3/3	0	0	0	0
3/+	0	0	0	0

between exposure and response; and whether allowance for any nuisance factors for which data were available altered conclusions concerning these relationships.

Attempts to answer these questions were made by the application of logistic regression methods, which offer a convenient way of describing expected values of observed frequencies as probabilities which are functions of linear regression relationships in explanatory variables, whose parameters, or regression coefficients, have to be estimated from the data (Cox, 1970; Dobson, 1983). Logistic regression is appropriate to binary responses, that is those which are equivalent to a classification which distinguishes only between presence or absence of some characteristic. Two separate sets of analyses were undertaken. In the first, a positive response was defined as the presence, in the variable holding the median profusions, of a median profusion of category 1/0 or greater; in the second, a positive response was defined as the presence of a median profusion of category 2/1 or greater (which are, by definition, a subset of those of 1/0 or higher).

Separate analyses were performed on the data from readings of films from the 4th, 5th and 6th PFR surveys and from the 1980 PXR survey. The results are summarised in tables which adopt a standardised format, and which are included here as exhibits 7.2.3 to 7.2.10. Looking for example at 7.2.3, it will be seen that each column of tabulated values contains the results of fitting one regression model to the data. Because this report necessarily contains a number of similar tabulations, the models have been labelled in a manner which helps to identify their nature. Each column is labelled with a code containing three elements separated by obliques: the first element represents the content of the response variable, here 1+ to show that the variable represents categories of 1 or higher on the 4-point ILO scale; the second identifies the survey, here the 4th; and the third is a sequential distinction between the different models applied.

The values in the body of the table show the estimated regression coefficients against those variables included in each model, and are blank where a variable was not included. Shown in parenthesis below each estimated coefficient is the absolute value of the calculated ratio of the estimate to its estimated standard error. This ratio may be taken as a partial t–statistic for testing the significance of including that term last, where such a test is valid. A more flexible method for testing differences between models is given by the deviance statistics. These are log-likelihood statistics measuring the distance of the fit of the model from a model which fits the data exactly (Nelder and Wedderburn, 1972). Under the null hypothesis that a term introduced does not improve the fit of the model, or equivalently that its true regression coefficient(s) is zero, the reduction in deviance is distributed approximately as chi-squared, with degrees of freedom equal to the number of additional parameters fitted for the term. Changes in the deviances at the foot of each column can thus be compared with critical values of the chi-squared distribution with degrees of freedom given by differences in the total available degrees of freedom tabulated below. Single degree of freedom tests based on the deviance are asymptotically equivalent to those based on the partial t–statistics, but may give different results when based on sparse data.

The rationale underlying the sequence of the model-fitting in this and all similar tables was to investigate which estimates of exposure to dust and quartz at various times up to the relevant survey, made the greatest, if any, contribution to the regression models, after allowance for age. In this and all the regression analyses presented here, the variables for estimates of exposure to respirable dust are those for the non-quartz fraction. They were obtained by subtracting the estimates of quartz exposure during each period from the estimates of exposure to whole respirable dust, of which the quartz was a constituent fraction. Since age was included as a continuous variable, accounting for a single degree of freedom, it was always possible to examine the effect of including a linear term in age last in the model, by inspection of the partial t–statistic as tabulated.

Particularly for the models for the later surveys, the choice of exposure variables for inclusion was from amongst several, and to display the results for all possible models would have used more space than was justified. All the logistic regression analyses were performed with the package GENSTAT (Alvey *et al*, 1983), and the sequences of models were often investigated using the facilities of that package for altering models interactively at the computer terminal. The results presented here are thus from models selected during that process, and many other possible models by implication have been found less interesting and omitted from the tabulations.

7.2.3 Logistic regression analyses of 4th survey films

Exhibit 7.2.3 presents some results from a logistic analysis of the binary variable constructed to have a value of 1 when the median profusion took the category 1/0 or higher on the ILO (1980) 12–point scale, and a value of 0 otherwise. Included in this analysis were 1104 films from the 4th PFR survey, in 1970. These were the films for which it was possible to calculate a median profusion, less one man who was excluded from all the logistic regression analyses because certain values of the variables estimating his dust exposures were not credible (see 5.4.4 above). The median profusion was of at least category 1/0 on the 12–point scale for 431 of these films.

The values in exhibit 7.2.3 show that, on its own, age was a very highly significant predictor of the probability of a 4th survey film achieving a median profusion of 1/0 or greater. The exposure variables considered for inclusion in these models were estimated exposures to respirable dust and to its quartz

Exhibit 7.2.3. Results of fitting different regression models to profusion of small opacities at 4th survey. Variable analysed is median profusion dichotomised between categories 0/1 and 1/0. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

		Parsimonious models			
Variables		1+/4/1	1+/4/2	1+/4/3	1+/4/4	1+/4/5
constant		-4.001 (12.98)	-3.661 (10.85)	-3.706 (10.95)	-3.974 (12.18)	-3.973 (12.43)
age at survey		0.07703 (12.23)	0.06266 (7.30)	0.06500 (7.65)	0.0772 (12.09)	0.07710 (12.14)
dust	previous		0.00759 (2.43)			
quartz	previous			0.1198 (2.09)		
dust	ISP 3				-0.00252 (0.29)	
quartz	ISP 3					-0.0218 (0.39)
deviance		1285.1	1279.1	1280.7	1285.0	1285.0
degrees of freedom		1102	1101	1101	1101	1101

Exhibit 7.2.4. Results of fitting different regression models to profusion of small opacities at 4th survey. Variable analysed is median profusion dichotomised between categories 1/2 and 2/1. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

		Parsimonious models			
Variables		2+/4/1	2+/4/2	2+/4/3	2+/4/4	2+/4/5
constant		-9.25 (7.48)	-9.36 (7.12)	-9.37 (7.12)	-9.17 (6.84)	-9.10 (6.93)
age at survey		0.1174 (5.29)	0.1213 (4.64)	0.1214 (4.90)	0.1170 (5.16)	0.1161 (5.11)
dust	previous		-0.00178 (0.29)			
quartz	previous			-0.033 (0.30)		
dust	ISP 3				-0.0044 (0.19)	
quartz	ISP 3					-0.060 (0.39)
deviance		307.3	307.2	307.2	307.3	307.2
degrees of freedom		1102	1101	1101	1101	1101

fraction, in two time periods: inter-survey period (ISP) 3, which was the period between the 3rd and the 4th survey; and the individual's whole working life (if any) up to the time of 3rd survey, which is labelled "previous". The greatest decrease in deviance arose from the inclusion of the previous dust exposure, with a deviance change of 1285.1 - 1279.1 = 6.0 with 1 degree of freedom. The 2.5% critical point of the chi-squared distribution is 5.02, so this term is statistically significant. (The same conclusion would be drawn from the partial t-statistic 2.43.) The equivalent deviance change for previous quartz exposure was somewhat smaller at 4.4, again for 1 degree of freedom, and neither of the variables for dust and quartz exposure during ISP 3 made a useful contribution to the fit of the model. Thus, exposure to dust up to 3rd survey appeared the best predictor, amongst those considered, of median profusion 1/0 or greater. With age and previous dust exposure in the model, none of the other exposure variables gave a further significant reduction in the deviance; these additional model fits are not shown here.

Results of similar analyses are shown in exhibit 7.2.4, where the response variable this time is the binary variable taking the value 1 if the median profusion at 4th survey achieved a category of 2/1 or greater; this is indicated by the first component of each column label, 2+. There were 41 positive responses at this level. In contrast to the analysis above of the less severe response, none of the variables considered made a significant contribution to the response of category 2/1 or greater. This did not seem to be due to confounding of long-term exposure with age, because inclusion of the exposure variables did not much alter either the estimated value or the statistical significance of the age coefficient; the correlation between age and, for example, the previous dust exposure was only 0.76. No combination of exposure variables improved the fit over that of the models with the variables taken singly, and results from these fits are not shown. It was concluded that none of the exposure variables available was a useful predictor of median profusions equal to or greater than category 2/1 at this survey.

7.2.4 Logistic regression analyses of 5th survey films

The median profusions calculated for the films from the 5th PFR survey, in 1974, were analysed by logistic regression, and the results are presented as described in 7.2.2. Exposure variables containing estimates of exposure to respirable dust and its quartz fraction were available for ISP 4, which was the period between the 4th and 5th surveys, and these variables were considered for inclusion in regression models, along with the ISP 3 and previous dust and quartz exposure variables.

Exhibit 7.2.5 presents some results from the logistic regressions of the binary variable taking the value 1 if the median profusion of the 5th survey films was category 1/0 or greater. Of the 870 films included in this analysis, 206 had median profusions qualifying them as positive responses on this classification. Neither of the previous exposure variables yielded much improvement when added to the model with only a constant and age, and these fits are not shown. Of the other exposure variables, the ISP 4 quartz exposures made the greatest contribution to the fit, with a highly significant deviance reduction of 27.4 on 1 degree of freedom. The ISP 4 dust exposure variable gave almost as large a contribution, with a deviance reduction of 23.4. The ISP 3 variables were much less significant, with deviance reductions of 6.3 for the dust and 8.7 for the quartz variables respectively. Models 1+/5/6 and 1+/5/7 show the results of adding the previous and ISP 3 quartz variables to the model containing age and ISP 4 quartz. Inclusion of ISP 3 quartz made no useful contribution to the fit. Adding previous quartz gave a deviance reduction of 2.8, which with 1 degree of freedom exceeds the formal 10% critical point of the chi-squared distribution; its inclusion reduced the magnitude and

Exhibit 7.2.5. Results of fitting different regression models to profusion of small opacities at 5th survey. Variable analysed is median profusion dichotomised between categories 0/1 and 1/0. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models							
	1+/5/1	1+/5/2	1+/5/3	1+/5/4	1+/5/5	1+/5/6	1+/5/7	1+/5/8
constant	-5.351	-5.578	-5.550	-6.466	-6.161	-5.763	-6.189	-6.314
	(11.22)	(11.16)	(11.20)	(11.38)	(11.34)	(9.87)	(11.36)	(10.79)
age at survey	0.08788	0.08457	0.08575	0.0982	0.0964	0.0828	0.0979	0.0976
	(9.43)	(8.83)	(8.99)	(9.68)	(9.51)	(6.45)	(9.39)	(9.42)
dust previous								
quartz previous						0.1341		
						(1.67)		
dust ISP 3		0.0265						
		(2.53)						
quartz ISP 3			0.2061				-0.0570	
			(2.99)				(0.62)	
dust ISP 4				0.0763				0.0231
				(4.83)				(0.78)
quartz ISP 4					0.3720	0.3655	0.4110	0.284
					(5.18)	(5.10)	(4.29)	(2.13)
deviance	825.7	819.4	817.0	802.3	798.3	795.5	797.9	797.7
degrees of freedom	868	867	867	867	867	866	866	866

increased the standard error of the age term, while the ISP 4 quartz coefficient was almost unchanged. It was concluded that there was no strong evidence that previous quartz was a useful predictor of median profusion of category 1/0 or greater, when ISP 4 quartz was already in the model.

Model 1+/5/8 shows the effect of including both dust and quartz from ISP 4 in the model. The inference to be drawn from the partial t–statistics is that after allowance for ISP 4 quartz, ISP 4 dust was not a useful addition to the model; but that after ISP 4 dust, which was highly significant on its own, ISP 4 quartz made a significant contribution. This is interpreted as evidence that the quartz variable contains information about hazard which is additional to that contained in the non–quartz fraction.

Similar analyses, not shown, showed that previous and ISP 3 dust exposure variables made even less useful contributions, in the presence of ISP 4 quartz, than their quartz counterparts.

Exhibit 7.2.6 presents results from logistic regression analyses of the binary variable taking the value 1 when the median profusion calculated for a 5th survey film was of category 2/1 or greater, indicated in the column labels as 2+. There were only 14 films from 870 so classified. The results were similar to those from the analysis of the 1+ variable, although there was hardly any difference between the deviance reductions for the ISP 4 dust and quartz variables, at 4.4 and 4.6 respectively. Neither the previous exposures (not shown) nor the ISP 3 exposures gave a significant deviance reduction, whether the ISP 4 quartz (or dust) exposure was in or out of the model. Fitting a model with both ISP 4 dust and ISP 4 quartz suggested that neither contributed information which was not contained in the other.

Exhibit 7.2.6 Results of fitting different regression models to profusion of small opacities at 5th survey. Variable analysed is median profusion dichotomised between categories 1/2 and 2/1. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models							
Variables	2+/5/1	2+/5/2	2+/5/3	2+/5/4	2+/5/5	2+/5/6	2+/5/7	2+/5/8
constant	-7.71	-8.10	-8.12	-9.60	-8.92	-8.61	-8.91	-9.37
	(5.05)	(4.83)	(4.91)	(4.51)	(4.78)	(4.19)	(4.66)	(4.39)
age at survey	0.0735	0.0699	0.0714	0.0938	0.0874	0.0776	0.0865	0.0920
	(2.56)	(2.30)	(2.36)	(2.57)	(2.61)	(1.80)	(2.48)	(2.54)
dust previous								
quartz previous						0.088		
						(0.37)		
dust ISP 3		0.0369						
		(1.12)						
quartz ISP 3			0.307				0.035	
			(1.62)				(0.13)	
dust ISP 4				0.0996				0.0481
				(2.17)				(0.57)
quartz ISP 4					0.412	0.405	0.387	0.250
					(2.58)	(2.39)	(1.66)	(0.76)
deviance	135.4	134.2	133.2	131.0	130.8	130.6	130.7	130.4
degrees of freedom	868	867	867	867	867	868	866	866

7.2.5 Logistic regression analyses of 6th survey films

The median profusions calculated for the films from the 6th PFR survey, in 1978, were analysed by logistic regression, and the results are presented as described in 7.2.2. Exposure variables containing estimates of exposure to respirable dust and its quartz fraction were available for ISP 5, which was the period between the 5th and 6th surveys, and these variables were considered for inclusion in regression models, along with the ISP 4, ISP 3 and previous dust and quartz exposure variables.

Exhibit 7.2.7 presents some results from the logistic regressions of the binary variable taking the value 1 if the median profusion of the 6th survey films was category 1/0 or greater, and 0 otherwise. Of the 617 films included in this analysis, 120 had median profusions qualifying them as positive responses on this classification. Neither of the previous exposure variables yielded much improvement when added to the model with only a constant and age, and these fits are not shown. Of the other exposure variables, the ISP 4 quartz exposure made the greatest contribution to the fit, with a highly significant deviance reduction of 27.3 on 1 degree of freedom. The ISP 4 dust exposure variable gave a smaller contribution, with a deviance reduction of 19.8. ISP 3 dust and quartz exposures gave reductions of 15.4 and 19.8, and ISP 5 dust and quartz gave reductions of 14.2 and 16.2 respectively. All of these reductions exceeded the 0.1% significance point of the chi-squared distribution with 1 degree of freedom. Models 1+/6/8 and 1+/6/9 show the results of adding the ISP 3 and ISP 5 quartz variables to the model containing age and ISP 4 quartz, chosen as a baseline because it gave the largest reduction of any single variable. The largest additional reduction in deviance was only 1.8 for either ISP 3 quartz or ISP 5 quartz, and this was well short of the 10% significance point of the chi-squared distribution. Similar analyses

Exhibit 7.2.7. Results of fitting different regression models to profusion of small opacities at 6th survey. Variable analysed is median profusion dichotomised between categories 0/1 and 1/0. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models									
Variables	1+/6/1	1+/6/2	1+/6/3	1+/6/4	1+/6/5	1+/6/6	1+/6/7	1+/6/8	1+/6/9	1+/6/10
constant	-5.340	-5.521	-5.421	-5.912	-5.869	-6.175	-6.371	-5.778	-6.199	-5.869
	(8.69)	(8.47)	(8.22)	(8.61)	(8.61)	(9.09)	(8.83)	(8.39)	(8.41)	(8.51)
age at survey	0.0831	0.0717	0.0730	0.0794	0.0831	0.0877	0.0954	0.0789	0.0881	0.0831
	(6.88)	(5.53)	(5.58)	(6.08)	(6.35)	(7.05)	(7.16)	(5.83)	(6.40)	(6.30)
dust previous										
quartz previous										
dust ISP 3		0.0496								
		(3.95)								
quartz ISP 3			0.3627					0.146		
			(4.45)					(1.33)		
dust ISP 4				0.0845						-0.0003
				(4.41)						(0.01)
quartz ISP 4					0.4310			0.333	0.356	0.432
					(5.16)			(3.00)	(3.56)	(2.65)
dust ISP 5						0.0832				
						(3.79)				
quartz ISP 5							0.425		0.175	
							(4.07)		(1.34)	
deviance	540.7	525.3	520.8	520.8	513.4	526.5	524.5	511.6	511.6	513.4
degrees of freedom	615	614	614	614	614	614	614	613	613	613

(not shown) adding previous dust and quartz, and ISP 3 and ISP 5 dust exposures showed that they made even less useful additional contributions than their quartz counterparts. It was concluded that the other variables were not useful additional predictors when age and ISP 4 quartz exposure were already in the regression model. Model 1+/6/10 includes both ISP 4 dust and ISP 4 quartz variables, and shows a pattern similar to 1+/5/10, indicating that the quartz variable contained information additional to that contained by the dust variable.

Exhibit 7.2.8 presents results from logistic regression analyses of the binary variable taking the value 1 when the median profusion calculated for a 5th survey film was of category 2/1 or greater, indicated in the column labels as 2+. There were only 15 films from 617 so classified. The results were similar to those from the analysis of the 1+ variable, with reductions in deviance of 19.8 for ISP 4 dust and 18.4 for ISP 4 quartz. Models 2+/6/8 and 2+/6/9 show the results for adding other variables to the model with age and ISP 4 dust; the smallest deviance, 110.5 with ISP 3 quartz added, gave a deviance reduction of 3.3 over the value of 113.8 achieved by the model with age and ISP 4 dust, which was less than the conventional 5% significance level of the chi-squared distribution with one degree of freedom. Model 2+/6/10 includes both ISP 4 dust and ISP 4 quartz. As can be deduced from the results for the individual variables, ISP 4 dust made a greater additional contribution after ISP 4 quartz than vice versa. However, in neither case did the additional deviance reduction exceed the 10% significance level.

In exhibit 7.2.8, it is notable that, for all the models 2+/6/2 to 2+/6/4, and 2+/6/8 to 2+/6/10, the partial t-statistic for the age coefficient was less than the formal 5% significance point of the t-distribution. For these models, the question of whether age is an essential term is less clear than than in the previous analyses of exhibits 7.2.3 to 7.2.7.

Exhibit 7.2.8. Results of fitting different regression models to profusion of small opacities at 6th survey. Variable analysed is median profusion dichotomised between categories 1/2 and 2/1. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models									
	2+/6/1	2+/6/2	2+/6/3	2+/6/4	2+/6/5	2+/6/6	2+/6/7	2+/6/8	2+/6/9	2+/6/10
constant	-7.15	-8.28	-7.54	-9.05	-8.35	-8.95	-9.21	-8.69	-8.18	-8.92
	(4.65)	(4.12)	(3.83)	(4.43)	(4.32)	(4.85)	(4.75)	(4.04)	(3.96)	(4.51)
age at survey	0.0721	0.0520	0.0473	0.0612	0.0723	0.0815	0.0962	0.0473	0.0565	0.0651
	(2.43)	(1.40)	(1.24)	(1.63)	(2.00)	(2.48)	(2.77)	(1.18)	(1.43)	(1.79)
dust previous										
quartz previous										
dust ISP 3		0.1176								
		(3.57)								
quartz ISP 3			0.762					0.458	0.472	
			(4.05)					(1.88)	(1.97)	
dust ISP 4				0.2126				0.1529		0.1419
				(4.02)				(2.40)		(1.67)
quartz ISP 4					0.690				0.450	0.289
					(4.38)				(2.26)	(1.04)
dust ISP 5						0.1545				
						(3.09)				
quartz ISP 5							0.691			
							(3.20)			
deviance	133.6	120.1	116.5	113.8	115.2	124.5	124.4	110.5	111.5	112.8
degrees of freedom	615	614	614	614	614	614	614	613	613	613

7.2.6 <u>Logistic regression analyses of 1980 films</u>

The median profusions calculated for the films from the NCB Medical Service's PXR survey in 1980 were analysed by logistic regression and the results are presented as described in 7.2.2. As for the films from the 6th survey, the exposure variables available for inclusion in the regression models were those estimating individuals' exposures to previous, ISP 3, ISP 4 and ISP 5 respirable dust and quartz.

Exhibit 7.2.9 presents some results from the logistic regressions of the binary variable taking the value 1 if the median profusion of the 1980 survey films was category 1/0 or greater. Of the 444 films included in this analysis, 119 had median profusions qualifying them as positive responses on this classification. Neither of the previous exposure variables yielded much improvement when added to the model with only a constant and age, and these fits are not shown. Of the other exposure variables, the ISP 3 quartz exposures made the greatest contribution to the fit, with a highly significant deviance reduction of 18.9 on 1 degree of freedom. The ISP 4 dust and quartz exposure variable each gave almost as large a contribution, with deviance reductions of 18.1 and 16.3 respectively. ISP 3 and ISP 5 dust and ISP 5 quartz exposures gave reductions of 10.3, 10.2 and 10.4 respectively. These reductions were all well in excess of the 1% significance level of the chi-squared distribution with 1 degree of freedom. Models 1+/80/8 and 1+/80/9 show the results of adding the ISP 4 and ISP 5 quartz variables separately to the model containing age and ISP 3 quartz, which was the exposure variable which gave the greatest individual deviance reduction. The additional reductions were 2.3 and 1.5 respectively, both of which were well short of the 10% significance level of chi-squared with 1 degree of freedom. Similar analyses fitting other combinations of variables produced even smaller deviance reductions. It was

Exhibit 7.2.9. Results of fitting different regression models to profusion of small opacities at 1980 survey. Variable analysed is median profusion dichotomised between categories 0/1 and 1/0. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models									
Variables	1+/80/1	1+/80/1	1+/80/1	1+/80/1	1+/80/1	1+/80/6	1+/80/7	1+/80/8	1+/80/9	1+/80/10
constant	-5.141	-5.239	-5.217	-5.652	-5.520	-5.826	-5.920	-5.379	-5.533	-5.179
	(7.72)	(7.21)	(7.17)	(7.58)	(7.56)	(7.71)	(7.73)	(7.12)	(6.94)	(7.12)
age at survey	0.0854	0.0746	0.0748	0.0823	0.0844	0.0886	0.0941	0.0774	0.0800	0.0798
	(6.59)	(5.18)	(5.24)	(5.84)	(6.04)	(6.37)	(6.63)	(5.26)	(5.23)	(5.46)
dust previous										
quartz previous										
dust ISP 3		0.0434								-0.0538
		(3.20)								(1.64)
quartz ISP 3			0.3839					0.265	0.321	0.702
			(4.28)					(2.22)	(3.13)	(3.27)
dust ISP 4				0.0766						
				(3.73)						
quartz ISP 4					0.3751			0.188		
					(3.95)			(1.50)		
dust ISP 5						0.0748				
						(3.19)				
quartz ISP 5							0.360		0.163	
							(3.26)		(1.24)	
deviance	458.1	447.8	439.2	444.0	441.8	447.9	447.7	436.9	437.7	436.4
degrees of freedom	442	441	441	441	441	441	441	440	440	440

concluded that the other variables were not useful additional predictors when age and ISP 3 quartz exposure were already in the regression model. Model 1+/80/10 includes both ISP 3 dust and ISP 3 quartz variables. In the presence of ISP 3 dust, the addition of ISP 3 quartz was highly significant. After ISP 3 quartz, the addition of ISP 3 dust just reached the 10% significance level, and produced a negative regression coefficient. It was concluded that, as in the analyses of the 5th and 6th survey median profusions of 1/0 and greater, the quartz exposure variable contained more information related to the response variable than did the equivalent non-quartz dust variable.

Exhibit 7.2.10 presents results from logistic regression analyses of the binary variable taking the value 1 when the median profusion calculated for a PXR 1980 survey film was of category 2/1 or greater, indicated in the column labels as 2+. There were only 15 films from 444 so classified. The results showed the greatest deviance reduction at 10.8 for ISP 4 dust, followed by 9.2 for ISP 4 quartz. Reductions due to ISP 5 dust and quartz variables were only slightly smaller, at 6.9 and 5.1, while the reductions for ISP 3 dust and quartz were also close at 5.4 and 6.8 respectively. Models 2+/80/8 and 2+/80/9 show the results of adding other variables to the model with age and ISP 4 dust. The smallest deviance, 110.5 with the addition of ISP 3 quartz, gave a deviance reduction of 3.3, which is well below the 10% significance level. Model 2+/80/10 included both ISP 4 dust and ISP 4 quartz, and neither made a significant contribution in the presence of the other.

In exhibit 7.2.10, even more than in 7.2.8, the age coefficient is seen to be of dubious necessity for the models, as evidenced by the small values of the partial t-statistics. However, conclusions from both of those exhibits concerning the effects of age or of exposures must be considered suggestive rather than conclusive, because of the small numbers of positive responses on which the analyses were

Exhibit 7.2.10. Results of fitting different regression models to profusion of small opacities at 1980 survey. Variable analysed is median profusion dichotomised between categories 1/2 and 2/1. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models									
Variables	2+/80/1	2+/80/1	2+/80/1	2+/80/1	2+/80/1	2+/80/6	2+/80/7	2+/80/8	2+/80/9	2+/80/10
constant	-5.12	-4.97	-4.84	-5.73	-5.35	-6.17	-6.15	-4.67	-5.46	-5.66
	(4.07)	(3.31)	(3.29)	(3.67)	(3.70)	(4.00)	(3.99)	(2.63)	(3.36)	(3.58)
age at survey	0.0369	0.0101	0.0124	0.0172	0.0261	0.0362	0.0460	-0.0180	0.0102	0.0184
	(1.48)	(0.31)	(0.40)	(0.55)	(0.90)	(1.26)	(1.61)	(0.42)	(0.30)	(0.58)
dust previous								0.0211		
								(1.25)		
quartz previous										
dust ISP 3		0.0731								
		(2.30)								
quartz ISP 3			0.502						0.171	
			(2.65)						(0.64)	
dust ISP 4				0.1536				0.1527	0.1297	0.1190
				(3.19)				(3.16)	(2.08)	(1.41)
quartz ISP 4					0.525					0.155
					(3.25)					(0.51)
dust ISP 5						0.1324				
						(2.67)				
quartz ISP 5							0.497			
							(2.46)			
deviance	128.8	123.4	122.0	118.0	119.6	121.9	123.7	116.5	117.7	117.8
degrees of freedom	442	441	441	441	441	441	441	440	440	440

based.

7.2.7 The effects of smoking on profusions of small opacities

All of the analyses shown in exhibits 7.2.3 to 7.2.10 were from logistic regression models in which variables representing the individuals' smoking habits had not been included. Indeed, the analyses were based on all men for whom a median profusion could be calculated and for whom, in addition, apparently reliable exposure estimates were available, and availability of data on smoking habits was not taken into account. It has already been noted that the tabulation of age and smoking habits in exhibit 5.4.1 showed the loss of questionnaires including smoking data from a few attendances at the 4th and 5th surveys, and the less trivial loss of a batch of questionnaires from the 6th, representing about 7% of those attending then. As a result, any analysis which includes smoking as a nuisance variable must necessarily be based on a subset of those already analysed; and the difference in the numbers for the 6th survey data is considerable.

Exhibit 7.2.11 shows a breakdown of the numbers of films available for analysis from each survey, by smoking habits where known and by grouped profusion categories. Although each man makes at most one contribution within each survey, many men attended more than one survey, and thus contribute to more than one section of the exhibit. Each man's classification by smoking habit is, for the 4th, 5th and 6th surveys, based on his answers at that survey, so for example the same man may be a smoker at one survey and an ex-smoker at a later survey. The classifications by median profusion of small opacities have been grouped into the major categories 0 (0/–, 0/0, 0/1), 1 (1/0, 1/1, 1/2) and 2 or higher (2/1 and higher).

Exhibit 7.2.11. Table showing breakdown of numbers of men included in logistic analyses of median profusion at each survey, by grouped median profusion and smoking habits. Smoking data used for 1980 survey are those collected at 6th survey.

Survey	Smoking habit	Median profusion of small opacities: up to 0/1	1/0 - 1/2	2/1 & higher	Total
	Nonsmoker	113	36	1	150
4th	Exsmoker	71	60	3	134
	Smoker	488	294	37	819
	Unknown	1	0	0	1
	Combined	673	390	41	1104
	Nonsmoker	121	21	1	143
5th	Exsmoker	114	25	1	140
	Smoker	428	145	12	585
	Unknown	1	1	0	2
	Combined	664	192	14	870
	Nonsmoker	97	7	0	104
6th	Exsmoker	88	15	0	103
	Smoker	292	68	6	366
	Unknown	20	15	9	44
	Combined	497	105	15	617
	Nonsmoker	63	11	0	74
1980	Exsmoker	48	14	0	62
	Smoker	164	57	6	227
	Unknown	50	22	9	81
	Combined	325	104	15	444

No questionnaire data were collected at Medical Service's PXR surveys, so the 1980 films had no corresponding smoking data. Of the 444 films for which analyses of median profusions have already been presented, 363 were from men for whom data on smoking habits were available from their 6th survey questionnaires. Because the elapsed time between the 6th and 1980 surveys was only about two years, the 6th survey smoking data were used for the analyses of the corresponding 1980 films.

All of the regression models which were fitted to the median profusion variables, from which the most interesting results have been presented in the preceding sections, were also fitted in alternative versions which included variables representing the individuals' smoking habits as well as the variables for age and dust and quartz exposures. For the most part, these expressed the smoking effects as differences between groups of men classified by their responses at the relevant survey as current smokers, ex-smokers, and lifelong non-smokers. Such comparisons can only be estimated as differences, and it is necessary to make an arbitrary choice of one of the groups to act as a reference level. In these analyses, the non-smoking group was chosen as the reference, and the comparisons expressed as differences between the non-smokers and the remaining two groups. Tests of statistical significance of the differences amongst the three smoking groups are associated with two degrees of freedom.

Exhibit 7.2.12 presents some results from the analyses of data from all four surveys. The variable analysed is the binary variable taking the value 1 where the median profusion of small opacities was of category 1/0 or greater on the 12-point scale. The tabulation shows direct comparisons between various parsimonious logistic regression models taken from amongst the results already presented, and the same model augmented by two parameters for the differences between the smoking groups. The column labels of the models without the smoking effects are the

Exhibit 7.2.12. **Results of fitting regression models with and without smoking variables to median profusions of small opacities, dichotomised between categories 0/1 and 1/0, from each survey. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.**

	Parsimonious models							
Variables	1+/4/2	1+/4/2S	1+/5/5	1+/5/5S	1+/6/5	1+/6/5S	1+/80/3	1+/80/3S
constant	-3.656	-3.929	-6.198	-6.512	-6.168	-6.928	-5.344	-5.588
	(10.82)	(10.45)	(11.35)	(10.90)	(8.20)	(8.16)	(6.49)	(6.49)
age at survey	0.06256	0.05956	0.0971	0.1000	0.0880	0.0884	0.0782	0.0790
	(7.29)	(6.85)	(9.53)	(9.51)	(6.11)	(5.97)	(4.81)	(4.74)
dust previous	0.00760	0.00825						
	(2.44)	(2.63)						
quartz previous								
dust ISP 3								
quartz ISP 3							0.283	0.268
							(2.80)	(2.62)
dust ISP 4								
quartz ISP 4			0.3715	0.3848	0.3416	0.3291		
			(5.17)	(5.28)	(3.74)	(3.57)		
dust ISP 5								
quartz ISP 5								
ex-smoker v. non		0.587		-0.483		0.408		-0.093
		(2.07)		(1.38)		(0.82)		(0.19)
smoker v. non		0.407		0.319		1.021		0.373
		(1.81)		(1.14)		(2.40)		(0.97)
deviance	1279.0	1274.3	794.1	783.0	441.3	437.5	347.1	344.9
degrees of freedom	1100	1098	865	863	570	568	360	358

same as in the tables in which these results were first presented, and the labels for the models with the smoking effects are the same, except for the addition of the suffix S indicating that smoking effects are included.

In the model for median profusion 1/0 or greater at 4th survey, predictor variables were age at survey and dust exposure prior to 3rd survey. Since only one man had no smoking data, the results are almost identical to those for the same model in exhibit 7.2.3. The addition of smoking effects produced a deviance reduction of 4.7, which just exceeds the 10% significance level of the chi-squared distribution with two degrees of freedom. The estimated coefficients suggested that most of the difference, if real, was concentrated in a contrast between the non-smokers on one hand and the smokers and ex-smokers on the other, with little difference between these two groups. With smoking in the model, the age coefficient was reduced by about 5%, while the coefficient of previous dust was increased by under 9%.

For median profusion 1/0 or greater at 5th survey, the baseline model had predictor variables age and ISP 4 quartz exposure. Again, the results were very similar to those already presented in exhibit 7.2.5, because the data sets differed by only two men. Addition of smoking effects gave the highly significant deviance reduction of 11.1. The estimates of group differences, in contrast to those for the 4th survey, had the most extreme difference between the ex-smokers and the smokers, with the non-smokers roughly halfway between the two groups. Inclusion of the smoking effects increased both the age and the ISP 4 quartz coefficients by about only 3%.

At 6th survey, the baseline model again included age and ISP 4 quartz exposure, and the addition of smoking terms here gave a deviance reduction of 3.8, which is well short of the 10% significance level. Again, the estimated coefficients showed the highest risk in the current smokers, but the ex–smokers' risk estimate was little greater than that of the non–smokers. Inclusion of the smoking effects had negligible influence on the age coefficient, and decreased the ISP 4 quartz coefficient by less than 4%. However, comparison of model 1+/6/5 with the same model in exhibit 7.2.7 shows that the estimated ISP 4 quartz coefficient was 21% smaller when calculated from the subset of the attendances for which smoking data were available.

For the analysis of the data from the 1980 survey, results are shown with the baseline model which includes age and ISP 3 quartz exposure. The deviance reduction due to smoking effects was only 2.2, well short of conventional significance levels. Even for the contrast between the smokers and non–smokers, the estimated effect was less than its standard error. Inclusion of smoking once more made little difference to the age and exposure coefficients. As for the 6th survey data, however, the estimate of the exposure coefficient was considerably smaller when based on the subset of the data which included smoking habits.

Results from many other models gave similar results, but are not shown here. The estimates of the smoking effects were almost constant for a particular set of data, regardless of the baseline model chosen or of which exposure variables were included or excluded, as were the sizes of the corresponding reductions in deviance. Furthermore, the inclusion or exclusion of smoking effects made hardly any difference to the estimates of the regression coefficients for age or exposure variables. In particular, no difference was observed in the ranked order of the deviances with different choices of exposure variables, in models with and without smoking effects. Additional models were fitted in which different estimates of the coefficients for the age and exposure variables were made in each of the three smoking groups. None of these produced any significant improvement in model fit over the models with parallel coefficients. There was therefore no evidence that the effects of age or occupational exposures on small opacities was influenced by the individuals' smoking habits.

It was not clear why smoking should appear to be a highly significant contributor to the models for the 5th survey data but not for the data from the other surveys. Nor was there an obvious interpretation of the magnitudes of the estimated coefficients, which suggested that the ex-smokers had a lower prevalence of small opacities than the non-smokers. However, for the present purposes, the most important conclusion from these analyses was that there was little or no partial confounding between smoking habits and exposures to dust or quartz, in that the conclusions to be drawn from the analyses already presented would not be altered by the inclusion of smoking variables.

The radical changes, on inclusion of smoking effects, to the exposure estimates from the 6th and 1980 surveys were likely to be related to some sort of selection effect in the lost batch of questionnaires. It can be seen from exhibit 7.2.11 that those for whom smoking habits are now unknown had a higher proportion of films for which the median profusion was 1/0 or greater, than the remainder whose data were analysed in relation to smoking. It was concluded that estimates of the effects of dust and quartz exposure from the reduced data set with questionnaires were probably subject to some distortion due to this selection effect, and that since smoking made little difference to the estimates in the reduced data set, it would be safer to base conclusions on the earlier analyses of the larger data set, without allowance for smoking.

7.2.8 Adequacy of the median as a summary of readings

All the regression results in the previous section have been derived from analyses of summaries of the film readings, obtained by calculating a median from the available assessments of profusion of small opacities, as described in 7.2.1. However, as is shown in 7.1.3 and 7.1.4, the readers did not record identical profusions for the same films, nor identical rankings for the profusions. And, although in general the agreement between the duplicate readings performed by each of the readers was better than between different readers, even the same reader did not produce identical rankings of the profusions at different readings.

This finding would not surprise anyone with experience of data from the subjective assessment of radiological abnormality, but it raises the question of the adequacy of the chosen method of summarising these data by a median. It is legitimate to enquire whether the use of the median has a stabilising effect over the variation in the data, or whether its use incurs an important loss of information over what might be retrieved by analysis of the separate assessments made by the individual readers.

Exhibit 7.2.13 is designed to shed some partial light on this question. To produce this tabulation, data were extracted for the five readers' assessments of small opacities profusion on the films from the fifth survey, which were read on two separate occasions. Attention was focussed on the 617 films for which duplicate readings existed from all five readers, and the ten individual assessments, along with their median, were converted into binary variables by dichotomising between categories 0/1 and 1/0. Regression analyses similar to those reported in exhibit 7.2.5 were performed on each of the binary responses; the initial model containing just a constant and a term for age at survey was augmented by the quartz and

Exhibit 7.2.13. Residual deviances obtained by fitting logistic regression models separately to each of five readers' duplicate assessments of profusion of small opacities at 5th survey, dichotomised between categories 0/1 and 1/0. Models fitted included a constant, age and the individual exposure variables one at a time.

Variables		Reader/occasion 101/1	104/1	105/1	106/1	112/1	101/2	104/2	105/2	106/2	112/2	median
age at survey		838.9	613.0	691.0	152.3	566.3	824.5	583.6	720.2	200.1	629.1	540.5
+dust	previous	838.9	613.0	687.1	151.5	564.3	816.8	583.3	717.7	200.1	626.9	539.9
+quartz	previous	838.9	613.0	687.0	151.8	563.8	817.6	582.6	716.8	200.1	624.9	539.8
+dust	ISP 3	834.5	606.1	668.3	139.6	548.3	817.0	572.1	707.5	188.8	614.1	525.1
+quartz	ISP 3	833.9	602.4	670.0	133.2	542.3	814.8	569.0	702.2	185.0	610.4	520.6
+dust	ISP 4	832.3	602.9	666.8	140.4	545.8	817.7	568.1	700.8	190.0	613.0	520.2
+quartz	ISP 4	832.9	596.6	670.2	140.9	543.6	815.9	563.8	702.5	190.2	609.9	512.3
+dust	ISP 5	831.7	608.4	673.5	144.3	557.4	820.3	574.4	704.6	192.7	619.3	526.2
+quartz	ISP 5	834.9	606.7	677.2	141.3	558.7	821.5	574.4	709.3	190.4	619.4	524.1

non-quartz dust exposure variables from ISPs 3, 4, 5 and the previous period, taken one at a time in turn. Exhibit 7.2.13 contains the residual deviances obtained from each of these models, organised in columns labelled with the identity of the response variable being analysed, so that 101/1 labels reader 101's first set of readings. The deviances are not readily comparable within rows, because they are in part a function of the number of positive responses, and these numbers differed widely between readers. It is possible, however, to compare across columns the pattern of the deviances within each column, and the conclusions which would be drawn from them, about the relative importance in the regression models of the various exposure variables.

The strongest impression from exhibit 7.2.13 is how well the pattern within each column resembles that for the median. Firstly, the addition of either previous dust or previous quartz produced only a little improvement in the regression model, with the exception of that for the assessments made by reader 101 at her second reading. And, again with that single exception, the ISP 3 and ISP 4 dust and quartz variables made rather more substantial contributions to the models than the previous dust and quartz. There was no clear unanimity in the choice between the ISP 3, 4 and 5 and previous exposure variables, in the matter of which produced the greatest reduction in the deviance. The only column in which the best model choice was from ISP 5 exposure was 101/1, and of the remainder the best fit was obtained by ISP 3 quartz in four of the columns, ISP 4 quartz in three, and ISP 4 non-quartz dust in two. In the majority of the columns, the quartz variables performed better than the corresponding non-quartz dust variables. Comparison of data from the same readers at different occasions showed that for readers 104, 105 and 106 the same variable appeared best at both occasions, which provides some general reassurance about the consistency of the panel members' reading habits over time. Even where the same variable was not the best both times, there appears close similarity between the rankings of the deviances, given that the results from different readers were based on widely differing numbers of positive responses.

It was concluded that, at least on the present subset of the data, the median had preserved and summarised the principal features of the individual readings, in the strength of their relationships with the exposure variables. This conclusion was drawn in the knowledge that, as with all summary measures, the use of the median may clarify any relationship underlying the joint distribution of the readings, but at the expense of the loss of information about the individual readings' deviations from that relationship.

7.2.9 Additional checks on adequacy of logistic models

All the results shown in exhibits 7.2.3 to 7.2.10 were from logistic models in which the logistic transform of the fitted frequencies was assumed to be a linear function of each of the explanatory variables, which in this case were age and the various exposure variables. Linearity on the logistic scale with respect to the explanatory variables is merely an assumption, which has the merit of simplicity, both in application and in expression. In some applications, the choice of a linear or other form of response may be made on theoretical considerations. In the present case, we have no such theoretical justification. Intuitively, and for consistency with previous findings on pneumoconiosis, we expect any relationships found to be at least approximately monotonically increasing with exposure to the relevant hazard. In such cases, models linear in the variables are often a good fit to the data, but alternatives using, for example, the logarithms of all or some of the variables may give improved fits.

One possible way to approach this question is to fit large numbers of models with different combinations of logarithms or polynomials of the explanatory variable, and to compare the fits of the different models. A different approach (see, e.g., McCullagh and Nelder, 1983) is to embed the linear term within a transformation family of which linearity is a special case, and to investigate whether the transformed scale produces a better fit than the linear scale. This usually involves estimating at least one parameter of the transformation family, in a separate computational step from that for the regression parameters. Both of these approaches may require a considerable amount of computation, although various short cuts and approximations can be employed.

An alternative approach has been employed here, which falls in spirit somewhere between those described in the last paragraph, but which requires less computation, and which may in some circumstances give a better picture of the adequacy or otherwise of linear effects. The ordered values of an explanatory variable are grouped, and different constants fitted for each group. The estimated effects from the groups can then be related, either arithmetically or graphically, to values chosen to be typical of the group; the arithmetic mean of the grouped values is usually a convenient choice. It should then be apparent whether any relationship which exists is linear, or curvilinear, and whether a curve is monotonic. If non-linear effects are appropriate, the grouped effects should give an improved fit which will be reflected in the deviance. Strictly, the linear and grouped models are not hierarchical, so that they do not correspond to an exact partition of the deviance; but this is unlikely to conceal serious non-linearity. (The deviance does partition exactly between a model with a separate effect for each group and one linear in the group means rather than in the actual data values.)

For the data from each survey, groups were constructed from the variables for age and for relevant dust and quartz exposures, by splitting each into six ordered groups containing the same numbers of data values (as nearly as possible). Models were then fitted for the response variable indicating a median profusion of at least 1/0, with the age and exposure variables being replaced, one at a time, by their grouped equivalents. Additionally, models were fitted which contained grouped age and the individual exposures, to investigate whether the estimates of exposure effects altered in the presence of non-linear age effects.

The results were almost entirely negative. For the exposure variables, there was no evidence of serious non-linearity in the relationships already noted. Little change was observed in the exposure estimates when non-linear age effects were fitted. At each of the 4th, 5th and 6th surveys, there was little evidence of non-linearity in age when exposure variables were not fitted. At the 1980 PXR survey, however, the decrease in the deviance when grouped age was fitted exceeded the formal 5% significance of chi-square with 4 degrees of freedom. Inspection of the data showed that the proportion of films for which the median profusion was 1/0+ increased over the first four groups (0%, 18%, 25%, 39%) but was relatively static thereafter (42%, 38%). This may be a chance occurrence due to multiple testing, or it may be due to some real phenomenon; a relative deficit of men with small opacities amongst the older workers may be due to selective losses from the colliery population during the contraction which was then taking place in the workforce. Nevertheless, inclusion of grouped instead of linear age effects did not alter the conclusions regarding the relative importance of the different exposure variables, and did not seriously alter the regression coefficients estimated for those variables.

7.2.10 Investigations of alternative models

In all the analyses of small opacities profusion presented so far, the approach to the analyses has been the same. The data for a particular radiograph have been summarised by taking the median value of the assessments made by a panel of readers; this summary median has then been analysed as a binary variable, dichotomised on the boundary between two adjacent categories on the ILO (1980) scale; the analysis has been by logistic regression, and there has been some investigation of the effects of introducing non-linear and non-parallel terms in the regression equation.

The step from ten (or five) independent assessments on a potentially twelve-point scale of abnormality to a single binary variable represents the maximum possible degree to which the data could be simplified and summarised. This or a similar approach is often used to analyse data of this sort, because the method is relatively easy to comprehend and software for fitting logistic models to binary data is widely available. However, it is legitimate to ask at what cost, in terms of lost information, are the gains of simplicity in analysis obtained, and are there alternative methods which, with a little more effort, might produce more informative analyses?

This section reports a limited case study in which the same data were analysed in a variety of ways, and compares the results. The data chosen were the assessments made by the five readers at their second examination of the films from the 6th survey. All films were included for which all five readers recorded valid assessments of the profusion of small opacities; these numbered 617. Data on age and exposure variables were extracted along with these readings, to form a basis for the investigation. The purpose was to answer the following question; would different methods of analysis lead to quantitatively different estimates of the effects of exposures on small opacities and/or different judgments on the statistical significance of those estimates?

Initially, four different approaches were tried. In the first, the five assessments were summarised by a median value, which was then converted to a binary variable taking the value 1 if the median was of category 1/0 or greater on the ILO (1980) scale, and 0 otherwise. This was then analysed by ordinary logistic regression. This method, which is analogous to the method used for up to ten readings for the films from the 4th, 5th and 6th surveys, and which produced the results already extensively tabulated and presented, is referred to as "median binary".

The second method of analysis used a slightly less severe process of summarisation to produce a response variable. The individual assessments, rather than the median, were each converted to binary variables with category 1/0 or greater again taken as a positive response. The five readers' results were then summarised as the number of positive responses for a film; these counts were then analysed as arising from a sample size of five, using logistic regression methods. This is referred to simply as analysis of "counts".

Both the third and fourth methods used the individual assessments, rather than a summary, and both once more converted these to binary variables taking positive values for categories 1/0 and greater. The third simply analysed these as if they were independent binary observations, giving a data set of 5*617 = 3085 data points; it was necessary to duplicate the explanatory variables for each film to the individual assessments. This is referred to as "individual binary".

The fourth method of analysis used the same data set as the third, but applied the model described in full by Bonney (1987) and more briefly above in 6.1.3, to allow for the dependence between results recorded for the same film by different readers. This involved estimating sets of logistic parameters describing the second- and higher order terms for lack of independence, simultaneously with the regression coefficients of the explanatory variables. This is referred to as "Bonney binary".

Each of these four methods used a variant of logistic analysis as commonly applied to binary or counted data. Some additional analyses were performed, using models which extend the use of logistic regression to data in ordered categories. The formulation of Anderson (1984) was employed, as described in 6.1.4, and two sets of analyses were performed.

The fifth set of analyses, and the first to use the ordered categorical methods, took the median which had been calculated from the five assessments; this variable took values on the ILO (1980) scale, and could therefore be treated as an ordered categorical variable in its own right, rather than being converted to a binary variable. This analysis is referred to as "median ordered".

The sixth and final set of analyses was motivated by the observation that, since the five readers did not have identical reading habits, the counts obtained over the individual binary variables might not be distributed like binomial variables, and it might be inappropriate to use logistic regression as if they were. However, the counts of positive results out of five readers created a six-category ordered response, and this was analysed using the same methods. This is referred to as "ordered counts".

Exhibit 7.2.14 shows some results from these six sets of analyses, labelled as described above. In each column are shown the estimated regression coefficients for the dust and quartz exposure variables when they were added to a baseline model one at a time. This table is unlike earlier presentations of regression results, in that a column shows only the exposure coefficients from a variety of different models, rather than the details of a particular model; the corresponding constant and age terms would be different for each model, and are not shown, since it is the effect on the exposure coefficients which is of primary interest.

The first column shows the results of the median binary analysis. Although based only on five rather than ten readings, this analysis produced results very similar to those already presented in exhibit 7.2.7, with the ISP 4 quartz making the most significant contribution, closely followed by ISP 4 dust and ISP 3 quartz.

The second column shows the corresponding results from the analysis of the counts as a binomial response. Since the denominator was greater than unity, the residual deviance supplied a test for goodness of fit of the binomial model, which was very highly significant, suggesting significant over-dispersion. This was not surprising, since the data would only be expected to follow the binomial distribution if the readers did not differ. A heterogeneity factor to adjust for this over-dispersion (Williams, 1982) was calculated as the ratio of the smallest residual deviance, 1306.8, to the degrees of freedom 614. Standard errors of the estimates were scaled upwards by the square root of this factor, that is by 1.46. Even after this scaling, the partial t-statistics for all the exposure variables were higher than for the median binary analyses, by up to about 40%, although in all cases the actual coefficients were smaller.

Exhibit 7.2.14. **Comparisons of regression coefficients for individual exposure variables fitted separately, in addition to baseline models,** in various alternative analyses of small opacities profusion at 6th survey. Tabulated are estimated regression coefficients and, in parentheses, absolute ratio of estimate to standard error.

Type of analysis (see text)

Variable added	Median binary	Counts	Individual binary	Bonney binary	Ordered median	Ordered counts
dust previous	0.00909	0.00740	0.00825	0.00494	0.01499	0.01383
	(1.64)	(2.10)	(3.17)	(1.72)	(2.44)	(2.07)
or						
quartz previous	0.182	0.1633	0.1839	0.1176	0.3397	0.3156
	(1.69)	(2.39)	(3.62)	(2.11)	(2.69)	(2.35)
or						
dust ISP 3	0.0382	0.03039	0.03607	0.02645	0.04341	0.05061
	(3.64)	(4.72)	(7.53)	(5.03)	(3.99)	(4.34)
or						
quartz ISP 3	0.3366	0.2587	0.3064	0.2252	0.4052	0.4325
	(4.31)	(5.33)	(8.39)	(5.64)	(4.95)	(5.11)
or						
dust ISP 4	0.0712	0.04573	0.05447	0.03973	0.08215	0.08674
	(4.67)	(5.03)	(8.03)	(5.34)	(4.93)	(5.12)
or						
quartz ISP 4	0.4129	0.2527	0.3028	0.2184	0.4706	0.4699
	(5.07)	(5.16)	(8.17)	(5.38)	(5.12)	(5.11)
or						
dust ISP 5	0.0664	0.04175	0.04961	0.03605	0.08006	0.08065
	(3.71)	(4.02)	(6.43)	(4.25)	(4.20)	(3.99)
or						
quartz ISP 5	0.376	0.2184	0.2606	0.1909	0.4369	0.4155
	(3.70)	(3.75)	(5.97)	(3.97)	(3.87)	(3.50)

Results from the individual binary analyses are shown in the third column; the baseline model here included overall reader effects along with a different age slope for each reader. The t–statistics for the exposure variables, shown in the third column, were much higher still, which is in line with the intuitive expectation that treating the data as if they were independent will overstate the information contained in the assessments of any one film in relationship to the explanatory variables.

The fourth column shows the results from applying the Bonney model to allow for the interdependence between the assessments of the same film. Testing of different orders of dependence showed that no improvement in the fit was obtained by adding cross–product terms of more than two variables at a time, and the results are therefore those where, in addition to non–parallel age effects by reader and the stated exposure, the model also contained ten terms for the cross–products of all pairs of assessments by the five readers. In fact, during the model fitting, one of these parameters was aliased in the data, and so the results presented are from models where one of the ten has been constrained to zero to permit convergence. These parameters made a highly significant contribution to the fit. In addition, the t–statistics were greatly reduced from those in the third column, and were similar in magnitude to those from the analyses of counts. The coefficient estimates from the Bonney model were the smallest of all from the four logistic analyses.

The remainder of the analyses were performed on ordered responses, the median category and the number of readers recording at least category 1/0. Exhibit 7.2.15 tabulates the 617 films by both variables, and shows the generally good agreement between the two orderings.

Exhibit 7.2.15. Distribution of numbers of films by calculated values of median profusion of small opacities as assessed by panel readers, by number of individual readers recording profusion of 1/0 or higher, on second readings of 617 films from 6th survey.

Median profusion	No of assessments >= 1/0: 0	1	2	3	4	5	Total
0/-	0	0	0	0	0	0	0
0/0	244	113	42	1	0	0	400
0/1	5	21	42	27	2	0	97
1/0	0	0	1	38	19	1	59
1/1	0	0	0	8	28	3	39
1/2	0	0	0	0	3	4	7
2/1	0	0	0	0	2	3	5
2/2	0	0	0	0	0	6	6
2/3	0	0	0	0	0	4	4
3/2	0	0	0	0	0	0	0
3/3	0	0	0	0	0	0	0
3/+	0	0	0	0	0	0	0
Total	249	134	85	74	54	21	617

Results from the application of Anderson's (1984) model for ordered categorical regression are not strictly comparable to those in the first four columns of exhibit 7.2.14. It is necessary, first of all, to fit various models to examine which sets of categories the explanatory variables are capable of distinguishing; Anderson notes that this adds a level of complexity to problems where the choice of explanatory variables is not clearly defined in advance. The analyses reported here were performed by first investigating the distinguishability of categories in models containing age and either ISP 3 or ISP 4 quartz, since these had been already seen to be the most strongly related to the response within other analysis frameworks. These subsets of categories were then used, with age, as a baseline model to which were added all the exposure variables in turn, as in the above logistic analyses. Where serious problems of convergence were encountered, it was sometimes necessary to reduce the number of categories by one, and this was then done for all the models for easy comparability.

In the analyses of the median as an ordered response, this approach led to the categories being distinguished only in the three groups 0/0, 0/1, (1/0 and higher). The fifth column of exhibit 7.2.14 shows the exposure estimates and their partial t–statistics in models based on these three categories and including age. The sixth column shows similar estimates from the analyses of the counts as an ordered, rather than a binomial, response. The ordered models for these analyses were based on four categories, corresponding to counts of none, one, two and three or more positives; again, each included age.

It was clear from this table that different methods of analysis were capable of producing quite different estimates of regression coefficients from the same data, and also different judgments regarding the significance of a range of explanatory variables. The latter was most notable in the case of previous quartz exposures, which had not appeared strongly significant in earlier analyses, but for which all the alternative analyses suggested a more important role.

Interestingly, in the earlier analyses, the ISP 4 quartz exposure had made the largest contribution to the regression model, and model 1+/6/8 in exhibit 7.2.7, where both ISP 3 and ISP 4 quartz were in the regression model simultaneously, had ISP 4 quartz significant after allowance for ISP 3, but not vice versa. In columns 2, 3 and 4 of exhibit 7.2.14, the ISP 3 quartz variable took the highest significance level. Other models were fitted (not shown here in detail) using the Bonney approach but with more than one exposure in the model at a time, and when ISP 3 and 4 quartz variables were both in the model, each was then significant in the presence of the other. The ordered median analyses had ISP 4 quartz the most significant, as with the median bianry; and the ordered counts were unique in that ISP 4 dust was most significant. In all of the alternative analyses, it was plain that there was little to choose between ISP 3 quartz, ISP 4 dust or ISP 4 quartz, so similar were their t–statistics within each set of analyses.

Comparisons across the analyses of the magnitudes of the estimated coefficients are interesting. It is not clear why the Bonney binary analyses should produce lower estimates for the exposure regressions than any of the other three logistic models, while producing judgments on statistical significance consistent with those of the larger coefficients from the analyses of the counts.

Comparisons with the magnitudes of the coefficients in the ordered models must be made with care, since they are not representing the same quantities. Reference to equation 6.13 shows that the estimated coefficient quantifies the regression of the log–odds of being placed in the highest category, as opposed to the lowest category; and that log–odds for comparisons of intermediate categories have regressions scaled by the parameters φ, which lie between 0 and 1 if the data fit the ordered model. In fact, the ordered median analyses with three groups produced analyses for the parameter φ_2 which ranged in the different models between 0.4 and 0.5. The ordered counts analyses, with four groups, produced estimates of φ_2 between 0.15 and 0.26 in the different models, and estimates of φ_3 between 0.42 and 0.5.

Exhibit 7.2.15 shows that of the 400 films for which the median category of small opacities profusion was 0/0, there were 244 for which no reader recorded a profusion as high as 0/1; and that of the 120 with medians 1/0 or greater, 119 had been recorded as 1/0 or greater by at least three readers. This considerable overlap between the extreme categories used in the two sets of ordered analyses is reflected in the similarity between the regression coefficients for the log-odds contrasting those categories, as seen in the fifth and sixth columns of exhibit 7.2.14.

Taking all these findings together, it was concluded that there is room for improvement over the common practice of summarising a set of profusion assessments by a median, and then dichotomising with a view to using binary logistic regression. If the findings in this data set are typical, and there seems no reason to expect that they will not be, there are more informative methods of analysis. Some of these, like the counts analyses, can be applied with the same software, while others are readily programmed in widely available packages. Since on this data set the estimates of regression coefficients were not all consistent across analysis methods, there is scope for further work to investigate the theoretical properties of the different estimates, and to compare results obtained from simulated data sets sampled pseudo-randomly from known distributions and conditions.

A further conclusion is that the judgments of statistical significance made from the binary analyses of medians throughout this report are likely to be consistently conservative.

7.3 Data from serial readings by medical readers

In this section are summarised the assessments made by the two medical readers; both were chest physicians experienced in the interpretation of chest radiographs, and in the use of the ILO Classification Scheme for the description of radiographic abnormalities. This reading exercise was specifically designed to examine progression of opacities between surveys, and the readings differed from those described in 7.1 in that the assessments were made with all the films from the same man visible at once. The protocol used was a shortened form of the ILO (1980) scheme, which provided only for the recording of film quality, of parenchymal abnormalities appearing as small or large opacities, and of any comments which the reader wished to make. Pleural abnormalities, and other abnormalities normally recorded by alphabetic codes, were excluded from the protocol for this exercise, and were not assessed.

Throughout this section, the readers are identified by the code numbers 010 and 011 by which their identities are routinely coded and distinguished on the IOM computer.

7.3.1 Film quality

Exhibit 7.3.1 shows a tabulation of the frequencies with which each of the readers assigned film quality scores to the films which they saw, broken down by survey. Since each reader saw each film only once, and film quality is scored for every

Exhibit 7.3.1. Table showing numbers of films from each survey in reading exercise XRES assigned to each grade of film quality by two medical readers.

Reader code	Survey	Film quality: Good	Acceptable	Poor	Unreadable	Total
010	4th	185	156	2	0	343
	5th	323	105	3	0	431
	6th	323	121	6	0	450
	1980	170	165	10	0	345
	Total	1001	547	21	0	1569
011	4th	186	5	1	0	192
	5th	237	7	0	0	244
	6th	240	10	1	0	251
	1980	151	47	6	1	205
	Total	814	69	8	1	892
Total	4th	371	161	3	0	535
	5th	560	112	3	0	675
	6th	563	131	7	0	701
	1980	321	212	16	1	550
	Total	1815	616	29	1	2461

film, it follows that the row totals in this table also represent the total numbers of films that the two readers saw. The imbalance between the readers is because, as described in 4.4.1, a third reader was unable to take his planned part in the exercise, and reader 010 read the extra batches to complete the reading. Batches contained around 112 films each, and the exercise consisted of 18 batches. Ten of these were read only by reader 010, four only by reader 011, and both readers read the remaining four batches.

Reader 010 appears to have been the more critical of film quality, and classified only 64% of the films he saw as of "good" quality, compared with 91% for reader 011. Most of the rest, however, were classified as "acceptable" for classification purposes, even if their quality was not of the highest standard. Both readers judged around 1% of the films as being of "poor" quality, but only for one film did reader 011 consider a film's quality so bad as to be unreadable.

Despite the overall differences in how critical they were of film quality, both readers recorded much lower proportions of "good" films from the 1980 PXR survey than in the 5th and 6th PFR surveys. There was a curious discrepancy in the films from the 4th survey, in which reader 010 recorded a relatively low proportion of "good" films compared with reader 011's recordings. However, it may be unwise to make too much of comparisons between the readers' results, since they were classifying different films, in the main; and the distinction between "good" and "acceptable" quality is unlikely to hold serious importance for comparisons of small opacities.

7.3.2 Progression of small opacities

Exhibit 7.3.2 displays, for each medical reader separately, comparisons between the profusions of small opacities recorded for film pairs from consecutive surveys. Since these films were read side by side, this was considered likely to provide a measure of the progression of disease within an individual which would be least susceptible to the random effects of within-reader variation.

Comparison of the first and second pages of this exhibit shows again the difference between the numbers of films seen by the two readers, with reader 010 assessing 14 batches and reader 011 assessing 8. Four of these were read by both readers, so that half of the films reader 011 saw were also read by reader 010.

Each section of exhibit 7.3.2 shows the relevant reader's assessment of the categories of profusion for each pair of films which came from a man's attendances at two successive surveys. While the breakdowns within each section are mutually exclusive and exhaustive, the different sections are not independent breakdowns; a man's 4th and 5th survey films can contribute to the comparisons of the first section, while the same 5th survey film can be compared in the second section to the 6th survey film, and so on, for men whose films covered more than two consecutive surveys.

Film pairs for which the reader recorded the same category for both the earlier and later films are enumerated along the diagonal of each tabular section of this exhibit. Pairs where the classification differed lie off the diagonal, and, since the row classification is in each case from the later survey, film pairs in which the profusion of opacities was assessed to have increased are enumerated below the diagonal.

Exhibit 7.3.2. Tabulation of assessments of profusion of small opacities by two medical readers in serial reading exercise. Included are all film pairs read for adjacent surveys.

Reader : 010

Profusions at 5th Survey	Profusions at 4th survey: 0-	00	01	10	11	12	21	22	23	32	33	3+	Total
0-	0	0	0	0	0	0	0	0	0	0	0	0	0
00	0	269	2	0	0	0	0	0	0	0	0	0	271
01	0	7	32	1	0	0	0	0	0	0	0	0	40
10	0	1	2	3	0	0	0	0	0	0	0	0	6
11	0	1	0	2	1	0	0	0	0	0	0	0	4
12	0	0	0	0	1	0	0	0	0	0	0	0	1
21	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	1	0	0	0	0	1
23	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0
3+	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	0	278	36	6	2	0	0	1	0	0	0	0	323

Reader : 010

Profusions at 6th Survey	Profusions at 5th survey: 0-	00	01	10	11	12	21	22	23	32	33	3+	Total
0-	0	0	0	0	0	0	0	0	0	0	0	0	0
00	0	331	5	0	0	0	0	0	0	0	0	0	336
01	0	5	29	0	0	0	0	0	0	0	0	0	34
10	0	2	2	3	0	0	0	0	0	0	0	0	7
11	0	2	3	2	3	0	0	0	0	0	0	0	10
12	0	0	1	0	0	0	0	0	0	0	0	0	1
21	0	0	1	1	0	1	0	0	0	0	0	0	3
22	0	0	1	0	1	1	0	1	0	0	0	0	4
23	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0
3+	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	0	340	42	6	4	2	0	1	0	0	0	0	395

Exhibit 7.3.2 continued.

Reader : 010

Profusions at 6th survey

P r o		0-	00	01	10	11	12	21	22	23	32	33	3+	Total
f	0-	0	0	0	0	0	0	0	0	0	0	0	0	0
u	00	0	256	2	0	0	0	0	0	0	0	0	0	258
s	01	0	4	19	1	0	0	0	0	0	0	0	0	24
i	10	0	0	3	2	1	0	0	0	0	0	0	0	6
o	11	0	0	0	4	8	0	0	0	0	0	0	0	12
n	12	0	0	0	0	0	0	0	0	0	0	0	0	0
s	21	0	0	0	0	0	1	1	0	0	0	0	0	2
	22	0	0	0	0	0	0	1	3	0	0	0	0	4
at	23	0	0	0	0	0	0	0	0	0	0	0	0	0
	32	0	0	0	0	0	0	0	0	0	0	0	0	0
'80	33	0	0	0	0	0	0	0	0	0	0	0	0	0
	3+	0	0	0	0	0	0	0	0	0	0	0	0	0
Survey Total		0	260	24	7	9	1	2	3	0	0	0	0	306

Reader : 011

Profusions at 4th survey

P r o		0-	00	01	10	11	12	21	22	23	32	33	3+	Total
f	0-	0	0	0	0	0	0	0	0	0	0	0	0	0
u	00	0	145	0	1	0	0	0	0	0	0	0	0	146
s	01	0	10	10	0	0	0	0	0	0	0	0	0	20
i	10	0	2	2	1	0	0	0	0	0	0	0	0	5
o	11	0	0	4	0	6	0	0	0	0	0	0	0	10
n	12	0	0	0	0	1	0	0	0	0	0	0	0	1
s	21	0	0	0	0	0	0	0	0	0	0	0	0	0
	22	0	0	0	0	0	0	0	0	0	0	0	0	0
at	23	0	0	0	0	0	0	0	0	0	0	0	0	0
	32	0	0	0	0	0	0	0	0	0	0	0	0	0
5th	33	0	0	0	0	0	0	0	0	0	0	0	0	0
	3+	0	0	0	0	0	0	0	0	0	0	0	0	0
Survey Total		0	157	16	2	7	0	0	0	0	0	0	0	182

Exhibit 7.3.2 continued.

Reader : 011

		Profusions at 5th survey												
P														
r		0-	00	01	10	11	12	21	22	23	32	33	3+	Total
o														
f	0-	0	0	0	0	0	0	0	0	0	0	0	0	0
u	00	0	164	0	0	0	0	0	0	0	0	0	0	164
s	01	0	12	8	0	0	0	0	0	0	0	0	0	20
i	10	0	2	4	0	1	0	0	0	0	0	0	0	7
o	11	0	1	9	2	6	0	0	0	0	0	0	0	18
n	12	0	0	0	2	1	0	0	0	0	0	0	0	3
s	21	0	1	1	0	0	0	0	0	0	0	0	0	2
	22	0	1	0	1	0	1	0	0	0	0	0	0	3
at	23	0	0	0	0	0	0	0	0	0	0	0	0	0
	32	0	0	0	0	0	0	0	0	0	0	0	0	0
6th	33	0	0	0	0	0	0	0	0	0	0	0	0	0
	3+	0	0	0	0	0	0	0	0	0	0	0	0	0
Survey														
Total		0	181	22	5	8	1	0	0	0	0	0	0	217

Reader : 011

		Profusions at 6th survey												
P														
r		0-	00	01	10	11	12	21	22	23	32	33	3+	Total
o														
f	0-	0	0	0	0	0	0	0	0	0	0	0	0	0
u	00	0	133	1	0	0	0	0	0	0	0	0	0	134
s	01	0	5	9	0	0	0	0	0	0	0	0	0	14
i	10	0	0	3	3	0	0	0	0	0	0	0	0	6
o	11	0	0	3	4	11	0	0	0	0	0	0	0	18
n	12	0	0	0	0	0	0	0	0	0	0	0	0	0
s	21	0	0	0	0	0	0	1	0	0	0	0	0	1
	22	0	0	0	0	0	2	0	0	0	0	0	0	2
at	23	0	0	0	0	0	1	0	0	0	0	0	0	1
	32	0	0	0	0	0	0	0	0	0	0	0	0	0
'80	33	0	0	0	0	0	0	0	0	0	0	0	0	0
	3+	0	0	0	0	0	0	0	0	0	0	0	0	0
Survey														
Total		0	138	16	7	11	3	1	0	0	0	0	0	176

By the same token, numbers above the diagonal represent film pairs where the later film was given a lower category of profusion than the earlier film. Reader 010 recorded such apparent regression in 12 assessments of film pairs, and reader 011 in only 3 pairs. Apart from one pair where reader 011 recorded regression of two categories, these regressions were all of a single category on the ILO 12-point scale.

Reader 010 saw 962 film pairs where the first and second films were judged to lie in the same category of profusion, and 50 in which the profusion was judged to have increased. The highest increase was a jump of five categories (from 0/1 to 2/2). Reader 011 classified 497 film pairs with the films in the same category, and 75 with increases in profusion, and a maximum increase of six categories (from 0/0 to 2/2). Both readers recorded fewer and smaller increases in profusion in the roughly two years between the 6th and 1980 surveys, and more frequent and larger increases in the roughly four years between 5th and 6th surveys.

7.3.3 Agreement between the medical readers

It is well known that even medically qualified readers with considerable experience in the application of the ILO classification schemes can show considerable between-reader variations (Jacobsen *et al*, 1984); it is widely believed that the use of side-by-side readings can help to reduce the variation in assessments of progression. It is legitimate to ask how good is the agreement on category and on progression between the two medical readers in this study, but such comparisons are not readily made from the information in exhibit 7.3.2, because this is not based on the same films for each reader.

Exhibit 7.3.3 shows, for the 445 films from the four batches which both readers examined, a comparison of the categories of profusion recorded by the two readers on the same films. As before, films assigned the same category by both readers appear on the diagonal. Since reader 011's classifications index the rows of the tabular sections, films enumerated below the diagonal are those where reader 011 recorded a higher category than 010, and vice versa.

In fact, a glance at exhibit 7.3.3 immediately reveals that the films for which reader 011 recorded higher categories than reader 010 considerably outnumbered those for which the reverse was true. The same category was recorded by both readers in 367 films, while reader 011 recorded a higher category for 60 films, and reader 010 was higher for only 18 films. Further, although reader 011's classification was at times as much as four categories in excess of 010's, reader 010 never exceeded 011 by more than one category. On the basis of the limited information from these four overlapping batches, then, it is concluded that, during this reading exercise, reader 011 tended to record higher categories for the same films than reader 010.

The information available on the readers' agreement on progression within series of films is necessarily even more limited. Exhibit 7.3.4 compares the results for the pairs of films from consecutive surveys which the readers had in common, each film pair being classified by the number of steps of progression on the ILO (1980) 12-point scale represented by the difference between the categories recorded for the earlier and later film by each reader. As in exhibit 7.3.3, reader 011's results index the rows, and numbers below the diagonal represent film pairs where reader 011 recorded categories showing more profusion than did reader 010.

Exhibit 7.3.3. Tabulation comparing assessments by the two medical readers of small opacities profusion on the same films, broken down by survey.

4th Survey

		Profusions : Reader 010												
P		0-	00	01	10	11	12	21	22	23	32	33	3+	Total
r														
o	0-	0	0	0	0	0	0	0	0	0	0	0	0	0
f	00	0	83	4	0	0	0	0	0	0	0	0	0	87
u	01	0	2	3	0	0	0	0	0	0	0	0	0	5
s	10	0	0	1	0	0	0	0	0	0	0	0	0	1
i	11	0	1	3	0	1	0	0	0	0	0	0	0	5
o	12	0	0	0	0	0	0	0	0	0	0	0	0	0
n	21	0	0	0	0	0	0	0	0	0	0	0	0	0
s	22	0	0	0	0	0	0	0	0	0	0	0	0	0
	23	0	0	0	0	0	0	0	0	0	0	0	0	0
Rdr	32	0	0	0	0	0	0	0	0	0	0	0	0	0
	33	0	0	0	0	0	0	0	0	0	0	0	0	0
011	3+	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total	0	86	11	0	1	0	0	0	0	0	0	0	98

5th Survey

		Profusions : Reader 010												
P		0-	00	01	10	11	12	21	22	23	32	33	3+	Total
r														
o	0-	0	0	0	0	0	0	0	0	0	0	0	0	0
f	00	0	95	7	0	0	0	0	0	0	0	0	0	102
u	01	0	8	2	0	0	0	0	0	0	0	0	0	10
s	10	0	0	2	0	0	0	0	0	0	0	0	0	2
i	11	0	1	4	0	1	0	0	0	0	0	0	0	6
o	12	0	0	0	0	0	0	0	0	0	0	0	0	0
n	21	0	0	0	0	0	0	0	0	0	0	0	0	0
s	22	0	0	0	0	0	0	0	0	0	0	0	0	0
	23	0	0	0	0	0	0	0	0	0	0	0	0	0
Rdr	32	0	0	0	0	0	0	0	0	0	0	0	0	0
	33	0	0	0	0	0	0	0	0	0	0	0	0	0
011	3+	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total	0	104	15	0	1	0	0	0	0	0	0	0	120

Exhibit 7.3.3 continued.

6th Survey

		Profusions : Reader 010												
P		0-	00	01	10	11	12	21	22	23	32	33	3+	Total
r														
o	0-	0	0	0	0	0	0	0	0	0	0	0	0	0
f	00	0	96	5	0	0	0	0	0	0	0	0	0	101
u	01	0	7	3	0	0	0	0	0	0	0	0	0	10
s	10	0	1	1	0	0	0	0	0	0	0	0	0	2
i	11	0	6	3	0	2	0	0	0	0	0	0	0	11
o	12	0	0	1	0	0	1	0	0	0	0	0	0	2
n	21	0	0	0	0	0	0	0	0	0	0	0	0	0
s	22	0	0	0	0	0	0	1	0	0	0	0	0	1
	23	0	0	0	0	0	0	0	0	0	0	0	0	0
Rdr	32	0	0	0	0	0	0	0	0	0	0	0	0	0
	33	0	0	0	0	0	0	0	0	0	0	0	0	0
011	3+	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total	0	110	13	0	2	1	1	0	0	0	0	0	127

1980 Survey

		Profusions : Reader 010												
P		0-	00	01	10	11	12	21	22	23	32	33	3+	Total
r														
o	0-	0	0	0	0	0	0	0	0	0	0	0	0	0
f	00	0	76	2	0	0	0	0	0	0	0	0	0	78
u	01	0	4	2	0	0	0	0	0	0	0	0	0	6
s	10	0	2	1	0	0	0	0	0	0	0	0	0	3
i	11	0	4	5	0	2	0	0	0	0	0	0	0	11
o	12	0	0	0	0	0	0	0	0	0	0	0	0	0
n	21	0	0	0	0	0	0	0	0	0	0	0	0	0
s	22	0	0	0	1	0	0	0	0	0	0	0	0	1
	23	0	0	0	0	0	0	1	0	0	0	0	0	1
Rdr	32	0	0	0	0	0	0	0	0	0	0	0	0	0
	33	0	0	0	0	0	0	0	0	0	0	0	0	0
011	3+	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total	0	86	10	1	2	0	1	0	0	0	0	0	100

Exhibit 7.3.4. Tabulation comparing assessments by the two medical readers of change in small opacities profusion on the same pairs of films, broken down by survey.

Period : 4th to 5th survey

		Steps of change in profusion, reader 010										
		-3	-2	-1	0	+1	+2	+3	+4	+5	+6	Total
C	-3	0	0	0	0	0	0	0	0	0	0	0
h	-2	0	0	0	1	0	0	0	0	0	0	1
a	-1	0	0	0	0	0	0	0	0	0	0	0
n	0	0	0	0	83	2	0	0	0	0	0	85
g	+1	0	0	0	6	1	0	0	0	0	0	7
e	+2	0	0	0	0	1	0	0	0	0	0	1
	+3	0	0	0	0	0	0	0	0	0	0	0
rdr	+4	0	0	0	0	0	0	0	0	0	0	0
	+5	0	0	0	0	0	0	0	0	0	0	0
011	+6	0	0	0	0	0	0	0	0	0	0	0
Total		0	0	0	90	4	0	0	0	0	0	94

Period : 5th to 6th survey

		Steps of change in profusion, reader 010										
		-3	-2	-1	0	+1	+2	+3	+4	+5	+6	Total
C	-3	0	0	0	0	0	0	0	0	0	0	0
h	-2	0	0	0	0	0	0	0	0	0	0	0
a	-1	0	0	0	0	0	0	0	0	0	0	0
n	0	0	0	1	89	0	0	0	0	0	0	90
g	+1	0	0	0	7	0	0	0	0	0	0	7
e	+2	0	0	0	5	1	1	1	0	0	0	8
	+3	0	0	0	1	0	0	0	0	0	0	1
rdr	+4	0	0	0	0	0	0	0	0	0	0	0
	+5	0	0	0	0	0	0	0	0	0	0	0
011	+6	0	0	0	0	0	0	0	1	0	0	1
Total		0	0	1	102	1	1	1	1	0	0	107

Exhibit 7.3.4 continued.

Period : 6th to 1980 survey

		Steps of change in profusion, reader 010										
		-3	-2	-1	0	+1	+2	+3	+4	+5	+6	Total
C	-3	0	0	0	0	0	0	0	0	0	0	0
h	-2	0	0	0	0	0	0	0	0	0	0	0
a	-1	0	0	0	1	0	0	0	0	0	0	1
n	0	0	0	0	77	1	0	0	0	0	0	78
g	+1	0	0	0	5	0	0	0	0	0	0	5
e	+2	0	0	0	2	1	0	0	0	0	0	3
	+3	0	0	0	0	1	0	0	0	0	0	1
rdr	+4	0	0	0	0	0	0	0	0	0	0	0
	+5	0	0	0	0	0	0	0	0	0	0	0
011	+6	0	0	0	0	0	0	0	0	0	0	0
Total		0	0	0	85	3	0	0	0	0	0	88

All three inter-survey periods combined

		Steps of change in profusion, reader 010										
		-3	-2	-1	0	+1	+2	+3	+4	+5	+6	Total
C	-3	0	0	0	0	0	0	0	0	0	0	0
h	-2	0	0	0	1	0	0	0	0	0	0	1
a	-1	0	0	0	1	0	0	0	0	0	0	1
n	0	0	0	1	249	3	0	0	0	0	0	253
g	+1	0	0	0	18	1	0	0	0	0	0	19
e	+2	0	0	0	7	3	1	1	0	0	0	12
	+3	0	0	0	1	1	0	0	0	0	0	2
rdr	+4	0	0	0	0	0	0	0	0	0	0	0
	+5	0	0	0	0	0	0	0	0	0	0	0
011	+6	0	0	0	0	0	0	0	1	0	0	1
Total		0	0	1	277	8	1	1	1	0	0	289

The last section of this exhibit shows the total numbers of film pairs, accumulated over the three inter-survey periods. Just as reader 011 read higher categories on individual films, so too did this reader record more progression than reader 010. For 252 of the 289 pairs of films, neither reader observed progression. Of the remainder, there were 26 pairs for which only reader 011 observed progression, 3 pairs where only reader 010 observed progression, and 8 pairs where both readers recorded progression.

Although the readers did not always record progression to the same extent, nor even on the same films, their agreement was better than would be predicted by chance. Reduction of the last section of exhibit 7.3.4 to a table of dimension 2x2, according to whether each of the readers recorded any progression, yielded a Pearson chi-squared statististic with one degree of freedom, for the test of independence between the readers, with a value of 41.1, which is highly significant. It was concluded that the readers did tend towards agreement on whether progression was present, even although such agreement was not complete.

7.3.4 Progression of large opacities

Exhibit 7.3.5 displays, for each of the two medical readers separately, comparisons between the profusions of large opacities recorded for the films within pairs from consecutive surveys. As with the tabulations of small opacities, there was some overlap of the batches read by the two readers, and the same film could be the second of a pair spanning one ISP and the first of a pair spanning the next ISP.

Tabulated are the numbers of pairs by the profusion of large opacities recorded for the first and second film in each pair. Here the classification is only by the recorded profusion of large opacities, and ignores information on profusions of

Exhibit 7.3.5. Tabulations of assessments of profusion of large opacities by two medical readers in serial reading exercise. Included are all film pairs read for adjacent surveys.

Reader : 010
Profusions of large opacities at 4th survey

a	0	A	B	C	Total
t					
0	324	0	0	0	324
5 A	0	0	0	0	0
t B	0	0	0	0	0
h C	0	0	0	0	0
Total	324	0	0	0	324

Reader : 011
Profusions of large opacities at 4th survey

a	0	A	B	C	Total
t					
0	182	0	0	0	182
5 A	0	0	0	0	0
t B	0	0	0	0	0
h C	0	0	0	0	0
Total	182	0	0	0	182

Reader : 010
Profusions of large opacities at 5th survey

a	0	A	B	C	Total
t					
0	393	0	0	0	393
6 A	2	0	0	0	2
t B	1	0	0	0	1
h C	0	0	0	0	0
Total	396	0	0	0	396

Reader : 011
Profusions of large opacities at 5th survey

a	0	A	B	C	Total
t					
0	216	0	0	0	216
6 A	0	0	0	0	0
t B	1	0	0	0	1
h C	0	0	0	0	0
Total	217	0	0	0	217

Reader : 010
Profusions of large opacities at 6th survey

a	0	A	B	C	Total
t					
0	304	1	0	0	305
1 A	1	0	0	0	1
9 B	0	0	0	0	0
8 C	0	0	1	0	1
0					
Total	305	1	1	0	307

Reader : 011
Profusions of large opacities at 6th survey

a	0	A	B	C	Total
t					
0	175	0	0	0	175
1 A	0	0	0	0	0
9 B	0	0	0	0	0
8 C	0	0	1	0	1
0					
Total	175	0	1	0	176

small opacities, or on any other abnormalities, present on the same films. Film pairs where the classification of large opacities profusion was the same for both films are enumerated along the diagonals of these tables, while entries below the diagonals signify pairs where the profusion was assessed to have increased.

Overall, there were hardly any recordings of the presence of large opacities, and none at all were recorded by either reader for films from the 4th or 5th surveys. Reader 010 recorded 2 films with a profusion of A and one with a profusion of B appearing from 5th to 6th survey, while reader 011 recorded only one instance of the appearance of a profusion of A over the same period. In the period between 6th survey and the 1980 PXR survey, reader 011 recorded one change of profusion from B to C and no other large opacities, and reader 010 recorded a change from B to C, one from 0 to A and one apparent regression from category A to 0 over the same period.

The number of cases where a non-zero profusion of large opacities was recorded was insufficient for analysis by regression modelling techniques, and no further analyses of the assessments of this variable were carried out.

7.4 Logistic regression analyses of serial readings

7.4.1 Choice of variable for analysis

The independent randomised readings made by the five members of the panel were summarised by the calculation of a median before analysis, but this approach could not be adopted for the serial readings, because a large proportion of the films were seen by only one of the two readers, while the remainder were seen by both. While it is possible to define the median of a single value as the value itself, and of two values as some value between those values if they differ, this does not avoid the fundamental difficulty that the data fall into three disjoint sets according as they were generated by reader 010, or 011, or both, and that despite some tendency towards agreement the readers were apparently operating at different levels of perception of the frequency of profusion and progression of small opacities.

It was decided, therefore, not to attempt to summarise the data for a film, but simply to analyse all the data from the conjunction of the readers' individual assessments. There were obvious disadvantages in this approach also. Firstly, it would be necessary to take into account reader differences. However, the overlap of the readers on four batches meant that the reader differences were not entirely confounded with the the batch differences, as they would have been if the readers had examined disjoint sets of batches. An additional, more subtle problem was anticipated, in that the assumption of independence of the random part of the data which underlies logistic regression would not be entirely realistic in the case of films read by both readers. On balance, however, this was felt to be the most practical approach, with the possibility that some additional investigations might make it possible to examine the effect of ignoring a partial lack of independence in the structure of the data set.

Because the focus of interest was in progression of disease, the analyses were of the progression between consecutive surveys. This was calculated as the number of steps of difference on the ILO (1980) 12-point scale of profusion of small opacities between the earlier and later surveys. Two variables were analysed, being those produced by dichotomising the variables according to whether they showed or did not show one step or more, or two steps or more, of progression.

Exhibit 7.4.1 summarises the numbers of film pairs showing various numbers of steps of progression on the 12-point scale in the relevant inter-survey periods. The logistic regression analyses reported in the following sections were based on these data.

7.4.2 The form and presentation of the analyses of progression

The rationale for the analyses of progression between pairs of the medical readings was similar to that for the panel's readings, as described in 7.2.2 above. The choice of explanatory variables was similar, except that it was necessary to include a dummy variable to summarise differences between the two readers' levels. The remaining variables considered were age, taken arbitrarily as at the later survey, and the variables summarising the individual's exposures to respirable coalmine dust and its quartz fraction, in the inter-survey periods 3, 4 and 5, and before the 3rd survey. Variables for the estimates of exposure to respirable dust were for the non-quartz fraction obtained by subtracting the estimates of quartz exposure from those for whole respirable dust.

The tabular layout of results is also similar to that employed earlier, as may be seen in exhibits 7.4.2 to 7.4.7. Each column contains the result of fitting one regression model to the data, and is labelled with a code containing three elements. The first element identifies the content of the response variable, here 1+ to show that the positive response represents one or more steps of progression; the second identifies the inter-survey period between the pair of surveys from which the films came, where I4 stands for ISP 4, between 4th and 5th survey; and the third element is a sequential distinction between the different models applied.

Exhibit 7.4.1. Table showing numbers of film pairs from consecutive surveys yielding data for logistic analysis of progression.

Period	Reader	-3	-2	-1	0	+1	+2	+3	+4	+5	+6	Total
		Steps of progression on 12-point profusion scale										
4th - 5th	010	0	0	3	306	12	1	1	0	0	0	323
	011	0	1	0	162	13	6	0	0	0	0	182
	Total	0	1	3	468	25	7	1	0	0	0	505
5th - 6th	010	0	0	5	367	10	6	5	1	1	0	395
	011	0	0	1	178	19	14	1	2	1	1	217
	Total	0	0	6	545	29	20	6	3	2	1	612
6th - 1980	010	0	0	4	289	13	0	0	0	0	0	306
	011	0	0	1	157	12	5	1	0	0	0	176
	Total	0	0	5	446	25	5	1	0	0	0	482

The values in the body of the table show the estimated regression coefficients against those variables included in each model, and are blank where a variable was not included. Shown in parenthesis below each estimated coefficient is the absolute value of the calculated ratio of the estimate to its estimated standard error, which may be taken as a partial t–statistic for a test of significance where valid. At the foot of each column is given the residual deviance and its degrees of freedom. For nested models, differences in the deviance provide statistics, distributed as chi–squared under an appropriate null hypothesis, for testing the significance of terms introduced to the model.

All the logistic regressions were fitted using the package GENSTAT (Alvey at al., 1983), often interactively. The results presented here are not exhaustive, but omit many from models judged to be less interesting and omitted from the tabulations.

7.4.2 <u>Regression analyses of progression 4th to 5th survey</u>

Exhibit 7.4.2 presents some results from a logistic regression analysis of the binary variable constructed to have a value of 1 when the reader's assessment of the profusion of small opacities on the 5th survey film was higher by one or more category on the 12–point ILO (1980) scale than the same reader's assessment of the 4th survey film, and a value of 0 otherwise. As already noted, some film series were read once each by both readers, and these series contribute twice to the data set.

Of 505 assessments of film pairs contributing to the analyses in exhibit 7.4.2, 33 implied progression of at least 1 step. There was a highly significant contribution from the term for the difference between readers, and this term has been included in all the models. The age coefficient was also significant.

Exhibit 7.4.2. **Results of fitting different logistic regression models to binary variable indicating one or more steps of progression of small opacities in ISP 4.** Tabulated values are estimated regression coefficients. Absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models							
Variables	1+/I4/1	1+/I4/2	1+/I4/3	1+/I4/4	1+/I4/4	1+/I4/6	1+/I4/7	1+/I4/8
constant	-5.20	-5.80	-5.47	-6.14	-5.93	-5.69	-5.34	-5.744
	(4.64)	(4.70)	(4.51)	(4.94)	(4.92)	(4.41)	(4.29)	(4.63)
reader 011	0.991	0.971	0.942	1.018	1.043	0.923	0.937	0.974
	(2.71)	(2.62)	(2.53)	(2.75)	(2.81)	(2.46)	(2.51)	(2.61)
age at survey	0.0458	0.0354	0.0329	0.0474	0.0490	0.0411	0.0339	0.0372
	(1.99)	(1.40)	(1.30)	(1.97)	(2.03)	(1.37)	(1.34)	(1.45)
dust previous								
quartz previous						-0.102		
						(0.51)		
dust ISP 3		0.0632					-0.0216	
		(2.66)					(0.45)	
quartz ISP 3			0.4501			0.461	0.548	0.340
			(3.47)			(3.51)	(2.19)	(2.06)
dust ISP 4				0.0847				
				(2.46)				
quartz ISP 4					0.394			0.188
					(3.17)			(1.15)
deviance	232.7	225.4	220.8	226.5	223.6	220.6	220.6	219.6
degrees of freedom	502	501	501	501	501	500	500	500

With a base model containing reader and age effects, inclusion of dust or quartz exposure previous to 3rd survey produced small deviance reductions, not shown here. The largest reduction was achieved with the inclusion of quartz exposure in ISP 3, with the highly significant deviance reduction of 11.9 on 1 degree of freedom. The equivalent deviance change for ISP 3 dust was 7.3, while those for ISP 4 dust and quartz were 6.2 and 9.1 respectively. With ISP 3 quartz in the model, none of the other candidate exposure variables gave a significantly improved model. Model 1+/I4/8 shows, for example, the model with both ISP 3 and ISP 4 quartz, which gave a deviance reduction of only 1.2 with 1 degree of freedom, over the model with ISP 3 quartz alone. Other model fits, which showed even smaller deviance changes, are not shown here. Model 1+/I4/7 included both ISP 3 dust and ISP3 quartz, and demonstrates that the quartz exposure variable made the greater contribution, with the addition of the dust variable after the quartz producing no significant improvement, but with the quartz remaining significant in the presence of the dust.

Although age was significant in model 1+/I4/1, the age coefficient was smaller and less significant in other models, notably those which included ISP 3 quartz. To ensure that conclusions about the exposure variables were not being distorted by the presence of age in the models, other models including the exposures but excluding age were also fitted. The estimated coefficients for the exposure variables were similar to those shown in exhibit 7.4.2, and the deviances retained the same rank order.

Exhibit 7.4.3 shows the results of logistic regression analyses where the response variable was the binary variable taking the value 1 if the progression over the film pair spanning ISP 4 was assessed to be two or more steps on the 12-point scale; this is indicated by the first element of each column label, 2+. Of the 33

Exhibit 7.4.3. Results of fitting different logistic regression models to binary variable indicating two or more steps of progression of small opacities in ISP 4. Tabulated values are estimated regression coefficients. Absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models 2+/I4/1	2+/I4/2	2+/I4/3	2+/I4/4	2+/I4/5	2+/I4/6	2+/I4/7
constant	-5.54	-6.66	-6.48	-5.59	-5.51	-6.11	-5.94
	(2.92)	(3.33)	(3.28)	(2.84)	(2.81)	(3.00)	(2.96)
reader 011	1.709	1.626	1.640	1.705	1.676	1.726	1.7465
	(2.13)	(1.98)	(2.09)	(2.08)	(2.11)	(2.16)	(2.13)
age at survey	0.0102	0.0591	0.0508	0.0075	-0.0012	0.0089	0.0091
	(0.26)	(1.24)	(1.09)	(0.18)	(0.03)	(0.22)	(0.22)
dust previous		-0.0383					
		(1.50)					
quartz previous			-0.614				
			(1.36)				
dust ISP 3				0.0114			
				(0.25)			
quartz ISP 3					0.272		
					(1.11)		
dust ISP 4						0.0633	
						(0.95)	
quartz ISP 4							0.311
							(1.30)
deviance	77.01	74.53	75.17	76.94	75.94	76.15	75.60
degrees of freedom	502	501	501	501	501	501	501

assessments showing one or more steps of progression in the 505 film pairs, 7 of these showed two, and 1 three, steps of progression, so that this analysis is based on only 8 positive responses.

None of the dust or quartz exposure variables made a significant contribution to the base model containing terms for reader and age. The largest deviance change was 1.41, for quartz exposure in ISP 4, but this fell well short of any conventional significance levels. No improvement was gained by fitting models with combinations of age and exposure variables, or with exposure variables without age, and these fits are not shown here.

7.4.3 Regression analyses of progression 5th to 6th survey

Logistic regressions for the variables representing the change in small opacities profusion in film pairs spanning inter-survey period 5, between 5th and 6th survey, were performed in a similar manner to those for ISP 4. Variables estimating the coalminers' exposures to dust and quartz during ISP 5 were considered for inclusion in these models, along with the ISP 4, ISP 3 and previous dust and quartz exposure variables.

Exhibit 7.4.4 presents some results from the logistic regression analysis of the variable taking the value 1 if a pair of films spanning ISP 5 was assessed as showing one or more step of progression on the 12-point scale. Of 612 assessments, 61 were positive. The labelling of the columns of this table follows the same scheme as in exhibit 7.4.2.

Both reader differences and the age coefficient were highly significant. Over a base model including age and reader, previous dust and quartz exposures did not

Exhibit 7.4.4. Results of fitting different logistic regression models to binary variable indicating one or more steps of progression of small opacities in ISP 5. Tabulated values are estimated regression coefficients. Absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models									
Variables	1+/15/1	1+/15/2	1+/15/3	1+/15/4	1+/15/5	1+/15/6	1+/15/7	1+/15/8	1+/15/9	1+/15/10
constant	-4.862	-5.442	-4.856	-6.717	-6.151	-6.700	-7.015	-5.694	-5.85	-6.56
	(6.20)	(5.97)	(5.50)	(7.00)	(6.67)	(7.12)	(7.14)	(5.88)	(5.83)	(6.56)
reader 011	1.287	1.372	1.362	1.473	1.532	1.354	1.335	1.500	1.381	1.523
	(4.59)	(4.71)	(4.59)	(4.96)	(5.05)	(4.67)	(4.61)	(4.81)	(4.62)	(4.94)
age at survey	0.0429	0.0183	0.0143	0.0388	0.0443	0.0565	0.0706	0.0258	0.0298	0.0544
	(2.85)	(1.03)	(0.80)	(2.25)	(2.60)	(3.48)	(4.09)	(1.36)	(1.53)	(2.94)
dust previous										
quartz previous										
dust ISP 3		0.1027								
		(5.69)								
quartz ISP 3			0.707					0.434	0.605	
			(6.42)					(3.13)	(5.05)	
dust ISP 4				0.1874						
				(6.40)						
quartz ISP 4					0.739			0.484		0.637
					(6.65)			(3.53)		(4.98)
dust ISP 5						0.1339				
						(4.75)				
quartz ISP 5							0.625		0.333	0.243
							(5.07)		(2.44)	(1.61)
deviance	367.9	332.6	322.7	320.5	319.5	344.7	342.8	309.8	317.1	317.1
degrees of freedom	609	608	608	608	608	608	608	607	607	607

Exhibit 7.4.4 continued.

Variables	Parsimonious models 1+/15/11	1+/15/12	1+/15/13	1+/15/14	1+/15/15
constant	-6.628	-6.17	-5.765	-6.03	-6.06
	(6.75)	(5.84)	(5.80)	(5.92)	(5.81)
reader 011	1.527	1.494	1.502	1.499	1.497
	(5.02)	(4.80)	(4.82)	(4.84)	(4.81)
age at survey	0.0421	0.0334	0.0251	0.0257	0.0292
	(2.40)	(1.67)	(1.31)	(1.35)	(1.52)
dust previous					
quartz previous					
dust ISP 3			0.0137		
			(0.38)		
quartz ISP 3		0.415	0.368	0.388	0.406
		(2.98)	(1.64)	(2.69)	(2.88)
dust ISP 4	0.1024			0.0641	
	(2.11)			(1.25)	
quartz ISP 4	0.414	0.424	0.482	0.315	0.447
	(2.23)	(2.95)	(3.52)	(1.67)	(3.19)
dust ISP 5					0.0353
					(1.00)
quartz ISP 5		0.186			
		(1.24)			
deviance	309.8	308.4	309.7	308.3	308.9
degrees of freedom	607	606	606	606	606

produce a significant deviance reduction, whereas all the other dust and quartz exposure variables did. The largest reduction was 48.4, highly significant with 1 degree of freedom, and due to ISP 4 quartz exposure. ISP 4 dust and ISP 3 quartz gave almost as large reductions, of 47.4 and 45.2 respectively. With ISP 4 quartz in the model, and considering the remaining dust and quartz exposure variables for inclusion, the best improvement in the fit was from the addition of ISP 3 quartz, giving a further deviance reduction of 9.7, which exceeds the 1% significance point of the chi-squared distribution with 1 degree of freedom. The partial t-statistics of 3.13 and 3.53 shown in model 1+/I5/8 for ISP 3 and ISP 4 quartz respectively imply that each variable made a significant independent contribution to the model. With ISP 4 quartz in the model, the addition of ISP 5 quartz failed to reach significance; however, in the model containing quartz variables from ISP 3 and ISP 5, both variables made significant contributions.

The continuation of exhibit 7.4.4 shows models 1+/I5/11 to 1+/I5/15. Model 1+/I5/11 contains both dust and quartz variables from ISP 4. Each made a significant contribution in the presence of the other. Each of the remaining four models in exhibit 7.4.4 contains three different exposure variables. In the model containing the quartz variables from ISPs 3, 4 and 5, that for ISP 5 failed to achieve significance, while those for ISPs 3 and 4 were both significant. The remaining models all included a dust variable in addition to the ISP 3 and 4 quartz variables. None made a significant improvement to the fit. Furthermore, where dust and quartz variables from the same ISP were included, the dust variable did not yield an additional improvement, while the improvement due to the quartz variable over the dust variable approached the 10% significance level. In model 1+/I5/14, where ISP 3 and 4 quartz variables were joined by ISP 4 dust, and this showed rather greater reduction in the significance of ISP 4 dust than in ISP 4 quartz, when compared with model 1+/I5/11.

As for progression in ISP 4, inclusion of ISP 3 quartz reduced the estimated age coefficient, but fitting models without age did not alter the conclusions regarding the exposure variables. No further inclusions or alterations to the model yielded an improvement in fit, and none of the alternatives are shown here.

Exhibit 7.4.5 shows the results of similar analyses for the variable which took the value 1 when the progression assessed on the film pair spanning ISP 5 was 2 or more steps on the 12-point scale. This variable had 32 positive values from the 612 assessments.

Over a base model containing reader and age terms, the introduction of previous dust and quartz exposures did not give significant deviance reductions. The largest deviance reduction was for ISP 3 quartz, at 19.7, closely followed by ISP 3 dust and ISP 4 dust and quartz at 17.0, 18.7 and 17.4 respectively. Addition of ISP 4 quartz to the model with ISP 3 quartz produced a further deviance reduction of only 3.3, which is less than the 5% significance level of the chi-squared distribution with one degree of freedom. As in the analysis of progression of 1+ steps, fitting of a model with dust and quartz from ISP 3 (not shown) showed very little improvement over that with only ISP 3 quartz.

Except in the models with only ISP 5 dust or quartz, the age coefficient was in general not significant. Again, models fitted without age gave very similar exposure coefficients and the same ranking of deviances.

7.4.5 Regression analyses of progression 6th to 1980 survey

The roughly two year period between the 6th PFR survey and the PXR surey in 1980 is here referred to for simplicity as ISP 6. Logistic regression analyses for

Exhibit 7.4.5. Results of fitting different logistic regression models to binary variable indicating two or more steps of progression of small opacities in ISP 5. Tabulated values are estimated regression coefficients. Absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models									
Variables	2+/15/1	2+/15/2	2+/15/3	2+/15/4	2+/15/5	2+/15/6	2+/15/7	2+/15/8	2+/15/9	2+/15/10
constant	-4.98	-5.39	-4.82	-6.277	-5.721	-6.26	-6.66	5.25	-5.54	-6.33
	(4.94)	(4.70)	(4.37)	(5.36)	(5.20)	(5.40)	(5.51)	(4.59)	(4.49)	(5.11)
reader 011	1.070	1.099	1.052	1.153	1.174	1.086	1.072	1.106	1.054	1.154
	(2.89)	(2.93)	(2.80)	(3.06)	(3.12)	(2.90)	(2.87)	(2.92)	(2.79)	(3.02)
age at survey	0.0333	0.0088	0.0066	0.0270	0.0311	0.0421	0.0545	0.0130	0.0178	0.0404
	(1.70)	(0.39)	(0.29)	(1.25)	(1.48)	(2.06)	(2.53)	(0.56)	(0.73)	(1.78)
dust previous										
quartz previous										
dust ISP 3		0.0927								
		(4.04)								
quartz ISP 3			0.590					0.410	0.512	
			(4.47)					(2.44)	(3.49)	
dust ISP 4				0.1496						
				(4.20)						
quartz ISP 4					0.539			0.300		0.451
					(4.39)			(1.89)		(3.10)
dust ISP 5						0.1003				
						(2.83)				
quartz ISP 5							0.515		0.255	0.230
							(3.36)		(1.49)	(1.24)
deviance	240.0	223.0	220.3	221.3	222.6	232.3	229.7	217.0	218.2	221.1
degrees of freedom	609	608	608	608	608	608	608	607	607	607

the variables representing the change in small opacities profusion in film pairs spanning this period were performed in a similar manner to those for ISP 4. As with the analyses of the changes in ISP 5, the variables estimating the coalminers' exposures to dust and quartz were those for ISPs 5, 4 and 3, and for the period prior to the 3rd survey. No additional exposure data were available to cover the period of ISP 6.

Exhibit 7.4.6 presents some results from the logistic regression analysis of the variable taking the value 1 if a pair of films spanning ISP 6 was assessed as showing one or more steps of progression on the 12–point scale. Of 482 assessments contributing to this analysis, 31 were positive. The labelling of the columns of this table follows the same scheme as in exhibit 7.4.2.

Reader differences were again highly significant, and the age term was significant in all models. Over a base model including terms for age and reader differences, the largest deviance reductions from the exposure variables were 14.2 for ISP 4 dust, and 13.7 for ISP 4 quartz. ISP 5 dust and quartz geve reductions of 7.7 and 5.5, while ISP 3 dust and quartz both gave reductions of 5.2. Of the models with more than one exposure variable, the lowest deviances were given by models where ISP 4 quartz was joined by either ISP3 or ISP 5 quartz, but neither of these variables gave significant improvements to the model fit. Model 1+/I6/10 shows the effect of fitting the model containing both dust and quartz variables from ISP 4. Neither variable improved the fit in the presence of the other.

There were only 6 assessments where progression of more than 2 steps was recorded, all by reader 011. Since reader 010 recorded no positives on this assessment, any logistic model including reader differences would produce estimates increasing without bound. Exhibit 7.4.7 presents, instead, the results of analysing

Exhibit 7.4.6. Results of fitting different logistic regression models to binary variable indicating one or more steps of progression of small opacities in ISP 6. Tabulated values are estimated regression coefficients. Absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models									
Variables	1+/I6/1	1+/I6/2	1+/I6/3	1+/I6/4	1+/I6/16	1+/I6/6	1+/I6/7	1+/I6/8	1+/I6/9	1+/I6/10
constant	-7.17	-7.34	-7.22	-8.31	-7.94	-8.17	-8.22	7.99	-8.10	-8.26
	(5.27)	(5.18)	(5.01)	(5.31)	(5.26)	(5.39)	(5.33)	(5.20)	(5.04)	(5.25)
reader 011	0.983	1.033	1.001	1.123	1.101	1.025	1.009	1.107	1.097	1.126
	(2.57)	(2.75)	(2.59)	(2.85)	(2.80)	(2.64)	(2.61)	(2.79)	(2.76)	(2.86)
age at survey	0.0815	0.0677	0.0715	0.0764	0.0811	0.0846	0.0929	0.0827	0.0834	0.0785
	(3.23)	(2.51)	(2.63)	(2.72)	(2.95)	(3.14)	(3.39)	(2.90)	(2.90)	(2.79)
dust previous										
quartz previous										
dust ISP 3		0.0534								
		(2.33)								
quartz ISP 3			0.307					-0.045		
			(2.33)					(0.23)		
dust ISP 4				0.1323						0.0813
				(3.64)						(1.32)
quartz ISP 4					0.505			0.534	0.475	0.244
					(3.80)			(2.92)	(2.98)	(1.03)
dust ISP 5						0.1011				
						(2.80)				
quartz ISP 5							0.415		0.075	
							(2.46)		(0.35)	
deviance	208.5	203.3	203.3	194.3	194.8	200.8	203.0	194.7	194.7	193.2
degrees of freedom	479	478	478	478	478	478	478	477	477	477

Exhibit 7.4.7. Results of fitting different logistic regression models to binary variable indicating two or more steps of progression of small opacities in ISP 6, using only assessments by reader 011. Tabulated values are estimated regression coefficients. Absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models									
Variables	2+/16/1	2+/16/2	2+/16/3	2+/16/4	2+/16/5	2+/16/6	2+/16/7	2+/16/8	2+/16/9	2+/16/10
constant	-7.52	-7.80	-7.55	-8.44	-7.74	-9.84	-10.32	-10.15	-9.67	-10.28
	(2.63)	(2.30)	(2.40)	(2.49)	(2.45)	(2.84)	(2.70)	(2.67)	(2.67)	(2.64)
age at survey	0.0839	0.0623	0.0746	0.0798	0.0840	0.0965	0.1192	0.1039	0.0968	0.1061
	(1.55)	(0.96)	(1.25)	(1.29)	(1.43)	(1.59)	(1.78)	(1.52)	(1.52)	(1.53)
dust previous										
quartz previous										
dust ISP 3		0.0830								
		(1.46)								
quartz ISP 3			0.280					-0.115		
			(1.04)					(0.29)		
dust ISP 4				0.1162						
				(1.43)						
quartz ISP 4					0.176				-0.282	
					(0.54)				(0.68)	
dust ISP 5						0.1841		0.1989	0.2047	0.143
						(2.25)		(2.02)	(2.41)	(0.98)
quartz ISP 5							0.757			0.233
							(2.15)			(0.35)
deviance	49.13	46.84	48.21	46.98	48.87	43.95	44.77	43.87	43.47	43.83
degrees of freedom	174	173	173	173	173	173	173	172	172	172

only the 176 assessments made by reader 011, from which the 6 positive assessments arose.

Over a base model containing only age, ISP 5 dust produced the largest reduction in deviance, at 5.2. ISP 5 quartz gave a reduction of 4.4, while ISP 3 and ISP 4 dust gave reductions of 2.3 and 2.2 respectively. The 5% significance point of the chi-squared distribution with one degree of freedom is 3.8. No combination of exposure variables was a significant improvement on ISP 5 dust alone. In the model including ISP 5 dust and quartz, neither variable made a significant contribution in the presence of the other.

The age term failed to achieve significance in any of the models, although the t-statistic for age was larger when ISP 5 exposures were in the model than when ISP 3 or 4 exposures were used. Fitting similar models without age produced the same rankings of deviances, and had little effect on the magnitudes of the exposure coefficients.

7.4.6 The effects of smoking on progression of small opacities

None of the analyses described in the previous sections and summarised in exhibits 7.4.2 to 7.4.7 included explanatory variables for individuals' smoking habits. In this section are described the principal results from some additional analyses intended to investigate the effects of smoking.

Exhibit 7.4.8 shows the results of augmenting some of the models which proved most interesting in the analyses of progression, by adding terms to distinguish ex-smokers and current smokers from lifelong non-smokers. The analyses are restricted to the variables representing one or more steps of progression over the

Exhibit 7.4.8. Results of fitting logistic regression models with and without smoking variables to binary variable indicating one or more steps of progression of small opacities profusion, in each ISP. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models							
Variables	1+/I4/3	1+/I4/3S	1+/I5/5	1+/I5/5S	1+/I5/8	1+/I5/8S	1+/I6/5	1+/I6/5S
constant	-5.47	-5.77	-6.184	-6.434	-5.716	-5.98	-9.87	-11.35
	(4.51)	(4.54)	(6.66)	(6.66)	(5.84)	(5.86)	(4.94)	(4.77)
reader 011	0.942	0.946	1.496	1.586	1.459	1.558	1.317	1.401
	(2.53)	(2.54)	(4.91)	(5.10)	(4.65)	(4.84)	(2.90)	(3.01)
age at survey	0.0329	0.0300	0.0452	0.0474	0.0259	0.0239	0.1134	0.1160
	(1.30)	(1.15)	(2.63)	(2.64)	(1.35)	(1.17)	(3.18)	(3.02)
dust previous								
quartz previous								
dust ISP 3								
quartz ISP 3	0.450	0.471			0.452	0.488		
	(3.47)	(3.55)			(3.25)	(3.43)		
dust ISP 4								
quartz ISP 4			0.741	0.764	0.478	0.488	0.463	0.459
			(6.65)	(6.72)	(3.48)	(3.54)	(3.04)	(3.01)
dust ISP 5								
quartz ISP 5								
ex-smoker v. non		-0.445		-0.855		-0.786		1.09
		(0.47)		(1.38)		(1.20)		(0.97)
smoker v. non		0.528		0.236		0.440		1.53
		(0.82)		(0.52)		(0.91)		(1.45)
deviance	220.8	218.3	315.1	309.2	304.7	297.5	153.3	150.0
degrees of freedom	501	499	603	601	602	600	442	440

relevant ISP, and the columns of exhibit 7.4.8 are numbered to correspond with the column labelling of exhibits 7.4.2, 7.4.4 and 7.4.6, with the addition of the suffix S to indicate models including smoking terms.

For each analysis, the smoking variable has been defined according to the data from the questionnaire from the survey at the start of the ISP. Thus, for example, progression in ISP 4 has been related to smoking status at 4th survey. It has already been noted that smoking data were not available for all the surveys at which films were taken, but this convention means that the serious loss of smoking data from 6th survey questionnaires affected only the numbers available for analysis for ISP 6. Of 482 men entering the analysis in exhibit 7.4.6, only 446 had data on smoking. For ISP 4, the analysis was of exactly the same group as in exhibit 7.4.2, and for ISP 5 the difference from the group contributing to exhibit 7.4.4 was only 5 men.

The three smoking groups did not show significant differences in the analyses of ISP 4 and ISP 6 data. For ISP 5, inclusion of the smoking effects gave deviance reductions in excess of the 5% point of the chi-squared distribution with 2 degrees of freedom, and the estimated coefficients suggested lower progression for ex-smokers, higher for current smokers, with the non-smokers in between but nearer the smokers.

Many other models were fitted and compared with and without smoking effects, and the overall conclusions were similar to those for the panel's readings of profusion. The estimates of smoking effects were of similar magnitude and significance for a particular set of data, regardless of which exposure variables were in the model. Estimates of the magnitudes and significance of exposure effects hardly differed between models with and without smoking terms, and the inclusion of smoking made no difference to the ranking of deviances across different choices of exposure variables. As with the panels' readings, it was concluded that there was no evidence that the effect of age or of exposure to respirable dust or quartz was influenced by the smoking habits of the individual coalworkers.

7.5 Summary and overview of findings on assessments of radiographs

This section contains an attempt to summarise the many analyses which have been presented, and to draw together their findings, for interpretation in context.

7.5.1 Findings from analyses of panel's readings

Between them, the five readers of the IOM panel made over 28000 assessments of over 3000 chest radiographs taken from more than 1400 men who were defined as belonging in the study population. Almost all of the films from the 4th, 5th and 6th PFR surveys were assessed, independently and in randomised order, on two separate occasions, while those obtained from the Medical Service 1980 survey were read once. Readings from a sixth reader who left before the exercise was complete were omitted from consideration.

The readers considered the vast majority of the films to be of adequate quality for assessment for signs of pneumoconiosis, and on only eight occasions did (one of) the readers consider the quality to be so bad that no assessment could be made. Readers differed in how critical they were of film quality.

The readers also differed considerably in the frequencies with which they recorded assessments of the profusion of small opacities for these films, but comparisons of readings of the same films by different readers showed association; although different readers did not place the films in exactly the same order of profusion, the tendency of readers to rank the films similarly was a great deal more consistent than could be expected to occur by chance. The readers also showed, although to different extents, a tendency to favour the central categories of the ILO (1980) classification of profusion.

The panel's assessments of the profusion of small opacities on each film were summarised by the calculation of a median value. Because the ILO (1980) profusion scale is ordered but not fully quantified, the principal analyses were performed on a binary variable formed by dichotomising the median variable between two major categories of profusion. Separate analyses were performed, by logistic regression methods, on binary variables split between categories 0 and 1, and between categories 1 and 2, although the low numbers showing disease sufficiently extreme to attract a median classification of 2/1 or greater meant that the latter analyses did not yield a great amount of information. Other abnormalities classifiable under the ILO (1980) scheme, such as large parenchymal opacities and abnormal features of the pleura, were not recorded frequently or consistently enough to merit similar analyses.

Separate analyses were performed of the assessments of films taken at different surveys. All analyses showed a highly significant relationship between the frequency of recordings of opacities at least 1/0 on the ILO (1980) 12-point scale, and the age of the individuals at survey. Taken over all the surveys, the increase in the log-odds per year of age was relatively constant at around 0.06 to 0.09. Age was included in all the regression models, before any of the exposure variables were added.

The exposure variables considered for inclusion in the regression models were based on individual records of times worked in various environments within the colliery, coupled with data from environmental measurements of the airborne concentrations of respirable dust in locations typical of those environments, and subsequent compositional analyses of the sampled dusts. These measurements were combined to produce time-weighted total concentrations of respirable quartz, and of the respirable non-quartz fraction of the dusts, for each man in each inter-survey period. Exposures up to the 3rd PFR survey were totalled, while those for ISPs 3, 4 and 5 were retained separately.

After allowance for age, the frequency at 4th survey of opacities of categories 1/0+ was found to increase significantly with increasing estimates of non-quartz dust exposure prior to 3rd survey. Association with quartz exposure over the same period was also significant, but less strongly. Frequency of opacities 2/1+ was much lower, and did not show any significant relationship with exposures either prior to 3rd survey or in ISP 3.

At 5th survey, frequency of small opacities 1/0+ showed significant associations with ISP 3 and ISP 4 exposures to both dust and quartz, but not with the previous exposures. The effect was considerably stronger with the ISP 4 exposures than with the ISP 3, and was a little stronger with quartz than with dust from both periods. Fitting combinations of variables suggested that once ISP 4 quartz was allowed for, no other exposure variable showed an association. Similar results were obtained from the analyses of opacities 2/1+, although the significance levels were rather lower in these analyses.

Frequency of small opacities at the 6th survey showed no associations with previous exposures, but significant relationships with ISP 3, 4 and 5 dust and quartz. Once again, ISP 4 exposures were most strongly associated, followed by ISP 3 and then ISP 5. At each survey, the association was stronger with the quartz than with the dust exposure. After allowance for ISP4 quartz, no other variable appeared to be significantly associated. Here, analyses of opacities 2/1+ showed similar results.

With the radiographs taken at the 1980 Medical Service survey, the frequency of small opacities 1/0+ showed similar associations, but in this case the strongest was with ISP 3 quartz; again, the association in any ISP was stronger with quartz than with dust. Opacities 2/1+ showed similar results, although here the ISP 4 exposures showed stronger associations than ISP 3, and dust in ISP 5 showed a stronger association than the quartz.

Additional analyses including variables to distinguish smokers and ex-smokers from lifelong non-smokers showed evidence of an effect of smoking on small opacities 1/0+ at 5th survey, but not at the other surveys. However, none of the conclusions regarding associations with dust and quartz exposures was altered by the inclusion of smoking habits, which had negligible effect on the magnitudes of the estimated coefficients. The estimates were likewise little affected by the inclusion of non-linear terms in age.

The conclusion that the frequencies with which the five readers recorded small opacities 1/0+ were associated more strongly with exposure to quartz than to dust was consistent over all the surveys, and was strengthened by the observation that, when both dust and quartz from the same ISP were in the model, the quartz exposure made a significant contribution after adjustment for the dust, but the reverse was not the case.

Consistency was also observed in the magnitudes of the coefficients of the exposures both between and within surveys; from the analyses at 6th survey, the coefficients for the increase in log-odds per $g.hr.m^{-3}$ of dust exposure in ISPs 3, 4 and 5 were 0.363, 0.431 and 0.425; while for the 1980 survey the corresponding estimates were 0.384, 0.375 and 0.360. At each of these surveys, however, inclusion of more than one quartz variable did not improve model fit over inclusion of the most strongly associated single variable.

The possibility was considered that these variables all represented the same information about individuals' exposures, a situation which could arise if the same men worked in the same conditions in each inter-survey period. Examination of the correlations between exposures, and of graphical displays, suggested that this was at most only partially true. Some further analyses were performed to aid interpretation. If, for example, two variables contained the same information about the association with small opacities, then inclusion of both of them, with their coefficients constrained to be equal, would be equivalent to forcing the same information into the regression twice, which in turn should produce a coefficient half that of either entered separately. In practice, regression with equality constraints on coefficients is straightforward to achieve, by adding together the explanatory variables concerned.

Regression of 6th survey opacities 1/0+ on combined ISP 4 and 5 quartz exposure yielded a coefficient of 0.280, which was less than either of the individual coefficients, but greater than a half of each. This combined exposure variable gave a slight improvement in model fit over either ISP 4 or ISP 5 quartz separately, as did combined ISP 3 and ISP 4 quartz. Of all the combinations tried, the lowest residual deviance was 511.0, with 614 degrees of freedom, with combined ISP 3, 4 and 5 quartz, for which the estimated regression coefficient was

0.192.

Analyses of opacities 1/0+ from the 1980 survey showed very similar results. In this case the smallest residual was given by the variable combining all of the previous, ISP 3, ISP 4 and ISP 5 quartz exposures, with a residual deviance of 435.1 on 441 degrees of freedom, and a regression coefficient of 0.167. A deviance of 435.6 was obtained from the combination of previous, ISP 3 and ISP 4 quartz exposures, for which the coefficient was 0.207.

These results from the 6th and 1980 surveys suggested that there was considerable overlap in the exposure information from different ISPs, and that the possibility of distinguishing separate effects from the different periods was very limited. Similar analyses for the 5th survey data yielded rather different results. Here, no combined exposure variable yielded a smaller residual deviance than had been obtained in the earlier analyses by adding ISP 4 quartz alone, although again the combined variables had considerably smaller coefficients than for ISP 4 quartz.

7.5.2 <u>Findings from medical readers' assessments of change.</u>

Two or more films per man were available from a subset of the study population. In a separate exercise, two medically-qualified readers made almost 2500 assessments of these series of films. Films from any one man were all seen at the same time, and in known temporal order, in order better to quantify changes in the degree of radiographic abnormalities taking place between the surveys. Of the abnormalities covered by the ILO (1980) classification, only the major parenchymal abnormalities, that is small and large opacities, were recorded. The readers considered that almost all of the films were of sufficient quality to allow such assessments to be made.

One reader assessed twice as many films as the other, and they assessed in common some 438 films. On the evidence of the assessments of this common subset, the reader who performed the majority of the serial readings tended to record small opacities at a lower level of frequency, on average, than the other. Recordings of large opacities were negligible in both readers' assessments.
Change in small opacities between adjacent surveys was quantified by calculating the numbers of steps on the ILO (1980) 12-point scale by which the side-by-side assessments of films differed. Again, the readers showed differences in the frequencies with which they recorded change on the same films, although the differences were less than in the profusion levels.

All the assessments of change were analysed by logistic regressions where the response variable was the recording of profusions showing at least one step of radiographic progression between adjacent surveys; and additional analyses were made for changes of two or more steps, of which the frequency was in general low. Most of the films had been seen by only one medical reader; allowance for which reader made the assessment was accomplished by including a term distinguishing the readers, to be estimated from those film pairs where asessments had been made by both readers.

Reader differences were significant, and were retained in the models, for all the analyses. Age was significant in the majority of the models, and was retained in all models for comparability. Frequency of observing one or more steps of progression over ISP 4 was not associated with exposures previous to 3rd survey, but was associated significantly with dust and quartz exposures in ISPs 3 and 4. The strongest associations were with ISP 3 quartz followed by ISP 4 quartz, but in models with more than one exposure variable only ISP 3 quartz remained significant in the presence of the other variables. Frequency of two or more steps was low, and showed no significant associations with any of the exposure variables.

Progression of at least one step over ISP 5 showed significant associations with all of the ISP 3, 4 and 5 dust and quartz exposure variables, the strongest being with ISP 4 quartz, followed by ISP 3 quartz, ISP 4 dust and ISP 3 dust. This was the only analysis of radiographic abnormalities in which, when more than one exposure variable was included in the model, more than one was significant. The strongest association was demonstrated by the model including both ISP 3 quartz and ISP 4 quartz, and the coefficients of these were similar, but somewhat greater than half the corresponding individual coefficients. Analyses of two or more steps of progression showed similar results, with reduced significance.

There was, as expected, a lower rate of change in the relatively short two-year period between the 6th and the 1980 survey (ISP 6). However, occurrences of progression again showed strongest association with ISP 4 quartz exposure, followed by ISP 4 dust, and ISP 5 dust and quartz in that order. With more than one variable included, only ISP 4 quartz was significant in the presence of other exposures. Analysis of two or more steps was based on only 6 positive results, all by one of the readers. It was unclear, on the evidence of such small numbers, whether or not these associations had occurred by chance.

The inclusion of terms for the smoking habits of the individuals had little effect on the conclusions. Smoking terms exceeded the 5% significance level only for progression in ISP 5; and allowance for smoking habits had only the smallest effect on either the estimated magnitudes or the statistical significance of the age and exposure coefficients. It was concluded that inferences drawn from analyses which did not take smoking into account would not be seriously distorted by that omission.

The analyses of progression in ISPs 4 and 6 gave results similar to those from the analyses of profusion, in that although the response was significantly associated with each of the dust and quartz exposures after 3rd survey, addition of more than one variable to the model at a time did not improve the model fit. In the same way as reported in 7.5.1 above, these analyses were supplemented by extra runs including variables with combined exposures obtained by adding together the individual exposure variables from different inter-survey periods. Analysis of progression 1+ in ISP 4 gave the lowest residual deviance with ISP 3 and 4 quartz conbined, at 219.8 with 501 degrees of freedom, only just less than with ISP 3 quartz alone. The estimated coefficient of the combined exposure was 0.263, compared with 0.450 for ISP 3 quartz. In the analysis of progression 1+ in ISP 6, no combined exposure variable improved the model over that containing just ISP 4 dust or ISP 4 quartz.

The analyses of progression in ISP 5 had been unique in that they showed evidence of independent effects of the quartz exposures in ISP 3 and ISP 4, but here the combined ISP 3, 4 and 5 quartz exposure gave the same residual deviance for one degree of freedom less, and an estimated coefficient of 0.356, where those for ISP 3 and 4 had been estimated at 0.434 and 0.484 respectively.

As with the results from the analyses of profusion, the inference that progression was associated with quartz, rather than with non-quartz exposure, was strengthened by regression models for progression in ISPs 4 and 5 with both quartz and dust from the same period, in which it was observed that the quartz remained significant in the presence of the dust, but that the reverse was not true.

7.5.3 Concluding remarks

The purpose of both of these exercises for the assessment of radiographic abnormalities was to examine any relationship with exposures, and in particular to examine any differences between quartz and non-quartz exposures in the strength of these associations. Both exercises have provided clear independent evidence of an association between small opacities and exposure history, and that the association was strongest with the individuals' exposures to quartz in ISPs 3 and 4.

These conclusion were drawn from analyses which converted responses on ordinal scales to binary variables for logistic regession analyses, which method, although widely used, may not be the most sensitive for analysis of such data. Limited analyses using alternative and more complex methods, while not contradicting these conclusions, suggested that the strength of the associations may have been underestimated by these analyses. These alternative methods demonstrated little unanimity in the magnitudes of their estimates of the regression coefficients, over which some uncertainty therefore remains.

8 RESULTS OF ANALYSES OF LUNG FUNCTION DATA

This chapter summarises the analyses made of lung function data from the medical surveys. These data were collected for almost all of the men who attended these surveys; occasionally, a man from whom a chest radiograph had been obtained would refuse to cooperate with the completion of questionnaires and lung function testing. In a very few cases the man was physically unable to perform the forced expiration necessary for the tests. In the remainder of cases, measurements were made from the spirometer traces, yielding estimates of the Forced Expiratory Volume in one second from the start of expiration (FEV), and the the whole Forced Vital Capacity (FVC) which the individual could expel from his lungs.

After one expiration for familiarisation, when the technicians could determine that the individual was able to make a satisfactory attempt, three further expirations were performed, with reruns for any which were technically unsatisfactory, and the variables analysed here are the arithmetic means over those three expirations. In a few cases, where the man was unable to complete three expirations, the means have been calculated over the number of expirations performed. From each survey, the analyses have included only those men for whom satisfactory FEV and FVC were both available, along with reliable data on smoking and exposures.

In 8.1, the data are considered cross-sectionally, taking separately the lung function data from the 4th, 5th and 6th surveys. (The 1980 survey was not a PFR survey, and medical questionnares and lung function tests were not applied). In 8.2, a longitudinal view is taken, and analyses are presented of the observed changes between consecutive surveys, in the lung function measurements of individuals.

8.1 Regression analyses of lung function data at each survey

8.1.1 The form and presentation of the regression analyses

In the cross-sectional analysis of the lung function data from each survey, the principal research question was whether the levels of lung function, on average, showed any relationship with the available data on the individuals' exposures to respirable dust and/or quartz up to the time of the survey. Lung function is related to body size, declines with increasing age, and is adversely affected by tobacco use, particularly by the inhalation of cigarette smoke (Fletcher *et al*, 1976; Cotes, 1979). Since age was known, and the medical questionnaires included a section on smoking habits, it was possible and desirable to allow for these factors in the course of the analyses, and this was done.

The dimension of the lung function measurements FEV and FVC is a volume; the unit employed throughout is litres. Analyses were performed by the application of standard multiple linear regression models (Draper and Smith, 1981), as widely used for the analysis of such variables. Initial analyses, performed interactively at a computer terminal and not shown here in detail, investigated the extent to which it was necessary to allow for age, body size and smoking habits. The results of these investigations clearly showed that age made a highly significant contribution to the model; that (standing) height was a better indicator of body size, and hence a better predictor of lung volumes, than weight; that when height was in the model, weight never gave a significant improvement; that smoking habits made a significant contribution to explaining differences between individuals in FEV; and that, in contrast, FVC showed no evidence of being related to smoking habits. All of these findings were consistent across the three surveys.

On the basis of these results, baseline models were defined relating FEV to age at survey, height and smoking habits (expressed as the mean differences between groups of ex–smokers and current smokers from the group of lifelong non–smokers); and FVC to age and height only. In addition, analyses of the data from the 5th survey included terms to allow for differences between measurements performed by two different technicians. Exhibits 8.1.1 to 8.1.6 show these baseline models, and the effects of adding to them dust and quartz exposure variables singly and in combinations.

The tabular layout of these exhibits is similar to that employed for the logistic analyses of the film–reading data, as described in 7.2.2. Each column is from the fitting of a different model, and is labelled with a combination of the variable being analysed (FEV or FVC), the survey number (4, 5 or 6), and a sequential number to differentiate the models. The entries in the tables are the estimated regression coefficients, quantifying the change in the response variable corresponding to a unit change in the explanatory variable, with all the other explanatories in the model held constant. Beneath each is the absolute value of the ratio of the coefficient to its estimated standard error, which in some cases is also appropriate for use as a partial t–statistic for the significance of the inclusion of the variable while all the other variables are included. In linear regression, these tests are algebraically identical with those using F–statistics obtained by testing partial mean squares against the residual mean square.

As in the earlier analyses, the exposure variables for ISPs 3 and 4 and prior to the 3rd survey were the sums of estimates calculated from "measured" and "unmeasured" time records. For all periods, the dust variable used was that representing the non–quartz fraction, obtained by subtracting the quartz exposure from the whole respirable dust exposure.

8.1.2 Regression analyses of 4th survey lung function data

Exhibit 8.1.1 presents some results from the linear regressions of FEV for 1090 men attending the 4th PFR survey. The initial baseline model included statistically significant contributions from linear terms in age and height, and differences between smoking groups. To this model were added in turn dust and quartz variables for the period prior to 3rd survey (previous), and between 3rd and 4th survey, that is inter-survey period (ISP) 3.

The estimates of the age, height and smoking effects varied little as different exposure variables were included. The average age effect was less than 40ml reduction for each year, while the height effect was a little over 40ml increase per cm. Ex-smokers and non-smokers did not differ significantly, but current smokers showed a significant deficit of around 130ml.

Neither previous dust nor previous quartz exposures made a significant contribution to the model. Both of the ISP 3 variables did bring highly significant improvements, but the model labelled FEV/4/6 shows that the inclusion of ISP 3 quartz was not an improvement over the model containing ISP 3 dust, which made the greater single contribution.

Both ISP3 dust and quartz had positive estimated coefficients, which would suggest that lung function was better in those who had been more heavily exposed to dust. The direction of this effect goes against expectation.

Exhibit 8.1.2 shows results of similar analyses of FVC at 4th survey for the same group of 1090 men. The baseline model included age and height, since smoking had been found not to make a significant contribution. Estimates of the average

Exhibit 8.1.1. Results of fitting different regression models to lung function at 4th survey. Variable analysed is forced expiratory volume (FEV) in one second, averaged over three (maximum) technically satisfactory expirations. FEV is expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models					
Variables	FEV/4/1	FEV/4/2	FEV/4/3	FEV/4/4	FEV/4/5	FEV/4/6
constant	-2.118	-2.123	-2.122	-2.225	-2.202	-2.222
	(4.39)	(4.40)	(4.40)	(4.65)	(4.60)	(4.64)
age at survey	-0.03676	-0.03583	-0.03638	-0.03834	-0.03766	-0.03819
	(27.15)	(17.64)	(18.22)	(27.52)	(27.65)	(26.92)
height (cm)	0.04206	0.04198	0.04204	0.04230	0.04229	0.04231
	(15.35)	(15.30)	(15.33)	(15.56)	(15.54)	(15.55)
ex-smokers v non	0.0005	-0.0036	-0.0011	0.0051	0.0053	0.0054
	(0.01)	(0.05)	(0.02)	(0.08)	(0.08)	(0.08)
smokers v non	-0.1399	-0.1417	-0.1406	-0.1312	-0.1326	-0.1313
	(2.75)	(2.78)	(2.76)	(2.60)	(2.62)	(2.60)
dust previous		-0.0005				
		(0.61)				
quartz previous			-0.0039			
			(0.26)			
dust ISP 3				0.00931		0.0067
				(4.28)		(1.32)
quartz ISP 3					0.0583	0.0188
					(4.10)	(0.57)
residual s.s.	337.50	337.38	337.48	331.90	332.34	331.81
degrees of freedom	1085	1084	1084	1084	1084	1083
residual m.s.	0.3111	0.3112	0.3113	0.3062	0.3066	0.3064

Exhibit 8.1.2. **Results of fitting different regression models to lung function at 4th survey. Variable analysed is forced vital capacity (FVC), averaged over three (maximum) technically satisfactory expirations.** FVC is expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models					
	FVC/4/1	FVC/4/2	FVC/4/3	FVC/4/4	FVC/4/5	FVC/4/6
constant	-4.611	-4.624	-4.626	-4.747	-4.714	-4.745
	(8.64)	(8.66)	(8.66)	(9.00)	(8.92)	(8.99)
age at survey	-0.03016	-0.02850	-0.02903	-0.03228	-0.03132	-0.03218
	(20.41)	(12.87)	(13.32)	(21.33)	(21.15)	(20.86)
height (cm)	0.06078	0.06064	0.06072	0.06109	0.06107	0.06110
	(19.95)	(19.88)	(19.91)	(20.29)	(20.25)	(20.29)
dust previous		-0.00092				
		(1.00)				
quartz previous			-0.0118			
			(0.70)			
dust ISP 3				0.01277		0.01102
				(5.30)		(1.97)
quartz ISP 3					0.0776	0.0127
					(4.93)	(0.35)
residual s.s.	418.58	418.19	418.39	408.01	409.42	407.97
degrees of freedom	1087	1086	1086	1086	1086	1085
residual m.s.	0.3851	0.3851	0.3853	0.3757	0.3770	0.3760

age effect were relatively constant in all the models, at about 30ml reduction per year, and the height effect was estimated at about 60ml difference per cm.

As with FEV, the inclusion of previous exposure variables produced negative estimates which were not statistically significant. Both ISP 3 exposure variables produced highly significant improvements, with dust the greater, and quartz not yielding an improvement when dust was in the model. Again, the estimated coefficients were positive, against expectation.

8.1.3 Regression analyses of 5th survey lung function data

Exhibit 8.1.3 presents some results from the linear regressions of FEV for 854 men attending the 5th PFR survey. The initial baseline model included statistically significant contributions from linear terms in age and height, and differences between smoking groups. Also included was a term to differentiate the results obtained by an experienced technician and a new technician, who had not previously performed these tests on survey. Examination of the completed questionnaires, on which the lung function measurements were recorded, had identified 100 where the experienced technician had performed the test, and 724 performed by the new technician. There were 45 questionnaires without indication, and these have been kept separate as a distinct group in the analyses. Of course, if both technicians contributed to the measurements of this group, their heterogeneity will slightly inflate the residual in the regression analyses. However, this option was chosen as preferable to the most obvious alternative, omission of the entire group and the potentially useful information it contained regarding the other relationships of interest.

Exhibit 8.1.3. Results of fitting different regression models to lung function at 5th survey. Variable analysed is forced expiratory volume (FEV) in one second, averaged over three (maximum) technically satisfactory expirations. FEV is expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models FEV/5/1	FEV/5/2	FEV/5/3	FEV/5/4	FEV/5/5	FEV/5/6	FEV/5/7
constant	-1.444	-1.490	-1.471	-1.539	-1.493	-1.534	-1.545
	(2.70)	(2.79)	(2.75)	(2.89)	(2.80)	(2.88)	(2.90)
age at survey	-0.03576	-0.03690	-0.03648	-0.03611	-0.03588	-0.03561	-0.03618
	(22.60)	(21.36)	(22.02)	(22.86)	(22.71)	(19.80)	(22.87)
height (cm)	0.03740	0.03766	0.03756	0.03763	0.03754	0.03754	0.03762
	(12.40)	(12.49)	(12.46)	(12.53)	(12.47)	(12.48)	(12.53)
ex-smokers v non	-0.0974	-0.0980	-0.0978	-0.1033	-0.1013	-0.1042	-0.1031
	(1.41)	(1.42)	(1.41)	(1.50)	(1.46)	(1.51)	(1.49)
smokers v non	-0.1731	-0.1697	-0.1702	-0.1697	-0.1711	-0.1708	-0.1697
	(3.18)	(3.11)	(3.12)	(3.13)	(3.15)	(3.15)	(3.13)
tech old v new	0.2149	0.2097	0.2123	0.1927	0.2012	0.1909	0.1928
	(3.53)	(3.44)	(3.48)	(3.15)	(3.29)	(3.12)	(3.15)
tech not known	0.1526	0.1376	0.1376	0.1216	0.1332	0.1231	0.1220
	(1.76)	(1.58)	(1.58)	(1.40)	(1.53)	(1.41)	(1.40)
dust ISP 3		0.00397				-0.00201	
		(1.64)				(0.59)	
quartz ISP 3			0.0237				
			(1.45)				
dust ISP 4				0.01045		0.01253	0.01515
				(2.94)		(2.51)	(2.21)
quartz ISP 4					0.0356		-0.0262
					(2.09)		(0.80)
residual s.s.	268.16	267.31	268.03	265.45	266.78	265.34	265.25
degrees of freedom	847	846	846	846	846	845	845
residual m.s.	0.3166	0.3160	0.3162	0.3138	0.3153	0.3140	0.3139

To the baseline model were added in turn dust and quartz variables for the period prior to 3rd survey (previous), and those for inter-survey periods (ISPs) 3 and 4. Estimates of the effects in the baseline changed little when exposure variables were included. The age effect was about 36ml per year, and the height effect about 38ml per cm. Current smokers differed significantly from non-smokers, with an estimated deficit of around 170ml on average, while the ex-smokers differed from the non-smokers by about 100ml. The new technician's measurements were on average about 200ml lower than those of the experienced (=old) technician, while those for whom the technician was not known were on average about 130ml higher than those of the new.

Previous dust and quartz made no significant contributions to any of the models tried, and are not included in exhibit 8.1.3. Estimated coefficients for both dust and quartz from ISP 3 were positive, but failed to reach conventional levels of statistical significance. ISP 4 dust made a highly significant contribution, as did ISP 4 quartz, although not quite as strongly. With ISP 4 dust in the model, neither ISP 3 dust nor ISP 4 quartz gave a significant improvement, while ISP 4 dust remained significant in the presence of the others.

The estimated coefficients of ISP 4 dust and quartz were positive, indicating better lung function in those with higher dust exposure. While this finding was against expectation, it was consistent with the results above for the data from 4th survey.

Exhibit 8.1.4 shows results from the regression analyses of FVC at 5th survey for the same group of men. The baseline model included age and height and a term for differences between the technicians, as described above. Smoking habits were not included.

Exhibit 8.1.4. Results of fitting different regression models to lung function at 5th survey. Variable analysed is forced vital capacity (FVC), averaged over three (maximum) technically satisfactory expirations. FEV is expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models FVC/5/1	FVC/5/2	FVC/5/3	FVC/5/4	FVC/5/5	FVC/5/6	FVC/5/7
constant	-3.218	-3.285	-3.256	-3.356	-3.293	-3.351	-3.363
	(5.53)	(5.65)	(5.60)	(5.82)	(5.69)	(5.81)	(5.83)
age at survey	-0.03169	-0.03353	-0.03285	-0.03227	-0.03191	-0.03150	-0.03235
	(18.73)	(18.13)	(18.55)	(19.21)	(18.95)	(16.42)	(19.22)
height (cm)	0.05077	0.05117	0.05101	0.05107	0.05097	0.05094	0.05107
	(15.40)	(15.55)	(15.50)	(15.65)	(15.55)	(15.59)	(15.65)
tech old v new	0.3367	0.3281	0.3323	0.3010	0.3134	0.2983	0.3011
	(5.04)	(4.92)	(4.98)	(4.52)	(4.69)	(4.47)	(4.52)
tech not known	0.3150	0.2912	0.2909	0.2667	0.2830	0.2689	0.2671
	(3.32)	(3.06)	(3.05)	(2.82)	(2.98)	(2.84)	(2.82)
dust ISP 3		0.00644				-0.00307	
		(2.43)				(0.83)	
quartz ISP 3			0.0390				
			(2.18)				
dust ISP 4				0.01673		0.01992	0.02208
				(4.33)		(3.66)	(2.96)
quartz ISP 4					0.0604		-0.0299
					(3.25)		(0.84)
residual s.s.	323.11	320.86	321.31	316.13	319.13	315.87	315.87
degrees of freedom	849	848	848	848	848	847	845
residual m.s.	0.3806	0.3784	0.3789	0.3728	0.3763	0.3729	0.3729

Again, the estimates of effects in the baseline model changed little when the exposure variables were included. The age effect was about 32ml reduction per year, and the height effect about 51ml increase per cm. The difference between the old and new technicians was about 300ml, about 50% greater than for the FEV data. The FVCs for whom the operating technician was not known were on average similar in level to those from the old technician, other things being equal.

Variables representing the men's previous dust and quartz exposures did not make a significant contribution to the fit of the models, and are not shown. However, the associations with the ISP 3 and 4 dust and quartz variables were somewhat stronger for FVC than for FEV, and all were highly significant. ISP 4 dust showed the strongest effect, and when this variable was in the model none of the others gave a significant improvement, although ISP 4 dust remained significant in their presence. As in the other analyses reported so far, the regression coefficients showed an association of higher lung function with higher exposures.

8.1.4 Regression analyses of 6th survey lung function data

Exhibit 8.1.5 presents results from the linear regressions of FEV for 569 men attending the 6th PFR survey. The initial baseline model included age, height and smoking habits, and the estimates of these effects changed little when exposure variables were added to the models. The age effect was about 36ml reduction per year, and the height effect about 37ml per cm. Smokers had a significantly lower lung function than non-smokers, by about 250ml on average. Ex-smokers differed from non-smokers by about 90ml.

Addition of the previous (not shown), ISP 3, 4 and 5 dust and quartz exposure variables failed to demonstrate any significant relationships of FEV with these

Exhibit 8.1.5. Results of fitting different regression models to lung function at 6th survey. Variable analysed is forced expiratory volume (FEV) in one second, averaged over three (maximum) technically satisfactory expirations. FEV is expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models									
	FEV/6/1	FEV/6/2	FEV/6/3	FEV/6/4	FEV/6/5	FEV/6/6	FEV/6/7	FEV/6/8	FEV/6/9	FEV/6/10
constant	-1.336	-1.336	-1.342	-1.358	-1.353	-1.377	-1.391	-1.380	-1.426	-1.357
	(2.02)	(2.02)	(2.02)	(2.05)	(2.04)	(2.07)	(2.09)	(2.09)	(2.16)	(2.05)
age at survey	-0.03631	-0.03639	-0.03595	-0.03749	-0.03685	-0.03652	-0.03624	-0.03594	-0.03640	-0.03763
	(18.82)	(15.72)	(16.61)	(17.92)	(18.50)	(18.74)	(18.77)	(15.51)	(16.82)	(17.66)
height (cm)	0.03731	0.03732	0.03731	0.03744	0.03741	0.03744	0.03748	0.03734	0.03758	0.03743
	(10.10)	(10.08)	(10.09)	(10.14)	(10.12)	(10.12)	(10.13)	(10.12)	(10.20)	(10.13)
ex-smokers v non	-0.0968	-0.0963	-0.0991	-0.0845	-0.0921	-0.0925	-0.0943	-0.0912	-0.0850	-0.0825
	(1.21)	(1.19)	(1.23)	(1.05)	(1.15)	(1.15)	(1.18)	(1.13)	(1.06)	(1.02)
smokers v non	-0.2463	-0.2461	-0.2462	-0.2422	-0.2454	-0.2424	-0.2426	-0.2429	-0.2366	-0.2411
	(3.89)	(3.89)	(3.89)	(3.83)	(3.88)	(3.82)	(3.83)	(3.85)	(3.75)	(3.81)
dust ISP 3		0.00018						-0.00716		
		(0.06)						(1.55)		
quartz ISP 3			-0.0078						-0.0556	
			(0.37)						(1.90)	
dust ISP 4				0.00647				0.01426	0.01452	0.00904
				(1.45)				(2.12)	(2.36)	(1.03)
quartz ISP 4					0.0228					-0.0142
					(1.07)					(0.34)
dust ISP 5						0.00368				
						(0.77)				
quartz ISP 5							0.0220			
							(0.99)			
residual s.s.	178.52	178.52	178.48	177.86	178.16	178.52	178.22	177.10	176.73	177.82
degrees of freedom	564	563	563	563	563	563	563	562	562	562
residual m.s.	0.3165	0.3171	0.3170	0.3159	0.3164	0.3168	0.3165	0.3151	0.3145	0.3164

variables. The largest improvement was with ISP 4 dust, but this was well short of even the 10% point of statistical significance. However, when ISP 3 quartz and ISP 4 dust were simultaneously added to the baseline model, the partial t-statistics indicated that each made a contribution after allowing for the other, although the overall reduction in the sum of squares was not significant; the coefficient of ISP 4 dust was positive, and that of ISP 3 quartz negative. ISP 4 dust and ISP 3 dust as a pair produced a similar but less pronounced pattern. The model with ISP 4 dust and quartz fitted less well, and other combinations, which are not shown, less well yet. It was concluded that there was little evidence of a dust or quartz effect on the 6th survey FEV data, in contrast with the strong positive relationships found in the equivalent 4th and 5th survey data.

Exhibit 8.1.6 shows results from the regression analyses of the 6th survey FVC data for the same men. The baseline model here included only age and height, and once more the estimates of their effects differed little, whether dust or quartz variables were included or not. The age effect was about 33ml reduction per year, and the height effect about 54ml increase per cm. Looking over exhibits 8.1.1 to 8.1.6, the consistency of these results is notable. In particular, the age effects seem particularly stable, at about 36ml per year for FEV and 32ml per year for FVC. The age effects varied a little more, from 37ml to 42ml per cm for FEV and from 51ml to 61ml per cm for FVC.

Neither ISP 3 dust nor quartz gave a significant reduction to the sum of squares, nor did those from the previous period, which are not shown. None of the ISP 4 and 5 dust and quartz variables made a contribution which exceeded the 5% level of statistical significance, but all except ISP 4 quartz exceeded the 10% point, with ISP 4 dust just ahead of ISP 5 dust. Addition of ISP 4 quartz or ISP 5 dust or quartz to a model including ISP 4 dust did not improve the fit. The model

Exhibit 8.1.6. **Results of fitting different regression models to lung function at 6th survey. Variable** analysed is forced vital capacity (FVC), averaged over three (maximum) technically satisfactory expirations. FEV is expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	FVC/6/1	FVC/6/2	FVC/6/3	FVC/6/4	FVC/6/5	FVC/6/6	FVC/6/7	FVC/6/8	FVC/6/9	FVC/6/10
					Parsimonious models					
constant	-3.445	-3.446	-3.450	-3.476	-3.466	-3.540	-3.540	-3.470	-3.513	-3.555
	(4.75)	(4.75)	(4.76)	(4.80)	(4.78)	(4.88)	(4.88)	(4.79)	(4.84)	(4.93)
age at survey	-0.03315	-0.03367	-0.03276	-0.03474	-0.03378	-0.03363	-0.03299	-0.03506	-0.03439	-0.03327
	(15.92)	(13.56)	(14.06)	(15.49)	(15.73)	(16.05)	(15.85)	(15.39)	(14.87)	(14.34)
height (cm)	0.05385	0.05390	0.05384	0.05404	0.05398	0.05416	0.05414	0.05401	0.05413	0.05419
	(13.25)	(13.25)	(13.24)	(13.33)	(13.29)	(13.35)	(13.34)	(13.31)	(13.33)	(13.41)
dust ISP 3		0.00130								
		(0.39)								
quartz ISP 3			-0.0086							-0.0749
			(0.37)							(2.33)
dust ISP 4				0.00918				0.01595	0.00571	0.02005
				(1.88)				(1.67)	(0.77)	(2.98)
quartz ISP 4					0.0279			-0.0375		
					(1.20)			(0.82)		
dust ISP 5						0.00951			0.00494	
						(1.83)			(0.62)	
quartz ISP 5							0.0410			
							(1.67)			
residual s.s.	217.54	217.48	217.48	216.18	216.98	216.26	216.46	215.92	216.03	214.12
degrees of freedom	566	565	565	565	565	565	565	564	564	564
residual m.s.	0.3843	0.3849	0.3849	0.3826	0.3840	0.3828	0.3831	0.3828	0.3830	0.3796

including ISP 3 quartz and ISP 4 dust produced a similar result to that for 6th survey FEV, in that the partial t-statistics indicated a significant negative association with ISP 3 quartz and a positive one with ISP 4 dust; but in this case the overall test for the inclusion of both variables over the baseline was significant at better than the 2% level.

One additional observation on the tables of results concerns the similarity of the residual mean squares, which hardly varied between surveys, although that for FVC was somewhat higher at about 0.38 than for FEV at about 0.31. This, coupled with the stability of the age and height coefficient estimates, suggests some inherent comparability of the data from the different surveys, as might be expected for data from (subsets of) the same men. The differences in the findings regarding lung function and exposures at 6th survey, from those at the 4th and 5th surveys, is thus more likely to be due to differences over time with regard to the working environment, in a way not directly related to age or height.

8.1.5 Checks on the adequacy of the regression models

The analyses of FEV presented in exhibits 8.1.1, 8.1.3 and 8.1.5 parameterised the effects of smoking as overall differences between the average FEV levels of groups of non-smokers, ex-smokers and current smokers. Such a parameterisation would not normally be expected to be adequate. Studies such as those of Fletcher *et al* (1976) and Love and Miller (1982) have demonstrated an increase in smokers in the rate of loss of FEV, which would lead to older (or longer-term) smokers showing greater differences from their non-smoking peers than would be observed in younger groups. In the absence of lifetime smoking histories, and on the assumption that age correlates well with length of time a man has smoked, we can model this process by fitting separate slopes on age in the three smoking groups.

Exhibit 8.1.7 shows the results from the three surveys, each of which consists of the appropriate baseline model plus the exposure which made the greatest individual contribution. The models in the first, fourth and seventh columns have appeared in earlier tabulations and bear the same model labels as before. Each is compared with a model containing the same terms, augmented by terms for the differences of the age slopes of the ex-smokers and smokers from the non-smokers, and bearing the same label with the suffix I.

For all three surveys, the estimated coefficients were negative, suggesting as expected a reduced FEV in smokers and non-smokers; only at the 5th survey was the F-test for inclusion of these terms significant at the 5% level, and it was at this survey that the estimated differences were largest. However, the inclusion of these terms made little difference to the size or statistical significance of the dust exposure effects at any of the surveys. Additional models (not shown here) included a term for the amount smoked, as reported by the current smokers at the relevant surveys, but this term did not make a significant contribution at any of the surveys. No representation of smoking habits had any significant effect on FVC at any survey.

The results in exhibits 8.1.1 to 8.1.6 were from regression models in which the relationship between lung function and the continuous variables representing age, height and exposure were assumed to be linear. There is no particular reason to expect linearity in these relationships, and indeed other studies have often shown a slightly curved relationship of lung function to age.

The validity of the assumption of linearity was checked in a similar manner to that described in 7.2.9 for the logistic regression models. Each of the variables age and height was split into six equal-sized groups, and a separate regression constant fitted for the level of each group; the residual sums of squares from these models were compared with those from the models which assumed linearity. The overwhelming conclusion from these tests was that a degree of non-linearity was present in the relationship with age, but not with height. It was decided to investigate the effect of non-linearity in age in a little more detail. The third, sixth and ninth columns of Exhibit 8.1.7 show the same models as before, augmented by quadratic and cubic terms in age, indicated by the suffix A. To minimise arithmetic problems, the mean age at each survey was subtracted before the squares and cubes were calculated; these values were 44.5, 44.6 and 43.0 years at 4th, 5th and 6th survey respectively.

The results from each survey showed a significant improvement in the fit, largely due to the cubic term. Little change was observed in the estimates of the height and smoking effects, nor did the technician effects at the 5th survey alter much. However, there was a considerable drop in the estimates of the dust exposure effects for the 4th and 6th surveys, although that for 5th survey dropped by only about 10%.

Exhibit 8.1.7. Results of fitting regression models with and without non-linear terms in age to lung function at 4th, 5th and 6th surveys. Variable analysed is forced expiratory volume (FEV) in one second, averaged over three (maximum) technically satisfactory expirations. FEV is expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models								
Variables	FEV/4/4	FEV/4/4I	FEV/4/4A	FEV/5/4	FEV/5/5I	FEV/5/4A	FEV/6/4	FEV/6/6I	FEV/6/4A
constant	-2.225	-2.400	-1.370	-1.539	-1.837	-1.041	-1.358	-1.526	-0.609
	(4.65)	(4.82)	(2.66)	(2.89)	(3.38)	(1.87)	(2.05)	(2.27)	(0.88)
age at survey	-0.03834	-0.03450	-0.04859	-0.03611	-0.02831	-0.04529	-0.03749	-0.03263	-0.05087
	(27.52)	(10.59)	(16.68)	(22.86)	(8.51)	(13.87)	(17.92)	(7.90)	(11.18)
(age-mean(age))			0.000040			0.000156			0.000237
			(0.30)			(1.10)			(1.04)
(age-mean(age))			0.000028			0.000026			0.000045
			(3.55)			(3.19)			(3.07)
height (cm)	0.04230	0.04249	0.04038	0.03763	0.03764	0.03709	0.03744	0.03734	0.03685
	(15.56)	(15.61)	(14.75)	(12.53)	(12.57)	(12.38)	(10.14)	(10.09)	(9.96)
ex-smokers v non	0.0051	0.372	0.0055	-0.1033	0.234	-0.0924	-0.0845	0.052	-0.0921
	(0.08)	(1.46)	(0.08)	(1.50)	(0.99)	(1.34)	(1.05)	(0.18)	(1.15)
smokers v non	-0.1312	0.028	-0.1310	-0.1697	0.237	-0.1575	-0.2422	0.025	-0.2353
	(2.60)	(0.18)	(2.60)	(3.13)	(1.47)	(2.88)	(3.83)	(0.13)	(3.74)
ex-smokers.age		-0.00844			-0.00865			-0.00384	
		(1.53)			(1.67)			(0.60)	
smokers.age		-0.00408			-0.01024			-0.00676	
		(1.15)			(2.68)			(1.46)	
tech old v new				0.1927	0.1912	0.1928			
				(3.15)	(3.13)	(3.16)			
tech not known				0.1216	0.1244	0.1287			
				(1.40)	(1.44)	(1.48)			
dust ISP 3	0.00931	0.00897	0.00611						
	(4.28)	(4.10)	(2.59)						
dust ISP 4				0.01045	0.01035	0.00944	0.00647	0.00652	0.00393
				(2.94)	(2.92)	(2.44)	(1.45)	(1.46)	(0.84)
residual s.s.	331.90	331.13	326.40	265.45	263.19	262.21	177.86	177.17	174.30
degrees of freedom	1084	1082	1082	846	844	844	563	561	561
residual m.s.	0.3062	0.3060	0.3017	0.3138	0.3118	0.3107	0.3159	0.3158	0.3107

Exhibit 8.1.8. Results of fitting regression models with and without non-linear terms in age to lung function at 4th, 5th and 6th surveys. Variable analysed is forced expiratory volume (FVC) in one second, averaged over three satisfactory expirations. FVC is expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models					
	FVC/4/4	FVC/4/4A	FVC/5/4	FVC/5/4A	FVC/6/4	FVC/6/4A
constant	-4.747	-3.577	-3.356	-2.405	-3.476	-2.402
	(8.92)	(6.32)	(5.82)	(4.02)	(4.80)	(3.19)
age at survey	-0.03228	-0.04521	-0.03227	-0.04806	-0.03474	-0.05285
	(21.33)	(14.22)	(19.21)	(13.79)	(15.49)	(10.70)
$(\text{age-mean(age)})^2$		-0.000163		0.000080		0.000072
		(1.12)		(0.53)		(0.29)
$(\text{age-mean(age)})^3$		0.000032		0.000042		0.000056
		(3.70)		(4.79)		(3.55)
height (cm)	0.06109	0.05843	0.05107	0.04994	0.05404	0.05284
	(20.29)	(19.40)	(15.65)	(15.49)	(13.33)	(13.21)
tech old v new			0.3010	0.2960		
			(4.52)	(4.50)		
tech not known			0.2667	0.2669		
			(2.82)	(2.85)		
dust ISP 3	0.01277	0.00735				
	(5.30)	(2.83)				
dust ISP 4			0.01673	0.01293	0.00918	0.00409
			(4.33)	(3.11)	(1.88)	(0.81)
residual s.s.	408.01	396.42	316.13	306.40	216.18	208.45
degrees of freedom	1086	1084	848	846	565	563
residual m.s.	0.3757	0.3657	0.3728	0.3622	0.3826	0.3702

Exhibit 8.1.8 shows results for a similar investigation of the effects of non-linearity in age on the analyses of FVC. The results were broadly similar, with a significant improvement due to the cubic term, and little difference in height or technician effects. Again, the estimates of the exposure effects at 4th and 6th survey were reduced considerably when the non-linear age terms were included, and this time the same was observed for the 5th survey as well.

Inclusion of non-linear terms in age, then, has been seen to reduce the estimates of most of the exposure effects. However, for 4th and 5th survey FEV and FVC, the reduced estimates were still highly significant and still positive, in the opposite direction to that expected.

8.2 Regression analyses of change in lung function between surveys

8.2.1 The form and presentation of the analyses

The cross-sectional analyses of lung function data at each survey, described in 8.1, showed, at least for the 4th and 5th survey data, a relationship between increasing exposure and increasing lung function which seemed implausible as a dose-response relationship. An alternative explanation may be that the relationship was a feature of the structure of the study population, perhaps due to selection effects during or prior to the period over which these surveys took place. We would expect such selection effects to have less influence on changes observed over time within individuals.

Exhibits 8.2.1 to 8.2.4 show results from fitting regression models to changes in lung function, obtained by calculating the difference between the values obtained at two consecutive surveys. Throughout, the difference has been calculated by subtracting the later from the earlier survey; the unit of the response is thus the drop (or gain, if negative) in FEV or FVC over a period of about 4 years. The regressions were carried out in the same way as the cross-sectional analyses reported in 8.1. Age, height and smoking habits were always taken from the earlier survey. Only men for whom changes in both FEV and FVC could be calculated were included in the analyses.

The labelling of the columns of results is similar to that employed in earlier exhibits; the response variables are labelled dFV for difference in FEV, and dVC for difference in FVC. The periods over which the differences were calculated were ISPs 4 and 5, and these are labelled I4 and I5 respectively.

8.2.2 Changes in lung function from 4th to 5th survey

Exhibit 8.2.1 presents results from the regression analysis of the change in FEV for 639 men with lung function data at both 4th and 5th surveys. The initial baseline model included age, height, smoking and a term for the differences between technicians which affected the lung function measurements at 5th survey.

The age coefficient was positive and significant, indicating a greater rate of loss in older men. Taller men also lost at a greater rate, presumably in some functional way as a result of their greater lung volumes. Smokers lost on average about 20ml per year (80ml per 4 years) more than non–smokers, from whom ex–smokers did not significantly differ. The estimates of these effects changed little when the exposure variables were added to the model.

None of the exposure variables for previous, ISP 3 or ISP 4 dust or quartz made a significant contribution to the model, either singly or in combinations. The largest reduction in the residual sums of squares was due to ISP 3 dust, but even this failed to reach the 10% significance level. However, it was notable that all the estimated coefficients for ISP 3 and 4 exposures were positive. Because the positive values of the response indicate a drop, any positive relationship with an exposure would, if real, indicate an association of greater loss of lung function with greater exposure.

Exhibit 8.2.1. Results of fitting different regression models to change in lung function between 4th and 5th survey, i.e. over about 4 years. Variable analysed is (positive) drop in FEV, expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models									
	dFV/I4/1	dFV/I4/2	dFV/I4/3	dFV/I4/4	dFV/I4/5	dFV/I4/6	dFV/I4/7	dFV/I4/8	dFV/I4/9	dFV/I4/10
constant	-0.777	-0.774	-0.779	-0.813	-0.792	-0.808	-0.790	-0.824	-0.811	-0.810
	(2.04)	(2.03)	(2.04)	(2.13)	(2.07)	(2.11)	(2.07)	(2.16)	(2.12)	(2.12)
age at survey	0.00587	0.00664	0.00623	0.00542	0.00567	0.00606	0.00598	0.00529	0.00636	0.00535
	(4.58)	(3.59)	(3.44)	(4.12)	(4.37)	(4.67)	(4.64)	(4.00)	(3.43)	(3.77)
height (cm)	0.00434	0.00423	0.00431	0.00441	0.00437	0.00436	0.00434	0.00443	0.00428	0.00441
	(2.02)	(1.96)	(2.00)	(2.06)	(2.04)	(2.03)	(2.02)	(2.07)	(1.98)	(2.06)
ex-smokers v non	0.0143	0.0107	0.0130	0.0159	0.0153	0.0138	0.0138	0.0157	0.0114	0.0161
	(0.27)	(0.20)	(0.24)	(0.30)	(0.28)	(0.26)	(0.26)	(0.29)	(0.21)	(0.30)
smokers v non	0.0818	0.0803	0.0812	0.0854	0.0836	0.0818	0.0813	0.0860	0.0836	0.0856
	(2.02)	(1.98)	(2.01)	(2.11)	(2.07)	(2.02)	(2.01)	(2.13)	(2.06)	(2.11)
tech old v new	-0.1089	-0.1087	-0.1086	-0.1088	-0.1085	-0.1128	-0.1119	-0.1094	-0.1085	-0.1081
	(2.81)	(2.80)	(2.80)	(2.81)	(2.80)	(2.90)	(2.88)	(2.83)	(2.80)	(2.76)
tech not known	-0.0255	-0.0263	-0.0256	-0.0329	-0.0309	-0.0307	-0.0290	-0.0308	-0.0343	-0.0325
	(0.46)	(0.48)	(0.46)	(0.59)	(0.56)	(0.55)	(0.52)	(0.56)	(0.43)	(0.59)
dust previous		-0.000415							-0.000528	
		(0.57)							(0.73)	
quartz previous			-0.0037							
			(0.28)							
dust ISP 3				0.00268				0.00572	0.00280	0.00287
				(1.56)				(1.46)	(1.62)	(1.22)
quartz ISP 3					0.0110			-0.0211		
					(1.02)			(0.86)		
dust ISP 4						0.00240				-0.00041
						(0.97)				(0.12)
quartz ISP 4							0.0093			
							(0.83)			
residual s.s.	69.740	69.704	69.731	69.472	69.625	69.636	69.664	69.390	69.414	69.472
degrees of freedom	632	631	631	631	631	631	631	630	630	630
residual m.s.	0.1103	0.1105	0.1105	0.1101	0.1103	0.1104	0.1104	0.1101	0.1102	0.1103

Exhibit 8.2.2 shows results from similar analyses of change over ISP 4 in FVC. The same baseline model was employed as for change in FEV, since the initial investigations had indicated that smoking made a (just) significant contribution to FVC change, which had not been the case for cross-sectional FVC. Conclusions regarding the effects of height, age and smoking were qualitatively similar to those for FEV in the same period, and changed little on the inclusion of exposure variables.

None of the exposure variables included made a significant contribution to the model. In contrast to the position with FEV change, the greatest contribution was from the previous exposures, with previous quartz just reaching the 10% significnce level. The coefficients of the previous exposures were negative, which is the direction of smaller losses of lung function in association with higher exposures; this seems implausible, although consistent with some of the findings from the cross-sectional analyses.

8.2.3 Changes in lung function from 5th to 6th survey

Exhibit 8.2.3 shows results of regression analyses of the change in FEV for 457 men with lung function data at both 5th and 6th surveys. The baseline model included the same terms as for the changes over ISP 4, that is age, height, smoking and a term for the technician differences at 5th survey.

Obviously, the analyses of ISP 5 and ISP 4 data were not based on the same men, although they overlapped by about 350. The estimate of the age coefficient was smaller than the corresponding estimate from ISP 4 by between 30% and 50% in the various models, while the height coefficient was larger than the ISP 4 estimate by about the same margin. Smoking effects also differed, with the

Exhibit 8.2.2. Results of fitting different regression models to change in lung function between 4th and 5th survey, i.e. over about 4 years. Variable analysed is (positive) drop in FVC, expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models									
Variables	dVC/I4/1	dVC/I4/2	dVC/I4/3	dVC/I4/4	dVC/I4/5	dVC/I4/6	dVC/I4/7	dVC/I4/8	dVC/I4/9	dVC/I4/10
constant	-1.523	-1.513	-1.538	-1.567	-1.540	-1.553	-1.531	-1.530	-1.592	-1.570
	(3.22)	(3.21)	(3.26)	(3.32)	(3.26)	(3.27)	(3.24)	(3.23)	(3.37)	(3.31)
age at survey	0.01000	0.01263	0.01259	0.00944	0.00978	0.01018	0.01007	0.01270	0.01232	0.01282
	(6.30)	(5.53)	(5.63)	(5.80)	(6.09)	(6.34)	(6.31)	(5.55)	(5.51)	(5.69)
height (cm)	0.00909	0.00871	0.00885	0.00918	0.00913	0.00912	0.00909	0.00879	0.00891	0.00887
	(3.42)	(3.27)	(3.33)	(3.46)	(3.43)	(3.43)	(3.42)	(3.29)	(3.36)	(3.34)
ex-smokers v non	-0.0143	-0.0266	-0.0238	-0.0124	-0.0132	-0.0147	-0.0146	-0.0252	-0.0229	-0.0244
	(0.22)	(0.40)	(0.36)	(0.19)	(0.20)	(0.22)	(0.40)	(0.38)	(0.34)	(0.37)
smokers v non	0.1002	0.0950	0.0961	0.1046	0.1022	0.1002	0.0999	0.0956	0.1007	0.0961
	(2.00)	(1.90)	(1.92)	(2.09)	(2.04)	(2.00)	(2.00)	(1.91)	(2.01)	(1.92)
tech old v new	-0.2399	-0.2392	-0.2375	-0.2397	-0.2394	-0.2435	-0.2418	-0.2380	-0.2369	-0.2413
	(5.00)	(4.99)	(4.96)	(5.01)	(4.99)	(5.05)	(5.02)	(4.96)	(4.95)	(5.01)
tech not known	-0.0428	-0.0458	-0.0439	-0.0519	-0.0488	-0.0478	-0.0451	-0.0446	-0.0546	-0.0492
	(0.63)	(0.67)	(0.64)	(0.76)	(0.71)	(0.70)	(0.66)	(0.65)	(0.80)	(0.72)
dust previous		-0.001434						-0.00051		
		(1.60)						(0.22)		
quartz previous			-0.0272					-0.0185	-0.0313	-0.0276
			(1.64)					(0.43)	(1.88)	(1.67)
dust ISP 3				0.00329					0.00384	
				(1.55)					(1.79)	
quartz ISP 3					0.0121					
					(0.91)					
dust ISP 4						0.00229				0.00243
						(0.75)				(0.80)
quartz ISP 4							0.0058			
							(0.42)			
residual s.s.	106.82	106.39	106.36	106.42	106.68	106.73	106.79	106.36	105.82	106.26
degrees of freedom	632	631	631	631	631	632	631	630	630	630
residual m.s.	0.1690	0.1686	0.1686	0.1686	0.1691	0.1691	0.1692	0.1688	0.1680	0.1687

Exhibit 8.2.3. **Results of fitting different regression models to change in lung function between 5th and 6th survey, i.e. over about 4 years. Variable analysed is (positive) drop in FEV, expressed in litres.** Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models									
	dFV/I5/1	dFV/I5/2	dFV/I5/3	dFV/I5/4	dFV/I5/5	dFV/I5/6	dFV/I5/7	dFV/I5/8	dFV/I5/9	dFV/I5/10
constant	-1.002	-1.047	-1.016	-1.132	-1.074	-1.193	-1.153	-1.191	-1.191	-1.171
	(2.27)	(2.37)	(2.30)	(2.56)	(2.42)	(2.69)	(2.59)	(2.68)	(2.68)	(2.64)
age at survey	0.00368	0.00251	0.00273	0.00289	0.00342	0.00426	0.00462	0.00412	0.00368	0.00379
	(2.29)	(1.43)	(1.60)	(1.79)	(2.12)	(2.65)	(2.78)	(2.20)	(2.12)	(2.23)
height (cm)	0.00593	0.00622	0.00607	0.00647	0.00627	0.00646	0.00637	0.00647	0.00654	0.00638
	(2.38)	(2.50)	(2.44)	(2.61)	(2.52)	(2.61)	(2.56)	(2.61)	(2.64)	(2.58)
ex-smokers v non	-0.0880	-0.0863	-0.0872	-0.0874	-0.0888	-0.0915	-0.0948	-0.0912	-0.0900	-0.0878
	(1.50)	(1.48)	(1.49)	(1.50)	(1.52)	(1.57)	(1.62)	(1.57)	(1.55)	(1.50)
smokers v non	0.0214	0.0236	0.0220	0.0243	0.0220	0.0285	0.0255	0.0285	0.0275	0.0290
	(0.47)	(0.52)	(0.48)	(0.54)	(0.48)	(0.63)	(0.56)	(0.63)	(0.61)	(0.64)
tech old v new	0.0916	0.0867	0.0895	0.0837	0.0864	0.0868	0.0915	0.0864	0.0846	0.0843
	(1.90)	(1.79)	(1.85)	(1.74)	(1.79)	(1.81)	(1.90)	(1.80)	(1.76)	(1.75)
tech not known	-0.268	-0.236	-0.238	-0.214	-0.243	-0.213	-0.236	-0.211	-0.205	-0.209
	(1.31)	(1.30)	(1.16)	(1.05)	(1.19)	(1.04)	(1.15)	(1.03)	(1.01)	(1.03)
dust ISP 3		0.00341						0.00036		
		(1.66)						(0.15)		
quartz ISP 3			0.0231							
			(1.66)							
dust ISP 4				0.00845					0.00409	
				(2.74)					(0.89)	
quartz ISP 4					0.0239					
					(1.68)					
dust ISP 5						0.00991		0.00958	0.00653	0.01529
						(2.89)		(2.35)	(1.27)	(2.11)
quartz ISP 5							0.0341			-0.0283
							(2.13)			(0.84)
residual s.s.	55.718	55.378	55.380	54.805	55.371	54.704	55.161	54.701	54.608	54.617
degrees of freedom	450	449	449	449	449	449	449	448	448	448
residual m.s.	0.1238	0.1233	0.1233	0.1221	0.1233	0.1218	0.1229	0.1221	0.1219	0.1219

coefficient for ISP 5 about one quarter that for ISP 4, and nowhere near statistical significance. The ex-smokers appeared to be losing lung function less quickly in ISP 5 than the non-smokers, although this difference was below the 10% significance level. The estimated difference between the old and new technicians was similar to that in ISP 4 (but of opposite sign, since the difference was at the start of ISP 5 but the end of ISP 4). The difference for those for whom the technician was not known looked very dissimilar from the ISP 4 value at a casual glance, but was in fact very poorly determined, since only three such men contributed to the ISP 5 analyses, compared to 39 for ISP 4.

The addition to the baseline model of the previous exposure variables made no useful contributions to the models, and they are not shown here. The addition of ISP 3, 4 and 5 dust and quartz variables gave estimates which were in all cases positive. Those for ISP 3 just exceeded the 10% level of significance, and were accompanied by the largest drop in the age coefficient. The largest individual contribution to the model was from ISP 5 dust, followed by ISP 4 dust, both of which were significant at better than the 1% level. In the models with two exposure variables, when both ISP 4 and ISP 5 dust variables were included, neither was significant in the presence of the other. However, model dFV/I5/10 shows that when ISP 5 dust and quartz were both included, quartz was not significant while dust was.

Exhibit 8.2.4 shows results of similar regression analyses of the change in FVC in the same men over ISP 5. As with FEV, the estimated age effect was smaller than for ISP 4, but unlike FEV the height effect in ISP 5 was also smaller, and in fact not statistically significant. Smoking effects were again not significant, but of similar magnitude to those for FEV. The estimated difference between the old and new technicians was also similar.

Exhibit 8.2.4. Results of fitting different regression models to change in lung function between 5th and 6th survey, i.e. over about 4 years. Variable analysed is (positive) drop in FVC, expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models									
Variables	dVC/I5/1	dVC/I5/2	dVC/I5/3	dVC/I5/4	dVC/I5/5	dVC/I5/6	dVC/I5/7	dVC/I5/8	dVC/I5/9	dVC/I5/10
constant	-0.755	-0.833	-0.781	-0.957	-0.890	-1.029	-0.986	-0.961	-0.953	-1.024
	(1.43)	(1.58)	(1.48)	(1.82)	(1.69)	(1.95)	(1.85)	(1.83)	(1.81)	(1.94)
age at survey	0.00595	0.00391	0.00423	0.00472	0.00545	0.00678	0.00739	0.00497	0.00458	0.00562
	(3.09)	(1.87)	(2.08)	(2.45)	(2.84)	(3.54)	(3.73)	(2.35)	(2.36)	(2.72)
height (cm)	0.00261	0.00311	0.00286	0.00345	0.00325	0.00337	0.00328	0.00344	0.00340	0.00353
	(0.88)	(1.05)	(0.97)	(1.17)	(1.10)	(1.14)	(1.11)	(1.17)	(1.15)	(1.20)
ex-smokers v non	-0.0816	-0.0788	-0.0802	-0.0807	-0.0831	-0.0867	-0.0920	-0.0811	-0.0797	-0.0837
	(1.16)	(1.13)	(1.15)	(1.17)	(1.19)	(1.25)	(1.32)	(1.17)	(1.15)	(1.21)
smokers v non	0.0367	0.0405	0.0378	0.0412	0.0378	0.0469	0.0430	0.0409	0.0421	0.0448
	(0.67)	(0.74)	(0.70)	(0.76)	(0.70)	(0.87)	(0.79)	(0.76)	(0.78)	(0.83)
tech old v new	0.2222	0.2137	0.2183	0.2098	0.2125	0.2153	0.2220	0.2102	0.2108	0.2109
	(3.84)	(3.71)	(3.80)	(3.67)	(3.69)	(3.77)	(3.87)	(3.67)	(3.68)	(3.69)
tech not known	-0.213	-0.156	-0.159	-0.128	-0.165	-0.133	-0.163	-0.130	-0.125	-0.118
	(0.87)	(0.64)	(0.65)	(0.53)	(0.68)	(0.55)	(0.67)	(0.54)	(0.51)	(0.49)
dust ISP 3		0.00592						-0.00101		
		(2.42)						(0.28)		
quartz ISP 3			0.0415							
			(2.50)							
dust ISP 4				0.01315				0.01427	0.01734	0.00819
				(3.58)				(2.63)	(2.50)	(1.49)
quartz ISP 4					0.0448				-0.0226	
					(2.64)				(0.71)	
dust ISP 5						0.01421				0.00743
						(3.47)				(1.22)
quartz ISP 5							0.0521			
							(2.73)			
residual s.s.	79.802	78.774	78.706	77.588	78.578	77.716	78.500	77.574	77.500	77.332
degrees of freedom	450	449	449	449	449	449	449	448	448	448
residual m.s.	0.1773	0.1754	0.1753	0.1728	0.1750	0.1731	0.1748	0.1732	0.1730	0.1726

As in ISP 4, addition of the dust and quartz variables produced stronger evidence of associations in FVC change than in FEV change. Previous quartz and dust were not significant, and are not shown. However, each of the ISP 3, 4 and 5 dust and quartz variables made a significant contribution, and all of them had positive estimated coefficients. The largest contributions were from ISP 4 and ISP 5 dust, both of which exceeded the 0.1% significance level.

In the models which included more than one exposure variable at a time, the ISP 3 variables were not significant in the presence of ISP 4 dust or ISP 5 dust. When ISP 4 dust and quartz were included together, the quartz effect was not significant, but the dust effect was still significant. With the inclusion of both ISP 4 and ISP 5 dust variables, both effects became not significant, although neither was entirely eliminated. However, none of the two-variable models was a significant improvement on the one-variable models with only ISP 4 or ISP 5 dust.

8.2.4 Checks on the adequacy of the regression models

The results in exhibits 8.2.1 to 8.2.4 were from analyses in which the effects of age, height and exposure were assumed to be linear. As with the cross-sectional analyses, there was no intrinsic reason why linearity should be a better description of any relationship found than might be given by some other curve. Non-linearity, particularly in the age-dependence, had already been observed in the cross-sectional analyses of lung function, and this provided an additional justification for examining the question for the longitudinal analyses.

In earlier analyses, the question of linearity in a variable had initially been investigated by forming a grouped version of the variable and examining the pattern of the estimates for the different groups. Since in this case there was ample justification, this process was omitted, and the analyses proceeded by examining the effects on regression models which had already been fitted, of including quadratic and cubic terms in age and in height.

In all cases, the inclusion of non-linear terms in height failed to improve the fit of the regression models. Linear adjustment for height was therefore accepted as adequate in the present data, and none of the results shown here includes any non-linear terms in height.

In contrast, the non-linear terms in age were significant in most of the models, and exhibits 8.2.5 and 8.2.6 show the results of including both quadratic and cubic terms in models which have already appeared in earlier tables of results. The column labelling uses the earlier model identifier, with the suffix A for the models in which the non-linear age terms have been included.

Exhibit 8.2.5 shows examples of comparisons from analyses of change in FEV over ISPs 4 and 5. In the first, the inclusion of the non-linear terms in one of the models fitted to FEV change was significant, but had little effect on the estimates of the effects of smoking or technician. The height effect was reduced by about 25%, and the ISP 3 dust effect, which had been positive but had not achieved significance, was annihilated.

The remaining four columns show similar comparisons between models for the change in FEV over ISP 5, with ISP 4 and ISP 5 dust variables respectively. With a linear term in age, ISP 4 dust was highly significant, and the addition of the non-linear terms gave a reduction in the residual sum of squares which fell short of the 5% significance level. With the non-linear terms added, there was some reduction in the height and technician terms, while the dust effect was reduced by

Exhibit 8.2.5. Results of fitting regression models with and without non-linear terms in age to change in lung function in ISPs 4 and 5. Variable analysed is (positive) drop in FEV expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

			Parsimonious models			
Variables	dFV/I4/4	dFV/I4/4A	dFV/I5/4	dFV/I5/4A	dFV/I5/6	dFV/I5/6A
constant	-0.813	-0.091	-1.132	-0.810	-1.193	-0.859
	(2.13)	(0.23)	(2.56)	(1.76)	(2.69)	(1.87)
age at survey	0.00542	-0.00393	0.00289	-0.00260	0.00426	-0.00173
	(4.12)	(1.71)	(1.79)	(0.92)	(2.65)	(0.61)
(age-mean(age))^2		-0.000260		0.000212		0.000147
		(2.00)		(0.67)		(0.46)
(age-mean(age))^3		0.000023		0.000024		0.0000230
		(3.19)		(1.68)		(0.58)
height (cm)	0.00441	0.00324	0.00647	0.00610	0.00646	0.00615
	(2.06)	(1.54)	(2.61)	(2.47)	(2.61)	(0.49)
ex-smokers v non	0.0159	0.0052	-0.0874	-0.0894	-0.0915	-0.0934
	(0.30)	(0.10)	(1.50)	(1.54)	(1.57)	(1.61)
smokers v non	0.0854	0.0745	0.0243	0.0253	0.0285	0.0280
	(2.11)	(1.87)	(0.54)	(0.56)	(0.63)	(0.62)
tech old v new	-0.1088	-0.1224	0.0837	0.0732	0.0868	0.0752
	(2.81)	(3.24)	(1.74)	(1.52)	(1.81)	(0.57)
tech not known	-0.0329	-0.0619	-0.214	-0.227	-0.213	-0.225
	(0.59)	(1.14)	(1.05)	(1.12)	(1.04)	(1.11)
dust ISP 3	0.00268	-0.00044				
	(1.56)	(0.25)				
dust ISP 4			0.00854	0.00667		
			(2.74)	(2.09)		
dust ISP 5					0.00991	0.00855
					(2.89)	(2.47)
residual s.s.	69.472	65.625	54.805	54.113	54.704	53.906
degrees of freedom	631	629	449	447	449	447
residual m.s.	0.1101	0.1043	0.1221	0.1211	0.1218	0.1206

over 20%, but was still significant.

In the model containing ISP 5 dust, the addition of the non-linear terms produced a reduction in the residual which was significant at the 5% level, although the partial t-statistics suggested that it was not necessary to include both quadratic and cubic terms. However, their inclusion reduced the height and technician coefficients, as well as that for ISP 5 dust, although this remained significant.

Exhibit 8.2.6 shows similar comparisons in models for FVC changes over ISPs 4 and 5. Non-linear age terms were highly significant in the model for change over ISP 4, and their inclusion again reduced the estimated height effect, but had only a small influence on the smoking and technician effects.

In model dVC/I4/3, with only a linear age term, the term for previous quartz exposure had produced a reduction in the residual which lay just on the 10% significance level. The estimate was negative, which is in the direction of less severe loss of lung function being associated with higher exposure. With quadratic and cubic terms in age in the model, however, the estimate was reduced by over a half, and no longer approached conventional significance levels. In model dVC/I4/4, where the ISP 3 dust variable was positive although not reaching significance, the estimate in the presence of non-linear age was near zero.

The non-linear terms were also highly significant in both models for the change in FVC over ISP 5, again producing only modest influences on the height, smoking and technician effects. The estimates of the ISP 4 and ISP 5 dust effects were reduced by around 20%, but both remained significant.

Exhibit 8.2.6. Results of fitting regression models with and without non-linear terms in age to change in lung function in ISPs 4 and 5. Variable analysed is (positive) drop in FVC expressed in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

	Parsimonious models							
Variables	dVC/I4/3	dVC/I4/3A	dVC/I4/4	dVC/I4/4A	dVC/I5/4	dVC/I5/4A	dVC/I5/6	dVC/I5/6A
constant	-1.513	-0.747	-1.567	-0.713	-0.957	-0.398	-1.029	-0.445
	(3.21)	(1.55)	(3.32)	(1.47)	(1.82)	(0.73)	(1.95)	(0.82)
age at survey	0.01259	0.00024	0.00944	-0.00109	0.00472	-0.00481	0.00678	-0.00368
	(5.63)	(0.07)	(5.80)	(0.38)	(2.45)	(1.45)	(3.54)	(1.09)
$(age-mean(age))^2$		-0.000392		-0.000415		0.000347		0.000254
		(2.42)		(2.58)		(0.92)		(0.68)
$(age-mean(age))^3$		0.000024		0.000024		0.0000416		0.0000400
		(2.61)		(2.60)		(2.43)		(2.33)
height (cm)	0.00885	0.00769	0.00918	0.00775	0.00345	0.00282	0.00337	0.00283
	(3.33)	(2.96)	(3.46)	(2.98)	(1.17)	(0.96)	(1.14)	(0.97)
ex-smokers v non	-0.0238	-0.0332	-0.0124	-0.0302	-0.0807	-0.0843	-0.0867	-0.0901
	(0.36)	(0.51)	(0.19)	(0.46)	(1.17)	(1.23)	(1.25)	(1.32)
smokers v non	0.0961	0.0861	0.1046	0.0863	0.0412	0.0427	0.0469	0.0459
	(1.92)	(1.75)	(2.09)	(1.75)	(0.76)	(0.79)	(0.87)	(0.86)
tech old v new	-0.2375	-0.2565	-0.2397	-0.2584	0.2098	0.1917	0.2153	0.1950
	(4.96)	(5.47)	(5.01)	(5.52)	(3.67)	(3.37)	(3.77)	(3.44)
tech not known	-0.0439	-0.0909	-0.0519	-0.0908	-0.128	-0.152	-0.133	-0.155
	(0.64)	(1.36)	(0.76)	(1.35)	(0.53)	(0.63)	(0.55)	(0.64)
quartz previous	-0.0272	-0.0126						
	(1.64)	(0.77)						
dust ISP 3			0.00329	-0.00055				
			(1.55)	(0.25)				
dust ISP 4					0.01315	0.01003		
					(3.58)	(2.67)		
dust ISP 5							0.01421	0.01183
							(3.47)	(2.89)
residual s.s.	106.39	100.52	106.42	100.60	77.588	75.497	77.716	75.288
degrees of freedom	631	629	631	629	449	447	449	447
residual m.s.	0.1686	0.1598	0.1686	0.1599	0.1728	0.1689	0.1731	0.1684

It was concluded that, at most, only a minor part of the significant effects of dust exposure variables on the changes in FEV and FVC over ISP 5 could conceivably be ascribed to confounding with non-linear effects of age differences. For ISP 4, however, the weak evidence from earlier models, of possible associations with exposures in ISP 3 and earlier, could just as readily be explained as an artefact caused by the omission of significant non-linear age effects.

8.3 Small opacities and lung function

All the analyses so far have treated separately the radiographic evidence of pneumoconiotic shadows and the measures of individuals' lung function. Although the patterns of the relationships have not been identical, both responses have shown associations with estimates of the amounts of dust and quartz to which the individuals had been exposed in the workplace. It is obviously appropriate to inquire whether these responses are independent, or whether they are manifestations of the same disease process, and this section describes a set of analyses directed at this question.

Care is always needed in the joint assessment of separate but possibly associated responses, particularly where they represent changes which take place over a period or periods of time. In the present data, it is reasonable to assume that, if the responses are directly and causally associated, the appearance of pneumoconiotic shadows will be accompanied or followed by a consequent loss of lung function. The possibility that the loss of lung function, or the damage mechanism underlying that loss, may cause pneumoconiotic shadows, is a much less plausible hypothesis. This investigation therefore adopted the strategy of including an indicator of pneumoconiotic status amongst the explanatory variables in regression analyses of the lung function variables. It was believed that useful insights into the nature and extent of any relationships would be obtained by examination of the regression coefficient for such a variable and its statistical significance, plus any effect of its inclusion on the regression coefficients of the other variables also present in the regression models.

8.3.1 Small opacities and lung function at individual surveys

Regression analyses which had already been performed, and reported in exhibits 8.1.1 to 8.1.6, were rerun with the inclusion of an extra explanatory variable which took the value 1 if the median category assigned to an individual's radiograph was 1/0 or higher, and 0 otherwise. In the analyses of small opacities profusions already reported, this was of course a response variable; but here, it was designed to represent any difference between the average values of lung function variables for men with and without small opcaities of at least category 1/0.

Some of the results are shown in exhibit 8.3.1, with the columns labelled as for the corresponding models in exhibits 8.1.1 to 8.1.6, with the suffix P to indicate that the model includes the variable for profusion of small opacities. Typical results are presented for both lung function variables, from each of the surveys. Numbers of men are very slightly smaller than in the earlier tabulations; median categories were not calculable for a few men for whom lung function data were analysed earlier.

The conclusions from each of these analyses was similar. After allowance for age, height and, where relevant, smoking habits and technician effects, the men with small opacities on their radiographs had lower FEV by between 100ml and 160ml, and lower FVC by between 180ml and 230ml. These differences were statistically significant in each of the models concerned. The inclusion of these differences

Exhibit 8.3.1. Results of linear regression analyses of lung function variables at 4th, 5th and 6th surveys, including indicator for small opacities median profusion >= 1/0 as explanatory variable. Variables analysed are FEV and FVC, in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models					
	FEV/4/4P	FEV/5/4P	FEV/6/4P	FVC/4/4P	FVC/5/4P	FVC/6/4P
constant	-2.286	-1.624	-1.298	-4.791	-3.445	-3.444
	(4.77)	(3.04)	(1.96)	(9.11)	(5.98)	(4.76)
age at survey	-0.03670	-0.03426	-0.03683	-0.02956	-0.02967	-0.03348
	(24.51)	(20.52)	(17.12)	(18.15)	(16.76)	(14.53)
height (cm)	0.04246	0.03783	0.03690	0.06109	0.05115	0.05367
	(15.60)	(12.58)	(9.98)	(20.37)	(15.71)	(13.23)
ex-smokers v non	0.0134	-0.1193	-0.0696			
	(0.20)	(1.73)	(0.87)			
smokers v non	-0.1250	-0.1695	-0.2197			
	(2.47)	(3.12)	(3.46)			
tech old v new		0.1957			0.3114	
		(3.20)			(4.68)	
tech not known		0.1476			0.2872	
		(1.68)			(3.02)	
dust ISP 3	0.00915			0.01245		
	(4.20)			(5.19)		
dust ISP 4		0.01148	0.00762		0.01840	0.01085
		(3.21)	(1.70)		(4.75)	(2.22)
opacities 1/0+	-0.1080	-0.1590	-0.1502	-0.1824	-0.2289	-0.2153
	(2.90)	(3.21)	(2.25)	(4.44)	(4.37)	(2.96)
residual s.s.	328.49	261.07	174.61	400.07	308.54	211.62
degrees of freedom	1077	839	559	1079	841	561
residual m.s.	0.3050	0.3112	0.3124	0.3708	0.3669	0.3772

made little difference to the estimates of age, height, smoking and technician effects. The estimates of the coefficients for dust and quartz exposure were also little changed; they remained positive, which as noted above was implausible in any directly causal sense; the likelihood that this was due to a large selection effect made it difficult to draw any useful inferences regarding the joint responses from these cross-sectional analyses. Models fitted with quartz instead of dust exposure variables led to very similar conclusions, and are not shown.

8.3.2 Small opacities and changes in lung function

Similar supplementary analyses were performed on the variables representing change in lung function over inter-survey periods 4 and 5. Here, because non-linear terms in age had made a significant contribution to the fit of the regression models, with a noticeable effect upon the regression coefficients of the exposure variables, the baseline model included polynomial terms up to the cubic in age, as well as terms for height, smoking habits, technician differences, and exposure variables. To these models was added the variable taking the value 1 if the median category of small opacities was of profusion 1/0 or greater, and the value 0 otherwise.

Selected results are shown in exhibit 8.3.2. These may be compared to the relevant columns of exhibits 8.2.5 and 8.2.6, in which the models with the cubic terms in age have column labels suffixed with A. (The analyses in exhibit 8.3.2 were based on slightly fewer men, but reruns of the earlier models on the reduced subgroups produced almost identical results.)

The reduction in the residual due to introducing the small opacities variable exceeded the 5% significance point only for the drop in FEV over ISP 4,

Exhibit 8.3.2. Results of linear regression analyses of changes in lung function variables in ISPs 4 and 5, including indicator for small opacities median profusion >= 1/0 at start of ISP as explanatory variable. Variables analysed are (positive) drop in FEV and FVC , in litres. Tabulated values are estimated regression coefficients; absolute value of ratio of estimate to standard error is in parenthesis.

Variables	Parsimonious models					
	dFV/I4/4P	dFV/I5/4P	dFV/I5/6P	dVC/I4/4P	dVC/I5/4P	dVC/I5/6P
constant	-0.106	-0.997	-1.035	-0.702	-0.556	-0.589
	(0.27)	(2.14)	(2.23)	(1.44)	(1.01)	(1.07)
age at survey	-0.00343	-0.00057	0.00029	-0.00024	-0.00305	-0.00196
	(1.45)	(0.20)	(0.10)	(0.08)	(0.89)	(0.56)
$(\text{age-mean(age)})^2$	-0.000259	0.000194	0.000122	-0.000404	0.000340	0.000242
	(1.98)	(0.61)	(0.39)	(2.51)	(0.91)	(0.65)
$(\text{age-mean(age)})^3$	0.000023	0.000021	0.000020	0.000024	0.000040	0.000038
	(3.15)	(1.48)	(1.39)	(2.60)	(2.29)	(2.20)
height (cm)	0.00327	0.00678	0.00678	0.00759	0.00338	0.00334
	(1.54)	(2.72)	(2.73)	(2.89)	(1.15)	(1.13)
ex-smokers v non	0.0049	-0.1006	-0.1049	-0.0302	-0.0962	-0.1020
	(0.09)	(1.74)	(1.81)	(0.46)	(1.40)	(1.48)
smokers v non	0.0756	0.0247	0.0274	0.0863	0.0399	0.0433
	(1.87)	(0.54)	(0.60)	(1.72)	(0.74)	(0.80)
tech old v new	-0.1234	0.0750	0.0773	-0.2618	0.1923	0.1958
	(3.22)	(1.57)	(1.62)	(5.54)	(3.39)	(3.45)
tech not known	-0.0554	-0.242	-0.242	-0.0775	-0.164	-0.168
	(1.01)	(1.20)	(1.20)	(1.14)	(0.68)	(0.70)
dust ISP 3	-0.00040			-0.00057		
	(0.23)			(0.26)		
dust ISP 4		0.00770			0.01091	
		(2.41)			(2.88)	
dust ISP 5			0.00932			0.01247
			(2.70)			(3.04)
opacities 1/0+	-0.0330	-0.1129	-0.1105	-0.0600	-0.0921	-0.0879
	(1.17)	(2.57)	(2.52)	(1.71)	(1.77)	(1.69)
residual s.s.	65.180	53.149	52.976	99.580	75.687	74.531
degrees of freedom	620	444	444	620	444	444
residual m.s.	0.1051	0.1215	0.1193	0.1606	0.1682	0.1679

although it exceeded the 10% point for the drop in FVC over both ISPs 4 and 5. In all cases, the estimated coefficient was negative; if real, this would suggest that the group showing small opacities on their radiographs had a lower rate of loss of lung function on average than those with no such radiological signs, but this does not seem plausible as a causal effect. Estimates of other terms in the models changed little in the presence of the small opacities variable, although the dust exposure effects increased slightly in all the models. Models fitted with quartz instead of dust showed similarly little change in the presence of the small opacities variable.

If loss of lung function were a sequel to parenchymal damage visible radiographically, and assuming that such damage were due to exposure to coalmine dust and/or quartz, we should have expected to observe in these analyses both a significant predictive role for the variable representing the observation of opacities, and a sizeable decrease in the coefficients of the exposures which preceded the damage. Therefore it was concluded that, while the effects observed might be due to population selection artefacts, these analyses had provided no evidence of a causal association between radiographic small opacities and level or loss of lung function in these men.

8.4 Summary and overview of findings on lung function data

This section summarises the findings from the cross-sectional analyses of differences between individuals' lung function variables at each survey, and from the analyses of change in individuals' lung function between adjacent surveys.

8.4.1 Findings from cross-sectional analyses

Lung function at each survey was typified by the mean values of Forced Expiratory Volume (FEV) in one second, and Forced Vital Capacity (FVC), both in litres, from usually three (and only occasionally fewer) forced expirations, taken from data already held on computer for the men studied in this project.

FEV and FVC at each survey were analysed separately, using techniques of multiple regression. Both variables showed at each survey, as expected, strong associations with age and with height, which appeared the most successful indicator of differences in individuals' body sizes. FEV was significantly lower in smokers than in non-smokers, but FVC showed no significant differences between groups according to their smoking habits. Measurements at the 5th survey were taken by two different technicians, and there was significant evidence of a difference, on average, between the values they recorded; this was allowed for in the analyses.

Inclusion of the dust and quartz exposure variables in the regressions produced unexpected results. After allowing for age, height, smoking habits and technician effects as appropriate, both FEV and FVC at 4th survey showed significant associations with ISP 3 dust and, less strongly, ISP 3 quartz. FEV at 5th survey showed similar associations with ISP 4 dust, and, less strongly, ISP 4 quartz, while 5th survey FVC showed much stronger associations with ISP 4 dust and quartz, and less strongly with ISP 3 dust and quartz. Lung function variables at 6th survey did not show significant associations with exposure variables. However, all of the significant associations with the 4th and 5th survey FEV and FVC were in the direction of better lung function with increased exposures, an effect which was considered biologically implausible. The effect may have been an artefact of the population structure of the study group, perhaps as a results of some selection process connected with fitness for work.

Additional analyses to test the adequacy of the regression models suggested that models with non-linear terms up to the cubic in age provided an improved fit. These reduced somewhat, but did not remove, the apparent positive associations mentioned above, and thus did not alter the conclusions regarding exposures.

8.4.2 Findings from analyses of change in lung function

Lung function data from men attending more than one survey were used to observe directly changes in lung function over inter-survey periods. By subtracting the later value from the earlier value, the amounts by which both FEV and FVC had dropped (or gained, if negative) in the ISP were calculated. This was done for those pairs of measurements which spanned ISP 4, and independently for ISP 5, and the changes were analysed by linear regression methods.

Regression models for the changes in both FEV and FVC in ISP 4 included significant terms for age, height, smoking habits and an effect due to which technician took the 5th survey measurements. Allowing for these factors, the strongest association of FEV with an exposure variable was with ISP 3 dust exposure, but this failed to reach even the 10% significance level. For FVC, previous quartz and previous dust exposures almost reached significance at 10%, but none of the other exposure variables reached this significance level.

Evidence of associations in the data from ISP 5 was more convincing. Highly significant associations were found with all the ISP 3, 4 and 5 exposure variables, the strongest being with ISP 4 and 5 dust exposures. Models including more than one exposure variable did not improve the fit, however, and no exposure variable was significant in the presence of one of the others.

As with the cross-sectional analyses of FEV and FVC, additional modelling suggested that the dependence on age was non-linear, and better modelled by the inclusion of quadratic and cubic terms. With the baseline models altered in this way, reanalysis of the ISP 4 lung function variables showed no evidence of associations with any of the exposure variables. For ISP 5, the inclusion of a cubic relationship with age reduced the significance of the exposure variables, but the associations with ISP 4 and ISP 5 dust were still the most strongly significant. The direction of these associations was of increasing rate of loss of lung function with increasing dust exposure in ISPs 4 and 5, which contrasted with the cross-sectional results but was consistent with published work, and biologically plausible.

Analyses which distinguished the individuals according to whether their radiographs had been assessed as showing small opacities of profusion at least 1/0+ failed to suggest that small opacities caused by exposure were a precursor of loss of lung function. Again, this was consistent with other published work.

9 DISCUSSION

To recap, the objectives of this project, stated in Chapter 2, were: to make an intensive study of chest radiographs taken from men working in Colliery P between 1970 and 1980; to relate any abnormalities on those radiographs to the men's histories of exposure to respirable dust and its components, particularly quartz; and to examine the relationships between those exposures, the progression of radiographic abnormalities, and spirometric measurements of lung function.

In analyses which examined the lung function variables FEV and FVC from each survey separately, as cross-sectional views of the study group at distinct points in time, the results showed the expected average increase in men of greater height, decrease in older men, and lower FEV in smokers. Lung function variables at 6th survey showed no evidence of a relationship with exposures, but the data from the 4th and 5th surveys showed evidence of associations with exposures in the inter-survey periods preceding these surveys; the direction of these associations, however, was of higher lung function values in men with higher exposures, and it seemed biologically implausible that the inhalation of coalmine dust could cause an increase in lung function, or even a decrease in the natural rate of age-related loss. Modelling of the age dependence by a cubic rather than a linear relationship diminished this relationship somewhat, but did not eradicate it.

A longitudinal view of these data produced a quite different inference. The losses experienced in FEV and FVC in inter-survey periods 4 and 5, by each individual man who attended consecutive surveys, were calculated by simple subtraction. These losses (or gains, where the values were negative) were analysed as responses, and initial analyses showed weak evidence of increased loss in ISP 4 of FVC in men with higher exposures to dust in ISP 3, and also some suggestion of decreased loss of FEV and FVC with higher exposures prior to 3rd survey. Allowing a cubic rather than linear adjustment for age had the effect of removing all suggestion of association with exposure. In contrast, the changes in lung function over ISP 5 showed rather stronger evidence of greater rates of loss in men with higher dust exposures in ISPs 4 and 5, and slightly less strong with ISP 3 exposures, but no relationship with previous exposures, even with linear adjustment for age. Allowing a cubic adjustment for age altered the regression estimates for the dependence on ISP 3, 4 and 5 exposures and reduced somewhat the apparent strength of the evidence, which however remained statistically significant.

One interpretation of these apparently contradictory findings is arrived at by arguing that the longitudinal analyses, depending on direct observations of change, are intrinsically more reliable as indicators of the effects of exposure; and that the results therefore point to a real effect of the exposures to dust between 3rd and 6th survey on the rates of loss of FEV and FVC in ISP 5, after allowance for other factors. There is biological plausibility in an assumption that some sort of damage has, in some men at least, occurred as a consequence of the exposure of the lungs to the insult of inhaled coalmine dust. It is also consistent with the observation of greater loss of FEV in men with higher exposures to dust which emerged from the analysis of change in FEV between 2nd and 4th surveys in five PFR collieries not including colliery P (Love and Miller, 1982). That finding was the first published demonstration of an effect of coalmine dust on rates of change of FEV, and was seen as confirmation of earlier work on cross-sectional comparisons by Rogan *et al* (1973), who showed lower FEV values in men with higher cumulative exposures to respirable dust and inferred that the differences were dust-related.

The reanalysis reported by Love and Miller (1982) was motivated by the the recognition that differences between FEV values from individuals examined cross-sectionally were less secure evidence of exposure-related changes than might be provided by examination of differences in longitudinally observed rates of change, but in that case the two analyses pointed to the same conclusion. In the present work, the cross-sectional analyses showing higher FEV in men with higher exposures appear implausible and contradictory as evidence of an effect of dust, but may simply be an aspect of differences within the population studied. In particular, they may point to selection effects within the population.

No population in any epidemiological study is entirely free from selection effects of one sort or another. In occupational epidemiology, apart from those explicitly introduced by the chosen study design, the selection effects may include the initial selection at recruitment of men fit enough to work in an industry, which may be physically demanding, as was traditionally the case in coal mining. Those who prove unequal to their tasks may be selected out of that industry, or to other less demanding tasks within it. In the coalmines, an additional factor was that, at least before the widespread introduction of mechanised coal-getting, the hardest physical tasks were those at the coal face, which were also associated with the highest concentrations of airborne respirable dust. It is thus easy to envisage situations where research could discern an association between high lung function and conditions of high exposure to dust. Situations may also be envisaged where these selection effects are extended over time by the transfer of men who show (possibly) exposure-related health effects to jobs where dust concentrations are lower, which would tend to reinforce such an association. A further factor in this particular colliery during the 1970s was the contraction in the workforce which preceded the closure in 1982, and again it is quite conceivable that the selection of men who would take redundancy or early retirement settlements would not have been random, and that the choices of individuals or of supervisors, or both, might have been influenced by the presence of respiratory problems.

Our ability to envisage such plausible mechanisms does not prove that they, or similar, were responsible for the findings from this study, nor that some quite different and hitherto unsuspected effect was not the cause. We believe, however, that the evidence of the longitudinal analyses points to a deleterious effect of dust exposure on lung function, and we note that the associations were stronger with the non-quartz fraction of the dust than with the quartz. We believe that these analyses have provided an interesting example of the possibility of drawing erroneous inferences about change from differences observed cross-sectionally.

Results from cross-sectional analyses of profusions of small opacities on the men's chest radiographs were more consistent with assessments of radiological changes between films taken at consecutive surveys, and viewed side-by-side. The IOM panel of readers are without medical qualifications, but have many years experience in the classification of abnormalities according to the ILO (1980) classification and its immediate predecessors. They classified the available films independently and in random order, and, with the exception of the films from the 1980 survey, on two separate occasions. Their assessments of profusion for each film were summarised by medians, which were analysed by converting them to binary variables and using logistic regression methods.

Exposures recorded for these men prior to the 3rd survey showed associations only with small opacities on the 4th survey films, and not at any of the later surveys. Profusions of small opacities at 4th and 5th surveys showed little evidence of association with the early exposures, but strong evidence of associations with exposures in ISP 3 and particularly in ISP 4; the association was stronger with the

variables estimating exposures to quartz than with those for the non-quartz fraction of the dust. The films taken in 1980 showed a similar pattern, although in this case the association was stronger with ISP 3 quartz than with ISP 4 quartz.
Similar results were obtained from the analysis of the side-by-side readings of series of films performed by two medical readers. The probability that the profusion of small opacities should have progressed by at least one step on the ILO (1980) 12-point scale between 4th and 5th surveys was strongly associated with the ISP 3 and 4 exposures, the strongest association being with ISP 3 quartz. Changes in ISP 5 were rather more strongly associated with exposures in ISPs 3, 4 and 5, the strongest being with ISP 4 quartz; ISP 5 was the only period over which changes of two or more steps could be shown to be associated with exposure variables. ISP 6 was employed as a label for the period between the 6th PFR and the 1980 Medical Service surveys, a period of about two years. Even over this short period, progression of one or more steps was shown to be associated with exposure variables, and most strongly with those in ISP 4.

These findings are, firstly, self-consistent. Despite the fact that the two reading exercises used different readers and entirely different protocols, they both pointed to an association between abnormalities observed on the films taken at 5th, 6th and 1980 surveys, and exposures to coalmine dust in the period between the 4th and 1980 surveys. Further, almost all of the strongest associations were with those variables representing exposure to the quartz fraction of the dust, rather than the non-quartz fraction. In addition, the assessments recorded independently by both medical readers, of change over ISP 5, that is between 5th and 6th surveys, included significant numbers of film series where the change was assessed at several, rather than one or two, steps on the ILO (1980) 12-point scale (see exhibit 7.3.2). In this assessment of unusual radiological changes in that particular period, and in the association of those changes with the quartz fraction of the men's exposures, these analyses are consistent with and reinforce the findings of the relatively small case control study from this colliery reported by Seaton *et al* (1981).

The present study has extended these findings; Seaton *et al* (1981) considered for their cases and controls only the film pairs spanning ISP 5. While our analyses have confirmed the element of unusual progression over that period, and the association with quartz, they have also shown similar associations for progression on a more limited scale observed over ISP 4 and ISP 6. While this may be seen as confirmation that the initial finding was not an artefact, it also suggests that the problem was not isolated in ISP 5, but was present if less obvious in the preceding and following periods. Again, this was consistent with the findings of a case-control study of radiological changes in less than 100 men between the 3rd and 5th PFR surveys (ISPs 3 and 4) at 10 collieries (of which colliery P was one) reported by Jacobsen and Maclaren (1982).

Reliable estimation of the quantitative aspects of these relationships may present more problems than simply demonstrating that they exist. In the first place, the cross-sectional analyses have depended upon converting the median profusion, by which the several assessments were summarised, into a simple binary variable, by dichotomising at a chosen point on the ILO (1980) scale. This common practice seems intuitively likely to sacrifice information. A limited exercise to compare the results for one section of the data with those obtained by applying alternative methods of greater complexity suggested that the binary-logistic analyses had less power than the others to detect associations; but there was no unanimity between the results, on the question of whether the process of dichotomy introduced bias in the estimation of the parameters of the association. The other methods produced estimates which were variously higher and lower than those from the binary-logistic approach.

There has been an upsurge recently in research interest in parametric models for the analysis of ordered categorical data (Andersen, 1980; McCullagh, 1980; Anderson, 1984). The topic is not new, and Ashford (1959) developed, in the context of radiological profusion, an ordered model which in many respects was the direct ancestor of McCullagh's (1980) model. Both earlier and more recent models required iterative calculations for their solution. There seems little doubt that the increased availability and hugely decreased cost of computer power, coupled with the important developments of the last two decades in the unification of families of statistical models, is leading to increases in the extent to which such models replace simpler, perhaps less realistic techniques in practical applications, or that the process is likely to continue. Further information on the comparability of results from simpler and more complex models would be welcome, and could obviously be obtained by computer simulations, but it has not been possible to perform such work within the life of this project.

Choice of an appropriate model is not the only factor in consideration of the quantification of risks from these data. Whichever model were adopted, it would still be necessary to consider how realistic were the measures of exposure as indicators of the effective dose to the individual, how biologically significant was the chosen response variable, and what was the temporal relationship between the accumulation of dose and the development of response. Data from radiological surveys carried out four years apart can show only the extent to which abnormalities have developed or progressed in an inter-survey period, but by their nature contain no information on short-term differences in rates of change in the periods before or between surveys. Even the most detailed measurements of respirable dust concentrations within working environments, coupled with the most painstaking recording of individual men's occupational activities, can stand only as a surrogate measure of the biological dose; differences in individuals' work patterns may be taken into account by adopting personal rather than static sampling strategies, but would not eliminate a number of imponderables, such as how much of the dust breathed is deposited in the lung, how much eliminated by mucociliary clearance, how much removed or deactivated by the lung's cellular defences, and other questions such as how long any noxious fraction remains active in the lung and how long a response takes to become manifest and visible on a radiograph.

There is, as always in epidemiology, the additional likelihood that any quantitative answers to these questions would hold only in the average, and that differences between individuals would be an important factor. Little direct evidence can ever be available on such questions, but animal experiments can yield some insights, even if inferences about the human context by extrapolation from animal experimentation must be treated with extreme caution.

Robertson *et al* (1984) reported on a series of experiments in which laboratory rats breathed airborne dusts sampled from two faces in a colliery adjacent to and working the same coal seam as in colliery P; the dusts had quartz contents 7% and 25% quartz, and separate groups of rats were exposed for 12 months to these dusts and to a third dust of quartz content 13% prepared by mixing the low quartz and high quartz dusts, all at the same airborne concentration. After dusting, or after a follow-up period, randomly chosen animals were sacrificed and their lung and associated lymph tissues were examined for pathological changes. Those exposed to the high quartz dust showed considerably higher profusions in the lungs of discrete cellular pigmented nodules, and produced nodules earlier. The low quartz group produced few pulmonary nodules. All exposed animals exhibited an extensive general reactive tissue response, as indicated by elevated lung tissue weights in comparison with those of control animals. All treatment groups displayed massive hypertrophy of the lymph nodes, but this was greatest in the animals exposed to the dust with high quartz. Analyses of weights and compositions of the dusts contained in the excised lungs and lymph nodes indicated that quartz was cleared from the lung after dusting ceased, and deposited in the lymph nodes, faster than the other components of the dust. However, the dusts retrieved from both lungs and lymph nodes contained higher proportions of quartz than had been measured in the dusts before the experiment started.

These findings were entirely consistent with the radiological observations which had been made at colliery P (Seaton *et al*, 1981), of greater profusions of radio-opaque pulmonary nodules in men with higher exposures to quartz in coalmine dust. The rat lung has been found to be a useful experimental model for dust-induced lung damage, but does not necessarily correspond exactly to the situation in the human lung; if the findings of Robertson *et al* (1984) are relevant to human subjects, they give some insight into the complexity of the mechanisms by which the lung tissues and their defence systems interreact with dusts of varying quartz contents, but they also confirm, by the considerable differences in responses exhibited by animals of the same strain exposed to the same conditions, that differences in human responses to the same environmental conditions are likely to be sizeable. Again, this is consistent with epidemiological evidence from men working side by side in coalmines, whose radiographs can show no evidence of abnormality in one man and progressive massive fibrosis in his neighbour.

Such individual differences in responses may have played a part in producing the observed relationships in time of the radiological abnormalities to the exposure histories, given that most of the analyses were based on relatively small numbers of positive responses. Selection effects such as those already mentioned in the context of lung function may also have had some influence, although the pathological changes detected as small opacities on radiographs may be less apparent to the individual concerned than might be the case with a loss of lung function. Nevertheless, coalminers are informed if their radiographs show signs of progression, and the effects of such information on decisions regarding, say, voluntary redundancy or early retirement in a colliery threatened by closure are open to speculation.

Population selection effects and individual differences in susceptibility may be part of the explanation for our inability to discern unambiguous patterns in the temporal relationships between the exposures and the radiological responses from surveys four years apart. Another factor likely to have had an effect is the well-recognised tendency of men in industries such as coal mining to retain similar jobs over periods of time, which induces at least some element of correlation between the exposures to which men are subject in different periods. This seems the likeliest explanation for the considerable reductions in the estimated regression coefficients of relationships with dust and quartz which were obtained from analyses in which the exposures from adjacent periods were added together. This phenomenon can be imagined at its most extreme in the hypothetical situation where each exposed man spends all his working life in the same exposure conditions, and where the risk of a particular response occurring is related to the total cumulative exposure up to the point when the response is measured. Then an estimate of the relationship of the response to precise measurements of the total exposure should give a representative picture of the underlying relationship; but a measurement of the exposure over a shorter period would be governed by the same information about the exposure conditions, and analysis with respect to such a measurement would produce a regression coefficient inflated by a factor which corresponded inversely to the ratio of the length of that period to the length of the man's working life.

The situation here, as in most real data sets, was nowhere near as simple as in this hypothetical case; exposure conditions changed over time, individuals joined and left employment, and there was a spread of ages and of lengths of exposures. Further, although examination of the exposure data showed a definite tendency for men with low exposures in one ISP to have low exposures in the next ISP also, the individual exposure estimates at adjacent ISPs were far from identical. For example, amongst the data used for the analyses of profusion at 5th survey, the raw linear correlation between ISP 3 and ISP 4 quartz exposure estimates was 0.63, and that between the corresponding dust exposures 0.68. These values are considerably less than the correlations between dust and quartz exposures in the same ISP, which were typically about 0.9, but they are certainly not negligible; similar values were obtained from the data analysed for other surveys.

However, with the exception of the analyses of progression in ISP 5, the analyses of data from both the panel's readings of profusion and the serial readings of progression showed associations individually with all the exposures after the 3rd PFR survey, but inclusion of more than one exposure variable in any model did not in general improve the fit. The introduction of variables which combined exposures over more than one ISP produced slight improvements in some analyses, but uniformly smaller coefficients for the exposure effects than were estimated for the single ISPs. It was therefore not possible from these analyses to draw unambiguous inferences about the relative importance of exposures in the different ISPs.

Radiological progression of one or more step in ISP 5 was exceptional in that it showed significant associations jointly and simultaneously with quartz exposures in ISP 3 and ISP 4, and the equivalent analysis of two or more steps had ISP 3 quartz significant and ISP 4 quartz just short of the 5% level when both were in the model. The data tabulations showed clearly that more progression was assessed by both medical readers for ISP 5 than for the other ISPs, and it is likely that the inability to show similarly detailed associations in the other ISPs is due to a combination of the lower numbers of men progressing and the element of correlation between the exposures at adjacent ISPs. Thus the analyses of progression at ISP 5 provide the strongest information regarding the magnitudes of the risks to which the men were exposed, while the results from the other ISPs can be taken as qualitative, if not quantitative, confirmation of the existence of a relationship between the development of radiological abnormalities and the dust, and in particular quartz, to which the men were exposed in the late 1960s and early 1970s. This conclusion was strengthened by the observation, in the analyses of profusion at 5th, 6th and 1980 survey, and of progression over ISP 4 and ISP 5, that when quartz and dust variables, from the ISP for which the association with exposure was strongest, were both in the regression model simultaneously, then the quartz made a significant contribution in the presence of dust, but dust made no significant contribution after adjustment for quartz.

Examination of the consistency of the magnitudes of estimates from these regression models with those from other PFR analyses is also not straightforward, because although the broad methodology was similar to that used in other studies, published work (e.g. Hurley *et al*, 1982) has often chosen different point on the 12-point scale at which to dichotomise the response, different sets of readers, and on occasions exposure variables on a logarithmic scale (which for the present data fitted less well than on the linear scale).

The cross-sectional analyses of lung function showed results which were ascribed to strong selection effects, but those which examined change were more easily related to published work. Love and Miller (1982) (L&M) reported on regression analyses of changes standardised to an eleven-year period in the five PFR collieries C, F, K, W and X. The smokers in L&M were losing FEV about 49ml in 11 years faster than the non-smokers, an excess in the rate of loss of about 4.5ml per year. In this study, the rates of loss for smokers were estimated as about 82ml in 4 years faster than the non-smokers, that is an excess of about 21ml per year, in ISP 4; for ISP 5, the excess was estimated at 28ml in 4 years, that is 7ml per year. These values were larger than, but of the same order of magnitude as, those in L&M. Estimates of coefficients of rate of loss on age and height from the present study were also about three to four times the size of those in L&M. Comparison of the age coefficients may not be entirely appropriate, since a linear model for age effects was replaced by a cubic. L&M's estimate of the effect of lifetime cumulative dust exposure on rate of loss was 0.033ml per $g.hr.m^{-3}$, whereas the estimate for ISP 4 or ISP 5 dust on loss of FEV in ISP 5, after adjustment for a cubic age effect, was about 1.7ml per $g.hr.m^{-3}$. This large discrepancy is probably partly due to the fact that L&M used lifetime cumulative exposures, while the present analyses were in terms of specific ISPs; the correlation between exposures in different ISPs would have a similar effect on the analysis of lung function as was observed in the analyses of radiological abnormalities.

Other factors may have contributed to these differences; none of the other collieries was Scottish, and regional differences in climate, environment and social habits may have played a part in addition to differences in working practices or composition of the coal mine dusts. An additional factor may have been differences in the extent of estimation error in the exposure estimates. One of the assumptions made in regression analysis is that the explanatory variables are known, that is measured without error, and it has long been known that if this requirement is not met the estimate produced is biased, towards zero if the variable involved is the only one in the regression model. Correlation between explanatory variables in models with more than one variable complicates matters considerably (Cochran, 1968).

Work by Heederik and Miller (in press; reproduced here as Appendix D) on a subset of the data from this project demonstrated how making adjustments for estimation error of different sizes could lead to the estimation of adjusted coefficients which were much larger than the unadjusted ones, and that such adjustments also had effects on the age estimates in the same model, even although age is known without estimation error. Further investigations would be required to assess the reliabilities of the components of our estimates of individuals' exposures, in order to fix a suitable level at which adjusted estimates could be produced. It may be noted, however, that the data studied by Love and Miller (1982) were from the 2nd to the 4th PFR surveys, and that a sizeable part of the estimation of exposures prior to this period depended on the mens' recall, and on assumptions about airborne concentrations for periods before the programme of detailed measurements began. The present data were from a later period, and most of the data contributing to the exposures of interest were derived from contemporary records of men's activities and occupational conditions. It is therefore reasonable to expect that the level of estimation error in the present data will have been less than in those used by Love and Miller (1982), with presumably a smaller effect on the coefficients. Much less is known about the effect of estimation errors in variables in non-linear regression models such as the logistic; statistical techniques to adjust for error do not seem to be readily available in this case.

Despite these difficulties in establishing the magnitudes of the effects involved, there seems little doubt of the main findings. The analyses both of profusion and progression of small opacities from the 5th PFR survey onwards showed associations with the dust to which men were exposed in a period which began after the 3rd PFR survey, and the associations appeared consistently stronger with estimates of the exposures expressed as quartz rather than as whole dust. Differences in lung function variables FEV and FVC did not appear to show plausible effects of dust, but selection effects were proposed as an explanation for the effects which were observed. Changes between surveys in these lung function variables did show an association between exposure to dust, rather than to quartz, and increased rate of loss of function. There was little evidence of a direct relationship between the quartz effect on small opacities and the dust effect on lung function.

Some suggestions for further work have arisen. It would be instructive to compare results of applying different statistical models for the analysis of ordered categorical data assessed by more than one judge, perhaps on a series of data sets simulated on a computer and with known distributional properties. Such a comparison would give guidance on the usefulness of the various models, which would assist in the choice of methods for analysing data from radiological assessments. In addition, further work on the influence of estimation error in the exposure variables on the estimated strength of the observed relationships could be informative, if based on realistic data-based assessments of the reliabilities of those estimates.

One important area in which further work is needed is in the reexamination of the men involved; these men, or at least some of them, were exposed to levels of quartz untypical of those covered by the rest of the PFR investigations. Although the evidence of an acute reaction to that exposure in these men is strong, little is known about whether the resulting radiological abnormalities might continue to develop, and whether their progress might depend on whether or not a man continued to work in conditions of dust exposure. Further exposure might add to the risks, but the mechanisms involved may be more complex. Robertson *et al*

(1984) discuss suggestions from Le Bouffant *et al* (1983), that other minerals such as illite may have a short-term protective effect against quartz-induced damage, which may vanish after exposure to these minerals ceases; the animal inhalation work of Robertson *et al* (1984) produced a small amount of data consistent with this hypothesis.

The data we have examined here were almost all from men who were still in employment in the coalmine, but only some of those transferred to other collieries when colliery P closed. It is hoped that, within the next few years, it will be possible to carry out a follow-up survey amongst the survivors from the population, which would allow examination of any progression of abnormalities which may have taken place since exposure ceased, in comparison with men who remained employed in coalmining. Such a follow-up could provide a perspective on the long-term risks of exposure to coalmine dust containing sizeable proportions of quartz which would complement well the information on short-term risk from this study and its immediate predecessors.

REFERENCES

ALVEY NG, BANFIELD CF, BAXTER RI, GOWER JC, KRZANOWSKI WJ, LANE PW, LEACH PW, NELDER JA, PAYNE RW, PHELPS KM, ROGERS CE, ROSS GJS, SIMPSON H, TODD AD, TUNNICLIFFE-WILSON G, WEDDERBURN RWM, WHITE RP, WILKINSON GN. (1983) Genstat: A general statistical program. Version 4.04 manual. Oxford: Numerical Algorithms Group.

ANDERSEN EB. (1980) Discrete statistical models with social science applications. Amsterdam: North-Holland.

ANDERSON JA. (1984) Regression and ordered categorical variables. Journal of the Royal Statistical Society Series B; 46: 1-30.

ASHFORD JR. (1959) A problem of subjective classification in industrial medicine. Applied Statistics; 8: 168-185.

BAKER RJ, NELDER JA. (1978) The GLIM system, release 3. Generalised Linear Interactive Modelling. Oxford: Numerical Algorithms Group.

BEBBINGTON AC. (1975) A simple method of drawing a sample without replacement. Applied Statistics; 24: 136.

BONNEY GE. (1987) Logistic regression for dependent binary observations. Biometrics; 43: 951-973.

COCHRAN WG. (1968) Errors of measurement in statistics. Technometrics; 10: 637-666.

COCHRAN WG, COX GM. (1957) Experimental designs. 2nd ed. New York: John Wiley.

COPLAND L, BURNS J, JACOBSEN M. (1981) Classification of chest radiographs for epidemiological purposes by people not experienced in the radiology of pneumoconiosis. British Journal of Industrial Medicine; 38: 254-261.

COTES JE. (1979) Lung function: assessment and application in medicine. 4th ed. Oxford: Blackwell Scientific Publications.

COX DR. (1970) The analysis of binary data. London: Methuen.

COX DR. (1972) The analysis of multivariate binary data. Applied Statistics; 21: 113-120.

DIXON WJ, BROWN MB, ENGELMAN L, FRANE JW, HILL MA, JENNRICH RI, TOPOREK JD, eds. (1983) BMDP statistical software. 1983 printing with additions. Berkeley (Calif): University of California Press.

DOBSON AJ. (1983) An introduction to statistical modelling. London: Chapman and Hall.

DODGSON J. (1963) Use of interference microscopy for the mineralogical analysis of samples of airborne dust obtained with the thermal precipitator. Nature; 199: 245-247.

DODGSON J, HADDEN GG, JONES CO, WALTON WH. (1971) Characteristics of the airborne dust in British coalmines. In: Walton WH, ed. Inhaled particles III. Proceedings of an International Symposium organized by the British Occupational Hygiene Society in London, 14-23 September 1970. Vol.2. Old Woking (Surrey): Unwin Bros: 757-782.

DODGSON J, WHITTAKER W. (1973) The determination of quartz in respirable dust samples by infrared spectrophotometry - I.The potassium bromide disc method. Annals of Occupational Hygiene; 16: 373-387.

DRAPER NR, SMITH H. (1981) Applied regression analysis. 2nd ed.. New York: John Wiley.

DUNMORE JH, HAMILTON RJ, SMITH DSG. (1964) An instrument for the sampling of respirable dust for subsequent gravimetric assessment. Journal of Scientific Instruments; 41: 669-672.

FAY JWJ. (1957) The National Coal Board's Pneumoconiosis Field Research. Nature; 180: 309-311.

FAY JWJ, ASHFORD JR. (1960) The study of observer variation in the radiological classification of pneumoconiosis. British Journal of Industrial Medicine; 17: 279-292.

FLETCHER C, PETO R, TINKER C, SPEIZER FE. (1976) The natural history of chronic bronchitis and emphysema. An eight-year study of early chronic obstructive lung disease in working men in London. Oxford: Oxford University Press.

GOUGH J. (1940) Pneumoconiosis in coal trimmers. Journal of Pathology and Bacteriology; 51: 277-285.

GREEN PJ. (1984) Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives. Journal of the Royal Statistical Society Series B; 46: 149-192.

HEEDERIK D, MILLER BG. (in press) Weak associations in occupational epidemiology: adjustment for exposure estimation error.

HURLEY JF, BURNS J, COPLAND L, DODGSON J, JACOBSEN M. (1982) Coalworkers' simple pneumoconiosis and exposure to dust at 10 British coalmines. British Journal of Industrial Medicine; 39: 120-127.

HURLEY JF, COPLAND L, DODGSON J, JACOBSEN M. (1979) Simple pneumoconiosis and exposure to respirable dust: relationships from twenty-five years' research at ten British coalmines. Edinburgh: Institute of Occupational Medicine. (IOM Report TM/79/13).

ILO. (1959) Meeting of experts on the International Classification of Radiographs of the Pneumoconioses, Geneva, 27 October to 7 November 1958. Occupational Safety and Health; 9(2): 63-69.

ILO. (1972) ILO U/C international classification of radiographs of pneumoconioses. Geneva: International Labour Office, 1972. (ILO Occupational Safety and Health series 22 (rev 1971)).

ILO. (1980) Guidelines for the use of ILO International Classification of Radiographs of Pneumoconiosis. Geneva: ILO. (ILO Occupational Safety and Health Series No.22 (rev.80)).

JACOBSEN M. (1975) Quantifying radiological changes in simple pneumoconiosis. Applied Statistics; 24: 229-249.

JACOBSEN M, BURNS J, ATTFIELD MD. (1977) Smoking and coalminers' simple pneumoconiosis. In: Walton WH, ed. Inhaled particles IV. Proceedings of an International Symposium organized by The British Occupational Hygiene Society, Edinburgh, 22-26 September 1975. Vol.2. Oxford: Pergamon Press: 759-772.

JACOBSEN M, MACLAREN WM. (1982) Unusual observations and exposure to coalmine dust: A case-control study. In: Walton WH, ed. Inhaled particles V. Proceedings of an international symposium organized by the British Occupational Hygiene Society. Cardiff, 8-12 September 1980. Oxford: Pergamon Press: 753-765 (Annals of Ocupational Hygiene; 26).

JACOBSEN M, MILLER BG, MURDOCH RM. (1984) Experience with the ILO (1980) Classification in an epidemiological study of asbestos workers' chest radiographs. In: Bergbau-Berufsgenossenschaft, eds. 6th International Pneumoconiosis Conference 1983, Bochum. Vol.2. Geneva: International Labour Organization: 868-880.

JACOBSEN M, RAE S, WALTON WH, ROGAN JM. (1971) The relation between pneumoconiosis and dust exposure in British coal mines. In: Walton WH, ed. Inhaled particles III. Proceedings of an International Symposium organized by the British Occupational Hygiene Society in London, 14-23 September 1970. Vol.2. Old Woking (Surrey): Unwin Bros: 903-919.

LE BOUFFANT L, DANIEL H, MARTIN JC, AUBIN C, LEHEUEDE P. (1983) Recherche communautaire sur le role du quartz dans la pneumoconiose des mineurs de charbon et sur l'influence des mineraux d'accompagnement. Verneuil-en-Halatte: CERCHAR.

LIDDELL FDK. (1974) Assessment of radiological progression of simple pneumoconiosis in individual miners. British Journal of Industrial Medicine; 31: 185–195.

LOUW SJ, COWIE HA, SEATON A. (1986) Epidemiologic studies of Scottish oil shale workers: II. Lung function in shale workers' pneumoconiosis. American Journal of Industrial Medicine; 9: 423–432.

LOVE RG, MILLER BG. (1982) Longitudinal study of lung function in coal-miners. Thorax; 37: 193–197.

MACLAREN WM. (1985) Using discriminant analysis to predict attacks of complicated pneumoconiosis in coalworkers. Statistician; 34: 197–208.

MACLAREN WM, SOUTAR CA. (1985) Progressive massive fibrosis and simple pneumoconiosis in ex-miners. British Journal of Industrial Medicine ; 42: 734–740.

MCCULLAGH P. (1980) Regression models for ordinal data. Journal of the Royal Statistical Society B; 42: 109–142.

MCCULLAGH P. (1983) Quasi-likelihood functions. Annals of Statistics; 11: 59–67.

MCCULLAGH P, NELDER JA. (1983) Generalized linear models. London: Chapman and Hall. (Monographs on Statistics and Applied Probability).

MEDICAL RESEARCH COUNCIL. (1942) Chronic pulmonary disease in South Wales coalminers. 1. Medical studies. London: HM Stationery Office. (MRC Special Report Series no.243).

MEDICAL RESEARCH COUNCIL. (1943) Chronic pulmonary disease in South Wales coalminers. II. Environmental studies. London: HM Stationery Office. (MRC Special Report Series No. 244).

MILLER BG, JACOBSEN M. (1985) Dust exposure, pneumoconiosis and coalminers' mortality. British Journal of Industrial Medicine; 42: 723–733.

NATIONAL COAL BOARD. (1949) The sampling of airborne dust for the testing of 'Approved Conditions'. London: NCB.

NATIONAL COAL BOARD. (1969) Approved conditions for airborne dust. London: NCB. (Standards and Procedures for Sampling, F 4040).

NELDER JA, WEDDERBURN RWM. (1972) Generalized linear models. Journal of the Royal Statistical Society A; 135: 370-384.

PETERS WL, REGER RB, MORGAN WKC. (1973) The radiographic categorization of coal workers' pneumoconiosis by lay readers. Environmental Research; 6: 60-67.

RAE S, WALKER DD, ATTFIELD M. (1971) Chronic bronchitis and dust exposure in British coal miners. In: Walton WH, ed. Inhaled particles III. Proceedings of an International Symposium organized by the British Occupational Hygiene Society in London, 14-23 September 1970. Vol.2. Old Woking (Surrey): Unwin Bros: 883-894.

ROBERTSON A, BOLTON RE, CHAPMAN JS, DAVIS JMG, DODGSON J, GORMLEY IP, JONES AD, MILLER BG. (1984) Animal inhalation experiments to investigate the significance of high and low percentage concentrations of quartz in coalmine dusts in relation to epidemiology and other biological tests. Edinburgh: Institute of Occupational Medicine. (IOM Report TM/84/5).

ROBINSON BN, ANDERSON GD, COHEN E, GADZIK WF, KARPEL LC, MILLER AH, STEIN JR. (1980) SIR scientific information retrieval. Users' manual version 2. Evanston (Ill): SIR Inc.

ROGAN JM, ATTFIELD MD, JACOBSEN M, RAE S, WALKER DD, WALTON WH. (1973) Role of dust in the working environment in development of chronic bronchitis in British coal miners. British Journal of Industrial Medicine; 30: 217-226.

SEATON A. (1983) Coal and the lung. Editorial. Thorax; 38: 241-243 (Correspondence: 1983; 38: 877-879 and 1984; 39: 397-398).

SEATON A, DICK JA, DODGSON J, JACOBSEN M. (1981) Quartz and pneumoconiosis in coal miners. Lancet; ii: 1272-1275 (Correspondence: 1982; i: 45-46).

SEATON A, LOUW SJ, COWIE HA. (1986) Epidemiologic studies of Scottish oil shale workers: I. Prevalence of skin disease and pneumoconiosis. American Journal of Industrial Medicine; 9: 409-421.

SOUTAR CA, GAULD S, LLOYD M, COPLAND LH, HURLEY JF. (1981) Epidemiological and clinical studies of polyvinylchloride workers. Edinburgh: Institute of Occupational Medicine. (IOM Report TM/81/08).

SOUTAR CA, MACLAREN WM, ANNIS R, MELVILLE A. (1986) Quantitative relations between exposure to respirable coalmine dust and coalworkers' simple pneumoconiosis in men who have worked as miners but have left the coal industry. British Journal of Industrial Medicine; 43: 29-36.

WATSON HH. (1936) The thermal precipitator. Transactions of the Institute of Mining and Metallurgy; 46: 155-240.

WILLIAMS DA. (1982) Extra-binominal variation in logistic linear models. Applied Statistics; 31: 144-148.

APPENDIX A Protocol for film readings by IOM panel

The procedures used by the IOM film panel for the assessment of chest radiographs, and for the recording of the assessments on locally designed forms, are standardised. Below are reproduced the standing instructions which describe the standard procedures, and which are available to the readers at all reading sessions for their assistance and guidance.

Completion of IOM forms for data from film readings

This note is intended to spell out the conventions to be used in the regular recording of data from the IOM's panel of film readers. It is hoped that this will help to ensure that all readers record their results in the same way, and to avoid any ambiguities.

By way of illustration, a copy of a form is attached with some fictitious data which are intended to show some combinations of codes which follow the rules below. You may notice that the forms have been revised slightly; in particular, BATCH NUMBER and DATE only appear on page 1 of a batch of forms, and CLERK has been omitted entirely. Forms have already been printed in the old style for the next few weeks' readings, but the principles of recording the data have not been changed, so it will only be necessary for you to ignore the irrelevant items in the headings.

At the start of a session

A set of forms is preprinted for each batch of films, for each reader, and will be supplied with the batch in the reading room. The heading on page 1 of a set identifies the exercise of which the batch is a part, which batch it is, and which reader the set was printed for. (The item [BATCH NUMBER] helps us to keep a track of what batches everyone has read, but is not informative unless you have the master diary.) Please check that the forms bear the correct reader id as per the list below, and enter the date.

Reader	Id
Mrs Scott	101
Mrs White	104
Mrs Duncan	105
Mrs Henderson	106
Miss Jeffreys	112
Mr Wilson	114

Check the identity of each film

The forms have the identities of the films preprinted in the order in which they appear in the batch. The identity is typically an alphabetic code for the source of the film followed by a numeric serial number assigned to a man, with the year the film was taken joined on by a hyphen. This is followed by a sequence number which runs from 001 up to the batch size, and thus gives the position of the film in the batch.

When you lift a film to the light box to read it, you should always first verify that you are looking at the correct film. It is easy to get out of step; and sometimes a film which was intended for reading goes missing and cannot be found at batching time, or is removed for some urgent clinical purpose after the batch is made up. There is plenty of room around the film id to make a tick or similar mark, or to mark films which were missing, if you find that this helps.

The identifying information on the film will not always be identical to that in the FILM ID column, but it should be sufficient to satisfy you that it is the correct film. The numeric part of the id and the year the film was taken will be present, and in the case of PFR films, the letter which follows PFR on the forms will also be present.

In what follows I will refer to the areas where recording takes place on the forms as fields; I have numbered them on the example form attached, for ease of reference here, and numbers in round brackets e.g. (3) will refer to that numbering.

All references to Left and Right will be with reference to the subject of the film. With a P/A film, the subject's right side is on the reader's left, and the layout of the fields on the form reflects this.

Film Quality (1),(2)

Record the quality of the film as 1,2,3 or 4 at (1). This assessment should always be present unless the film is missing from the batch, in which case it (and every other field) should be left blank.

If the quality is scored as other than 1, i.e. less than perfect, the defects should be recorded at (2), using the IOM's standard in-house codes. Up to three defects can be recorded in this way. If the quality is as bad as 4, i.e. unreadable, then fields (3) to (14) will be left blank , but a comment would still be permissible at (15) (see below).

CPA Off (3)

Sometimes a film has been taken in such a position that one or both of the costophrenic angles is below the bottom of the film and therefore not visible. This is really another aspect of the quality of the film. Such an occurrence is recorded at (3) using one of the symbols R, L, B for Right, Left or Both. If both angles are visible leave this field blank.

Film Normal (4)

By convention, if you observe no abnormalities at all on a film, we allow you to record this with an X in field (4). In this case fields (5) to (14) will be left blank, although a comment would still be permissible at (15) (see below). It is important to note that if any abnormality covered by the Classification is observed, then it is not appropriate to use this field.

Small opacities (5), (6), (7)

This assessment is compulsory unless the film is unreadable. It may only be omitted where an X has been entered at field (4), which will be taken to imply a small opacities profusion score of 0/0.

Enter the profusion score assigned after comparison with the ILO standard films at field (5), using one of the 12 possible codes 0/-, 0/0, 0/1, 1/0, 1/1, 1/2, 2/1, 2/2, 2/3, 3/2, 3/3, 3/+. If the profusion is scored 0/- or 0/0 then fields (6) and (7) will be left blank, otherwise they must be completed.

When a score above 0/0 is assigned, mark the zones involved in diagrammatic form at (6), using x's. If all zones are involved a convenient short cut is to put a large X through the whole box at (6), but in other cases the zones must be marked individually. Enter the alphabetic codes for the predominant and secondary size and shape at (7). Since each can be one of the six letters P, Q, R, S, T, U, there are 36 possible 2-letter codes.

Large Opacities (8)

This assessment is also compulsory unless the film is unreadable. An X entered at (4) will be taken to imply a large opacities profusion of 0. Otherwise, enter the score for the profusion at (8) according to the definition in the Classification, using one of the codes 0, A, B, C.

Small and Large Opacities are the only compulsory assessments of abnormality; the remainder are all optional.

Pleural Thickening : Plaques (9)

Assess and record each lung separately. If there are no plaques in a lung then leave the field blank in the recording positions for that lung. If plaques are observed in a lung, their total extent must be recorded as 1, 2 or 3 as defined in the classification. The section marked FACE is used to indicate whether any of the plaques in that lung is seen wholly or partially face-on; if any are face-on, then a Y is entered. When all plaques are seen side-on, an N is entered. The maximum width is coded A, B or C as defined in the Classification; it is noted there that if all of the plaques can be seen only face-on, it may not be possible to measure a width, and in this case the section marked WID may be left blank.

Pleural Thickening : Diffuse (10)

Diffuse pleural thickening is assessed for extent, whether or not face-on, and width in exactly the same way as for pleural plaques; the rules for recording are identical to those for (9).

Pleural Thickening : Diaphragm (11)

The occurrence of thickening of the diaphragm is scored and recorded R,L or B, according to whether it is observed on Right, Left or Both lungs.

Costophrenic Angle Obliteration (12)

If obliteration of the costophrenic angle can be seen on Right, Left or Both lungs, record this observation by R,L or B; otherwise leave blank. It is possible that either or both angle may be off the film; that is recorded at (3), but obviously limits the site at which obliteration might be observed.

Pleural Calcification (13)

Pleural calcification is assessed independently for each lung. Record with an X the site(s) where it is observed (chest wall, diaphragm, other) and record the total extent over the whole lung as 1,2 or 3 as defined in the Classification.

Other Abnormalities (14)

Record the presence of any other abnormalities observed using the 2-letter codes in the standard list. You may record up to 6 abnormalities in this way.

Comments (15)

A Comments book will be kept in the reading room for noting features not covered by the Classification. If you write a comment about a film, put an X in this field. It is always possible to comment, whatever other results have been recorded. (This is used mainly by medically qualified readers when they observe some clinically significant feature which is unrelated to the pneumoconioses.)

Brian G Miller
July 9, 1986

INSTITUTE OF OCCUPATIONAL MEDICINE - I.O.M. - FILM READING PANEL PAGE: 1

READING ID: XRTY BATCH ID: 02 READER ID: 104 [BATCH NUMBER: 1396] DATE: 25/12/85

FILM ID	SEQ	SMALL OPS: QUAL 12/34	QUAL 1-/9	CPA OFF RLB	N X	PROF 012/3-+	ZONES X RL	TYPE P.Q.R/S.T.U	LGE OPS O.A/B.C	PLEURAL THICKENING: PLAQUES FACE Y,N RL	PLAQUES WID ABC RL	PLAQUES EXT 123 RL	DIFFUSE FACE Y,N RL	DIFFUSE WID ABC RL	DIFFUSE EXT 123 RL	DIA RLB	CPA RLB	PLEURAL CALC: CHEST WALL X RL	DIA X RL	OTHER X RL	EXT 123 RL	OTHER ABN	C X
PFRT1207-66	001	(1)	(2)	(3)	(4)	(5) /	(6)	(7) /	(8)		(9)			(10)		(11)	(12)		(13)			(14)	(15)
PFRW0015-66	002					/		/															
PFRP0607-64	003	1				1/0	x x x	P/R	0														
PFRY0184-65	004	2	4			2/3	x x x x x x x	R/R	A				N	A	2	R	R			X	1	ID	
PFRC0250-65	005	4	2 5			/		/															
PFRF2366-67	006	1			X	/		/															X
PFRC1778-65	007	1				2/1	X	R/Q	B	Y Y	A	2 2						X	X		1 2		
PFRW0718-66	008					/		/															
PFRQ1256-65	009					/		/															
PFRF0484-67	010					/		/															

Artificial examples — NOT real data

APPENDIX B Protocol for medical readers' film readings

The medical readers who assessed the series of radiographs used a shortened form of the ILO (1980) classification, with which they assessed small and large parenchymal opacities, but not pleural or other abnormalities. The form on which they recorded their assessments was IOM-designed. Reproduced below are the detailed instructions for this assessment, and for completion of the forms, with an explanatory covering note.

FILE NOTE

Project CEC025 - Serial film reading exercise

This note details some aspects of the serial reading trial and has been compiled for information and background. Readers are requested to read carefully the attached protocol for the exercise before reading commences, and to address any queries to Brian Miller.

The principal aim of this project is the investigation of the relationships between pneumoconiotic signs on radiographs taken from men working at colliery P between 1970 and 1980, and measures of the men's exposures to airborne dust and its constituents. Particular interest lies in the role of the quartz fraction of the dusts, as measured from the samples taken to assess concentration.

The study population is 1416 men who were all recorded as having attended at least one of the 1970, '74, and '78 PFR surveys. All available films for these men, including any taken at the PXR survey in 1980, have now been read, independently and randomised into batches, by the IOM's film-reading panel of five readers.

The current reading is intended to be the last for this project, and aims to gain a different perspective on progression of pneumoconiotic signs over the period from that which will emerge on analysis of the independent readings. Three medically qualified readers experienced in the classification of radiographs have agreed to read, for each individual, the series of films available. It is hoped that simultaneous viewing of up to four films per man will reduce the effects of within-reader variation on the assessment of progression of pneumoconiosis.

Because we already have readings by the panel of all the available films (duplicate readings with the systematic exception of the 1980 PXR films), it has been decided to limit the reading burden for the present exercise by concentrating on those combinations of films most likely to yield useful data on progression. Firstly, all men for whom only one film was available have been omitted. Secondly, amongst those men for whom only two films were available, those with films only at the 1970 and '74 surveys have also been omitted.

The sub-group of 632 men whose films are to be read is thus defined as:

a) all men with 3 or 4 films
b) all men with 2 films at least one of which was taken after 1974.

(Throughout the above, availability of a film has been defined in terms of our ability actually to find the film, rather than depending on the information on computer records of what is supposed to exist.)

The current reading will include 2035 films from these 632 men (with the possible exception of a few films which cannot be found at time of batching, as often happens). These have been allocated to 18 batches of about 113 films each. Men with 4, 3 and 2 films have been allocated to batches in rough proportion to the sample, and each batch contains films for about 35 men, allocated to positions within the batch by a randomisation procedure. The films for each man have been placed together in a temporary film envelope, and any earlier films for these individuals will not be present during the reading exercise.

Each reader is being asked to assess eight of these batches, four of which will be read only by him; the other two readers will also see two of the remaining batches each. Thus some information will be available on reader differences in assessing prevalence and progression.

Brian G Miller

PROJECT CEC025 - PROTOCOL FOR SERIAL FILM READING

This exercise has been allocated the identification code XRES and the 18 batches within it are numbered 01-18.

Each batch will consist of temporary film envelopes for around 35 men, each envelope containing from 2 to 4 films for the same man. A batch will be accompanied by a set of recording forms with the batch number, the reader's identification code and the identities of the films preprinted on them, similar to the attached sample. The order of the envelopes will correspond to the order printed on the forms.

Since some of the batches are destined to be seen by another reader, it would be appreciated if the order of the envelopes within the batch could be preserved during and after the reading.

The films for each man are listed together on the form, in ascending time order. The form itself is an adaptation of the recording form used routinely by the IOM's own reading panel, retaining only the elements relating to film quality, small and large opacities, and comments. It is not intended for this exercise to record codes for pleural or other abnormalities. However, features which the reader considers particularly important or relevant could be recorded as comments.

On starting to read a batch, please check that the identifying information on the top recording form refers to the correct batch number and reader code. The codes in use are:

010	Dr Pern
011	Dr Soutar
013	Dr Seaton

Please enter date of starting to read the batch in the appropriate space at the top of the form. (This date is used for reference, even if it is not possible to complete the reading of the batch within one day.)

Each envelope contains the relevant films for one man, and the identities of the films are printed on the form as a group, seperated from another man's by a blank line. The film identity consists of a location code (here always PFRP), an X-ray number, then a hyphen followed by the approximate year of the survey. (The year is approximate in a few cases where a replacement film was taken a little while after the survey.) The sequence number is the position of the film within the batch and is used for locating a film on the form when the punched data are being validated.

The reader should satisfy himself that the film being read bears the correct PFR identity, referring any queries to John

Allan, Mrs Murray or Brian Miller. Please note that 1980 PXR films will have been relabelled with the PFR X-ray number in chinagraph pencil.

Readers are asked to assess each film in the context of other films for the same man, so the use of a four-film box is essential. The illumination in the Institute's boxes has recently been standardised.

All classifications are to be made according to the IOM usage of the 1980 ILO classification scheme, shortened here to exclude pleural and other non-parenchymal abnormalities. The codes to be used to record results are detailed below.

FILM QUALITY: Film quality is recorded in the first column as 1,2,3 or 4. The codes are 1-Good, 2-Acceptable, 3-Poor, 4-Unacceptable. If the film quality is recorded as other than 1, the defects should be recorded in the next column. If the quality is recorded as 4, no assessment of small or large opacities will be made.

FILM QUALITY DEFECTS: Defects requiring a film to be scored other than 1 are recorded by the following (IOM-specific) codes:

1 Dark film
2 Pale film
3 Too much contrast
4 Lack of contrast
5 Patient movement
6 Poor inspiration
7 Marks on film
8 Ill-defined lung markings
9 Other

Up to three defects may be recorded in this field. If any of these is 9 (Other), the defect should be noted as a comment.

(A further aspect of film quality normally recorded is the inclusion of the costophrenic angles within the frame of the film. Since obliteration of the angles is not being assessed in this reading, this aspect of quality is not relevant, and has been omitted.)

FILM NORMAL: By local convention, if no abnormalities are observed on a film, the third column is marked with an X. This saves the reader time in recording, and is treated as if the reader had recorded 00 for small opacities and 0 for large opacities. If either small or large opacities are being recorded as present, this field should be left blank.

SMALL OPACITIES: This asssessment is compulsory unless film quality is recorded as 4, but may be omitted if the third column (FILM NORMAL) has been marked X, when a score of 00 will be implied.

Enter the profusion score assigned after comparison with the ILO standard films, using one of the 12 possible codes 0/-, 0/0, 0/1, 1/0, 1/1, 1/2, 2/1, 2/2, 2/3, 3/2, 3/3, 3/+.

ZONES: When a score for profusion above 0/0 is assigned, mark the zones involved in diagrammatic form, using x's. If all zones are involved it is convenient to put a large X through the whole box; otherwise the zones should be marked individually.

TYPE: Enter the alphabetic codes for the predominant and secondary shape and size. Since each can be one of the six letters P,Q,R,S,T,U, there are 36 possible 2-letter codes.

LARGE OPACITIES: This assessment is also compulsory unless the film is unreadable, but may be omitted if the third column (FILM NORMAL) has been marked X, when a score of 0 will be implied. Otherwise, enter the score for the profusion according to the definition in the Classification, using one of the codes 0,A,B,C.

COMMENTS: If you wish to make a comment about the film, other than about some aspect of technical quality, mark an X in the last column. It is always possible to comment, whatever other results have been recorded.

Comments on technical quality or other aspects should be written on the form, in the blank space to the right of the recording grid, in line with the film concerned.

The attached sample form has been completed with some wholly fictitious data designed to give examples of the use of the above codes, and the valid combinations in which they can coexist.

On completing a batch, please return the batch and the completed forms to the member of staff who supplied them.

Brian G Miller

I. O. M SERIAL FILM READING EXERCISE PAGE: 1

READING ID: XRES BATCH ID: 00

READER ID: 000 DATE: 25/12/86

FILM ID	SEQ	SMALL OPS: QUAL 12 34	QUAL 1- 9	N X	PROF 012 3-+	ZONES X R L	TYPE P. Q. R S. T. U	LGE OPS O. A B. C	C X	
PFRP3660-70	001	2	1	X	/		/			
PFRP3660-74	002	1		X	/		/			
PFRP3660-78	003	2	2		1/1	X X X	P/Q	0		
PFRP3660-80	004	1			2/1	X	Q/R	A		
-------------					/		/			
PFRP0887-70	005	1			0/0		/	A		
PFRP0887-74	006	2	1 9		0/0		/	A		Film upside down
PFRP0887-78	007	4	1 5 6		/		/			
PFRP0887-80	008	2	1		1/0	X X X	P/Q	A	X	Rheumatoid lung
-------------					/		/			

Artificial data

Reproduced below are the SIR (Robinson *et al*, 1980) schemas which defined the structure of the project database and of the separate types of data records which were designed to hold the different types of data in a logically structured manner.

Schema page 1

```
RUN NAME          FILE LVREAD   CODEBOOK DEFINITION
TASK NAME         INITIALIZATION COMMANDS
NEW FILE          LVREAD
TASK NAME         CASE DEFINITION
N OF CASES        1600
RECS PER CASE     1166
MAX INPUT COLS    104
RECTYPE COLS      50
MAX REC TYPES     30
MAX REC COUNT     1193
CASE ID           CASEID
COMMON VARS       CASEID   (A,8)    /   YEAR4    (I,2)    /
                  YEAR5    (I,2)    /   YEAR6    (I,2)    /
                  YEARPXR  (I,2)    /   DOBIRTH  (DATE)   /
COMMON SECURITY 20  20
TEMP VARS         BLDUM      EXERCISE  LOCATION  TEST
DOCUMENT
                           This database is set up to store film
                           reading data from several exercises connected with
                           project CEC025, reexamining radiographs from colliery
                           P in the light of detailed exposures to respirable
                           dust and quartz. It is intended that the development
                           of this database will be useful in setting up further
                           databases for film reading.

                           There are three distinct types of cases in the database,
                           but the case id always has the same format. It is an
                           A8 variable, of the form AAAAnnnn where the A's stand
                           for (uppercase) letters and the n's for digits.

                           The first type of case is a man; there are 1416 in the
                           study, and have ids PFRPnnnn where the digits are the
                           man's usual PFR xray no.

                           The second type of id indexes the various batches set
                           up in the different film-reading exercises. In these,
                           the digits represent the batch number within the
                           exercise, and the letters are the batch identifier
                           (the old card-class is the forerunner of this). Thus
                           batch 1 in exercise XRLV is case XRLV0001.

                           The last case type corresponds to occupational groups
                           for which environmental data have been collected.
                           These are of the form OGSNnnnn.

                           The different record types have one or the other case
                           types, according to context, and are described below.
```

```
TASK NAME         RECORD 1 (MANLIST )  SCHEMA DEFINITION
RECORD SCHEMA     1  MANLIST
DOCUMENT
                         This record type is used to set up the sample
                         population for the study and define the set of case
                         ids for the individuals.

                         Data on the availability of each individual's films
                         are entered and then stored in the CIR. Once this has
                         been done, the actual records of type 1 will be
                         deleted from the database, since the CIR will contain
                         the necessary information.

                         It is anticipated that further information on the
                         current availability of films will come to light
                         during the process of extracting and batching the
                         films. This information will be used to update the
                         CIR, using either SIR/FORMS or RETRIEVAL UPDATE.

SEQUENCE CHECK    OFF
MAX REC COUNT     1
DATA LIST         FIXED (1)
                  /1          CASEID          1 -    8     (A)
                  /1          YEAR4          10 -   11     (I)
                  /1          YEAR5          13 -   14     (I)
                  /1          YEAR6          16 -   17     (I)
                  /1          YEARPXR        19 -   20     (I)
                  /1          LOCATION        1 -    4     (A)
CAT VARS          LOCATION    ( 'PFRP'   )/
VAR RANGES        YEAR4       (70 70)/
                  YEAR5       (74 75)/
                  YEAR6       (78 79)/
                  YEARPXR     (80 80)/
VAR SECURITY      CASEID      (20,20 ) / YEAR4       (20,20 ) /
                  YEAR5       (20,20 ) / YEAR6       (20,20 ) /
                  YEARPXR     (20,20 ) /
MISSING VALUES    YEAR4       ( BLANK   )/
                  YEAR5       ( BLANK   )/
                  YEAR6       ( BLANK   )/
                  YEARPXR     ( BLANK   )/
VALID VALUES      YEAR4       ( 70   )/
                  YEAR5       ( 74,75    )/
                  YEAR6       ( 78,79    )/
                  YEARPXR     ( 80   )/
VALUE LABELS      YEAR4       (70)'1970' /
                  YEAR5       (74)'1974',(75)'1975' /
                  YEAR6       (78)'1978',(79)'1979' /
                  YEARPXR     (80)'1980' /
                  LOCATION    ('PFRP')'PFRP' /
VAR LABELS        CASEID        CASE IDENTIFIER/
                  YEAR4         Year of film from 4th survey/
                  YEAR5         Year of film from 5th survey/
                  YEAR6         Year of film from 6th survey/
                  YEARPXR       Year of film from PXR survey/
END SCHEMA
```

```
TASK NAME         RECORD 2 (BATCHES )  SCHEMA DEFINITION
RECORD SCHEMA     2  BATCHES
DOCUMENT
                         This record type stores the batch lists for the
                         readings. The case id is that of the batch. The sort
                         order for the data is in ascending sequence within the
                         batch, i.e. the order in which the films appear in the
                         batch. The data items of man id and year of film allow
                         networking of man and their films to the reading order.

SORT IDS          SEQNO (A)
SEQUENCE CHECK    OFF
MAX REC COUNT     205
DATA LIST         FIXED (1)
                  /1          CASEID           16 -    23    (A)
                  /1          SEQNO            13 -    15    (I)
                  /1          MANID             1 -     8    (A)
                  /1          YEAR             10 -    11    (I)
                  /1          EXERCISE         16 -    19    (A)
                  /1          LOCATION          1 -     4    (A)
CAT VARS          EXERCISE   ( 'XRLV','XRVL','XREX','XRES'   )/
                  LOCATION   ( 'PFRP'   )/
VAR RANGES        YEAR       (70 80)/
VAR SECURITY      CASEID    (20,20 ) /
VALID VALUES      YEAR       ( 70,74,75,78,79,80   )/
VALUE LABELS      YEAR       (70)'1970',(74)'1974',(75)'1975',(78)'1978',
                             (79)'1979',(80)'1980' /
                  EXERCISE   ('XRLV')'XRLV',('XRVL')'XRVL',('XREX')'XREX',
                             ('XRES')'XRES' /
                  LOCATION   ('PFRP')'PFRP' /
VAR LABELS        CASEID       CASE IDENTIFIER/
                  SEQNO        Position of film in batch/
                  MANID        PFR identity/
                  YEAR         Year film taken/
END SCHEMA
TASK NAME         RECORD 3 (FILMS    )  SCHEMA DEFINITION
RECORD SCHEMA     3  FILMS
DOCUMENT
                         This record type stores data on the batches in which a
                         man's films are placed. Together with record type 2
                         (BATCHES), it defines a networking of men's films with
                         positions in batches. The case id is that of the man,
                         and the sort order is of batch id within increasing
                         film year. The data items are the positions in which
                         films for that man appear. Although an input data
                         list is given, it is straightforward and possibly more
                         efficient to create these records by running the
                         retrieval stored in the data base as SETUP3:T. This
                         also means that the duplicate information stored in
                         these records doess not have to be present unless it
                         is needed, after which it can be deleted by the
                         retrieval stored as DELETE3:T.

SORT IDS          YEAR (A)        BATCHID (A)
SEQUENCE CHECK    OFF
MAX REC COUNT     20
DATA LIST         FIXED (1)
                  /1          BATCHID          16 -    23    (A)
                  /1          SEQNO            13 -    15    (I)
                  /1          CASEID            1 -     8    (A)
                  /1          YEAR             10 -    11    (I)
                  /1          EXERCISE         16 -    19    (A)
                  /1          LOCATION          1 -     4    (A)
CAT VARS          EXERCISE   ( 'XRLV','XRVL','XREX','XRES'   )/
                  LOCATION   ( 'PFRP'   )/
VAR RANGES        YEAR       (70 80)/
VAR SECURITY      CASEID    (20,20 ) /
VALID VALUES      YEAR       ( 70,74,75,78,79,80   )/
VALUE LABELS      YEAR       (70)'1970',(74)'1974',(75)'1975',(78)'1978',
                             (79)'1979',(80)'1980' /
                  EXERCISE   ('XRLV')'XRLV',('XRVL')'XRVL',('XREX')'XREX',
                             ('XRES')'XRES' /
                  LOCATION   ('PFRP')'PFRP' /
VAR LABELS        BATCHID      Batch identity/
                  SEQNO        Position of film in batch/
                  CASEID       CASE IDENTIFIER/
                  YEAR         Year film taken/
END SCHEMA
```

```
TASK NAME        RECORD 4 (READDATE)  SCHEMA DEFINITION
RECORD SCHEMA    4  READDATE
DOCUMENT
                        This record type holds the information on when a batch
                        was read by each reader, and when the data from that
                        reading were input to the database. There is one
                        record per reader per batch, and the reader is the
                        sort id. The case id is the batch.

SORT IDS         READER (A)
SEQUENCE CHECK   OFF
MAX REC COUNT    6
DATA LIST        FIXED (1)
                 /1       CASEID       1 -   8    (A)
                 /1       READER      10 -  12    (I)
                 /1       DATEREAD    14 -  19    (A)
                 /1       DATEINPT    21 -  26    (A)
                 /1       EXERCISE     1 -   4    (A)
DATE VARS        DATEREAD  ('DDMMYY')/
                 DATEINPT  ('DDMMYY')/
CAT VARS         EXERCISE  ( 'XRLV','XRVL','XREX','XRES'    )/
IF               (EXISTS(DATEINPT)EQ 0)DATEINPT=TODAY(0)
VAR RANGES       READER     (10 112)/
VAR SECURITY     CASEID    (20,20 ) /
MISSING VALUES   DATEREAD  ( '      '  )/
                 DATEINPT  ( '      '  )/
VALID VALUES     READER    ( 101,104,105,106,111,112,010,011,013  )/
VALUE LABELS     READER    (101)'Mrs Scott',(104)'Mrs White',(105)'Mrs Duncan',
                           (106)'Mrs Henderson',(111)'Mrs Chalmers',
                           (112)'Miss Jeffreys',(10)'Dr Pern',(11)'Dr Soutar',
                           (13)'Dr Seaton' /
                 EXERCISE  ('XRLV')'XRLV',('XRVL')'XRVL',('XREX')'XREX',
                           ('XRES')'XRES' /
VAR LABELS       CASEID      CASE IDENTIFIER/
                 READER      Reader identity/
                 DATEREAD    Date this batch was read/
                 DATEINPT    Date of input to database/
END SCHEMA

TASK NAME        RECORD 5 (XRLV    )  SCHEMA DEFINITION
RECORD SCHEMA    5  XRLV
DOCUMENT
                 This record type contains data on all chest radiographs
                 read by the IOM panel during the first reading exercise for
                 this project.

                 For more information on codes used - see ILO 1980
                 classification booklet, and IOM reading protocols.

SORT IDS         YEAR (A)        READER (A)
SEQUENCE CHECK   OFF
MAX REC COUNT    24
DATA LIST        FIXED (1)
                 /1       YEAR        26 -  27    (I)
                 /1       READER       8 -  10    (I)
                 /1       EXERCISE     1 -   4    (A)
                 /1       CASEID      17 -  24    (A)
                 /1       QUAL        32          (I)
                 /1       DEFECT1     33          (A)
                 /1       DEFECT2     34          (A)
                 /1       DEFECT3     35          (A)
                 /1       CPAOFF      36          (A)
                 /1       NORM        37          (A)
                 /1       SOPRO       38 -  39    (A)
                 /1       ZONE1       40 -  41    (A)
                 /1       ZONE2       42 -  43    (A)
                 /1       ZONE3       44 -  45    (A)
                 /1       SOTYP       46 -  47    (A)
                 /1       LOPRO       48          (A)
                 /1       PLFAC       49 -  50    (A)
                 /1       PLWID       51 -  52    (A)
                 /1       PLEXT       53 -  54    (A)
                 /1       DPTF        55 -  56    (A)
                 /1       DPTW        57 -  58    (A)
                 /1       DPTE        59 -  60    (A)
                 /1       DIAPH       61          (A)
                 /1       CPAOBL      62          (A)
                 /1       PCCW        63 -  64    (A)
                 /1       PCDIA       65 -  66    (A)
```

```
                /1          PCOTH        67 -   68    (A)
                /1          PCEXT        69 -   70    (A)
                /1          OA1          71 -   72    (A)
                /1          OA2          73 -   74    (A)
                /1          OA3          75 -   76    (A)
                /1          OA4          77 -   78    (A)
                /1          OA5          79 -   80    (A)
                /1          OA6          81 -   82    (A)
                /1          COM          83           (A)
CAT VARS        EXERCISE    ( 'XRLV'  )/
                DEFECT1     ( '1','2','3','4','5','6','7','8','9',' '   )/
                DEFECT2     ( '1','2','3','4','5','6','7','8','9',' '   )/
                DEFECT3     ( '1','2','3','4','5','6','7','8','9',' '   )/
                CPAOFF      ( ' ','R','L','B'  )/
                NORM        ( ' ','X'  )/
                SOPRO       ( '0-','00','01','10','11','12','21','22',
                      '23','32','33','3+'   )/
                ZONE1       ( '  ','X ',' X','XX'   )/
                ZONE2       ( '  ','X ',' X','XX'   )/
                ZONE3       ( '  ','X ',' X','XX'   )/
                SOTYP       ( '  ','PP','PQ','PR','PS','PT','PU','QP',
                              'QQ','QR','QS','QT','QU','RP','RQ','RR',
                              'RS','RT','RU','SP','SQ','SR','SS','ST',
                              'SU','TP','TQ','TR','TS','TT','TU','UP',
                              'UQ','UR','US','UT','UU'   )/
                LOPRO       ( '0','A','B','C'   )/
                PLFAC       ( '  ','N ','Y ',' N',' Y','NN','YN','NY','YY'   )/
                PLWID       ( '  ','A ','B ','C ',' A',' B',' C','AA','BA','CA',
                              'AB','BB','CB','AC','BC','CC'   )/
                PLEXT       ( '  ','1 ','2 ','3 ',' 1',' 2',' 3','11','21','31',
                              '12','22','32','13','23','33'   )/
                DPTF        ( '  ','N ','Y ',' N',' Y','NN','YN','NY','YY'   )/
                DPTW        ( '  ','A ','B ','C ',' A',' B',' C','AA','BA','CA',
                              'AB','BB','CB','AC','BC','CC'   )/
                DPTE        ( '  ','1 ','2 ','3 ',' 1',' 2',' 3','11','21','31',
                              '12','22','32','13','23','33'   )/
                DIAPH       ( ' ','R','L','B'   )/
                CPAOBL      ( ' ','R','L','B'   )/
                PCCW        ( '  ','X ',' X','XX'   )/
                PCDIA       ( '  ','X ',' X','XX'   )/
                PCOTH       ( '  ','X ',' X','XX'   )/
                PCEXT       ( '  ','1 ','2 ','3 ',' 1',' 2',' 3','11','21','31',
                              '12','22','32','13','23','33'   )/
                OA1 to OA6( '  ','AX','BU','CA','CN','CO','CP','CU','DI','EF',
                              'EM','ES','FR','HI','HO','ID','IH','KL','OD','PI',
                              'PX','RP','TB'   )/
                COM         ( ' ','X'   )/
COMPUTE         BLDUM=1/0
COMPUTE         TEST=0
IF              (QUAL EQ 4 AND SUM(CPAOFF,NORM,ZONE1,ZONE2,ZONE3,SOTYP,
                PLFAC,PLWID,PLEXT,DPTF,DPTW,DPTE,DIAPH,CPAOBL,PCCW,PCDIA,
                PCOTH,PCEXT,OA1,OA2,OA3,OA4,OA5,OA6)EQ 24 AND MISNUM(
                SOPRO)EQ 1 AND MISNUM(LOPRO)EQ 1)TEST=1;
IF              (TEST EQ 1)CPAOFF=BLDUM;
IF              (TEST EQ 1)NORM=BLDUM;
IF              (TEST EQ 1)SOPRO=BLDUM;
IF              (TEST EQ 1)ZONE1=BLDUM;
IF              (TEST EQ 1)ZONE2=BLDUM;
IF              (TEST EQ 1)ZONE3=BLDUM;
IF              (TEST EQ 1)SOTYP=BLDUM;
IF              (TEST EQ 1)LOPRO=BLDUM;
IF              (TEST EQ 1)PLFAC=BLDUM;
IF              (TEST EQ 1)PLWID=BLDUM;
IF              (TEST EQ 1)PLEXT=BLDUM;
IF              (TEST EQ 1)DPTF=BLDUM;
IF              (TEST EQ 1)DPTW=BLDUM;
IF              (TEST EQ 1)DPTE=BLDUM;
IF              (TEST EQ 1)DIAPH=BLDUM;
IF              (TEST EQ 1)CPAOBL=BLDUM;
IF              (TEST EQ 1)PCCW=BLDUM;
IF              (TEST EQ 1)PCDIA=BLDUM;
IF              (TEST EQ 1)PCOTH=BLDUM;
IF              (TEST EQ 1)PCEXT=BLDUM;
IF              (TEST EQ 1)OA1=BLDUM;
IF              (TEST EQ 1)OA2=BLDUM;
IF              (TEST EQ 1)OA3=BLDUM;
IF              (TEST EQ 1)OA4=BLDUM;
IF              (TEST EQ 1)OA5=BLDUM;
IF              (TEST EQ 1)OA6=BLDUM
IF              (NORM EQ 2 AND MISNUM(SOPRO)EQ 1)SOPRO=2
IF              (NORM EQ 2 AND MISNUM(LOPRO)EQ 1)LOPRO=1
```

```
VAR RANGES        YEAR        (70 80)/
                  READER      (101 112)/
                  QUAL        (1 4)/
VAR SECURITY      CASEID     (20,20 ) /
MISSING VALUES    SOPRO       ( BLANK  )/
                  LOPRO       ( BLANK  )/
VALID VALUES      READER      ( 101,104,105,106,111,112  )/
VALUE LABELS      YEAR        (70)'1970',(74)'1974',(75)'1975',(78)'1978',
                              (79)'1979',(80)'1980' /
                  READER      (101)'Mrs Scott',(104)'Mrs White',(105)'Mrs Duncan',
                              (106)'Mrs Henderson',(111)'Mrs Chalmers',
                              (112)'Miss Jeffreys' /
                  EXERCISE    ('XRLV')'XRLV' /
                  QUAL        (1)'good',(2)'acceptable',(3)'poor',(4)'unacceptable'/
                  DEFECT1 to('1')'dark film'
                    DEFECT3 ('2')'pale film'
                              ('3')'too much contrast'
                              ('4')'lack of contrast'
                              ('5')'patient movement'
                              ('6')'poor inspiration'
                              ('7')'marks on film'
                              ('8')'ill defined marks'
                              ('9')'other'
                              (' ')'no entry' /
                  CPAOFF      (' ')'neither'
                              ('R')'right'
                              ('L')'left'
                              ('B')'right & left'
                              (UNDEFINED)'not assessed' /
                  NORM        (' ')'not coded normal'
                              ('X')'coded normal'
                              (UNDEFINED)'not assessed' /
                  SOPRO       ('0-')'0/-',('00')'0/0',('01')'0/1',('10')'1/0',
                              ('11')'1/1',('12')'1/2',('21')'2/1',('22')'2/2',
                              ('23')'2/3',('32')'3/2',('33')'3/3',('3+')'3/+',
                              (UNDEFINED)'not assessed' /
                  ZONE1 to    ('  ')'blank',('X ')'right',
                    ZONE3     (' X')'left',('XX')'right & left',
                              (UNDEFINED)'not assessed' /
                  SOTYP       ('  ')'blank'
                              (UNDEFINED)'not assessed' /
                  LOPRO       (UNDEFINED)'not assessed' /
                  PLFAC       ('  ')'blank',(UNDEFINED)'not assessed' /
                  PLWID       ('  ')'blank',(UNDEFINED)'not assessed' /
                  PLEXT       ('  ')'0.0',('1 ')'1.0',('2 ')'2.0',('3 ')'3.0',
                              (' 1')'0.1',(' 2')'0.2',(' 3')'0.3',('11')'1.1',
                              ('21')'2.1',('31')'3.1',('12')'1.2',('22')'2.2',
                              ('32')'3.2',('13')'1.3',('23')'2.3',('33')'3.3',
                              (UNDEFINED)'not assessed' /
                  DPTF        ('  ')'blank',(UNDEFINED)'not assessed' /
                  DPTW        ('  ')'blank',(UNDEFINED)'not assessed' /
                  DPTE        ('  ')'0.0',('1 ')'1.0',('2 ')'2.0',('3 ')'3.0',
                              (' 1')'0.1',(' 2')'0.2',(' 3')'0.3',('11')'1.1',
                              ('21')'2.1',('31')'3.1',('12')'1.2',('22')'2.2',
                              ('32')'3.2',('13')'1.3',('23')'2.3',('33')'3.3',
                              (UNDEFINED)'not assessed' /
                  DIAPH       (' ')'neither',('R')'right',('L')'left',
                              ('B')'right & left',(UNDEFINED)'not assessed' /
                  CPAOBL      (' ')'neither',('R')'right',('L')'left',
                              ('B')'right & left',(UNDEFINED)'not assessed' /
                  PCCW        ('  ')'neither',('X ')'right',(' X')'left',
                              ('XX')'right & left',(UNDEFINED)'not assessed' /
                  PCDIA       ('  ')'neither',('X ')'right',(' X')'left',
                              ('XX')'right & left',(UNDEFINED)'not assessed' /
                  PCOTH       ('  ')'neither',('X ')'right',(' X')'left',
                              ('XX')'right & left',(UNDEFINED)'not assessed' /
                  PCEXT       ('  ')'0.0',('1 ')'1.0',('2 ')'2.0',('3 ')'3.0',
                              (' 1')'0.1',(' 2')'0.2',(' 3')'0.3',('11')'1.1',
                              ('21')'2.1',('31')'3.1',('12')'1.2',('22')'2.2',
                              ('32')'3.2',('13')'1.3',('23')'2.3',('33')'3.3',
                              (UNDEFINED)'not assessed' /
                  OA1 to OA6('  ')'no entry'
                              ('AX')'ax - coalescence'
                              ('BU')'bu - bullae'
                              ('CA')'ca - cancer lung'
                              ('CN')'cn - calcification'
                              ('CO')'co - cardiac abn.'
                              ('CP')'cp - cor pulmonale'
                              ('CU')'cv - cavity'
                              ('DI')'di - organ distort.'
```

```
                      ('EF')'ef - effusion'
                      ('EM')'em - definite emph.'
                      ('ES')'es - eggshell calc.'
                      ('FR')'fr - fractured ribs'
                      ('HI')'hi - enlarged nodes'
                      ('HO')'ho - honeycomb lung'
                      ('ID')'id - ill. def. dia.'
                      ('IH')'ih - ill. def. heart'
                      ('KL')'kl - septal lines'
                      ('OD')'od - other sig. abn.'
                      ('PI')'pi - pl. thickening'
                      ('PX')'px - pneumothorax'
                      ('RP')'rp - rheum. pneum.'
                      ('TB')'tb - tuberculosis'
                      (UNDEFINED)'not assessed' /
           COM        (' ')'no comment present',('X')'comment present' /
VAR LABELS YEAR        year of film/
           READER      reader identity/
           CASEID      CASE IDENTIFIER/
           QUAL        film quality/
           DEFECT1     film defects (1)/
           DEFECT2     film defects (2)/
           DEFECT3     film defects (3)/
           CPAOFF      costophrenic angle off/
           NORM        film normality indicator/
           SOPRO       small opacities - profusion/
           ZONE1       zones of profusion - upper/
           ZONE2       zones of profusion - middle/
           ZONE3       zones of profusion - lower/
           SOTYP       small opacities - type/
           LOPRO       large opacities - profusion/
           PLFAC       plaques - face on/
           PLWID       plaques - width/
           PLEXT       plaques - extent/
           DPTF        diffuse pleural thickening - face on/
           DPTW        diffuse pleural thickening - width/
           DPTE        diffuse pleural thickening - extent/
           DIAPH       diaphragm diffuse pleural thickening/
           CPAOBL      costophrenic angle obliteration/
           PCCW        pleural calcification - chest wall/
           PCDIA       pleural calcification - diaphragm/
           PCOTH       pleural calcification - other/
           PCEXT       pleural calcification - extent/
           OA1         other abnormalities (1)/
           OA2         other abnormalities (2)/
           OA3         other abnormalities (3)/
           OA4         other abnormalities (4)/
           OA5         other abnormalities (5)/
           OA6         other abnormalities (6)/
           COM         comment/
REJECT REC IF ((SUM(CPAOFF,NORM,ZONE1,ZONE2,ZONE3,SOTYP,PLFAC,PLWID,
           PLEXT,DPTF,DPTW,DPTE,DIAPH,CPAOBL,PCCW,PCDIA,PCOTH,PCEXT,
           OA1,OA2,OA3,OA4,OA5,OA6)GT 0)AND QUAL EQ 4)
REJECT REC IF (QUAL EQ 1 AND(DEFECT1 NE 10))
REJECT REC IF (QUAL GT 1 AND(DEFECT1 EQ 10))
REJECT REC IF ((DEFECT1 EQ 10)AND(DEFECT2 NE 10))
REJECT REC IF ((DEFECT2 EQ 10)AND(DEFECT3 NE 10))
REJECT REC IF (NORM EQ 2 AND SOPRO GT 2.5)
REJECT REC IF (NORM EQ 2 AND LOPRO GT 1.5)
REJECT REC IF (NORM EQ 1 AND(EXISTS(SOPRO)EQ 0))
REJECT REC IF (NORM EQ 1 AND(EXISTS(LOPRO)EQ 0))
REJECT REC IF (SOPRO GT 2.5 AND(SUM(ZONE1,ZONE2,ZONE3)EQ 3 OR SOTYP EQ
           1))
REJECT REC IF (SOPRO LT 2.5 AND(SUM(ZONE1,ZONE2,ZONE3)GT 3 OR SOTYP GT
           1))
REJECT REC IF (MISNUM(SOPRO)EQ 1 OR MISNUM(LOPRO)EQ 1)
REJECT REC IF (((PLEXT GT 4.5 AND PLEXT LT 7.5)OR(PLEXT EQ 1))AND((
           PLFAC GT 1.5 AND PLFAC LT 3.5)OR(PLFAC GT 5.5)OR(PLWID
           GT 1.5 AND PLWID LT 4.5)OR(PLWID GT 7.5)))
REJECT REC IF (((PLEXT GT 1.5 AND PLEXT LT 4.5)OR(PLEXT GT 7.5))AND(
           PLFAC EQ 1 OR 4 OR 5))
REJECT REC IF (((PLEXT GT 1.5 AND PLEXT LT 4.5)OR(PLEXT GT 7.5))AND(
           PLFAC EQ 2 OR 6 OR 8)AND(PLWID EQ 1 OR 5 OR 6 OR 7))

REJECT REC IF ((PLEXT LT 4.5)AND((PLFAC GT 3.5)OR(PLWID GT 4.5)))
REJECT REC IF ((PLEXT GT 4.5)AND(PLFAC LT 3.5))
REJECT REC IF ((PLEXT GT 4.5)AND(PLFAC EQ 4 OR 6 OR 7)AND(PLWID LT 4.5)
           )
REJECT REC IF (((DPTE GT 4.5 AND DPTE LT 7.5)OR(DPTE EQ 1))AND((DPTF
           GT 1.5 AND DPTF LT 3.5)OR(DPTF GT 5.5)OR(DPTW GT 1.5 AND
           DPTW LT 4.5)OR(DPTW GT 7.5)))
```

```
REJECT REC IF     (((DPTE GT 1.5 AND DPTE LT 4.5)OR(DPTE GT 7.5))AND(DPTF
                  EQ 1 OR 4 OR 5))
REJECT REC IF     (((DPTE GT 1.5 AND DPTE LT 4.5)OR(DPTE GT 7.5))AND(DPTF
                  EQ 2 OR 6 OR 8)AND(DPTW EQ 1 OR 5 OR 6 OR 7))
REJECT REC IF     ((DPTE LT 4.5)AND((DPTF GT 3.5)OR(DPTW GT 4.5)))
REJECT REC IF     ((DPTE GT 4.5)AND(DPTF LT 3.5))
REJECT REC IF     ((DPTE GT 4.5)AND(DPTF EQ 4 OR 6 OR 7)AND(DPTW LT 4.5))

REJECT REC IF     ((CPAOFF EQ 2 OR 4)AND(CPAOBL EQ 2 OR 4))
REJECT REC IF     ((CPAOFF EQ 3 OR 4)AND(CPAOBL EQ 3 OR 4))
REJECT REC IF     (((PCEXT EQ 2 OR 3 OR 4)OR(PCEXT GT 7.5))AND(PCCW EQ 1
                  OR 3)AND(PCDIA EQ 1 OR 3)AND(PCOTH EQ 1 OR 3))
REJECT REC IF     ((PCEXT EQ 1 OR 5 OR 6 OR 7)AND((PCCW EQ 2 OR 4)OR(PCDIA
                  EQ 2 OR 4)OR(PCOTH EQ 2 OR 4)))
REJECT REC IF     ((PCEXT GT 4.5)AND(PCCW EQ 1 OR 2)AND(PCDIA EQ 1 OR 2)
                  AND(PCOTH EQ 1 OR 2))
REJECT REC IF     ((PCEXT LT 4.5)AND((PCCW EQ 3 OR 4)OR(PCDIA EQ 3 OR 4)OR(
                  PCOTH EQ 3 OR 4)))
REJECT REC IF     ((OA1 EQ 1 AND OA2 GT 1)OR(OA2 EQ 1 AND OA3 GT 1)OR(OA3
                  EQ 1 AND OA4 GT 1)OR(OA4 EQ 1 AND OA5 GT 1)OR(OA5 EQ 1
                  AND OA6 GT 1))
END SCHEMA

TASK NAME         RECORD 6 (XRVL     ) SCHEMA DEFINITION
RECORD SCHEMA     6  XRVL
DOCUMENT
                  This record type contains data on all chest radiographs
                  read during the second reading exercise for this project.

                  Note that the batch lists for this reading were the same as
                  for XRLV, and they have not been stored separately in the
                  database.

                 [The schema for this record type is almost identical to
                  that for record type 5 XRLV]

TASK NAME         RECORD 7 (XREX     ) SCHEMA DEFINITION
RECORD SCHEMA     7  XREX
DOCUMENT
                  This record type contains data on all chest radiographs
                  read during the third, supplementary, reading exercise for
                  this project.

                 [The schema for this record type is almost identical to
                  that for record type 5 XRLV]

TASK NAME         RECORD 8 (XRES     ) SCHEMA DEFINITION
RECORD SCHEMA     8  XRES
DOCUMENT
                  This record type contains data on all chest radiographs
                  read during the serial reading exercise undertaken by
                  three medical readers.

                  The reading used a shortened protocol which covered
                  only small and large opacities.

                  For more information on codes used - see ILO 1980
                  classification booklet and IOM local protocols.

SORT IDS          YEAR (A)      READER (A)
SEQUENCE CHECK    OFF
MAX REC COUNT     24
DATA LIST         FIXED (1)
                  /1          YEAR          26 -   27    (I)
                  /1          READER         8 -   10    (I)
                  /1          EXERCISE       1 -    4    (A)
                  /1          CASEID        17 -   24    (A)
                  /1          QUAL          32           (I)
                  /1          DEFECT1       33           (A)
                  /1          DEFECT2       34           (A)
                  /1          DEFECT3       35           (A)
                  /1          NORM          36           (A)
                  /1          SOPRO         37 -   38    (A)
                  /1          ZONE1         39 -   40    (A)
                  /1          ZONE2         41 -   42    (A)
                  /1          ZONE3         43 -   44    (A)
```

```
                /1          SOTYP           45 -    46      (A)
                /1          LOPRO           47              (A)
                /1          COM             48              (A)
CAT VARS        EXERCISE    ( 'XRES'  )/
                DEFECT1     ( '1','2','3','4','5','6','7','8','9',' '   )/
                DEFECT2     ( '1','2','3','4','5','6','7','8','9',' '   )/
                DEFECT3     ( '1','2','3','4','5','6','7','8','9',' '   )/
                NORM        ( ' ','X'  )/
                SOPRO       ( '0-','00','01','10','11','12','21','22','23',
                              '32','33','3+'  )/
                ZONE1       ( '  ','X ',' X','XX'   )/
                ZONE2       ( '  ','X ',' X','XX'   )/
                ZONE3       ( '  ','X ',' X','XX'   )/
                SOTYP       ( '  ','PP','PQ','PR','PS','PT','PU','QP','QQ','QR',
                              'QS','QT','QU','RP','RQ','RR','RS','RT','RU','SP',
                              'SQ','SR','SS','ST','SU','TP','TQ','TR','TS','TT',
                              'TU','UP','UQ','UR','US','UT','UU'   )/
                LOPRO       ( '0','A','B','C'   )/
                COM         ( ' ','X'  )/
COMPUTE         BLDUM=1/0
COMPUTE         TEST=0
IF              (QUAL EQ 4 AND SUM(NORM,ZONE1,ZONE2,ZONE3,SOTYP)EQ 5 AND
                MISNUM(SOPRO)EQ 1 AND MISNUM(LOPRO)EQ 1)TEST=1;
IF              (TEST EQ 1)NORM=BLDUM;
IF              (TEST EQ 1)SOPRO=BLDUM;
IF              (TEST EQ 1)ZONE1=BLDUM;
IF              (TEST EQ 1)ZONE2=BLDUM;
IF              (TEST EQ 1)ZONE3=BLDUM;
IF              (TEST EQ 1)SOTYP=BLDUM;
IF              (TEST EQ 1)LOPRO=BLDUM;
IF              (NORM EQ 2 AND MISNUM(SOPRO)EQ 1)SOPRO=2
IF              (NORM EQ 2 AND MISNUM(LOPRO)EQ 1)LOPRO=1
VAR RANGES      YEAR        (70 80)/
                READER      (10 13)/
                QUAL        (1 4)/
VAR SECURITY    CASEID     (20,20 ) /
MISSING VALUES  SOPRO       ( BLANK  )/
                LOPRO       ( BLANK  )/
VALID VALUES    READER      ( 010,011,013 )/
VALUE LABELS    YEAR        (70)'1970',(74)'1974',(75)'1975',(78)'1978',
                            (79)'1979',(80)'1980' /
                READER      (10)'Dr Pern',(11)'Dr Soutar',(13)'Dr Seaton' /
                EXERCISE    ('XRES')'XRES' /
                QUAL        (1)'good',(2)'acceptable',(3)'poor',(4)'unacceptable'/
                DEFECT1 to('1')'dark film'
                  DEFECT3 ('2')'pale film'
                            ('3')'too much contrast'
                            ('4')'lack of contrast'
                            ('5')'patient movement'
                            ('6')'poor inspiration'
                            ('7')'marks on film'
                            ('8')'ill defined marks'
                            ('9')'other'
                            (' ')'no entry' /
                NORM        (' ')'not coded normal',('X')'coded normal',
                            (UNDEFINED)'not assessed' /
                SOPRO       ('0-')'0/-',('00')'0/0',('01')'0/1',('10')'1/0',
                            ('11')'1/1',('12')'1/2',('21')'2/1',('22')'2/2',
                            ('23')'2/3',('32')'3/2',('33')'3/3',('3+')'3/+',
                            (UNDEFINED)'not assessed' /
                ZONE1 to    ('  ')'blank',('X ')'right',
                  ZONE3     (' X')'left',('XX')'right & left',
                            (UNDEFINED)'not assessed' /
                SOTYP       ('  ')'blank',(UNDEFINED)'not assessed' /
                LOPRO       (UNDEFINED)'not assessed' /
                COM         (' ')'no comment present',('X')'comment present' /
VAR LABELS      YEAR          year of film/
                READER        reader identity/
                CASEID        CASE IDENTIFIER/
                QUAL          film quality/
                DEFECT1       film defects (1)/
                DEFECT2       film defects (2)/
                DEFECT3       film defects (3)/
                NORM          film normality indicator/
                SOPRO         small opacities - profusion/
                ZONE1         zones of profusion - upper/
                ZONE2         zones of profusion - middle/
                ZONE3         zones of profusion - lower/
                SOTYP         small opacities - type/
                LOPRO         large opacities - profusion/
                COM           comment/
```

```
REJECT REC IF      ((SUM(NORM,ZONE1,ZONE2,ZONE3,SOTYP)GT 0)AND QUAL EQ 4)
REJECT REC IF      (QUAL EQ 1 AND(DEFECT1 NE 10))
REJECT REC IF      (QUAL GT 1 AND(DEFECT1 EQ 10))
REJECT REC IF      ((DEFECT1 EQ 10)AND(DEFECT2 NE 10))
REJECT REC IF      ((DEFECT2 EQ 10)AND(DEFECT3 NE 10))
REJECT REC IF      (NORM EQ 2 AND SOPRO GT 2.5)
REJECT REC IF      (NORM EQ 2 AND LOPRO GT 1.5)
REJECT REC IF      (NORM EQ 1 AND(EXISTS(SOPRO)EQ 0))
REJECT REC IF      (NORM EQ 1 AND(EXISTS(LOPRO)EQ 0))
REJECT REC IF      (SOPRO GT 2.5 AND(SUM(ZONE1,ZONE2,ZONE3)EQ 3 OR SOTYP EQ
                   1))
REJECT REC IF      (SOPRO LT 2.5 AND(SUM(ZONE1,ZONE2,ZONE3)GT 3 OR SOTYP GT
                   1))
REJECT REC IF      (MISNUM(SOPRO)EQ 1 OR MISNUM(LOPRO)EQ 1)
END SCHEMA


TASK NAME          RECORD 10 (EXPSOTO4)  SCHEMA DEFINITION
RECORD SCHEMA      10  EXPSOTO4
DOCUMENT           This record type holds the TOTAL time and associated exposures
                   for isps 0 - 2, with the measflag set to 1; the MEASURED time
                   and associated exposures for isps 3 and 4  by isp, with the
                   measflag set to 2; and the UNMEASURED time and associated
                   exposures for isps  3 and 4 by isp with measflag set to 3.
SORT IDS           MEASFLAG (A)  ISP (A)
SEQUENCE CHECK     OFF
MAX REC COUNT      7
DATA LIST          FIXED (1)
                   /1          CASEID           1 -    8     (A)
                   /1          ISP             10            (I)
                   /1          CONTIME         12 -   19     (F2)
                   /1          EXDUST          20 -   29     (F2)
                   /1          EXASH           30 -   39     (F2)
                   /1          EXQUARTZ        40 -   49     (F2)
                   /1          MEASFLAG        51            (I)
VAR RANGES         ISP         (0 4)/
                   MEASFLAG    (1 3)/
VAR SECURITY       CASEID    (20,20 ) /
VALUE LABELS       ISP         (0)'pre 1st pfr survey',
                               (1)'between 1st and 2nd pfr survey'
                               (2)'between 2nd and 3rd pfr survey'
                               (3)'between 3rd and 4th pfr survey'
                               (4)'between 4th and 5th pfr survey' /
                   MEASFLAG    (1)'total - isps 0,1,2'
                               (2)'measured - isps 3,4'
                               (3)'unmeasured - isps 3,4' /
VAR LABELS         CASEID        CASE IDENTIFIER/
                   ISP           inter survey period/
                   CONTIME       contributing time/
                   EXDUST        exposure to dust/
                   EXASH         exposure to ash/
                   EXQUARTZ      exposure to quartz/
                   MEASFLAG      total, measured or unmeasured data/
REJECT REC IF      (contime eq 0)
END SCHEMA


TASK NAME          RECORD 11 (ISP5ATT )  SCHEMA DEFINITION
RECORD SCHEMA      11  ISP5ATT
DOCUMENT           This record type holds the MEASURED time and the ogsn to which
                   it refers for isp 5 held by quarter. The time is held as shifts
                   and there are 7.25 hours per shift.
SORT IDS           QUARTER (A)    OGSN (A)
SEQUENCE CHECK     OFF
MAX REC COUNT      128
DATA LIST          FIXED (1)
                   /1          CASEID           1 -    8    (A)
                   /1          QUARTER         10           (A)
                   /1          OGSN            11 -   18    (A)
                   /1          SHIFTS          19 -   23    (F1)
CAT VARS           QUARTER     ( 'A','B','C','D','E','F','G','H','I','J','K','L',
                                 'M','N','O','P'  )/
VAR SECURITY       CASEID    (20,20 ) /
VALUE LABELS       QUARTER     ('A')'A',('B')'B',('C')'C',('D')'D',('E')'E',
                               ('F')'F',('G')'G',('H')'H',('I')'I',('J')'J',
                               ('K')'K',('L')'L',('M')'M',('N')'N',('O')'O'
                               ('P')'P' /
VAR LABELS         CASEID        CASE IDENTIFIER/
                   QUARTER       quarter to which data refers/
                   OGSN          occupational group serial no/
                   SHIFTS        number of shifts worked/
REJECT REC IF      (ogsn gt'OGSN2000')
END SCHEMA
```

```
TASK NAME         RECORD 12 (ISP5ENV ) SCHEMA DEFINITION
RECORD SCHEMA     12 ISP5ENV
DOCUMENT          This record type holds the dust composition data for each ogsn
                  in each quarter. The caseid is the ogsn
SORT IDS          QUARTER (A)
SEQUENCE CHECK    OFF
MAX REC COUNT     16
DATA LIST         FIXED (1)
                  /1        QUARTER        1             (A)
                  /1        CASEID         2 -    9      (A)
                  /1        DMC            10 -   15     (F2)
                  /1        AIND           16 -   21     (F3)
                  /1        QINA           22 -   27     (F4)
CAT VARS          QUARTER   ( 'A','B','C','D','E','F','G','H','I','J','K',
                             'L','M','N','O','P'  )/
VAR SECURITY      CASEID    (20,20 ) /
VALUE LABELS      QUARTER   ('A')'A',('B')'B',('C')'C',('D')'D',('E')'E',
                            ('F')'F',('G')'G',('H')'H',('I')'I',('J')'J',
                            ('K')'K',('L')'L',('M')'M',('N')'N',('O')'O',
                            ('P')'P' /
VAR LABELS        QUARTER     quarter to which data refers/
                  CASEID      CASE IDENTIFIER/
                  DMC         dust mass concentration/
                  AIND        percentage ash in dust/
                  QINA        percentage quartz in ash/
END SCHEMA

TASK NAME         RECORD 13 (ISP5INV ) SCHEMA DEFINITION
RECORD SCHEMA     13 ISP5INV
DOCUMENT          This record type holds measured time information for the ogsns
                  for which there are no dust data.  The information held is the
                  number of shifts worked in the ogsn, the quarter in which it
                  was worked and the id of the man who did the work.
SORT IDS          QUARTER (A)    OGSN (A)
SEQUENCE CHECK    OFF
MAX REC COUNT     3
DATA LIST         FIXED (1)
                  /1        CASEID         1 -    8      (A)
                  /1        QUARTER        9             (A)
                  /1        OGSN           10 -   17     (A)
                  /1        HOURS          18 -   26     (F4)
CAT VARS          QUARTER   ( 'A','B','C','D','E','F','G','H','I','J','K','L'
                             'M','N','O','P'  )/
VAR SECURITY      CASEID    (20,20 ) /
VALUE LABELS      QUARTER   ('A')'A',('B')'B',('C')'C',('D')'D',('E')'E',
                            ('F')'F',('G')'G',('H')'H',('I')'I',('J')'J',
                            ('K')'K',('L')'L',('M')'M',('N')'N',('O')'O',
                            ('P')'P' /
VAR LABELS        CASEID      CASE IDENTIFIER/
                  QUARTER     quarter to which data refers/
                  OGSN        occupational group serial no/
                  HOURS       time worked in hours/
END SCHEMA
```

```
TASK NAME        RECORD 14 (ISP5PH   )  SCHEMA DEFINITION
RECORD SCHEMA    14  ISP5PH
DOCUMENT         This record type holds the unmeasured time for isp 5, that is
                 the time collected from the occupational histories taken at 6th
                 survey.
SORT IDS         JOBCAT (A)
SEQUENCE CHECK   OFF
MAX REC COUNT    1
DATA LIST        FIXED (1)
                 /1          CASEID          1 -    8    (A)
                 /1          JOBCAT         10           (I)
                 /1          HOURS          11 -   14    (I)
VAR RANGES       JOBCAT     (1 6)/
VAR SECURITY     CASEID    (20,20 ) /
VALID VALUES     JOBCAT     ( 1,5,6  )/
VALUE LABELS     JOBCAT     (1)'face',(5)'surface',
                            (6)'elsewhere underground' /
VAR LABELS       CASEID       CASE IDENTIFIER/
                 JOBCAT       job category code/
                 HOURS        number of hours worked outside survey pit/
END SCHEMA

TASK NAME        RECORD 15 (PANDAS   )  SCHEMA DEFINITION
RECORD SCHEMA    15  PANDAS
DOCUMENT         this record type holds the information collected at the pfr
                 surveys on lung function values and smoking habits for each
                 man.
SORT IDS         SURVEY (A)
SEQUENCE CHECK   OFF
MAX REC COUNT    3
DATA LIST        FIXED (1)
                 /1          CASEID          1 -    8    (A)
                 /1          MSUR            9 -   11    (I)
                 /1          YSUR           12 -   14    (I)
                 /1          Q8             43 -   44    (I)
                 /1          Q8A            45 -   46    (I)
                 /1          Q8B            47 -   49    (I)
                 /1          Q8C            50 -   52    (I)
                 /1          Q8D            53 -   55    (F1)
                 /1          Q8E            56 -   57    (I)
                 /1          HEIGHT         62 -   65    (I)
                 /1          WT             70 -   73    (I)
                 /1          FEV2           74 -   77    (I)
                 /1          FEV3           78 -   81    (I)
                 /1          FEV4           82 -   85    (I)
                 /1          FVC2           86 -   89    (I)
                 /1          FVC3           90 -   93    (I)
                 /1          FVC4           94 -   97    (I)
                 /1          SURVEY         99           (I)
                 /1          TECHFLAG      100           (I)
VAR RANGES       SURVEY     (4 6)/
                 TECHFLAG   (0 3)/
VAR SECURITY     CASEID    (20,20 ) /
MISSING VALUES   MSUR       ( -99    )/
                 YSUR       ( -99    )/
                 Q8         ( -9   )/
                 Q8A        ( -9   )/
                 Q8B        ( -99    )/
                 Q8C        ( -99    )/
                 Q8D        ( -99    )/
                 Q8E        ( -9   )/
                 HEIGHT     ( -999    )/
                 WT         ( -999    )/
                 FEV2       ( -999,-888    )/
                 FEV3       ( -999,-888    )/
                 FEV4       ( -999,-888    )/
                 FVC2       ( -999,-888    )/
                 FVC3       ( -999,-888    )/
                 FVC4       ( -999,-888    )/
VALUE LABELS     Q8         (1)'BLANK',(2)'NO',(3)'YES' /
                 Q8A        (1)'BLANK',(2)'CIGS',(3)'PIPE',(4)'BOTH' /
                 Q8E        (1)'BLANK',(2)'NO',(3)'YES' /
                 TECHFLAG   (0)'NOT APPLICABLE',(1)'NEW TECHNICIAN',
                            (2)'EXPERIENCED TECHNICIAN',(3)'NOT KNOWN' /
```

```
VAR LABELS      CASEID      CASE IDENTIFIER/
                MSUR        MONTH OF SURVEY/
                YSUR        YEAR OF SURVEY/
                Q8          SMOKING/
                Q8A         CIGS, PIPE OR BOTH/
                Q8B         NO OF CIGS PER WEEK DAY/
                Q8C         NO OF CIGS PER WEEKEND DAY/
                Q8D         OZS TOBACCO PER WEEK/
                Q8E         EVER SMOKED 1 CIG PER DAY FOR 1 YR/
                HEIGHT      HEIGHT/
                WT          WEIGHT/
                TECHFLAG    TECHNICIAN WHO TOOK FEV + FVC READINGS/
END SCHEMA

TASK NAME       RECORD 16 (PERSONAL)  SCHEMA DEFINITION
RECORD SCHEMA   16  PERSONAL
DOCUMENT        This record type holds the name and national insurance number
                of the man, for identification purposes.
SEQUENCE CHECK  OFF
MAX REC COUNT   1
DATA LIST       FIXED (1)
                /1          CASEID       1 -   8   (A)
                /1          NATINSNO    13 -  20   (A)
                /1          NAME        22 -  39   (A)
                /1          DOBIRTH     40 -  45   (A)
DATE VARS       DOBIRTH     ('DDMMYY')/
VAR SECURITY    CASEID      (20,20 ) / DOBIRTH  (20,20 ) /
VAR LABELS      CASEID      CASE IDENTIFIER/
                NATINSNO    national insurance number/
                NAME        surname and initials/
                DOBIRTH     date of birth/
END SCHEMA
```

APPENDIX D Copy of Heederik D and Miller BG (in press).

In 1987, the Dutch epidemiologist Dick Heederik visited the IOM from Wageningen Agricultural University, the Netherlands. Part of the work he carried out was a collaborative investigation with the author on effects on regression of estimation error in explanatory variables, using as an example data from the present study. Heederik presented some results at a meeting of the Dutch Epidemiological Society in April 1988, and the final text of the presentation has been accepted for publication in the proceedings by the International Epidemiological Association. The final draft is reproduced below with the permission of Dick Heederik.

weak associations in occupational epidemiology: adjustment for exposure estimation error

Dick Heederik,
University Wageningen,
the Netherlands

Brian G. Miller,
Institute of Occupational Medicine,
Edinburgh, Scotland

Correspondence address: Departments of Environmental and Tropical Health and Air Pollution, Agricultural University PO Box 238, 6700 AE Wageningen, The Netherlands.

SUMMARY

Epidemiological studies often estimate the health effects of occupational exposures by multiple regression techniques. The standard theory of regression analysis is based on the assumption that the explanatory variables are known without error, and it has long been known that departures from this assumption will lead to underestimation of the true regression coefficients. In reality, there may be considerable imprecision in the measurement of individuals' exposures to hazards in the workplace, but this is seldom taken into account in analyses. We therefore studied the effect of allowing for imprecision in the exposure estimate with more sophisticated statistical methods, using lung function data from a sample of 348 British miners exposed to mixed coal dust, over an eight year period.

Change in lung function over an eight year period was regressed on the cumulative dust exposure, weight, age and smoking habits. The error in the exposure estimation was assumed to be up to 30% of the total variance of the distribution of the exposure. Adjustment of the regression coefficients of lung function change on dust exposure for the estimation error using linear structural relationships increased the regression coefficient more than threefold compared with those calculated by standard regression analysis. The adjustment led to a change of the coefficients of age as well.

The results indicate that a serious underestimation of the relation between lung function change and occupational exposure may occur, which might lead to false interpretations about the relative importance of the occupational exposure as a determinant

of disease.

INTRODUCTION

Cross-sectional epidemiological surveys of British coal-miners have demonstrated an inverse relationship between forced expired volume in one second (FEV_1) and cumulative exposure to respirable mixed coal-mine dust, independent of the presence of pneumoconiosis[1], suggesting that respirable dust exposure might be related to rate of decline of lung function in excess of that attributable to ageing and smoking. Results have been consistent with similar studies in other coalmining countries[2,3]. The relationship between loss of lung function and cumulative dust exposure has been confirmed in longitudinal analyses[4].

There is much concern about the acceptability of health risks which arise from occupational exposures. Discussion on the magnitude of relative risks of smoking and of exposure to coalmine dust on lung function have aroused particularly heated controversy[5]. However, conclusions based on the magnitude of regression coefficients can be misleading if the independent variables in the regression analysis are subject to random estimation error. This type of error leads to a bias towards zero or underestimation of the regression coefficient, often referred to as attenuation[6]. There are many potential sources of error in the estimation of individuals' exposures to dust in the workplace. Sampling schemes based on periodic measurements of occupationally exposed workers will not usually incorporate sufficient allowance for day-to-day variation in concentrations, nor spatial variations away from the sampling points, nor inhomogeneity of occupational groups. Long term variations may be missed if the sampling points or occasions are sparse. The data on individuals' work patterns may not relate exactly to sampling locations, and work histories taken by interview may suffer from faulty recall. Variations in individuals' patterns of breathing, or lung anatomy, may produce differences in the quantities of dust deposited in the lung, even if ambient concentrations are the same.

Estimates of individual exposures may thus be subject to considerable error from any or a mixture of these sources, and the attenuation of the regression coefficient increases with the amount of error in the estimates[6,7]. In the case of a simple regression with one independent variable measured with error, a simple formula is available to adjust for the error and to provide a "disattenuated" estimate of the regression coefficient. Extensions to adjust for this in the coefficients of other variables in more complicated regression equations, and to allow for error in more than one independent variable, rapidly become algebraically complex[7]. However, such situations are fairly easy to specify as extended regression models using so-called linear structural equations[8].

We describe here an example of such an analysis, where an estimate of the relationship between mixed dust exposure and lung function was adjusted for estimation error in the exposure.

METHODS

Available survey and exposure data

The data analyzed were from one British colliery studied as part of the National Coal Board's Pneumoconiosis Field Research (PFR)[9]. From the records of medical surveys which took place at that colliery in 1970 and 1978, lung function measurements and information on smoking habits were extracted. The highest values from at least two technically satisfactory forced expirations of forced vital capacity (FVC) and forced expiratory volume in one second (FEV) were analyzed.

Data on standing height without shoes, age at survey and weight were available. Men were classified as non-smoker (who had never smoked as much as one cigarette a day for a year), ex-smoker, or smoker on the basis of consistent replies at each survey in 1970, 1974 and 1978. Persons who changed their smoking habits during the time of observation, were defined as intermittent smokers. A few workers whose replies were seriously inconsistent were omitted from the analysis.

Estimates of individual workers' exposures to respirable dust in the PFR were based on environmental measurements of concentrations in defined occupational groups, and on records of the amounts of time each man spent working in each of these groups[9]. For the period before the start of the PFR, or when men transferred from another colliery, employment histories were obtained by an interview at the medical survey. Exposure estimates were calculated for each miner as the product of respirable dust concentrations and times worked, cumulated over various time periods.

The change in lung function was calculated by subtracting the lung function at the first survey from that at the last survey. To exclude possible disturbing effects of lung growth among the younger workers, persons aged 25 or less at the last survey were excluded from the analysis. Suitable data on lung function, smoking and exposure to dust were available for 348 men.

Data Analysis

Loss of lung function was initially analyzed by the application of detailed graphical and tabular analysis. Multiple linear regression analysis was used to relate decline in lung function to respirable dust exposure estimates, with adjustment as necessary for age, height, weight and smoking habits. Models with interaction terms between the smoking habits and the exposure variables were tested, as well as models with quadratic and logarithmic expressions for the exposure. Residuals were examined in detail. These analyses were performed with the Statistical Analysis Software (SAS) System on a Prime computer[10].

The regression analyses was supplemented and extended by further work using the package LISREL[11], to adjust for the attenuation due to estimation error in the exposure variables. L I S R E L allows estimation, by maximum likelihood, of coefficients in linear relationships more complex than that of multiple regression; it includes facilities for adjustment of error in some or all of the independent or explanatory variables. This is achieved by supplying some estimate of the error variances of the variables affected. In using LISREL to examine the effects of random errors in the exposure measurements, we assumed a simple model for the changes in the lung function variables as suggested by ordinary regression analysis. For example for the change in FVC (dFVC) the model eventually selected was:

$$dFVC=b_0 + b_{1j} + b_2x_2 + b_3x_3 + b_4x_4 + e$$

where b_{1j} is the coefficient of the dummy variable identifying the appropriate smoking group, and x_2, x_3, and x_4 are the variables age, weight and the chosen dust exposure. This was augmented by the equation:

$$x'_4= x_4 + f_4$$

indicating that the exposure variable in the regression equation was not observed, but measured with error f_4. A solution to these simultaneous equations can be obtained provided that information is supplied on the magnitudes of the components of variance of x'_4; that is, on how much of the variation in x'_4 is due to the variation in x_4 in the population, and how much due to the random variation represented by the measurement error f_4.

Little hard information is available on the magnitude of the measurement error variance, but previous published work[12] and discussions with occupational hygienists suggested that an assumption that it may contribute between 20% and 40% of the total variance of the observed exposures, as estimated by the cumulations of the products of recorded time worked and ambient measured concentrations, would not be unreasonable in many studies.

The analyses reported below were based on assumptions consistent with these rough guesses.

RESULTS

Details of the study population are presented in table 1. The cumulative dust exposure estimates are shown separately for the periods up to the 1970 survey and that between the 1970 and 1978 surveys, which is labelled "concurrent" to indicate that it coincides with the period over which the losses in FVC and FEV were calculated. These losses and the previous and concurrent cumulative exposures are given in table 2 grouped according to age and smoking habits. The Pearson correlation between the previous and concurrent cumulative exposure for respirable dust was 0.22. Table 3 shows the results of fitting regression models for the losses of FVC and FEV, in terms of the variables age and weight, and dummy variables expressing membership of the four smoking groups (with non-smokers as the reference). Table 4 shows the results of adding either the previous or the concurrent dust exposure to the models in table 3. Models which used the logarithms of the exposures fitted better than linear or quadratic terms in the original exposures. As can be seen from table 4, the previous exposure made the more significant contribution to the models for both dependent variables, with evidence of a stronger effect on FVC than on FEV. Further modelling did not result in better fits than those presented for the models in the tables.

The negative signs of the regression coefficients on dust exposure and age (table 3) imply an increasing rate of loss in lung function in older men, and additionally in those with a higher dust exposure, adjusted for age (table 4). Smokers showed a greater rate of loss than non-smokers and ex-smokers (table 3).

Table 5 shows some of the results obtained from separate analysis of FVC when either previous or concurrent dust exposure was included as x'_4 in the linear structural model. Analyses were performed assuming reliabilities of 1.0, 0.9, 0.8, and 0.7 for the exposure variables, corresponding to assumptions that the variance of the random error contributed 0% to 30% of the total variance of x'_4. A reliability of 1.0 implies the absence of any estimation error and in this case the model reduces to an ordinary multiple regression analysis. Table 5 shows that the regression coefficients of lung function on previous and concurrent cumulative exposure increase as reliability decreases (or as error variance increases) in the LISREL model. For the model containing concurrent dust exposure, decreasing the assumed reliability by 30% resulted in a 44% increase in the magnitude of the concurrent dust exposure coefficient, with only a 2% increase in the age coefficients. The results from a similar analysis including previous dust exposure were even more dramatic, with an assumed reliability of 70% resulting in a 233% increase in the exposure coefficient and a 87% reduction in the age coefficient.

The values of the corresponding t-statistics for the statistical significance of previous dust exposure did not increase with the adjusted estimates, indicating that the standard error was increasing along with the estimate. This reflects the increase in residual error which is always traded-off for bias adjustments. However, there was an interesting drop in the magnitude and significance of the age coefficient as these adjustments took place, suggesting that the imperfect measurement of previous exposure was excessively confounded with age when their effects were estimated without adjustment.

Some further modelling showed that the effect of allowing for estimation error was much greater for the exposures before 1966 than between 1966 and 1970 (table 6).

For all these models, the adjustment for the reliability of the exposures had negligible effect on the coefficients of weight and smoking, and is not shown in tables 5 and 6. Similar results were obtained in the analysis of change in FEV.

DISCUSSION

The purpose of the work reported here was to gain some insight into the effects of errors in the measurement of occupational exposures on estimates of their health effects. This was accomplished using models based on linear structural relationships. The data used were, for convenience, from a single colliery, and originated from less than 350 miners. Nevertheless, multiple regression analyses, of the sort commonly applied to similar epidemiological data, showed a clear effect of mixed coal dust exposure on rate of lung function decline after allowing for smoking, age and weight. The change in FVC showed a more profound effect of dust exposure than the change in FEV, and the relationship was stronger with previously accumulated exposure to dust than with that accumulated during the period over which the change was observed. The results were qualitatively similar to those reported by Love and Miller[4] for FEV.

As expected from the theory, the introduction of an adjustment for error in the measurement of individuals' exposures resulted in an increase in the regression coefficient of the dust exposure variable, with the size of the adjustment becoming greater as lower reliability in the exposure was assumed. The largest adjustment was seen for change in FVC with an exposure variable representing exposures before 1966, when the coefficient after adjustment for an assumed reliability of 70% was three times the unadjusted coefficient.

The effect of the adjustment was strongest on the exposure for the period before 1966, that is the furthest before the observed lung function changes. It is to be expected that these earlier exposure estimates were less reliable than those for the later periods, both because later accumulation of data on times worked was extracted more frequently and more directly from a computerised payroll system, and also because the estimated exposures for the earlier periods include estimates based on the men's recall of employment before they entered the research programme. In addition, it is these exposures which are most highly correlated with age; and the apparent reduction in the importance of the age variable in the adjusted analyses with previous exposure, suggests that measurement error in the exposure variable makes it more difficult, using ordinary regression methods, to disentangle the partial confounding between historical exposure estimates and age. Further work is needed, possibly on other data sets, to identify and disentangle the relative importance of these and other aspects of historical exposures.

We do not underestimate the task of attempting a realistic quantification of the measurement error attached to explanatory variables in epidemiological studies, in order to make the sort of adjustment employed here. An alternative approach to the problem is to characterise individuals' exposures more precisely, but this requires an increase in effort to be sustained throughout the measurement process. The option of making adjustments through techniques such as linear structural equations will often be more feasible in practice. Several computing packages provide solutions to such equations[8], LISREL[11] being perhaps the most general.

Although these findings were based on a small data set, it is clear that they have wider implications. Assessment of the severity of health effects judged from unadjusted regression coefficients may underestimate the strength of the effects, if there is considerable error in the estimation of individuals' exposures. It is possible that statutory limits on such exposures may, in turn, not be sufficiently stringent, although a full discussion of this aspect would have to consider the sources of variation in the processes of measurement by which control is maintained.

Comparison of estimated regression coefficients between the results of different studies should to take into account whether the studies are measuring the exposures with the same precision. Adjustment for measurement error may alter the researcher's perception of the extent to which the effects being investigated are confounded with each other or with nuisance variables.

Even within the same study, comparison of, say, the effects of smoking and of occupational exposure may need to allow for the different precision with which these variables are measured. The situation is made more complicated when more than one relevant explanatory variable is measured with error, because then even the direction of the biasses in the coefficients cannot be known with certainty from general considerations. We see a need for further practical investigations of the magnitudes of attenuation in effects in other data sets.

ACKNOWLEDGEMENTS

The work was carried out with financial aid from the British Coal Corporation and the European Coal and Steel Community (Contract No. 7260-04/025/08).

We want to thank Fintan Hurley and Dr. Michael Jacobsen for their valuable comments and discussions and British Coal for access to the data set used for this study.

REFERENCES.

1. Rogan JM, Attfield MD, Jacobsen M, Rae S, Walker DD, Walton WH. Role of dust in the working environment in development of chronic bronchitis in British coalminers. Br J Ind Med 1973; 30: 217 - 26.

2. Reichel G & WT Ulmer. Results obtained by the Bochum research group. In: Chronic bronchitis and occupational dust exposure. Bonn. Deutsche Forschungsgemeinschaft 1978; 224 - 91.

3. Hankinson JL, RB Reger, RP Fairman, NL Lapp, WKC Morgan. Factors influencing expiratory flow rates in British coal miners. In: Walton WH, ed. Inhaled Particles IV. Oxford, Pergamon Press, 1977; 735 - 37.

4. Love R & B Miller. Longitudinal study of lung function in coal-miners. Thorax 1982; 37: 193 - 7.

5. Morgan WKC. On dust, disability and death. Am Rev Resp Dis 1986; 134: 639- 41.

6. Snedecor GW & WG Cochran. Statistical methods. Seventh edition. IOWA, State University Press, 1980.

7. Cochran WG. Errors of measurement in statistics. Technometrics 1968; 10: 637 - 66.

8. Everitt BS. An introduction to latent variable models. London, Chapman and Hall, 1984.

9. Hurley JF, WP Alexander, DJ Hazledine, M Jacobsen & WM Maclaren. Exposure to respirable coalmine dust and incidence of progressive massive fibrosis. Br J Ind Med 1987; **44**: 661-672.

10. SAS Institute Inc. SAS User's Guide: Statistics. Version 5 Edition. Cary NC, SAS Institute Inc. 1985.

11. Jöreskog & Sörbom. LISREL VI. Analysis of linear Structural Relationships by the method of Maximum Likelihood. Sweden, University of Uppsala, 1986.

12. Heederik DJJ, H Kromhout, JSM Boley. Variability of exposure measurements -concequences in occupational epidemiology. 5th International Symposium on Epidemiology in Occupational Health. Los Angeles, 1986, 9 - 11 September, University of California.

Table 1. General characteristics and dust exposure of men at the end of the follow-up period. Percentages and standard deviations between parentheses (n=348)

	mean	(sd)	n	(%)
age	48.2	(9.0)		
standing height (cm)	170.3	(6.4)		
weight (kg)	75.6	(11.5)		
non-smokers			57	(16.4)
ex-smokers			20	(5.7)
intermittent smokers			60	(17.3)
smokers			211	(60.6)
dust exposure until 1970 (ghr/m^3)	48.4	(26.7)		
concurrent dust exposure	19.3	(11.2)		

Table 2. FVC and FEV_1 losses over an eight year period and dust exposure according to age and smoking habits

smoking habit	age	<40	40-49	>50	all
non-smokers	n	18	18	21	57
	dFVC	-414	-367	-543	-446
	dFEV	-416	-356	-400	-391
	PD	14.7	45.8	62.7	42.2
	CD	21.8	21.7	17.5	20.2
intermittent smokers		18	16	26	60
		-263	-825	-898	-688
		-219	-747	-621	-534
		18.2	44.8	60.1	43.4
		21.5	20.3	19.1	20.2
ex-smokers		1	5	14	20
		-250	-470	-375	-393
		-200	-330	-361	-345
		41.2	44.5	51.7	49.4
		27.0	22.4	14.5	17.1
smoker		35	60	116	211
		-304	-553	-544	-506
		-373	-512	-469	-465
		19.0	45.5	64.3	51.4
		19.4	19.7	18.6	19.1
all men		72	99	177	348
		-321	-559	-582	-522
		-343	-512	-474	-458
		18.0	45.4	62.5	48.4
		20.7	20.3	18.2	19.3

n number of observations
dFVC mean change in Forced Vital Capacity (l)
dFEV mean change in Forced Expiratory Volume in one second (l)
PD mean previous cumulative dust exposure (ghr/m^3)
CD mean concurrent cumulative dust exposure (ghr/m^3)

Table 4. Regression of the change in FVC (ml) and FEV_1 (ml) on exposure. All exposure variables were tested in a separate analysis

logarithm of exposure	FVC b*	FVC t⁺	FEV b*	FEV t⁺
previous exposure	-81.5	-2.33	-57.7	-2.00
concurrent exposure	-40.9	-1.38	-11.5	-0.47

* regression coefficient after allowing for age, weight and smoking habits
\+ t-value

Table 3. Regression models for the change in FVC (ml) and FEV_1 (ml) on age, weight and smoking habits (n=348)

	FVC		FEV_1	
adjusted R-square (%)	12.6		5.8	
Residual SD (ml)	413		340	
	coefficient	t-value	coefficient	t-value
age (yr)	-13.6	-5.4**	-6.7	-3.2**
weight (kg)	-7.6	-3.8**	-5.0	-3.1**
mean difference from non-smoker for:				
ex-smoker	-130.1		-81.5	
intermittent smoker	-254.7		-155.6	
smoker	-60.9		-82.3	
constant	780.0		316.7	

* $0.05 < p < 0.10$
** $p < 0.05$

Table 5. Regression of FVC-change on exposure calculated with LISREL for different reliabilities of the exposure.

Reliability	logarithm previous mixed dust exposure b*	t+	age b*	t+	logarithm concurrent mixed dust exposure b*	t+	age b*	t+
1.0	-81.5	-2.3	-8.2	-2.4	-41.0	-1.4	-14.0	-5.6
0.9	-101.0	-2.3	-6.9	-1.8	-45.7	-1.4	-14.1	-5.6
0.8	-131.7	-2.3	-4.9	-1.1	-51.6	-1.4	-14.2	-5.6
0.7	-189.5	-2.3	-1.1	-0.2	-59.2	-1.4	-14.3	-5.6

* regression coefficient after allowing for age, weight and smoking habits
\+ t-value

Table 6. Regression of FVC-change on previous exposure by means of LISREL with different reliabilities of the exposure.

Reliability	logarithm of the exposure: before 1966 b*	t+	from 1966 - 1970 b	t
1.0	-51.0	-2.1	-77.0	-2.3
0.9	-66.0	-2.1	-86.4	-2.3
0.8	-91.7	-2.1	-96.7	-2.3
0.7	-151.6	-2.0	-115.0	-2.3

* regression coefficient after allowing for age, weight and smoking habits
\+ t-value