

https://dspace.org/

Combined effect of dust exposure and tobacco smoking on the occorence of lung disease in South African gold miners

Hnizdo, Eva

1991-04
https://hdl.handle.net/10539/26132

Downloaded from a DSpace repository. https://dspace.org/

COMBINED EFFECT OF DUST EXPOSURE AND TOBACCO SMOKING ON THE OCCURRENCE OF LUNG DISEASES IN SOUTH AFRICAN GOLD MINERS

Eva Hnizdo

Epidemiology Research Unit

Medical Bureau for Occupational Diseases

Department of National Health and Population Development

A Thesis submitted to the Faculty of Medicine

University of the Witwatersrand, Johannesburg

for the Degree of Doctor of Philosophy

April 1991

Johannesburg

ABSTRACT

The objective of this thesis is to investigate whether tobacco smoking modifies the effect of gold mining dust exposure on the occurrence of lung diseases known, or suspected, to be causally associated with the gold mining occupation. In particular, the purpose is to ascertain whether the combined effect of dust and smoking on these diseases' occurrences is additive (i.e. the effects are independent), or synergistic (i.e. smoking increases the effect of dust), or antagonistic (i.e. smoking decreases the effect of dust).

To establish the presence of synergism or antagonism, the observed occurrence of diseases is tested for deviances from the expected, derived from a statistical model which assumes absence of synergism or antagonism, namely the additive relative risk model. The general relative risk model, which includes the additive model as a special case, is applied to establish the best functional relationship between the effect measures (e.g. the relative risks).

The lung diseases considered in the present thesis are: simple chronic bronchitis, lung function impairment, death from chronic obstructive lung disease (COLD), and death from lung cancer.

In the study of simple chronic bronchitis two study samples are used: 1) A sample of 827 subjects who were selected from a mixed non–mining and mining population of Carletonville on the Witwatersrand for a study of chronic bronchitis. 2) A sample of 2209 white gold miners and 483 non–miners who were studied for respiratory impairment in 1968–71 when they were 45–54 years old. The miners started their mining exposure in the 1940s and worked on average 24

years in dusty occupations in gold mines and had minimal exposure in other type of mines.

Results from both study samples indicate that the mining dust exposure and tobacco smoking are associated with a significant increase in the prevalence of chronic bronchitis, and that the combined effect is synergistic. The age adjusted odds ratio, estimated from the first study sample, from the comparison of men not exposed to either exposure with those exposed to dust only is 2.77 (95% confidence interval (CI) 1.4–5.5), with those exposed to smoking only 4.26 (2.0–9.2), and with those exposed to dust and smoking 11.8 (1.8–73.4). In the second study sample, the corresponding odds ratios are 9.6 (1.3–72), 5.2 (0.7–39) and 41.5 (5.7–299) for the response to respiratory questionnaire, and 5.5 (1.7–18), 8.9 (2.7–29), and 17.2 (4.7–62) for measured mucus.

In the study of lung function impairment, the data from 2209 white gold miners and 483 non–miners are re–analyzed. The method of analysis and the results obtained are as follows:

1) The dose–response trend between individual spirometric measurements and cumulative dust are examined. The results indicate that the effects of dust and smoking on the individual lung function measurements are additive.

2) The pattern and degree of ventilatory impairment as reflected in the spirogram is examined. The subjects are grouped according to their lung function profiles as 'normal', minimal, moderate and marked obstructive, marked obstructive with restriction, and pure restrictive. Each profile group is compared to the normal group for exposure prevalences. The two marked obstructive impairment profiles were mainly associated with the effect of smoking and the combined effect of dust and smoking.

To study the mortality from COLD and lung cancer, the 2209 gold

miners, studied in 1968–71 for respiratory impairment, were followed up to 31 December 1986, at which date their vital status was established. During the follow–up period, 945 miners died. According to the 9th revision of the International Classification of Diseases (ICD), there were 66 deaths from COLD and 77 deaths from lung cancer. The results for the associations with dust and smoking are as follows:

1) Death from COLD is most strongly associated with cumulative dust exposure before 1950 and with years of cigarette smoking as established at the 1968–71 examination. Cumulative dust exposure, measured as shifts weighted by surface area of respirable dust, calculated to 1950 ranges from 0 to 34 thousands, and the adjusted relative risk (RR) associated with one unit of cumulative dust of one thousand is 1.097 (1.044–1.153). The combined effect of dust and smoking is found to be synergistic in miners exposed to high dust and to 20 or more years of smoking. All those who died of COLD were tobacco smokers.

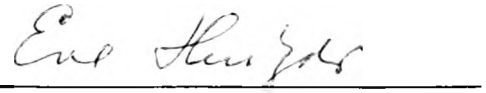
2) Death from lung cancer is most strongly associated with cumulative dust exposure to the start of follow–up in 1968–71, and with cigarette equivalent pack–years. The CDI calculated to the start of follow–up ranges from 7.3 to 79.44, and the adjusted RR associated with one unit of CDI is 1.023 (1.005 – 1.042). The combined effect of dust and smoking is found to be synergistic in the highest exposure levels for dust and smoking.

In conclusion, the results indicate that for the lung diseases studied in this thesis the combined effect of gold mining dust exposure and tobacco smoking is synergistic.

DECLARATION

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy in the Faculty of Medicine in the University of the Witwatersrand, Johannesburg.

This thesis has not been submitted for any degree or examination in any other University.



April, 1991

I declare that this research project has been approved by the Committee for Research on Human Subjects of the University of the Witwatersrand, and the Protocol No. is 5/8/89.

DEDICATION

To the memory of my father Bohuslav Smalcl (1905–1990) and my mother Vlasta Smalclova (1910–1990).

ACKNOWLEDGEMENT

The most important factor which allowed this work was the availability of unique epidemiological data sets collected on the South African gold miners. I am greatly indebted to Dr F.J. Wiles and Dr M.H. Faure for collecting the dataset on 2209 white gold miners, which provided the primary source of data in the present work. Dr F.J. Wiles also initiated the collection of the mortality data, but from the beginning I was responsible for the collection, preparation and analysis of the mortality data.

For the supervision of the thesis I thank:

Dr G.K. Sluis–Cremer for his advice, critical comments and constant encouragement through out the whole period of my work;

Professor Paul Fatti for his critical comments on Chapters 3, 6 and 7;

Professor Anthony Davies for his critical comments on the clarity of presentation.

I would also like to thank:

Dr P.A. Hessel for giving me the initial impetus to enroll for the Ph.D. degree;

Dr E. Baskind for his advice with the lung function impairment profiles and for his critical comments on Chapter 5;

Professor I.T.T. Higgins for his critical comments on Chapters 5 and 7;

The referees who reviewed the papers based on the work of Chapters 5, 6 and 7, as their criticism of the various shortcomings of each paper helped to improve the content of this thesis.

During the time that I worked on the present thesis I was employed by

the Epidemiology Research Unit associated with the Medical Bureau for Occupational Diseases, Department of National Health and Population Development.

PUBLICATIONS

Publications based on the work of the individual chapters are:

1) Chapter 5:

Eva Hnizdo, E. Baskind, G.K. Sluis–Cremer.

The combined effect of silica dust exposure and tobacco smoking on the prevalence of respiratory impairments in gold miners.

Scan J Work Environ Health 1990; 16:411–422.

2) Eva Hnizdo, P.A. Hessel, G.K. Sluis–Cremer.

Chronic airflow limitation in South African gold miners.

Letter to the editor. Am Rev Respir Dis (submitted).

3) Chapter 6:

Eva Hnizdo

Combined effect of silica dust and tobacco smoking on mortality from chronic obstructive lung disease in gold miners.

Br J Ind Med 1990; 47:656–664.

4) Chapter 7:

Eva Hnizdo and G.K. Sluis–Cremer

Silica exposure, silicosis and lung cancer: a mortality study of South African gold miners.

Br J Ind Med 1991; 1:53–60.

5) Chapters 2, 5, and 6:

Eva Hnizdo

Health risk factors among white South African gold miners: dust,

smoking and chronic obstructive lung disease.

South African Medical Journal (submitted).

Previous publications in which the combined effect of occupational dust and tobacco smoking was investigated in South African asbestos miners, using similar methodology to that described in this thesis are:

6) Eva Hnizdo and GK Sluis–Cremer.

Effect of tobacco smoking on the presence of asbestosis at postmortem and on the reading of irregular opacities on roentgenograms in asbestos–exposed workers.

Am Rev Respir Dis 1988; 138:1207–1212.

7) Eva Hnizdo and Sluis–Cremer GK. Response to effect of tobacco smoking on the presence of asbestosis at postmortem and on the reading of irregular opacities on roentgenograms in asbestos–exposed workers.

Am Rev Respir Dis 1989; 139:1568–1569.

CONTENTS Page No.

Abstract i

Declaration v

Dedication vi

Acknowledgement vii

Publications ix

List of tables xvi

List of figures xxi

Chapter 1. Introduction 1

1.1 General introduction 1

1.2 Objectives of the thesis 6

1.3 Epidemiological context 8

1.4 General scientific perspective 11

Chapter 2. Gold mining and associated lung diseases in South Africa 12

2.1 Chronological developments of the exposures on the mines 12

2.2 Lung diseases associated with gold mining – definition and epidemiological evidence 15

2.2.1 Simple chronic bronchitis 16

2.2.2 Lung function impairment 19

2.2.3 Death from COLD 24

2.2.4 Death from lung cancer 26

Chapter 3. Description of the analytical methods used 30

3.1 Measures of disease occurrence, measures of association, and study designs 30

3.1.1 Measures of disease occurrence 30

3.1.2 Basic measures of association 31

3.1.3 Basic study designs and methods of analysis 32

3.2 Interaction and synergism – definitions 40

3.3 Assessment of interaction 48

3.3.1 Test for additivity of relative risks 49

3.3.2 Test for multiplicativity of relative risks 51

3.3.3 Visual assessment of presence of interaction 52

3.4 General Relative Risk Models (GRRM) 54

3.4.1 Specification of the GRRM 54

3.4.2 GRRM with power transformation 56

3.4.3 Problems associated with GRRM 58

3.5 Estimation of regression coefficients 59

3.5.1 Unmatched case–referent study 60

3.5.2 Matched case–referent study 60

3.5.3 Follow–up (cohort) study 61

3.6 Statistical inference 62

3.7 General modelling strategy 63

3.8 Estimation of attributable risk 65

Chapter 4. Simple chronic bronchitis 67

4.1 Objectives of chapter 4 67

4.2 The Carletonville study 68

4.2.1 Study subjects 68

4.2.2 Combined effect of mining dust and smoking on the presence of simple chronic bronchitis 69

4.2.3 Conclusions from the analysis 73

4.3 The Wiles and Faure study 77

4.3.1 Study subjects and methods 77

4.3.2 Combined effect of mining dust and smoking on the presence of simple chronic bronchitis 84

4.3.3 Combined effect of mining dust and smoking on abnormal sputum production 94

4.4 Discussion 101

Chapter 5. Lung function impairment 103

5.1 Objectives of chapter 5 103

5.2 Study subjects and methods 104

5.3 Results 111

5.3.1 Analysis of individual lung function measurements 111

5.3.2 Analysis of lung function profiles 115

5.4 Discussion 131

Chapter 6. Mortality from chronic obstructive lung disease 142

6.1 Objectives of chapter 6 142

6.2 Study subjects and methods 143
6.3 Nested case–referent study 146
6.4 Cohort data analysis 155
6.5 Discussion 167

Chapter 7. Mortality from lung cancer 172
7.1 Objectives of chapter 7 172
7.2 Study subjects and methods 172
7.3 Selection of best predictors of death from lung cancer 176
7.4 Assessment of the combined effect of dust and smoking 179
7.5 Association between lung cancer and silicosis 188
7.6 Association between silicosis and tobacco smoking 189
7.7 Discussion 190

Chapter 8. General conclusion 195
8.1 The changing pattern of lung diseases in the gold miners 195
8.2 Findings of this thesis in the context of gold miners health 197
8.2.1 Simple chronic bronchitis 197
8.2.2 Lung function impairment 198
8.2.3 Mortality from COLD 201
8.2.4 Mortality from lung cancer 201
8.2.5 Other evidence from mortality studies 202
8.3 Implication for preventive measures 204

8.4 Observations made from using the GRRM 206

References 208

Appendix A 226

LIST OF TABLES

Table 3.1 Data layout for a cohort study p. 33

Table 3.2 Data layout for a case–control study p. 37

Table 4.1 Prevalence of chronic bronchitis, by presence of gold mining dust and smoking status, stratified by age p. 75

Table 4.2 Predicted frequencies of chronic bronchitis, estimated from different relative risk models, fitted to age stratified data p. 76

Table 4.3 Comparison of miners with different reasons for attending annual examination, for cumulative dust exposure and lung function tests p. 82

Table 4.4 Characteristics of miners and non–miners p. 83

Table 4.5 Relations for chronic bronchitis, smoking (yes/no) and gold mining dust exposure (non–miner/miner) p. 86

Table 4.6 Dose–response relations for chronic bronchitis p. 90

Table 4.7 Frequencies and odds ratios for chronic bronchitis, by smoking and cumulative dust exposure p. 92

Table 4.8 Relations for abnormal sputum production, smoking (yes/no) and gold mining dust exposure (non–miners/miners) p. 96

Table 4.9 Dose–response relations for abnormal sputum production p. 97

Table 4.10 Frequencies and odds ratios for abnormal sputum production, by smoking and cumulative dust exposure p. 99

Table 5.1 Estimated regression coefficients for prediction of FEV_1, VC and FEF_{25-75} p. 112

Table 5.2 Estimated loss of lung function attributable to silica dust and to

tobacco smoking, according to percentiles of each exposure p. 113

Table 5.3 Characteristics of the lung function profile groups p. 116

Table 5.4 Lung function profile groups frequencies (N) and the observed ORs, according to smoking habits and cumulative dust exposure p. 118

Table 5.5 Results from the additive RR model, Eq. (5.1), fitted with indicator smoking variables and an ordinal dust variable to the marked obstructive profiles p.120

Table 5.6 Results from the additive RR model fitted with the categorized exposure variables, for each obstructive profile p. 123

Table 5.7 Summary table of the log likelihood ratio tests of no interaction p. 124

Table 5.8 Etiological fractions for minimal obstructive profile p. 125

Table 5.9 Etiological fractions for moderate obstructive profile p. 126

Table 5.10 Etiological fractions for marked obstructive profile p. 127

Table 5.11 Etiological fractions for marked obstructive profile with restriction p. 128

Table 5.12 Etiological fractions for marked obstructive (4+5), benefit examinations were excluded p. 129

Table 5.13 Summary of the etiological and preventable fractions p. 130

Table 6.1 Comparison of the COLD cases and referents p. 147

Table 6.2 Goodness–of–fit statistic for the multiplicative RR model, RR_m, fitted with different cumulative dust variables, adjusted for smoking (CYSM) p. 149

Table 6.3 Goodness–of–fit statistic for the multiplicative RR model, RR_m, fitted with different smoking variables, adjusted for dust (CD40) p. 149

Table 6.4 Two–way frequency table for death from COLD, by categorized cumulative dust to 1949, and years of cigarette smoking p. 151

Table 6.5 Goodness–of–fit statistic, parameter estimate, and X^2–test of significance, estimated from RR model, RR_p, witted with changing λ, and categorized or continuous cumulative dust variables p. 152

Table 6.6 Goodness–of–fit statistic, parameter estimate, and X^2–test of significance, estimated from RR model, RR_p, fitted with changing λ, and categorized or continuous smoking variables p. 152

Table 6.7 Goodness–of–fit statistic, parameter estimates and X^2–test of significance, estimated from RR model, RR_p, fitted with different λ, and categorized or continuous exposure variables p. 154

Table 6.8 Number of deaths from COLD (d) and person–years (P–Y), by age and year at risk p. 157

Table 6.9 Number of deaths from COLD (d) and person–years (P–Y), by cumulative dust exposure and years of cigarette smoking p. 158

Table 6.10 Estimated regression coefficients for dust and smoking, adjusted for age and year, estimated from the multiplicative RR model, by Poisson regression, and using the program GLIM p. 160

Table 6.11 Comparison of parameter estimates obtained from programs GLIM and MINUIT from fitting the mortality data p. 161

Table 6.12 Parameter estimates and goodness–of–fit statistic estimated from the RR model, RR_p, fitted with different λ p. 162

Table 6.13 Observed death rates per 1 000, h, and death rates predicted from RR model, RR_p, for given λ p. 163

Table 7.1 Comparison of lung cancer cases with the rest of the cohort p. 177

Table 7.2 Results from the Cox's proportional hazards model applied to select best predictors of lung cancer p. 178

Table 7.3 Number of deaths (d) from lung cancer and person–years (P–Y), by cumulative dust exposure and cigarette equivalent pack–years p. 180

Table 7.4 Number of deaths (d) from lung cancer and person–years (P–Y), by age and year at risk p. 181

Table 7.5 Results from RR model, Eq. 6.1, for changing λ. Parameter estimates for dust and smoking are adjusted for age and year at risk p. 183

Table 7.6 Observed and predicted death rates/1 000, h, estimated from the general relative risk model with different value of λ p. 184

Table 7.7 Association between death from lung cancer and silicosis of hilar glands p. 189

Table 8.1 Exposure profile of the categorized causes of death p. 203

Table A.1 Simple chronic bronchitis – Carletonville study.
Parameter estimates (E) and standard errors (±SE) estimated from the GRRM, Eq. (4.1) p. 227

Table A.2 Simple chronic bronchitis – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the GRRM, Eq. (4.2), with dichotomous exposure variables p. 227

Table A.3 Simple chronic bronchitis – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the GRRM, Eq. (4.2), with individual exposure variables p. 228

Table A.4 Simple chronic bronchitis – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the GRRM, Eq. (4.2), with categorized or continuous exposure variables p. 229

Table A.5 Simple chronic bronchitis – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the Thomas' modified model, RR_{MT} p. 230

Table A.6 Abnormal sputum production – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the

Table 7.4 Number of deaths (d) from lung cancer and person–years (P–Y), by age and year at risk p. 181

Table 7.5 Results from RR model, Eq. 6.1, for changing λ. Parameter estimates for dust and smoking are adjusted for age and year at risk p. 183

Table 7.6 Observed and predicted death rates/1 000, h, estimated from the general relative risk model with different value of λ p. 184

Table 7.7 Association between death from lung cancer and silicosis of hilar glands p. 189

Table 8.1 Exposure profile of the categorized causes of death p. 203

Table A.1 Simple chronic bronchitis – Carletonville study.
Parameter estimates (E) and standard errors (±SE) estimated from the GRRM, Eq. (4.1) p. 227

Table A.2 Simple chronic bronchitis – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the GRRM, Eq. (4.2), with dichotomous exposure variables p. 227

Table A.3 Simple chronic bronchitis – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the GRRM, Eq. (4.2), with individual exposure variables p. 228

Table A.4 Simple chronic bronchitis – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the GRRM, Eq. (4.2), with categorized or continuous exposure variables p. 229

Table A.5 Simple chronic bronchitis – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the Thomas' modified model, RR_{MT} p. 230

Table A.6 Abnormal sputum production – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the

GRRM, Eq. (4.2), with dichotomous exposure variables p. 231

Table A.7 Abnormal sputum production – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the GRRM, Eq. (4.2), with individual exposure variables p. 232

Table A.8 Abnormal sputum production – Wiles and Faure study.
Parameter estimates (E) and standard errors (±SE) estimated from the GRRM, Eq. (4.2), with categorized exposure variables p. 232

LIST OF FIGURES

Fig 3.1 Hypothetical set of sufficient causes arising from two known causes p 45

Fig 3.2 Trends in the RR that can be observed with two risk factors p 53

Fig 4.1 Observed odds ratios for chronic bronchitis and cumulative dust exposure, by smoking habits p. 93

Fig 4.2 Observed odds ratios for abnormal sputum production and cumulative dust exposure, by smoking habits p. 100

Fig 5.1 Individual lung function measurements, plotted against cumulative dust exposure (RSA x years / 10 000), according to smoking habits p. 114

Fig 5.2 Observed and predicted odds ratios for individual lung function profiles, plotted against cumulative dust exposure (RSA x years / 10 000), according to smoking habits p. 121

Fig 5.3 The risk approportionment for the cases with marked obstructive impairment profile (n=402) p. 138

Fig 5.4 Percent surviving during the follow–up period, according to lung function impairment profile p. 139

Fig 5.5 Percent surviving during the follow–up period, according to smoking habits p. 140

Fig 5.6 Percent surviving during the follow–up period, according to cumulative dust exposure quartiles p. 141

Fig 6.1 SMRs for death from COLD by cumulative dust exposure, and by years of smoking p. 159

Fig 6.2 Observed death rates per 1000 person–years for COLD p. 164

Fig 6.3 Predicted death rates per 1000 person–years for COLD, estimated from the GRRM with λ=0 p. 165

Fig 6.4 Predicted death rates per 1000 person–years for COLD, estimated from the GRRM with λ=0.2 p. 165

Fig 6.5 Predicted death rates per 1000 person–years for COLD, estimated from the GRRM with λ=0.5 p. 166

Fig 6.6 Predicted death rates per 1000 person–years for COLD, estimated from the GRRM with λ=1 p. 166

Fig 7.1 SMRs for death from lung cancer by cumulative dust exposure, and by cigarette equivalent pack–years p. 182

Fig 7.2 Observed death rates per 1000 person–years for lung cancer p. 185

Fig 7.3 Predicted death rates per 1000 person–years for lung cancer, estimated from the GRRM with λ=0 p. 186

Fig 7.4 Predicted death rates per 1000 person–years for COLD, estimated from the GRRM with λ=0.2 p. 186

Fig 7.5 Predicted death rates per 1000 person–years for COLD, estimated from the GRRM with λ=0.5 p. 187

Fig 7.6 Predicted death rates per 1000 person–years for COLD, estimated from the GRRM with λ=1 p. 187

CHAPTER 1

INTRODUCTION

1.1. General introduction

Tobacco smoking is a major risk factor in the development of chronic lung diseases and lung cancer (US DHHS. 1982; US DHHS. 1984). In addition. a number of occupational exposures have been also established as being associated with increased risk of chronic lung diseases and lung cancer (reviewed by Selikoff and Lee. 1978; Higgins. 1973; Becklake. 1985). Tobacco smoking may potentiate or add to the effect of workplace exposure, and the combined effects of cigarette smoking and occupational exposure are therefore a matter of clinical as well as public health significance (US DHHS, 1985).

The combined effect of two exposures is considered to be synergistic in the public health sense when the risk of the disease from the joint effects of two exposures increases above the additive effect of independent exposures (Rothman, 1974; Blot and Day, 1979).

Traditionally. in epidemiological studies. the presence of interaction was assessed by the logistic regression model. which assumes that the risk factors combine in a multiplicative fashion to estimate the risk of the disease (Breslow and Day. 1980; Kleinbaum *et al.*, 1982). However. as more knowledge was gained. mainly through epidemiological studies of lung cancer. the possible underlying pattern of disease causation became of interest (Armitage and Doll, 1961; Peto. 1977; Saracci. 1977). It then became apparent that there are limitations to the

application of the logistic model for the assessment of the joint effect of two or more risk factors in epidemiological studies (Greenland, 1979; Rothman, 1974 and 1976).

The recent development of the generalized relative risk theory, which enables one to apply a wider range of mathematical models of the relative risk to the data, extended the choice of the mathematical tools available to the epidemiologist (Thomas, 1981; Prentice and Self, 1983). Mathematical models are now being proposed in the literature which are designed for increased flexibility in modelling the functional relation between the response variable and each of the individual risk factors, and also of the combined effect of two or more risk factors (Thomas, 1981; Guerrero and Johnson, 1982; Breslow and Storer, 1985; Lubin and Gaffey, 1988).

In addition, methods for quantifying the risk attributable to a specific exposure and to the combined effect of two exposures have been proposed (Miettinen, 1974; Walker, 1981). In recent years, the methods have been extended to the estimation of the population attributable risk for multiple risk factors, using case–control data (Whittemore, 1983; Denman and Schlesselman, 1983; Bruzzi *et al.*, 1985; Kjuus *et al.*,1986).

An integral part of the development of the analytical techniques is their application to different types of diseases and exposures in order to assess their practical value and to uncover their shortcomings. Most studies of the joint effect of mining exposure and tobacco smoking have been done on lung cancer in asbestos miners (Doll, 1955; Berry *et al.*, 1972; Hammond *et al.*, 1979; Liddell *et al.*, 1984) and uranium miners (Whittemore and McMillan, 1983), and on respiratory dysfunction in coal miners (as reviewed by Marine *et al.*, 1988). There is very little epidemiological evidence on the combined effects of silica dust and tobacco smoking on chronic lung diseases occurring in gold miners (Holman *et al.*,1987).

The gold mining occupation generally entails an exposure to occupational dust with a high content of free crystalline silica. The diseases reported to be associated with the mining occupation are simple chronic bronchitis (Sluis–Cremer *et al.* 1967; Wiles and Faure, 1977), chronic airway obstruction (Wiles and Faure, 1977), emphysema (Becklake *et al.*, 1987; Hnizdo *et al.*, 1991), silicosis (Beadle, 1971) and lung cancer (Wyndham *et al.*, 1986; Goldsmith and Guidotti, 1986). The chronic nonmalignant pulmonary conditions characterized by airway dysfunction are often grouped together as the chronic obstructive lung disease (COLD) (Burrows, 1981; Pride, 1984). Tobacco smoking is also prevalent among the gold miners, and is known to be the main causal risk factor for chronic obstructive lung disease and lung cancer. Thus, the question arises whether these two harmful exposures act synergistically. Epidemiological evidence on the presence of synergism can be related to the public health and to individual patient health management, and may contribute to the understanding of the underlying biological mechanism of disease causation. The knowledge of attributable risk for each exposure is important for the development of intervention programmes and compensation policies.

In addition, the lung diseases occurring in the gold miners provide useful data for the assessment and testing of the above mentioned newly developed epidemiological methods as they differ in several important aspects. For example:
(1) The diseases differ in the way in which the mining dust and tobacco smoking are involved in the disease causation process. In the case of chronic obstructive lung disease, tobacco smoking is considered to be the main causal risk factor. Opinion is divided, however, on the role that the present day occupational dust exposures on the mines plays in the aetiology of these diseases. Some researchers argue that dust on its own can cause a clinically significant disability (Soutar *et al.*, 1986; Marine *et al.*, 1988), whereas others maintain that occupational dust is

merely adding to the damage already done by smoking, and that dust on its own does not produce a serious respiratory disability leading to increased mortality (Burrows, 1981; Morgan, 1986; Higgins, 1973). On the other hand, pneumoconiosis is caused by an occupational exposure to mining dust, and the question of whether tobacco smoking modifies the effect of the mining exposure, through the impairment of lung clearance mechanisms or the immunological system, is still discussed in the literature (Kilburn, 1981; Weiss, 1984; Hnizdo *et al.*, 1988 and 1989, in connection with asbestosis), (Schuler and Rutner, 1986, in connection with silicosis).

(2) Another aspect of the chronic lung disease occurring in miners is ascertainment. For example, the presence of simple chronic bronchitis is evaluated by a questionnaire, the response to which depends to a large extent upon the subject's perception of his lung health. The measure of the response is a dichotomous variable, which reflects whether a syndrome of an underlying disease is present or absent. On the other hand, the decision whether a lung function impairment is present is usually based upon the result of several lung function measurements which are measured on a continuous scale. The presence of pneumoconiosis is ascertained either by a radiological reading or by an autopsy examination. In general, the radiological reading may involve a large misclassification error, which may also be influenced by tobacco smoking.

(3) The prevalence of disease or death in a population may also determine whether the epidemiological methods can be applied to determine the joint action of dust and tobacco smoking.

The most important factor which allowed this work was the availability of epidemiological databases collected over many years on the South African mining population. South Africa has a large workforce of miners employed on gold mines. Since about 1920, each gold miner who works in a scheduled dusty

occupation has been compelled by law to undergo an annual medical examination. The Medical Bureau for Occupational Diseases, Johannesburg (MBOD) keeps a medical file on each miner, containing information relating to each man's health status, smoking history (recorded annually since 1960) and detailed occupational history. The average dust exposures associated with different occupational categories in the gold mines were measured by Beadle (1971). In addition, all the mine workers have been offered an autopsy for compensation purposes, if their next of kin agreed, since 1954; these have been computerized since 1975 (Hessel *et al.*, 1987).

The Epidemiology Research Unit (ERU), associated with the MBOD, has been concerned over many years with studies of the diseases associated with occupational exposure on South African mines. During the past twenty years, several large studies have been conducted to assess the dose–response relation between mining exposure and occupational lung diseases. The availability of the data on smoking, mining exposure and diagnosis of the diseases for compensation or research purposes obtained during miners' lives and at autopsy examinations, provide a unique opportunity for the assessment of the combined effect of tobacco smoking and mining exposure on the chronic lung diseases occurring in miners. However, the medical and occupational records for black miners, who until recently were mostly migrant workers, are not very reliable and thus most of the past studies were done on white miners. This situation is changing, however.

1.2 Objectives of the thesis

The objective of the present thesis are:

(1) To establish whether the joint exposure to gold mine dust with a high content of free crystalline silica and tobacco smoking act synergistically in the public health sense. to increase the occurrence of the following lung diseases or conditions:

a) simple chronic bronchitis

b) lung function impairment

c) death from chronic obstructive lung disease

d) death from lung cancer.

In preliminary analyses of several independent data sets the association between the occurrence of silicosis and tobacco smoking was not found to be statistically significant and thus a specific analysis of silicosis will not be included in this thesis, but when an opportunity arises the association will be examined.

2) To establish the attributable risk for mining dust and tobacco smoking for the above diseases.

3) To evaluate the applicability of the statistical techniques, developed for the assessment of the combined effect of two exposures to the problem of chronic lung diseases occurring in gold miners. In particular, the general relative risk model proposed by Breslow and Storer (1985) for the assessment of the combined exposures will be applied and its applicability assessed.

To assess the combined effect of the two exposures use is made of existing data bases collected at the MBOD for the purpose of assessing the dose–response relationship between the individual diseases and gold mining exposure.

The databases from the following studies will be re–analyzed:

1) A population–based study of simple chronic bronchitis, which was conducted in 1965 in a mining town (Carltonville) on the Witwatersrand, in order to determine whether mining exposure increases the prevalence of simple chronic bronchitis (Sluis–Cremer *et al.*, 1967).

2) A study of 2209 white gold miners and 483 population controls conducted from 1968 to 1971 in order to establish the presence of a dose–response relation between simple chronic bronchitis and chronic obstructive lung disease, and underground gold mining exposure (Wiles and Faure, 1977).

3) A follow–up study of 2209 white gold miners included in study 2. The miners were followed up to 31 December 1986, when their vital status and the cause of death were established. The purpose of this study was to establish an association between the presence of simple chronic bronchitis at the time of the first examination and mortality from chronic obstructive lung disease during the follow–up period (Wiles and Hnizdo, 1991).

These studies will be used for analysis of various lung diseases as follows:

Studies 1) and 2) for analysis of simple chronic bronchitis;

Study 2) for analysis of lung function impairment;

Study 3) for analysis of mortality from chronic obstructive lung disease;

Study 3) for analysis of mortality from lung cancer;

The data sets used in the present thesis were specifically collected for research purposes and the results of the studies were published in international medical journals. The ascertainment of diseases, smoking and dust exposure was done as accurately as possible and the data were extensively checked for errors at the time of the original data analysis. For this thesis the data will be re–checked for any inconsistencies prior to the analysis of the combined effect.

1.3 Epidemiological context

Most of the reported epidemiological studies of combined effect phenomena have been done on the risk of lung cancer in asbestos workers. In this section, the evidence from relevant lung cancer studies is summarized, and implications for the present study are pointed out.

The increased occupational exposure to asbestos fiber coincided with the increased prevalence of tobacco smoking in Western countries. In particular, workers born between 1910 and 1930 were likely to have both a higher smoking prevalence and a higher cumulative asbestos exposure (US DHHS, 1985). The extensive use of asbestos in shipbuilding in the 1940s, and during the World War II industrial expansion, resulted in increased incidence of asbestos related diseases during the post war period (Corn *et al.*, 1987). The first rigorous epidemiological study of the incidence of lung cancer in relation to asbestos exposure and smoking appeared in 1955 (Doll, 1955) and was followed by several other studies. As both the smoking data and asbestos exposure data were collected on large numbers of asbestos workers, the risk of individual and combined exposures could be evaluated with a smaller error. This contributed to the recognition of the public health significance of the combined effect and to the development of epidemiological and statistical tools for investigating its presence (Rothman, 1974; Thomas, 1981; Breslow and Storer, 1985). The studies done on mortality from lung cancer in asbestos exposed workers have been reviewed for the asbestos fiber and tobacco smoking interaction and several possible models of interaction have been suggested (Berry *et al.*, 1972; Saracci, 1977; Steenland *et al.*, 1986). The studies are also instructive as to their design and the results obtained.

In a study by Hammond *et al.* (1979), the mortality of 12 051 male insulator workers, who had at least 20 years of asbestos exposure, was compared

with that of 73 763 blue–collar workers enrolled in a large prospective study conducted by the American Cancer Society. The calculated relative risks were 5.2 for asbestos exposure alone, 10.6 for smoking alone and 52.2 when both exposures were present. Thus the relative risk of the combined exposure is approximately given by the product of the two risks of the independent exposures. In a more recent study by Liddell *et al.* (1984), a cohort of 10 939 chrysotile asbestos miners was followed from the year 1926 to 1975. The association between smoking and asbestos exposure was studied by a nested case–control study of 223 lung cancer cases and three controls matched to each case for date of birth and smoking habit. Exposures during the nine years prior to death of the case were ignored for the case and its matched controls. The calculated relative risks were 2.98 for nonsmokers with a high asbestos exposure, 4.9 for smokers with a low asbestos exposure and 8.2 for smokers with a high asbestos exposure.

In two studies of lung cancer mortality of asbestos factory workers, the mortalities of the asbestos workers were compared to the expected smoking–specific population death rates from lung cancer. In a study by Selikoff *et al.* (1980), a cohort of 582 men who began working with asbestos from 1941 to 1945 was followed up to 1977, and the mortality rates for lung cancer were compared to a population of blue–collar workers in the large prospective study conducted by the American Cancer Society. The estimated relative risks were 17.5 for nonsmoking asbestos workers, 7.0 for smokers without asbestos exposure, and 32.7 for smoking asbestos workers. In a study by Berry *et al.* (1972), the mortality rates from lung cancer in 1250 male and 420 female asbestos factory workers were compared to population smoking–specific death rates of London. The estimated relative risks were 10.0 for nonsmoking asbestos workers, 11.7 for smokers without asbestos exposure, and 27.8 for smoking asbestos workers. Thus the relative risks for the combined exposure, in both studies, were greater than the sum of the

relative risks due to the individual effects.

The results obtained from the above studies may not be comparable as there are many different aspects related to smoking habits, asbestos exposure, and their temporal patterns that could have been different. Nevertheless, it was observed that when the asbestos workers are compared with the population based controls, the relative risks are higher and show a greater deviation from the additivity of the individual effects than those obtained from the cohort–nested case–control study. It was suggested that using an internal non–exposed or low–exposed control group may fail to detect the presence of effect modification, whereas studies using non–occupational external comparison populations may reveal such modification (Steenland and Thun, 1986).

Several conceptual models of the underlying biological mechanisms of lung cancer causation by smoking and asbestos fiber have been proposed. The carcinogenic process, it was suggested, was a multistage process, in which individual agents may act at different stages to initiate or promote the disease. Thus, for example, a tumor initiator is required to start the carcinogenic process and another agent may act on the initiated tissue to promote the disease (Armitage and Doll, 1961; Peto, 1977; Day and Brown, 1980). Such a conceptual model would then be consistent with the resulting risk being a multiple of the individual effects involved in the causation process (Rothman, 1976). Some studies have suggested that smoking may cause impairment in asbestos fiber clearance from the lung and thus increase the carcinogenic effect of the asbestos fiber (Cohen *et al.*, 1979). Others have shown that the carcinogens in cigarette smoke are adsorbed by asbestos fibre, and the combination of asbestos and polycyclic aromatic hydrocarbons can induce tumors in animal systems at concentrations at which neither agent alone causes a tumor (Mossman and Craighead, 1981). The scientific evidence thus suggests that the carcinogenic process may involve a

complex underlying biological mechanism through which smoking and asbestos fiber may interact.

The epidemiological studies of lung cancer in relation to asbestos fiber exposure provide an example of how important the effect of combined exposure can be in terms of public health. Such studies can also provide epidemiological evidence of the interaction effect which complements experimental and clinical evidence, and contribute to the development of models of the underlying biological mechanism of disease causation.

1.4 General scientific perspective

The study of "epidemiological interactions" can be placed in the broader context of a general scientific inquiry, which involves a search for fundamental unifying laws of nature. The modern scientific methodology entails the formation of hypotheses, the construction of models and eventually theories, and their testing against quantitative empirical data.

An epidemiological inquiry is primarily concerned with the fundamental laws of disease occurrence and causation. An important part of epidemiology is the study of interactions between causative factors of disease. The presence of an interaction is detected by observing deviations from the expectations determined by mathematical models, assuming the interaction to be absent. However, such models are necessarily a superficial representation of reality, and more representative models must be searched for continually (Weed *et al.*, 1988).

CHAPTER 2

GOLD MINING AND ASSOCIATED LUNG DISEASES IN SOUTH AFRICA

Literature review

2.1 Chronological development of exposure on the mines

South African gold mines made their first impact on the health and wealth of the country just over one hundred years ago in 1886. The gold mines were modern in the sense that they introduced the major use of drilling machines which facilitated the mining of gold on a large scale. However, the health result of the initial mining methods was a large number of miners disabled or dying at a relatively young age from lung disease – miner's phthisis (Gordon, 1954).

The Miner's Phthisis Act of 1912 established the statutory position of the first Miner's Phthisis Commission which identified the factors responsible for the high prevalence of the disease as the practice of dry drilling, blasting during working shifts and lack of ventilation. Since then. strict dust control measures have been enforced through improved ventilation, axial water–feed drills, blasting between working shifts and regular dust measurements. By the 1930s respirable dust on the gold mines had reached levels which have remained almost unchanged until the 1970s (Page–Shipp and Harris, 1972). Since then the dust levels have dropped slightly more due to the establishment of a compensation fund into which each mine contributes a levy based on the dust

levels found by the Government Mining Hygiene Inspectors (Du Toit, personal communication).

Since the 1970s mechanized mining and high speed drilling have been introduced gradually on some of the large new mines. The machinery used for drilling and transportation of the ore is powered by diesel fuel. As these are relatively recent developments there are no safe limits prescribed for the diesel fumes levels by any government legislation, and even though the fume levels are monitored there has been no epidemiological study done to identify the long term health effect of these newer exposures.

Current exposure levels

The respirable dust contains a high level of crystalline silica in the quartz form. About 60–80% of the virgin rock is silica and approximately 30% of free silica is present in respirable dust. The current dust levels for underground air representative of the gold mining industry are: total dust 0.33–1.38 mg/m^3, repirable dust 0.23–0.97 mg/m^3, and the silica content of total dust ranges from 20–62%, i.e. 0.05–0.84 mg/m^3 (Beadle, 1970). The average respirable dust concentration to which the miners are exposed in high dust occupations is 0.48 mg/m^3, in medium dust occupations 0.30 mg/m^3, and low dust occupations 0.20 mg/m^3 (Page–Shipp and Harris, 1972).

Radon daughters concentrations are generally considered to be low: the average working level (WL) was estimated as 0.4, ranging from 0.1 to 3.0 for different mines. The average working level months (WLM) accumulated by the gold miners was estimated to be 36 (S.D. 0.4) (Wyndham, 1970). However, miners with more than 20 years of gold mining may (in theory) accumulate over 100 WLM. The level of radon daughters is generally determined by the uranium content of the ore, which also determines whether uranium is produced on the

mine. There are some gold mines in which uranium is the main product and gold mining is secondary. The WL is defined in terms of the energy released by alpha decay of the radon daughters. Normally WL is applied in terms of WLM, or the product of the WL and the number of working months (i.e., nominally 170 hours) of exposure at that level (Kathren and Petersen, 1989). At present, the measurements of radon daughters exposure that are available are not reliable enough to conduct an epidemiological study. The prime reason for this lack of availability of reliable data is the past perception that the radiation levels on South African gold mines were safe. This perception is changing, however, and monitoring of radiation levels will in future become a part of routine procedures carried out by the Government Mining Hygiene Inspectors.

No published figures on the diesel fume concentration on the gold mines are available.

Tobacco Consumption

Tobacco use on large scale started during World War I and increased during the subsequent decades untill a peak consumption was reached in the 1950s and 1960s (McCusker, 1988). Tobacco consumption is prevalent amongst white gold miners. In two epidemiological studies where smoking was recorded approximately 70% of white gold miners were found to be current smokers and only 12% had never smoked (Sluis–Cremer *et al.*, 1970: Wiles and Faure, 1977). Tobacco consumption was traditionally less prevalent in the black miners. However, this situation has been changing with the improved financial situation of these miners.

Financial compensation of occupational lung diseases

Until 1952 silicosis (pneumoconiosis caused by inhalation of silica particles) and

tuberculosis were the only compensable lung diseases. The acceptance of chronic obstructive lung disease as an occupational disease in the gold miners has come about later; unofficially in 1952, as a result of compensation demands by miners (du Toit, 1969), and officially in 1973, when supporting epidemiological evidence became available (Wiles and Faure, 1977). Lung cancer is not a compensable disease in South African gold miners.

Labour force

The gold mines employ around 500 000 black miners and some 50 000 white miners. The physical work is done by the black miners and the white miners are mostly supervisors (Cowie, 1989). The majority of the black miners are migrant workers with contract periods of about 12 months, after which they return to their homes for a certain time. Thus large time gaps without dust exposure are characteristic of the black miners. Some 80% of the labor force is recruited from the territories outside the borders of South Africa. The high labor turnover of the black miners has advantages from the pneumoconiosis and chronic obstructive lung disease point of view (Wyndham, 1970).

In terms of epidemiological studies, the exposure records for the black miners are unreliable because of gaps in mining exposure and no extant centralized records. In addition, the age distribution is highly skewed, as the miners who are over 40 years of age usually do not return or are no longer recruited to the mines unless they have been promoted into less physically demanding jobs. Thus, there is very strong selection in favor of the fitter miners.

2.2 Lung diseases associated with gold mining – definition and epidemiological evidence

The main occupational disease known to be associated with gold mining in South Africa is silicosis. In several preliminary analyses using radiological and autopsy data we have found no evidence that tobacco smoking is associated signigicantly with the presence of silicosis of the lung. Since the objective of this thesis is to evaluate whether tobacco smoking modifies the effect of silica dust on the diseases of the lung associated with gold mining dust exposure, silicosis is not considered in the major analysis.

The diseases considered in this thesis are: chronic obstructive lung disease (COLD) and lung cancer. The disorders that are incorporated in COLD are: emphysema, peripheral airways disease and chronic bronchitis (American Thoracic Society, 1987). In the present study chronic bronchitis, lung dysfunction and death from COLD will be considered.

2.2.1 **Simple chronic bronchitis**

This syndrome is characterized by abnormal secretion of mucus, impaired ciliary function, and a reduction in the effectiveness of the lung defence mechanism (Burrows, 1981: Cotes, 1975; Thurlbeck, 1970; Bignon, 1976). The structural basis for mucus hypersecretion may be related to inflammatory processes in the central airways (Mullen, 1987). The resulting symptoms are productive cough and increase in susceptibility to respiratory infections. An additional problem, often associated with simple chronic bronchitis is complaint of wheezing.

Epidemiological definition is solely on the basis of sputum production and cough on most days for three or more months of the year for at least two successive years (American Thoracic Society, 1987).

There are large number of potential risk factors. Host susceptibility factors, such as age, sex, past health experience and the genetic characteristics

related to immunological mechanisms are the most important risk factors for simple chronic bronchitis (Burrow, 1981). Tobacco smoking is considered to be the most important environmental risk factor for the development of simple chronic bronchitis, and environmental pollution may also be a contributory factor. Occupational exposures, especially mining exposures which involve a high exposure to dust, have also been implicated as causative risk factors (reviewed by Becklake, 1985).

Simple chronic bronchitis is prevalent in the general population, especially among tobacco smokers. Very little can be said about the time of the onset of the disease. The incidence of new diagnoses can be examined, but this may have little real meaning in the case of diseases which appear to be very gradual in onset. There also appears to be a major problem with diagnostic bias when the prevalence and incidence of the condition are examined (Burrows, 1981).

Epidemiological evidence on the effect of mining exposures

Higgins (1973) and Becklake (1985) reviewed the published studies of chronic non–malignant conditions characterized by airway dysfunction, namely simple chronic bronchitis and chronic airflow limitation, of men engaged in dusty occupations, mostly mining. It is now generally accepted that asthma and simple chronic bronchitis may be caused by occupational exposures (Becklake, 1985; Pride, 1984; Pepys, 1982). However, the confounding effect of cigarette smoking on the association makes it difficult to assess the relative importance of the occupational dust on the disease occurrence (Becklake, 1985; Morgan, 1984 and 1986; Higgins, 1973; Elmes, 1981).

The two studies of chronic bronchitis done on South African gold miners have demonstrated increased prevalence of chronic bronchitis associated with

gold mining exposure. In the first, a population–based cross–sectional study, the percentages with chronic bronchitis were found to be 2% for non–smoking non–miners, 5.6% for smoking non–miners, 3.6% for non–smoking miners and 13% for smoking miners (Sluis–Cremer, 1967). Results from the second cross–sectional study of South African gold miners have shown a significant dose–response relationship between the prevalence of chronic bronchitis and underground gold mining exposure measured in particle years (Wiles and Faure, 1977). Irwig *et al.* (1978) observed that the symptoms of chronic bronchitis were equally common in men with radiologically diagnosed silicosis and those without silicosis, in a case–control study nested within the cross–sectional study done by Wiles and Faure (1977).

Two studies of prevalence of chronic bronchitis and associated lung dysfunction, done on Western Australian gold miners, have also shown an association with underground gold mining. In the first study (Mc Nulty, 1968), the prevalence of chronic bronchitis was found to be 10% in the surface workers and 27% in the underground workers, and increased with the number of years of mining. In the second study (Holman *et al.*, 1987), the adjusted odds ratios and the 95% confidence intervals for chronic bronchitis were 1.8 (1.0–3.3) for one to nine years, 2.5 (1.2–5.2) for 10–19 years, and 5.1 (2.4–10.9) for 20 or more years of underground gold mining.

Several studies have suggested that chronic bronchitis and airflow obstruction may be two separate diseases (American Thoracic Society, 1987). In a more recent study done on South African gold miners, airways obstruction was strongly related to mortality from all causes and from COPD. However, after standardization for airways obstruction mucus hypersecretion was not related to mortality from COPD (Wiles and Hnizdo, 1991).

Epidemiological evidence on the combined effect of smoking and gold mining

In the study of the Western Australian gold miners (Holman *et al.*, 1987), the possibility that smoking and underground gold mining act synergistically in causing simple chronic bronchitis and chronic bronchitis with lung dysfunction was explored. Positive, but statistically not significant, interaction between smoking and underground mining on prevalence of chronic bronchitis was observed on the additive and multiplicative scale. The index of synergism on the multiplicative scale, S(m), was estimated as 1.4 (95% confidence interval (CI), 0.3–7.3), and on the additive scale, S(a), was estimated as 2.0 (0.4–9.6).

2.2.2 Lung function impairment

The two main syndromes of abnormal lung function are the obstructive and restrictive lung function impairments (Cotes, 1975).

Obstructive lung function impairment is characterized by widespread narrowing of the lung airways, which is a prominent feature in diseases such as asthma, chronic obstructive bronchitis and emphysema. In all three cases, the narrowing leads to increased airflow obstruction. When the abnormalities of the airways are quite extensive, the expiratory flow, as measured by forced expiratory volume in one second (FEV1), or the fraction of the forced vital capacity that can be expired in the first second (FEV1/FVC), is reduced. With mild diffuse airways obstructive disease of any type, the first change in the expiratory flow will be noted in the middle and towards the end of exhalation. The mid–expiratory flow ($FEF_{25-75\%}$), and the prolongation of the mid–expiratory flow–time (MEFT), are more sensitive indicators of airways obstruction, as they are less effort dependent(Cotes, 1975).

The restrictive lung function impairment arises from changes in the lung resulting in a diminution in the lung compliance. The reduction in the lung

compliance can be due to diffuse changes in the lung parenchyma often associated with interstitial fibrosis. The change causes reduction in the vital capacity or the other subdivisions of the total lung capacity, and may also be associated with an increase in the FEV1 measurement (Cotes, 1975; Snider, 1986).

Emphysema leads to increased static lung volumes (total lung capacity (TLC), function residual capacity (FRC) and residual volume (RV)) and reduced rates of expiratory airflow as measured by FEV1, FEF_{25-75}, and FEV_1/FVC. Combination of two or more patterns of lung injury may alter an individual function measures in opposite direction. Thus the presence of interstitial fibrosis and emphysema in the lung may result in a normal TLC despite extensive lung damage (Burns *et al.*, 1988).

The risk factors related to the occurrence of lung impairment are much the same as those listed for simple chronic bronchitis. Tobacco smoking is the most prominent environmental risk factor for diseases of the airways and emphysema. Environmental pollution and many occupational exposures are also implicated as risk factors for the above diseases. Some researchers, however, regard occupational and environmental exposures as simply adding to the burden already caused by smoking and not as inherent risk factors (Burrow, 1981; Morgan, 1986).

Epidemiological evidence on the effect of mining exposures

Several large studies done on coal miners have observed that occupational exposure to coalmine dust is related to impairment of expiratory flow (Rogan *et al.*, 1973; Soutar *et al.*, 1986; Love *at al.*, 1982; Attfield, 1985). It was stated that coal mining dust exposure has a clinically important effect on loss of FEV1 which is independent of smoking (Marine *et al.*, 1988). However, as more

evidence emerges on the role that tobacco smoking plays in the causation of non–malignant respiratory diseases in smokers, scepticism is voiced as to whether the present levels of occupational dust exposure can cause a disabling degree of respiratory disorder and the associated increased mortality arising from it (Morgan, 1986; Foxman *et al.*, 1986; US DHHS, 1985).

Several studies of South African gold miners have demonstrated increase in lung function impairment associated with gold mining dust exposure. One of the first studies of respiratory function in silicosis was that of Dowds, who investigated 500 Rand gold miners with varying radiological degree of silicosis. He showed that with radiological progression of silicosis there was a progressive reduction of vital capacity (Irvine *et al.*, 1930). In another study comparing lung function in silicosis among Witwatersrand gold miners (Becklake *et al.*, 1958), the pattern of functional disturbances was examined in two samples of miners 41–50 and 51–60 years of age. Various lung function tests were considered. Vital capacity was found to be reduced in both age groups studied, and residual volume was increased. In a study of a random sample of miners and non–miners in a Witwatersrand town, done by Sluis–Cremer *et al.* (1967), no difference in ventilatory function (FEV1 and peak expiratory flow rate, PEFR) was found between miners and non–miners other than that due to the excess of chronic bronchitis in miners. In a cross–sectional study of 2209 white gold miners 45–54 years of age (Wiles *et al.*, 1977), significant negative dose–response relationships between the FEV1 and FEF_{25-75} measurements and underground gold mining exposure, measured as surface area of respirable dust, were observed. The decreasing trend in the FEF_{25-75} measurement was found to be statistically significant for non–smokers ($p<0.05$), ex–smokers ($p<0.05$), and current smokers ($p<0.001$).

To find out whether silicosis per se causes lung function disability two

case–control studies were conducted. A nested case–control study done on the 2209 gold miners (Irwig *et al.*, 1978) compared the radiologically diagnosed silicotics with non–silicotics for lung function impairments. Forced vital capacity did not differ between the cases and controls. FEV1 and FEF_{25-75} were found to be lower in the cases, but this difference was almost entirely accounted for by the higher dust exposures of the cases. The conclusion drawn from the study was that men with silicosis have the same, or only slightly more, airway obstruction than men without silicosis who have had similar total exposure to dust, but that the effect of dust on lung function, independent of silicosis, is important. This conclusion was confirmed by a second case–control study, in which 61 cases with radiological silicosis were each matched with a non silicotic control, on age, dust exposure and tobacco smoking. No statistically significant differences were observed in any of the extensive lung function tests that were performed, with the exception that the cases were found to have significantly increased value for aleveolar plateau measurement (Wiles *et al.*, 1991). A possible physiological reason for the higher alveolar plateau is being investigated by examining whether in miners who had an autopsy examination the lung alveoli of the silicotics contain more fibrosis than of the non–silicotics.

More recent studies, done on deceased South African gold miners, have found a statistically significant association between the occurrence of emphysema diagnosed at autopsy, and gold mining dust exposure measured in terms of number of shifts worked in high dusty occupations, standardized for tobacco smoking (Becklake *et al.*, 1987; Hnizdo *et al.*, 1991).

A study of Western Australian gold miners has shown an association between underground gold mining exposure and a simple chronic bronchitis with a restrictive lung function impairment (Holman *et al.*, 1987).

Epidemiological evidence on the combined effect of smoking and mining

The evidence from epidemiological studies done in coal miners is mostly in agreement with the hypothesis that respirable dust on coal mines and tobacco smoking act additively in increasing the prevalence of obstructive airway impairments (Becklake, 1985; Soutar *et al.*,1986; Leigh *et al.*, 1986; Marine *et al.*, 1988). In the more recent study of British coal miners (Soutar *et al.*, 1986), the results obtained suggest that the functional effects of dust induced damage are different from those of smoking. Smoking was found to be related to a classical obstructive effect, while exposure to coal mining dust was related to a parallel reduction of FEV1 and FVC, which was interpreted as indicating damage at the level of the respiratory bronchiole or alveolus. This pattern was illustrated by the non–smokers, but in smokers and ex–smokers the resulting type of abnormality was predominantly of the obstructive pattern.

In asbestos cement workers, the asbestos dose and smoking were found to be independent and additive contributors to decreased pulmonary function, FEV1 and FVC (Finkelstein *et al.*, 1986).

In the study of the Western Australian gold miners (Holman *et al.*, 1987), the possibility that smoking and underground gold mining act synergistically in causing simple chronic bronchitis with lung dysfunction was explored. An indication of a positive interaction between smoking and underground mining on prevalence of chronic bronchitis with associated restrictive impairment was reported on the additive and the multiplicative scale. The adjusted prevalence odds ratios were found to be 1.3 (0.1–17.9) for smoking, 1.9 (0.2–17.0) for dust, and 4.1 (0.4–38.2) for dust and smoking. The index of synergism on the multiplicative scale, S(m), was estimated as 1.6, and on the additive scale, S(a), was estimated as 2.5.

A more recent study done on hard–rock miners has shown that hard–rock

mining exposures affected smokers and non–smokers differently. In smokers the airflow limitation and overdistension were greater, and single breath carbon monoxide diffusion capacity (DLCO) divided by alveolar volumes (VA) was lower. In non–smokers lung volume was lower, flow rates at absolute lung volumes were higher and DLCO/VA was higher (Kreiss *at al.*,1989).

Manfreda *et al.* (1982) suggested that mining exposure may be associated with a restrictive effect in non–smokers and aggravated hyperinflation and airflow limitation in smokers.

The epidemiological evidence from coal miners cannot be generalized to the gold miners, particularly those in South Africa, where the respirable dust contains about 30% free silica (measured after treatment by heat and acid), and there is little epidemiological evidence in the literature to indicate whether the joint exposure to gold mining dust with a high level of free silica and to tobacco smoking act additively in producing obstructive or restrictive respiratory impairments, or whether tobacco smoking modifies the effect of silica exposure.

It has been pointed out that in order to assess the combined effect of tobacco smoking and occupational exposure, the pattern of effects of lung injury on pulmonary function, rather than a single lung function measure, should be examined (Burns *et al.*, 1988; Harber *et al.*, 1987). Thus in the present thesis, a number of individual lung function measurements and the spirogram pattern are examined.

2.2.3 Death from chronic obstructive lung disease

The epidemiological evidence supports the hypothesis that silica dust exposure is associated with increased prevalence of respiratory impairment (Wiles and Faure, 1977) and of emphysema at autopsy (Becklake *et al.*, 1987; Hnizdo *et al.*, 1991). The experimental evidence also indicates that silica dust exposure induces

emphysema in rats (Wright *et al.*, 1988; Churg *et al.*, 1989). The question then arises whether the exposure to silica dust is associated with increased mortality from chronic obstructive lung disease. The causes of death, according to the 9th revision of the International Classification of Diseases (ICD) (1977), that are included as death from COPD includes death from chronic bronchitis (ICD=491), emphysema (ICD=492), chronic airways obstruction (ICD=496) and chronic pulmonary heart disease (ICD=416).

Epidemiological evidence on the effect of mining exposures

Mortality studies of silica–exposed cohorts have consistently shown increased mortality rates of nonmalignant respiratory disease, accounted for largely by silicosis (US DHHS, 1985). Davis *et al.* (1983) did not find an excess mortality from bronchitis in Vermont Granite workers. Finkelstein *et al.* (1987) found nonmalignant respiratory disease (8th revision ICD 460–519) to be the most frequent cause of death, with standardized mortality ratio (SMR) over 600, among Ontario miners receiving compensation for silicosis. In a mortality study of Western Australian gold miners the overall death rate from nonmalignant respiratory disease was slightly less than expected, but the death rate from silicosis was high (Armstrong *et al.*, 1979).

In a previous study of mortality in middle–aged South African gold miners (Wyndham *et al.*, 1986), a cohort of 3971 white miners, born between 1 January 1916 and 31 December 1930, who were alive on 1 January 1970 and currently working on the gold mines, was followed up to December 1978. Standardized mortality ratios were calculated using as a reference population the total white male population of the Republic of South Africa. Increased SMR for chronic respiratory diseases, SMR=165.6, (95% CI, 108–243) was reported. Of the 530 deaths found, 29 (5.5%) miners had died of chronic respiratory diseases,

which included bronchitis, emphysema, asthma, and pulmonary heart disease. The diseases were coded according to best available information drawn from the death certificates, the MBOD's medical files and autopsy findings. A nested case–referent study was performed in order to establish the presence of dose–response trend with dust, after adjustment for smoking. The relative risks of 2.48 per 10 000 particle–years was reported (log likelihood ratio statistic, $X^2_{LR}=4.51$, $p<0.03$).

2.2.4 Death from lung cancer

Death from lung cancer is registered on the death certificate as ICD 162. In our study the diagnosis can be confirmed by the information given in the MBOD medical file if a biopsy or necropsy was done. There are, however, problems associated with a diagnosis of primary lung cancer (Hessel *et al.*, 1990).

Epidemiological evidence on the effect of mining exposures

The issue of carcinogenicity of crystalline silica has been extensively discussed in the literature (Goldsmith *et al.*, 1986; McDonald, 1989; IARC, 1987). A recent evaluation by the International Agency for Research on Cancer is that there is sufficient evidence for the carcinogenicity of crystalline silica to experimental animals, but that the evidence for humans is limited (IARC, 1987).

The epidemiological evidence on the association between underground gold mining exposure and mortality from lung cancer is conflicting. A cohort study of 6 972 Ontario underground gold miners was reported to have an increased standardized mortality ratio for lung cancer, SMR=145 (Muller *et al.*, 1983). A cohort study of Western Australian gold miners also indicated an excess of lung cancer, SMR=140 ($p<0.01$) (Armstrong *et al.*, 1979). Several studies of South Dakota (U.S.A.) gold miners have been done. Gillam *et al.*,

(1976) reported an SMR of 370 calculated on 440 underground gold miners. McDonald *et al.* (1978), however, did not find convincing evidence of an increase in respiratory cancer in a cohort of 1321 gold miners from South Dakota. A more recent historical cohort mortality study of 3 328 South Dakota gold miners reported the SMR for respiratory cancer (including malignant tumors of the trachea and bronchus) to be 100 with 95% confidence interval (73–135) (Brown *et al.* 1986).

The epidemiological evidence from South African gold miners is also conflicting. In a first mortality study of a cohort of 3 971 white, middle aged South African gold miners, followed over 9 years period, 39 deaths due to lung cancer were found and the SMR of 161.2 with 95% confidence interval (114.6–220.3) was reported (Wyndham *et al.*, 1986). A nested case–referent design was used to estimate the association with dust, the relative risk adjusted for smoking was 1.77 (0.94–3.31) per 10 000 particle years (X^2_{LR}=3.42, p<0.06). However, the results from two case–control studies done on deceased, white, South African gold miners showed no association between lung cancer and cumulative gold mining dust exposure, nor any association with radiological or autopsy identified silicosis (Hessel *et al.*, 1986, Hessel *et al.*, 1990).

The conflicting epidemiological evidence on the association between lung cancer and gold mining exposure may be due to varying levels of the potential risk factors. There is still uncertainty as to which of the potential risk factors is associated with the increased risk of lung cancer found on some mines. Silica has been shown to produce tumors in the respiratory epithelium as well as lymphoma in rats, after intratracheal instillation in several studies (Saffiotti, 1986). However, in hamsters no such tumors were observed in several independent studies, implying inter–species variation in susceptibility to silica induced lung cancer (Saffiotti, 1986). Studies done on workers compensated for

silicosis have invariably reported increased SMRs due to lung cancer (Kurppa *et al.*,1986; Schuler and Ruttner, 1986; Zambon *et al.*, 1986). However, in these studies the confounding effect of smoking was not considered and it has been suggested by some of the authors that the increased standardized mortality ratio due to lung cancer may be because smoking may increase the risk of silicosis. If smokers develop silicosis more readily than nonsmokers, then the silicotic cases would be heavier smokers and consequently develop lung cancer at a higher rate. Archer *et al.* (1986) compared the rates of silicosis and lung cancer among mining or milling groups selected on the basis of their exposure to siliceous dust and radon daughters. When quartz dust levels were high, pneumoconiosis rates were also high, but lung cancer rates were high only when radon daughters levels were also elevated. The authors concluded that the two effects of mining exposure (silica dust and radon daughters) are largely independent.

Goldsmith and Guidotti (1986) reviewed some of the published studies for a possible synergistic effect of silica exposure and cigarette smoking. However, satisfactory epidemiological studies, in which both smoking and dust levels have been measured, were noted to be lacking. Armstrong *et al.* (1979) reported that smoking underground gold miners had a relative risk for lung cancer increased to 60.0 compared to non–smoking miners. Larsson and Damber (1982) reported a synergistic odds ratio of 25.4 for men who both smoked and worked on the mines. Mastrangelo *et al.*, (1982) also observed a synergistic odds ratio of 13.0 for lung cancer among smoking silicotics compared to either dust exposed non–smokers without fibrosis or unexposed non–smokers as the referent denominator. Goldsmith and Guidotti concluded that the combined effect of silica dust and smoking is synergistic rather than additive. However, they indicated that their calculations lacked in precision due to the insufficient number of studies with good exposure data and indicated the need for further

studies of the combined effect.

CHAPTER 3

DESCRIPTION OF THE ANALYTICAL METHODS USED

Literature review

3.1 Measures of disease occurrence, measures of association, and study designs

3.1.1 Measures of disease occurrence

Epidemiological studies are concerned with the occurrence of disease, or death from a disease, in a human population. The three indices, usually employed to quantify the disease occurrence in a population are risk, rate and prevalence (Kleinbaum *et al.*, 1982).

Risk is the probability of an individual developing a specific disease within a given time period (conditional on the individual not dying from any cause during that period). Risk is estimated as a ratio between the number of new cases of a disease developed during a follow–up period and the number of disease–free subjects at the start of the follow–up. This measure, also called cumulative incidence, is dimensionless and requires a reference time.

Incidence rate is an instantaneous potential for a change in disease status per unit of time, relative to the size of the disease–free (i.e. candidate) population at a given time. Rate is estimated as a ratio between the number of new cases of a disease and the total population time of follow–up (e.g. person–years) of the disease–free population. It is expressed in units of 1/time and is also called

incidence density.

Prevalence is the measurement of disease state in a population at a particular point in time. It is estimated as the ratio of number of subjects with a disease and the total number of subjects in the study population at the time.

3.1.2 Basic measures of association

The primary concern of most epidemiological studies is to identify factors that increase the probability (i.e. risk) of a disease occurrence. Such factors are called risk factors.

To measure the strength of association between a risk factor and a disease, measures of association are applied. These measures compare the disease occurrence in the ith exposure level (i=1,....no. of exposure levels) to the disease occurrence in the lowest exposure level (i=0). The change in the disease occurrence can be measured as a ratio or as a difference, between the measures of disease occurrence (i.e. risk, rate or prevalence). The most frequently used measures of association are:

a) Incidence density ratio, which is a ratio between incidence rates.

b) Risk ratio (or relative risk), which is a ratio between disease risks.

c) Exposure odds ratio (or odds ratio). In case–control studies the exposure specific risk or rate cannot be estimated as the marginal number of cases and controls are fixed by the study design. The strength of association between disease and exposure can be measured by comparing the odds of exposure in the cases to the odds of exposure in the controls. The ratio of the two odds is called exposure odds ratio (or odds ratio).

d) Mortality rate ratio, which is a ratio between mortality rates. Mortality rates standardized for the effect of age and a year of observation are usually compared with those observed in a comparison group and the term standardized mortality

ratio (SMR) is used.

e) Incidence density difference (or incidence rate difference), which is the difference between incidence rates.

f) Risk difference (or attributable risk), which is a difference between risks.

3.1.3. Basic study designs and methods of analysis

The risk factors are usually identified through observing a given population over a specific follow–up period, or at a particular point in time, and relating the disease occurrence to a set of potential risk factors. The basic study designs employed in etiological studies are: follow–up (or cohort) studies, cross–sectional studies and case–control (or case–referent) studies.

a) Cohort study

The basic principle of a cohort study is that a study population 'at risk' of disease or death is followed over a time period during which new cases of disease or deaths are ascertained. The ratio between the number of new cases with a disease or deaths and the total time of follow–up of the population at risk, usually estimated in person–years, estimates the morbidity or mortality rate, respectively. The morbidity or mortality rates can be compared to those in an external population or used to estimate a dose–response trend with increasing level of exposure. Before the rates can be compared, the differences in the age and time at risk structure between the comparison groups have to be taken into consideration.

The data layout from a hypothetical follow–up study with n levels of an exposure factor and m levels of a confounding factor is shown in Table 3.1.

Table 3.1 Data layout for a cohort study

Stratum (j=1,...,m)	Outcome	Exposure (k=0,1,...,n) 0	1	.....	n	Total
j	Deaths	d_{j0}	d_{j1}	.....	d_{jn}	D_j
	Person–time	n_{j0}	n_{j1}	.....	n_{jn}	N_j

Let d_{jk} denote the number of deaths, n_{jk} the person–years of follow–up, and $h_{jk}=d_{jk}/n_{jk}$ the death rate in the jkth cell. Under the assumption of a constant mortality rate over all the exposure levels, the d_{jk} are assumed to follow an independent Poisson distribution with mean $h_{j0} \cdot n_{jk}$, where h_{j0} denotes the death rate at the baseline exposure level, k=0. For the Poisson distribution the denominators n_{jk} are assumed to be fixed.

In studies of the effect of occupational exposure on cause specific mortality the statistic of interest is the standardized mortality ratio (SMR), which compares the observed number of deaths in the cohort's jth stratum, D_j, with an expected number of deaths obtained by applying the standardized death rates, $^*h_j$, to the cohort's age and year at risk structure as

$$SMR = \Sigma_j D_j \,/\, \Sigma_j N_j {}^*h_j.$$

The standardized death rates can be obtained from an external population and then the resulting SMR indicates whether a particular study group has an increased mortality from a specific cause in comparison to an external population. Since the SMR statistic assumes no effect of exposure, the expected value for SMR is 1 or 100%. A better indication of causality is to demonstrate a trend in the mortality ratios with increasing exposure level. This can be done by using an external comparison, or using comparison with internally standardized rates.

If applying the internal standardization, the stratum–specific death rates calculated without regard to exposure are $h_j=D_j/N_j$. Thus, the expected number

of deaths in the kth exposure level, assuming that exposure has no effect on the rates, is

$$E_k = \Sigma_j\, n_{jk}\, h_j = \Sigma_j\, n_{jk}\, (D_j/N_j).$$

The expected number of deaths is used to calculate an approximate estimate of relative risk for each exposure category through estimation of the internally standardized mortality ratio

$$SMR_k = \Sigma_j d_{jk} / E_k.$$

The relative risk associated with kth exposure level is then estimated as

$$\psi_k = SMR_k / SMR_0.$$

Thus estimated relative risks can be used to test for their heterogeneity and trend with exposure by applying the chi–square test and chi–square test for trend (Breslow, 1984; Breslow and Day, 1987).

When several risk factors are involved in the cohort data analysis, then the multifactorial methods of analysis are more appropriate. For the multifactorial analysis several approaches have been proposed:

(1) Cox's proportional hazards model (Cox, 1972). The basic assumption behind the proportional hazards model is that after an instantaneous exposure of intensity x at time t, the hazards rate h[T,x(t)] at time T is proportional to the baseline rate $h_0(T)$ multiplied by a relative risk that is dependent on the intensity of exposure and some probability density function for times–to–disease (Thomas and Whittemore, 1988). The proportional hazards model is often regarded as the "gold standard" of analytical methods for cohort data.

(2) Poisson regression approach (Holford, 1980; Frome and Checkoway, 1985). This method is used for the analysis of grouped data. The number of deaths within each cross–tabulated category of exposure is assumed to have an independent Poisson distribution. Theoretical analysis has shown an inherent link between Poisson regression and the proportional hazards model (Holford, 1980).

Let us assume, for example, that our data are crosstabulated into I categories for dust, J categories for smoking and K categories for the nuisance effect (eg age and year at risk). Then the data in the (i,j,k)th cell consist of: the observed number of deaths d_{ijk} from the disease, person–years N_{ijk}, the mortality rates $h_{ijk}=d_{ijk}/N_{ijk}$, and the mean values for the exposure variables denoted as $\underline{X}_{ijk}=(X_{1ijk}, X_{2ijk})$, where X_{1ijk} represents the dust variable and X_{2ijk} represents the smoking variable. The expected mortality rate within each ijk cell, h_{ijk}, is estimated by a function $h_{ijk}= h_{0k} \cdot R(\underline{X}_{ijk};\beta)$, where h_{0k} denotes the age–and–year specific disease rate at $R(0)=1$, and $R(\underline{X}_{ijk};\beta)$ can be any relative risk function parameterized by β (Breslow and Day, 1987). The person–years in each (ijk)th cell can be calculated according to a method described by Frome and Checkoway (1985), and Pearce and Checkoway (1987).

(3) Logistic regression (Liddell *et al.*, 1977). Nested case–referent data can be sampled from the cohort data and analyzed by logistic regression using conditinal maximum likelihood method of estimation.

Alternatively, unconditional logistic regression can be applied to the whole cohort data. This method was shown to be comparable to Cox's proportional hazards model only when a follow–up period is sufficiently small or the survival rate is high (Annesi *et al.*, 1989).

The efficiency of each of the proposed method was recently evaluated (Pearce *et al.*, 1988). Using the whole cohort, the Poisson regression and the proportional hazards model yielded very similar findings, whereas the logistic regression tended to yield effect estimates that were too large. Case–referent sampling resulted in a loss of precision of the rate ratio estimates.

To evaluate the combined effect of two risk factors the analysis of grouped data is considered to be preferable, as it allows the trends in death rates and the actual exposure levels involved to be examined (Breslow and Day, 1987). The

application of generalized relative risk models with the Poisson regression approach have been described by several authors (Whittemore *et al.* 1983; Lubin and Gaffey, 1988; Breslow and Day, 1987).

b) **Case–control study** (or case–referent study)

In this type of study the subjects with a specific disease (the cases) are compared to those without the disease (the controls) in order to identify the risk factors related to the presence of disease. The cases are either incident (i.e. new) and have been ascertained over a defined study period, or prevalent (i.e. existing) and have been ascertained at some specific time. The controls to which the cases are compared are assumed to represent a random sample of subjects who do not have the disease and who are selected from a candidate population from which the cases originated. The cases can be identified retrospectively or prospectively. The control sample may be stratified or individually matched on the basis of some confounding factor, e.g. age. Under these circumstances the controls are assumed to constitute a random sample from each of the subpopulations formed by the stratification or matching (Breslow and Day, 1980).

The relative risk attributable to a specific risk factor can be estimated from case–control studies provided that certain assumptions of study design are fulfilled. Cornfield (1951), has shown that the ratio of the odds of being exposed for the cases against the odds of exposure for the controls, i.e. the exposure odds ratio, is equal to the ratio of the odds of having the disease in the exposed against the unexposed subjects, i.e. the disease odds ratio, and that the odds ratio is approximately equal to the risk ratio if the disease is rare.

It has been shown that a case–control study can be nested within a cohort study (Mantel, 1973), combining elements of cohort and case–control study design. A single candidate population is defined at the begining of the follow–up period and all the incidence of disease or death are identified during the follow–up period.

The nested case–control study will yield unbiased estimates of the relative risk from a full cohort analysis (Lubin *et al.*, 1984). The sampling method employed in this type of study is called incidence density sampling (Pearce, 1989). Each case is matched to one or more controls selected at random from a set of subjects who are still at risk (also called risk set) at the time that the case was identified (Liddell *et al.*, 1977). The risk set includes subjects who may subsequently develop the disease (Lubin at *al.*, 1984). Another approach is to sample non–cases without replacement from the remainder of the risk set (Robins *et al.*, 1986).

Table 3.2 shows the frequency data layout from a hypothetical case–control study involving an exposure factor with n levels and a confounding factor with m levels.

Table 3.2 Data layout for a case–control study

Stratum	Outcome	Exposure (k=1.2,...,n)				
(j=1,...m)		0	1	.....	n	Total
j	Cases	d_{j0}	d_{j1}	.....	d_{jn}	D_j
	Controls	c_{j0}	c_{j1}	.....	c_{jn}	C_j
	Total	n_{j0}	n_{j1}	.....	n_{jn}	N_j

The basic unit on which the statistical tests are based is a 2 x 2 table in which the kth level of exposure is compared to the lowest level of exposure, i.e. the baseline. In the frequency data, obtained from a case–control or cohort study, the probability of observing the data in any 2 x 2 table can be regarded as a product of two binomial distributions with parameters (p_{j0}, n_{j0} and p_{jk}, n_{jk}) where p_{j0} and p_{jk} are either the disease or exposure probabilities and n_{j0} and n_{jk} are the marginal totals for disease or exposure for cohort and case–control studies, respectively.

The key parameter of interest in a case–control study is the odds ratio, OR, which, as mentioned earlier, has the same value regardless whether calculated from

the disease or exposure probabilities. The odds ratio, OR_{jk} estimates the increase in the odds of disease for the kth exposure level relative to the baseline level, k=0, as $OR_{jk}= (d_{jk}c_{j0})/(c_{jk}d_{j0})$. The probability distribution which best satisfies the conditional distribution of the data is the hypergeometric distribution which assumes all marginal totals to be fixed. To test that the OR is statistically different from unity, a large–sample test based on the hypergeometric distribution, the Mantel–Haenszel chi–square test (Mantel and Haenszel, 1959), can be applied. Alternatively, another large–sample test, based on the comparison of two binomial proportions, the chi–square test, gives good approximation of the above test (Kleinbaum *et al.*, 1982; Breslow and Day, 1980). A small–sample statistical test based on hypergeometric distribution can be derived by assuming fixed marginal totals, and is known as Fisher's exact test. A $100(1 - \alpha)\%$ large sample confidence interval for OR based on logarithmic transformation, called logit confidence limits, gives reasonable approximation for large samples (Breslow and Day, 1980).

The usual approach to the dose–response analysis is to choose one exposure level as a baseline against which to compare each of the other levels using the methods for 2 x 2 table (Breslow and Day, 1980). To test whether there is a significant trend in the prevalence with increasing exposure level, the Mantel–Haenszel statistic for trend (Mantel, 1963) is used.

To examine the joint effect of two exposures with k ordinal levels, the same basic techniques as used in a 2 x 2 table can be used in a k x k table. The odds ratios are obtained by using as a baseline the category corresponding to the combination of baseline levels of each individual factor. The marginal odds ratios for one factor, adjusted for the effect of the other, can be used to determine whether each factor operates independently of the other for certain combinations of the exposure levels. The trend in the size of the odds ratio within a given category of the other factor can be compared to the marginal estimates (Breslow and Day,

1980).

The multifactorial analysis usually involves applying the unconditional or conditional maximum likelihood method depending upon the study design (Breslow and Day, 1980).

c) Cross–sectional studies

This type of study usually involves a study population that has been selected by random (i.e. equal probability) sampling from a large target population to which inference is to be made. The prevalent (i.e. existing) cases of disease are ascertained and the estimable measure of disease occurrence is prevalence. This type of study is usually used to study chronic diseases which have long duration and in which the disease onset is difficult to determine (e.g. chronic bronchitis and lung function impairment) and is often used to study the effect of occupational exposure on chronic lung diseases. To estimate the exposure relevant to the disease a certain latency period should be assumed (Rothman *et al.*, 1985). The statistical inference is done in the same way as if the data were collected from a case–control study (Kleinbaum *et al.*, 1982).

Relationship between the indices

In studies in which disease etiology is being investigated, the prime interest is to relate rates of disease occurrence in a population, or the risk of developing the disease, to a set of potential risk factors. The ideal effect measure is incidence. In a steady–state population (i.e. that in which the age distribution and size does not change or when incidence and prevalence remain constant) a mathematical relationship between prevalence odds ratio and incidence density ratio can be derived and thus to some extent justify the use of prevalent cases and mortality studies. If the mean duration of the disease is identical for the exposed and the unexposed cases and if the disease does not affect the exposure status, then the prevalence odds ratio estimated from either a cross–sectional or a case–control

study is equal to the incidence density ratio, provided that the disease is rare (Miettinen, 1976; Morrison, 1979).

It has also been shown (Morrison, 1979) that the mortality ratio is approximately equal to incidence density ratio (the parameter of etiological interest) under steady–state conditions, if the disease is irreversible and does not affect exposure status (before selection), and if the disease is rare and/or rapidly fatal.

3.2 Interaction and synergism – definitions

In observational epidemiological studies, the concept of interaction is more difficult to define than in experimental studies, for which the statistical interaction concept was originally developed. The difficulties arise for two main reasons.

Firstly, the etiology of a disease is often very complex, with many causative and confounding factors involved in the intricate biological mechanism of the disease causation. For example in the case of cancer, the carcinogens may act at different stages independently of each other, but each stage is required to occur before the next one can start (Peto *et al.*, 1977). The question then arises as to whether the epidemiological studies that use probability theory and statistically identify the risk factors, can provide an insight into the biological mechanism of disease causation. In particular, there is the question of whether the presence of an interaction term in a statistical model can be interpreted also on the biological level, and if so, what is the definition of biological interaction.

The second factor which complicates the issue lies in the mathematical models used to analyze epidemiological data. The most frequently applied model is the logistic regression, where the odds of disease, transformed by the logarithmic

function, called log odds or logit, are related to a linear combination of the risk factors. The logistic model assumes that the risk factors combine on a multiplicative scale to estimate the risk of disease. Thus, even though no interaction term, i.e. a product term, is present in the model, the effect of the individual terms is additive on the logarithmic scale, i.e. multiplicative on non–logarithmic scale. This leads to the question of whether there is independence of the risk factors if additivity on a logarithmic scale applies.

Rothman (1974) proposed that the evaluation of the combined effect of two causes is pertinent to the study of a causal mechanism of disease causation, and that the aim of most epidemiogical studies is to try to reveal this mechanism. He argued that statistical analyses which use transformation of the variables tend to obscure the effects, and that only a "natural scale" that most directly measures the effects is appropriate. He then used probability theory to show that a disease caused by either of two independent causes, A and B, will occur with the probability, P,

$$\begin{aligned} P(\text{disease}) &= P(A \text{ or } B) \\ &= P(A) + P(B) + P(A)\cdot P(B) \\ &= P(A) + P(B) \end{aligned}$$

if P(A) and P(B) are small, which is usually the case in epidemiological studies. This simple relation is the basis of his description of interaction among causative factors which he called synergy :

"If the risk attributable to a combined exposure exceeds the sum of the risks attributable to each exposure separately, we are observing positive interaction or synergy " (Rothman, 1974).

Rothman's proposition was criticized by several authors, but it lead to an increased interest in the subject and consequently also to new developments. Walter and Holford (1978) argued that any definition of interaction is, in the

absence of a specific physiological model, arbitrary and that the choice of the definition should depend to some extent on statistical convenience. They substantiated their argument by employing two possible physiological models, using the probability theory for the Poisson distribution, and showed that one model leads to additivity on additive scale, whereas the other to additivity on multiplicative scale. Kupper and Hogan (1978) pointed out that the index S, suggested by Rothman as a measure of interaction, is not always applicable, and suggested the use of a log–linear model for identifying the presence of interaction. However, the log–linear model measures interaction on the log, i.e. multiplicative, scale.

Blot and Day (1979) supported Rothman's proposition and suggested that the term synergism be used in the public health context, i.e. when departure from the additive model is present, then it can be said that the interaction between two risk factors is of public health significance.

In 1980 Rothman *et al.* (1980) suggested that interaction in epidemiological studies be defined in four broad contexts: statistical, biological, public health and individual decision making, as follows:

Statistical interaction is dependent on the mathematical model fitted to the data. If the regression coefficient for the product term of the risk factors in question is significantly different from zero and the goodness–of–fit statistic shows that presence of the interaction term improves the fit, then statistical interaction is present and should be included in the model.

Biological interaction is present when two or more causative factors act interdependently to produce the disease. The authors argue that two categories of general interest are: 1) those in which the etiological factors act interchangeably in the same step in a multiple process and thus the effect of each factor adds to the effect of the other; or 2) those which act at different steps in the process. (For

example, tobacco smoking may cause a certain amount of damage to the tissue cells and subsequent exposure to dust acts on those cells to increase the probability of lung cancer. Thus the risk of lung cancer is a product of the individual probabilities of exposure.) These two broad categories then correspond generally to mathematical models in which the combined effect of exposure factors is assumed to be additive or multiplicative, respectively. The best fitting of the two models can give an indication of the nature of the interrelation between the causal factors.

Public–health interaction is present when the effect of one risk factor on the disease occurrence is modified by the presence of another and when departure from the additivity of the individual effects is observed. In terms of the mathematical models, applied to test the combined effect, it means that a significant interaction term is present in the mathematical model which assumes additivity of relative risks, and that the mathematical model which assumes multiplicativity of the effects may fit significantly better the data than the model which assumes additivity of the effects.

Individual–decision making context entails evaluating personal risk of disease considering that the causal factors acting in combination lead to an increased/decreased risk of disease above/below the sum of the individual risks.

Siemiatycki and Thomas (1981) presented examples from multistage carcinogenesis to demonstrate that the use of statistical analysis of epidemiological data to infer biological processes can often be misleading. They showed a hypothetical example of two independently acting carcinogenic factors which are fitted best by a multiplicative model. Thus, they argued that elucidation of biological interaction by means of statistical models requires knowledge of the underlying biological model and reasoning, which cannot be done mechanically.

Saracci (1980) proposed to keep the first three levels of interaction as proposed by Rothman *et al.* (1980). He also stressed the importance of the

knowledge of the public health interaction, which he called synergism.

Miettinen (1982, 1985) expanded on the issues regarding causal and preventive interdependence. He defined an instance of synergism of two deterministic causal factors as one in which both exposures are needed to bring about the effect. This means that in some instances of disease occurrence the two exposures jointly constitute a sufficient cause. Analogously, an instance of antagonism of two deterministic factors may be when one factor prevents the action of another causal factor taking place. Independence of the effects of two factors could then be defined as the absence of both synergism and antagonism.

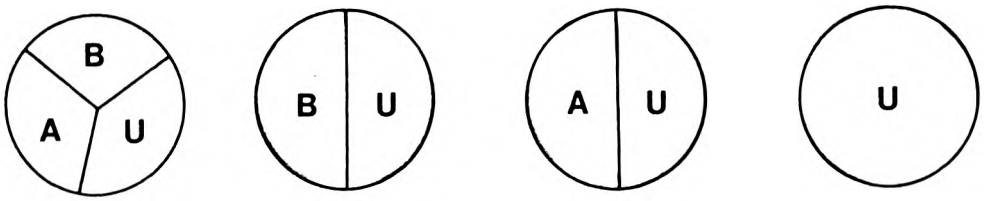
Greenland and Poole (1988) have shown that the causal or preventive interdependence of effects, in particular the concepts of synergism and antagonism as defined by Miettinen, varies with changes in the exposure reference categories. Without a deep knowledge of the individual effects it is difficult to decide on a reference category to study presence of synergism or antagonism. For example, in case of sexually bounded diseases where males may get the disease and females not, it would be difficult to claim that the male sexual hormones increase the risk of the disease, i.e. synergism between male sexual hormones and the other causal effects is implied, or that the female sexual hormones prevent the disease, i.e. antagonism is implied. Nevertheless, using all possible combinations of causal and preventive interdependence for two deterministic exposures, the authors identified patterns of synergistic and antagonistic responses under which the joint effect can be studied even in the absence of "natural" reference category. They also concluded that additivity of risks or rate differences is an elementary criterion for assessing the independence of effects.

Miettinen (1982) has shown, using two deterministic causal factors, that detection of synergism depends upon the degree of correlatedness of the susceptibilities to the two factors. A causation model proposed by Rothman (1976)

is briefly described below and applied to explain the concept of correlated susceptibilities.

Causal models and correlated susceptibilities

The underlying biological processes of disease causation that manifest themselves in a certain risk pattern are generally very complex. A simple model of disease causation was first put forth by Rothman (1976), and later on used by Koopman (1981) to evaluate the measure of deviation from additivity. In the causation model, a sufficient cause was defined as a set of minimal conditions and events that inevitably produce the disease. A completion of the sufficient cause results in the disease onset. The sufficient cause consists of one or more component causes, and each component cause may be present in one or more sufficient causes, which all result in the disease. The component causes may be attributes of the host or attributes of the environment including the presence of an agent (Koopman, 1981). A component cause present in all sufficient causes is called a necessary cause. Fig. 3.1 shows a hypothetical set of four sufficient causes which, when completed, will produce a particular disease.



sufficient cause I | sufficient cause II | sufficient cause III | sufficient cause IV

Fig. 3.1. Hypothetical set of sufficient causes arising from two known causes.

The component causes involved in the causation model are

U – an unknown factor

A – exposure to A

B – exposure to B.

The unknown component in each sufficient cause can be always broken down into more component causes as the knowledge of the disease causation improves.

The cases of the disease in question resulting from the sufficient cause I are produced by the synergistic effect of the two exposures, A and B, i.e. both exposures are necessary to produce the disease. The question then arises of whether we can detect, and if so, quantify the synergistic effect of the two exposures, i.e. the risk of disease resulting from the sufficient cause I. Several factors may determine whether we can identify or quantify synergistic effect. Miettinen (1982) has shown how detection of synergism depends upon the degree of correlation of the susceptibilities to the exposures. For example two extreme cases of correlatedness of susceptibilities to two deterministic (all–or–none effect) exposure factors, A and B, would be as follows:

a) **Maximum correlatedness** – when each effect, A or B, are deterministic causes and those susceptible to B are a subset of those susceptible to A, i.e. the susceptibilities to A and B are not mutually exclusive. Then the expected rate difference (or risk ratio) under the joint exposure is equal to the maximum rate difference (or risk ratio) of the two single effects. In our example it would be the expected rate or risk for the exposure A adjusted for the baseline risk.

b) **Minimal correlatedness** – when individuals susceptible to A are not susceptible to B, but both effects can cause the same disease, i.e. susceptibilities to the disease are mutually exclusive. The expected rate difference (or risk ratio) under the joint exposure is then equal to the sum of the two rate differences (or risk ratios).

Miettinen has shown that when the rate difference exceeds the expected value, derived on the basis of the degree of correlatedness of susceptibilities, presence of synergism should be assumed. However, without the knowledge of the strength of correlatedness we may, for example, assume that a negative interaction

term indicates the presence of antagonism, even though some synergism may be at play, because the susceptibilities to the factors are correlated.

The above concept can be illustrated for the exposures A and B, and applying the above causal model. Let us assume that the unknown component, U, in Figure 3.1, can be broken down into four component causes representing susceptibilities to disease from different combinations of exposure. For simplicity we assume that these are determined by a single gene, as follows:

S_g susceptibility to the background exposure

S_a susceptibility to exposure A

S_b susceptibility to exposure B

S_{a*b} susceptibility to combined effect of A and B.

Let $R_{00}=pr(D=1 \mid A_0, B_0)$

$R_{10}=pr(D=1 \mid A_1, B_0)$

$R_{01}=pr(D=1 \mid A_0, B_1)$ and

$R_{11}=pr(D=1 \mid A_1, B_1)$

represent the risks that disease is present (D=1) given four possible combinations of exposure to A and B, where A_0 and B_0 represent the absence of exposure and A_1 and B_1 the presence exposure.

The above risks of disease can be expressed in terms of probabilities, P, of being susceptible to which ever combination of exposure as follows:

$R_{00}= P(S_g)$ is the risk that a person will get the disease when A and B are absent:

$R_{10} = P (S_g) + P (S_a) - P (S_g \cap S_a)$ is the risk that a person will get the disease when exposure A is present and B is absent;

$R_{01} = P (S_g) + P (S_b) - P (S_g \cap S_b)$ is the risk that a person will get the disease when exposure A is absent and B is present;

$R_{11} = P (S_g) + P (S_a) + P (S_b) - P (S_g \cap S_a) - P (S_g \cap S_b) -$

$P (S_g \cap S_b) + P (S_g \cap S_a \cap S_b)$ is the risk that a person will get the disease

when both exposures are present.

Assuming that no cases of the disease arise from the combined effect of A and B, i.e. that there are no subjects who have susceptibility S_{a*b}, the following index, T, (Hogan *et al.*, 1978)

$$T = R_{11} - R_{10} - R_{01} + R_{00},$$

is expected to be zero. However, when expressed in terms of the above probabilities the expected value is

$$T = P (S_g \cap S_a \cap S_b) - P (S_a \cap S_b).$$

Since $P (S_a \cap S_b)$ will always be greater than $P (S_g \cap S_a \cap S_b)$ the T value will always be negative unless the susceptibilities are mutually exclusive, i.e.

$$P (S_g \cap S_a) = P (S_g \cap S_b) = P (S_g \cap S_b) = P (S_g \cap S_a \cap S_b) = 0.$$

Thus only when the susceptibilities are mutually exclusive would T statistic estimate the correct amount of synergism, i.e. the risk of disease caused by a sufficient cause I in Figure 3.1.

3.3 Assessment of interaction

Assessment of the presence of interaction between two or more risk factors can be made by testing for a departure from a statistical model formulated on the basis of a "null hypothesis" for the absence of interaction. The two commonly postulated "null hypotheses" are the additive and multiplicative hypotheses (Greenland, 1979; Thomas, 1981; Breslow and Storer, 1985; Thomas and Whittemore, 1988). To assess presence of synergism of public health importance, the departure from the additive model should be tested (Rothman, 1974; Blot and Day, 1980; Greenland and Poole, 1988). In this thesis the term additivity or multiplicativity will be used

to refer to the pattern on which the relative risks combine, whereas the mathematical models used to assess the pattern will be referred to as linear (or additive) and exponential (or multiplicative).

3.3.1 Test for additivity of relative risks

Let us define the risk of disease for individual levels of two dichotomous exposures factors A_i and B_j to be $R_{ij}=Pr\ (D=1 \mid A_i\ B_j)$ for i=0,1 and j=0,1. Here D=1 denotes the presence of disease, A_0 and B_0 denote the absence of exposure to factors A and B, respectively, and A_1 and B_1 denote the presence of exposure to factors A and B, respectively. The relative risk of disease for the groups exposed to exposure A only, or to exposure B only, can be written in terms of the background risk R_{00} and in terms of the additional risk due to each exposure as

$$R_{10}=R_{00}+\alpha \tag{3.1}$$

and

$$R_{01}=R_{00}+\beta$$

where α and β are the excess risks due to the exposure A and B, respectively. For the effects of exposures A and B to act additively, i.e. to be constant across each other level, the risk of disease under the joint exposure to both factors is expected to be equal to

$$\begin{aligned} R_{11} &= R_{00} + \alpha + \beta \qquad (3.2)\\ &= (R_{00}+\alpha) + (R_{00}+\beta) - R_{00}.\\ &= R_{10} + R_{01} - R_{00} \end{aligned}$$

When Eq. 3.2 is true than the state of "no interaction on an additive scale" applies. Hogan *et al.*, 1978 proposed an index, T, to be used to assess departures from the state of no interaction on an additive scale,

$$\begin{aligned} T &= R_{11} - R_{10} - R_{01} + R_{00}, \qquad (3.3)\\ &= 0 \text{ if no interaction is present.} \end{aligned}$$

In Eq. 3.2 the risk of disease for the group exposed to both exposures is equivalent to the sum of the risk attributable to the individual exposures and to the background risk. Given that the relative risk of disease in the ijth exposure category, RR_{ij}, is

$$RR_{ij}=R_{ij}/R_{00},$$

Eq. (3.2) can be rewritten in terms of relative risk and excess relative risk as

$$RR_{11}=1 + \alpha/R_{00} + \beta/R_{00}, \tag{3.4}$$

where α/R_{00} and β/R_{00} are the excess relative risks attributable to A and B exposures, respectively.

The presence of an interaction effect between A and B factors is then measured as a departure from the additive model given in Eq. (3.4). The presence of the interaction effect can be assessed by the additive model which includes the interaction term as

$$RR_{ij}= 1 + \alpha' A_i + \beta' B_j + \gamma' A_i B_j,$$

for i and j=0,1 with $A_0=B_0=0$ and $A_1=B_1=1$. The coefficients α', β' and γ' estimate the excess relative risks for the effect of exposures A, B and their combined effect, respectively. It can be shown that the coefficient γ' measures the same effect as the index for interaction T for the relative risk formula. The general form of the additive relative risk model, RR_a, given a vector of explanatory variables $\mathbf{z}=(z_1, z_2,...,z_p)$ which can represent the risk factors or any product terms between the risk factors and can include continuous or discrete variables, is

$$RR_a = 1 + \sum_{j}^{p} \beta_j z_j \tag{3.5}$$

where β_j is the regression coefficient for the jth variable and measures the excess relative risk.

For dichotomous exposure variables there are two indices which can be used to assess the departure from the additivity of relative risks. The T index, Eq.

(3.3), can be used to test departures from the additive "null hypothesis" by using the Z statistic (Hogan *et al.*, 1978)

$$Z= (T - 0)/\sqrt{\text{Var}\ (T)}$$

where T is given in Eq 3.3 and

$$\text{Var}\ (T)= \sum_{i}\sum_{j} RR_{ij}(1-RR_{ij})/N_{ij}, \qquad i,j=0,1$$

where $RR_{ij}= R_{ij}/R_{00}$.

Another index of synergism was developed by Rothman (1976) and under the assumption of independence and rare disease this index is equal to

$$S=(R_{11}-R_{00})/(R_{10}+R_{01}-2R_{00})$$

$$=(RR_{11}-1)/(RR_{10}+RR_{01}-2).$$

The estimated value of index $S>1$ indicates the presence of a synergistic effect and $S<1$ the presence of antagonism. Hogan *at al.* 1978 have shown advantages in using the index T instead of S to assess deviations from additivity in cohort studies.

3.3.2 Test for multiplicativity of relative risks

Under the multiplicative hypothesis of no interaction the expected risk of disease for the group exposed to both exposures A and B. $R_{11}=Pr\ (D=1 \mid A_1B_1)$, is equal to

$$R_{11}= (R_{10}\ R_{01})/R_{00}$$

which can be rewritten in terms of the relative risk as

$$RR_{11}= RR_{10} \cdot RR_{01}. \qquad (3.6)$$

For the two dichotomous exposures A_i and B_j, i=j=0,1, with $A_0=B_0=0$ and $A_1=B_1=1$, the multiplicative relative risk model for testing the departure from the multiplicative hypothesis will be

$$RR_{ij}=\exp(\beta_1 A_i+\beta_2 B_j)$$

and the fully saturated model with the interaction term will be

$$RR_{ij}=\exp(\beta_1 A_i+\beta_2 B_j+\gamma A_i B_j).$$

It follows from Eq (3.6) that the test of the null hypothesis $H_0: \gamma=0$ is equivalent to the test of the null hypothesis (Kleinbaum *et al.*,1982)

$$H_0: RR_{11}=RR_{10}\cdot RR_{01}.$$

The general form of the multiplicative relative risk model, RR_m, is given by

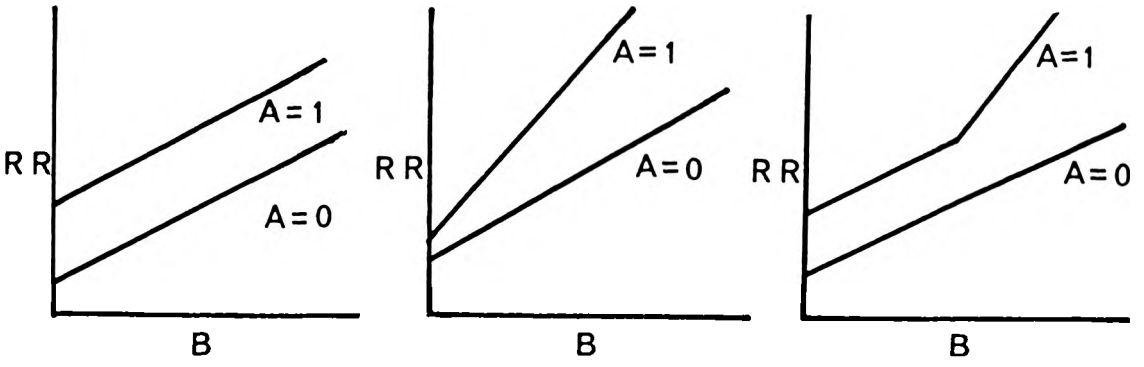
$$RR_m = \exp\left(\sum_{j=1}^{P} \beta_j z_j\right) \qquad (3.7)$$

where β_j estimates the logarithm of relative risk of disease associated with the jth explanatory variable relative to the subjects with $z_j=0$, i.e. the subjects not exposed to the effect of the variable z_j. The set of explanatory variables $\mathbf{z}=(z_1,z_2,...,z_p)$ represents risk factors and any of their product terms and can include continuous, categorized or indicator type of variables.

The so called mixture relative risk models have been proposed as a convenient tool which can distinguish the scale on which interaction is present (Thomas, 1981; Breslow and Storer, 1985: Lubin and Gaffey, 1988). In section 3.5, a brief theoretical background of the theory of the generalized relative risk models is given and the various models proposed specified. The method used for estimation of the relative risk model parameters, by means of the maximum likelihood method, is also presented.

3.3.3 Visual assessment of presence of interaction

By plotting the calculated relative risks against the exposure of interest, a trend in the dose–response relationship can be detected. Consider a continuous variable for an exposure B and a (0,1) indicator variable for an exposure A. By plotting the relative risks for each exposure A against the exposure B, an indication of whether an interaction affect is present or not can be detected. If, for example, the exposure A modifies the effect of the exposure B, then effect modification, i.e. an interaction, is present. Fig. 3.2 shows some of the possible relations.



a) b) c)

Fig. 3.2 The trends in RR that can be observed with two risk factors

The detection of presence of an effect modification depends to some extent upon the transformation scale applied when analyzing the data. For example, when we try to infer the presence of an interaction from the data plotted in Fig. 3.2 in terms of the additive or the multiplicative models, the following facts can be found:

i) the lines are parallel, therefore there is no evidence of an interaction on the additive scale, but there is a negative interaction on the multiplicative scale (Fig 3.2a);

ii) there may be no interaction on the multiplicative scale, but a positive interaction on the additive scale (Fig. 3.2b);

iii) an interaction is present, but may be detected only by means of indicator variables for the exposure B, in which case an additive relative risk model is more likely to detect the interaction (Fig. 3.2c).

3.4 General Relative Risk Models

The general relative risk models (GRRM's) are used in epidemiological studies to assess an association between disease incidence ratio (or relative risk) and a set of explanatory variables, also called risk factors. In particular, these models are useful for study of forms of dose–response relationships with a single continuous or categorical exposure variable, and of forms of interaction between multiple exposure variables (Thomas, 1981).

3.4.1 Specification of the GRRM

Let us write the probability of a binary outcome $y_i=1$, $i=1,2,...,N$, given the ith combination of the explanatory variables $\mathbf{x}_i= (x_{i1}, x_{i2},...,x_{ip})$ as

$$p_i=Pr(y_i=1 \mid \mathbf{x}_i)$$

where y_i, $i=1,2,...,N$ are independent random variables with possible values 0 or 1 and N is the total number of observations. Let $p_0=Pr(y_0=1 \mid \mathbf{x}_0=0)$ denotes the probability of disease for the subjects unexposed to any of the exposure factors, $\mathbf{x}_0=0$. In epidemiological studies it is of interest to estimate the relative risk of disease $R(x_i)$ for a person with exposure $\mathbf{x}_i$, relative to a person who is unexposed to any of the factors, $\mathbf{x}_0$, (Breslow and Storer, 1985). In terms of the probability, p_i, we can define the relative risk as $R(\mathbf{x}_i)=p_i/p_0$, and when disease is sufficiently rare, i.e. $p_0<<1$ and $p_i<<1$, the relative risk can be approximated by the odds ratio

$$OR(\mathbf{x}_i)=p_i(1-p_0)/\{p_0(1-p_i)\}.$$

General linear models relate the probabilities of disease, p_i, to a set of explanatory variables x_{ij} by expressing a "link" function of the probabilities

$$\eta_i=g(p_i)$$

as linear combinations of x_{ij}

$$\eta_i = g(p_i) = \alpha + \sum_{j=1}^{p} x_{ij}\beta_j, \qquad i=1,2....,N$$

with unknown coefficients α and $\beta_1,.....,\beta_p$. The function $g(p_i)$ is in general a non–linear transformation mapping the unit interval (0,1) onto the real line $(-\infty,+\infty)$. The most common link function used in epidemiological studies is the so–called logit function

$$\text{logit}(p_i) = \log[p_i / (1-p_i)], \qquad i=1,2,...,N$$

where log stands for log to the base *e*, i.e. the natural log.

The logit, or logistic, function has number of desirable properties (Thomas, 1981):

a) it goes from $-\infty$ to $+\infty$ as p_i goes from 0 to 1;

b) the linear logistic model $\log[p_i/(1-p_i)] = \alpha + \sum_{j=1}^{p} x_{ij}\beta_j$ is monotonic in each component x_{ij} and changes of equal intervals in any component produce equal change in logit;

c) maximum likelihood estimate (MLE) of β always exists

d) as relative risks tend to be approximately log normally distributed (Thomas,1981), confidence regions estimated from the information matrix at β using approximate Normal theory tend to be close to the true size.

There are, however, situations where many plausible biological models may not be adequately represented by the logistic function. For example, some dose–response relationships between the risk of disease and a single exposure may not be of exponential nature. The logistic model may also not be adequate for assessing the form of interaction between multiple exposure variables, or when fitting complex models (Rothman, 1974; Thomas, 1981).

The general relative risk models (GRRM) relate the relative risk $R(x_i)$ to a linear combination of explanatory variables and their regression coefficients via a linking function, g, as

$$R(\mathbf{x}_i,\beta)=g\left[\sum_{j=1}^{p} x_{ij}\beta_j\right]. \qquad (3.8)$$

For stratified samples of cases and controls a general relative risk model can be fitted to the data as

$$\text{logit}(p_i)= \alpha + \log[R(\mathbf{x}_i;\beta)]$$

where

$\alpha=\text{logit}(p_0)$ is the stratum parameter and

$R(\mathbf{x}_i)$ is given by (3.8) (Breslow and Storer, 1985).

In the special case of the link function g being the exponential function, the traditional logistic regression model $\text{logit}(p_i)= \alpha + \Sigma_j\, x_{ij}\, \beta_j$ is applied. The reason why the logistic model and the multiplicative relative risk models are equivalent derives from the following argument:

$$R(\mathbf{x}_i;\beta) \cong [p_i/(1-p_i)]/[p_0/(1-p_0)]=g\left[\sum_{j=1}^{p} x_{ij}\beta_j\right]$$

as in 3.8. This can be fitted to the data as

$$\log[p_i/(1-p_i)] = \log[p_0/(1-p_0)] + \log\{g\left[\sum_{j=1}^{p} x_{ij}\beta_j\right]\}.$$

When g is a logarithmic function the model becomes equivalent to the logistic model.

3.4.2 GRRM's with power transformation

A power transformation has been proposed as a tool that can discriminate between the additive and the multiplicative models (Thomas, 1981; Breslow and Storer, 1985). The models with the power transformation parameter are also called mixture models as they include the additive and the multiplicative models as special cases. In these models the functional relationship between the relative risk $R(\mathbf{x}_i)$ and the set of risk factors $\mathbf{x}_i$ is defined in terms of the vector of regression coefficients β and a scalar λ. The parameter λ determines the general shape of the

function, and when two or more exposure variables of interest are in the model then λ determines the scale on which they combine. The scale can range from sub–additive to supra–multiplicative, and the additive and the multiplicative models are included as special cases.

The mixture models proposed in the literature are:

1) Breslow and Storer (1985):

$$\log R_{B-S}(\mathbf{x}_i) = \begin{cases} \{ (1 + \mathbf{x}_i'\boldsymbol{\beta})^{\lambda} - 1 \} / \lambda, & \lambda \neq 0 \\ \log (1 + \mathbf{x}_i'\boldsymbol{\beta}), & \lambda = 0 \end{cases} \qquad (3.9)$$

The function reduces to the additive model in the limit $\lambda \rightarrow 0$, and to the multiplicative model when $\lambda=1$.

2) Thomas (1981):

$$R_T(\mathbf{x}_i) = [\exp(\mathbf{x}_i'\boldsymbol{\beta})]^{\lambda} [\, 1 + \mathbf{x}_i'\boldsymbol{\beta}]^{1-\lambda} \qquad (3.10)$$

In this model $\lambda=0$ corresponds to the additive model, and $\lambda=1$ corresponds to the multiplicative model.

3) Guerrero and Johnson (1982):

$$R_{G-J}(\mathbf{x}_i) = \begin{cases} [1 + \gamma(\mathbf{x}_i'\boldsymbol{\beta})]^{1/\gamma} & \text{if } \gamma \neq 0 \\ \exp (\mathbf{x}_i'\boldsymbol{\beta}) & \text{if } \gamma = 0 \end{cases} \qquad (3.11)$$

Here, $\gamma=0$ corresponds to the multiplicative relative risk model and $\gamma=1$ corresponds to additive relative risk model.

4) Lubin and Gaffey (1988):

For any of the relative risk functional forms given above under 1)–3), a large class of models have been proposed such that for any $R(x_{i_1})$ and $R(x_{i_2})$ (which may be

of the R_{B-S} , R_T ,or $R_{(i-J}$type) the multiplicative or additive association can be assumed. Two such models have been suggested:

a) $$R_{TM}(x_{i1},x_{i2})= [R(x_{i1})R(x_{i2})]^{\theta} [R(x_{i1})+R(x_{i2}) - 1]^{1-\theta} \quad (3.12)$$

When θ=1, the exposures combine multiplicatively and when θ=0, additively.

b) $$R_e(x_{i1},x_{i2}) = 1 + e(x_{i1}) + e(x_{i2}) + \theta e(x_{i1})e(x_{i2}) \quad (3.13)$$

where $e(x_{ij})= R(x_{ij}) - 1$ for j=1,2 are the relative excessive risks and θ defines the scale on which the individual risk functions combine. When θ=1 the function is multiplicative and when θ=0 it is additive.

3.4.3 Problems associated with GRRMs

Recently certain shortcomings of the mixture models have been pointed out in the literature (Moolgavkar and Venzon, 1986, 1987; Lubin and Gaffey, 1988).

a) The models 1 and 2 of section 3.4.2 were found not to be invariant under an arbitrary choice of coding of a binary covariate. In the additive model, a change of coding for one of the x variables leads to a nonlinear transformation of the parameters for all the x variables. Thus the interpretation of the regression parameters depends upon the choice of coding.

b) Problems with inference may arise with additive relative risk functions. It was pointed out that when the relative risk function is not exponential then the Wald test and the Wald confidence intervals derived form the asymptotic covariance matrix of the estimates are poorly behaved and, in addition, are dependent upon the coding scheme. Moolgavkar and Venzon have shown that the maximum likelihood–based 95 per cent confidence intervals are independent of the coding. In addition, as the strength of the association increases, the confidence interval derived for the additive model based on the Wald statistics becomes too wide and

includes unity. The likelihood based confidence interval was shown to be more narrow and less likely to include unity with the increasing strength of the association. Moolgavkar and Venzon concluded that unless the sample size is very large, inference based on the asymptotic standard errors may be seriously flawed when the relative risk function is not multiplicative, and recommended that likelihood–based procedures be used for inference.

c) Lubin and Gaffey pointed out that the assumption of the same functional relationship for all the covariates included in the mixture model may be a shortcoming. It was suggested by them that the mixture models 1 – 3 of section 3.4.2 be used for modelling of relative risk pattern of the individual exposures, and the large class of models, as for example models specified under 4) in section 3.4.2, be used to model the combined effects of two or more covariates.

The above criticisms do not invalidate the mixture approach as a test for comparing alternative models, but they should caution against overinterpreting the point estimate of the mixture parameter.

3.5 Estimation of regression coefficients

The general relative risk models parameters, θ_j, are estimated by the maximum likelihood method. The maximum likelihood estimators θ_j of these parameters are found by solving simultaneous equations

$$\frac{\partial \ln L}{\partial \theta_j} = 0 \qquad j=1,2,...,k$$

where L is the likelihood function based on the sample and k is the total number of the parameters of the model. In our case θ_j represent coefficient vectors α and β. and the power parameter λ.

The derivation of the maximum likelihood equations depends on the study

design employed. The three study designs employed in the present thesis are unmatched case–referent study, matched case–referent study and follow–up (or cohort) study and the likelihood functions derived for these studies are described.

3.5.1. Unmatched case–referent study

For the unmatched case–referent studies the unconditional likelihood function, L, can be derived directly by assuming independent, binomially distributed response variables y_i, i=1,2,...,N, with parameters p_i and n_i (McCullagh and Nelder, 1983). Under this assumption, the probability of observing the values y_i's, or the likelihood of the observed data, is given by

$$L = \prod_{i=1}^{N} \binom{n_i}{y_i} p_i^{y_i} (1-p_i)^{n_i - y_i} \qquad (3.14)$$

By making the probability parameter p_i dependent on a set of explanatory variables x_i and on their coefficients θ_j, such estimators θ_j can be found which maximize the value of L, i.e. maximize the likelihood of the occurrence of the observed data.

The likelihood function derived from (3.14) suitable for analysis of grouped data, i.e. when $n_i>1$ is then as

$$\ln L = \sum_{i=1}^{N} y_i \ln p_i + (n_i - y_i) \ln (1-p_i) + C \qquad (3.15)$$

and for ungrouped data, when $n_i=1$, the log likelihood is derived from (3.14) as

$$\ln L = \sum_{i=1}^{N} y_i \cdot \ln p_i - (1-y_i) \cdot \ln (1-p_i) \qquad (3.16)$$

with $0 <= (y_i/n_i) <= 1$.

3.5.2 Matched case–referent study

For the matched case–referent studies the β estimators are obtained from the conditional probability (likelihood) that the first n_i subjects out of N are in fact

cases, given all the $\binom{n}{n_1}$ possible ways of dividing n subjects into n_1 cases and n_0 controls (Holford *et al.*, 1978; Breslow *et al.*, 1978, Prentice and Breslow, 1978; Thomas, 1981)

$$L = \prod_{i=1}^{n} \{ \exp(\beta' z_{i0}) / \sum_{k \in R_i} \exp(\beta' z_{ik}) \}$$

where z_{i0} denotes the vector of explanatory variables for the ith case, i=1,2,...,n, R_i the risk set consisting of ith case and its set of k−1 matched controls, z_{ik} is the vector of explanatory variables for the kth member of the ith risk set and β is the vector of regression coefficients. The likelihood function for the general relative risk model was derived (Thomas, 1981) by replacing $\exp(\beta' z)$ by any positive function $R(\theta, z)$ differentiable with respect to components of θ as

$$L = \prod_{i=1}^{n} \frac{R(\theta, z_{i0})}{\sum_{k \in R_i} R(\theta, z_{ik})}.$$

3.5.3 Follow–up (cohort) study

In a follow–up study the time over which each individual or the whole cohort is followed–up is also considered. Holford (1980) has shown that the number of deaths, d, occurring in a particular age–time–exposure cell has a Poisson distribution and that the general theory of survival distribution applies to it. Suppose there are J cells with associated death rates h_j. Let $\delta_{ij}=1$ or $\delta_{ij}=0$ denote whether an ith individual in the jth cell dies or not, respectively and let t_{ij} denote his contribution to the observational time n_j in that cell. According to the theory of survival distribution the log–likelihood function of the data may be written as

$$\ln[L(h)]=\ln[L(h_1,...,h_j)]=\Sigma_i \Sigma_j(\delta_{ij} \log h_j - t_{ij}h_j)=\Sigma_j(d_j \log h_j - n_j h_j),$$

and the likelihood also arises when d_j are independent Poisson variables (Holford, 1980). The Poisson assumptions are applicable for large–sample likelihood inferences in which the number of deaths in each cell is small.

3.6 Statistical inference

Measures of goodness–of–fit

To find the best fitting model, two statistics are most commonly used:

1) Log likelihood ratio statistic. The log likelihood function, lnL, can be used directly to calculate the log likelihood ratio as

$$-2\ln (L_1 / L_2)$$

where L_1 is the likelihood function for the reduced model (reduced by one or more explanatory variables to be tested for statistical significance) and L_2 is the likelihood function for the full model which includes the full set of the explanatory variables.

The log likelihood ratio is asymptotically distributed as the chi–square distribution with p degrees of freedom, where p is the number of parameters by which the full and the reduced model differ.

2) Deviance statistic. This statistic measures the discrepancy between the observed proportion, p_i, and the fitted value, Π_i, by the formula

$$\sum_{i=1}^{N} \ln (p_i / \Pi_i).$$

By substituting the binary formula for proportion into the above formula and writing it in terms of the observed y_i and fitted counts μ, we obtain the formula for deviance (McCullagh and Nelder, 1983) as

$$D (y_i,\Pi_i) = 2 \Sigma \{y_i \ln (\frac{y_i}{\mu}) + (n_i - y_i) \ln[(n_i-y_i)/(n_i-\mu)]\}.$$

Test of hypothesis and estimation of confidence interval

To test the hypothesis that the parameters θ_i are different from some null value $\overline{\theta}_i$ two tests are commonly used; the Wald test and the likelihood ratio test; both tests are asymptotically equivalent. However, it has been shown that the

confidence limit derived from the Wald statistics may have a large error for the mixture models and that the likelihood–based confidence limit is more reliable (Moolgavkar and Venzon, 1987).

Wald test of $H_0 : \theta_j = \overline{\theta}_j$ is given as

$$X^2_1 = [(\hat{\theta}_j - \overline{\theta}_j) / \sqrt{\quad Var \quad (\hat{\theta}_j)}$$

where Var ($\hat{\theta}_j$) is derived from the asymptotic covariance matrix of the estimates. The Wald 95 percent confidence interval for θ_j is given by

$$\hat{\theta}_j \pm 1.96 \sqrt{\text{Var}}\ (\hat{\theta}_j).$$

Likelihood–based test of H_0: $\theta_j = \overline{\theta}_j$ is based on the likelihood ratio statistic defined by

$$LR(\hat{\theta}) = 2[\ln \varphi(\hat{\theta}) - \ln \varphi(\overline{\theta})],$$

where $\hat{\theta}$ is the maximum likelihood estimate of θ and under H_0 the LR is distributed as a chi–square variate with 1 degree of freedom. To construct the 95 percent confidence interval for a single parameter θ_1 in a vector of parameters $(\theta_1,\theta_2,...,\theta_k)$, a method based on a profile likelihood (Cox, 1970) is defined by

$$\{\theta_1 |\ 2\ln \varphi\ (\hat{\theta}_1,\hat{\theta}_2,...,\hat{\theta}_k) - 2\ln \varphi\ (\theta_1,\ \hat{\theta}_{-1}(\theta_1)\} \leq 3.84,$$

the function $\varphi(\theta_1,\hat{\theta}_{-1}(\theta_1))$ is such that the $\hat{\theta}_{-1}(\theta_1) = (\hat{\theta}_2(\theta_1),...,\hat{\theta}_k(\theta_1))$ is that value of $(\theta_2,...,\theta_k)$ which maximizes the likelihood for a given value of θ_1.

3.7 General modelling strategy

It is recommended that firstly the nature of the interaction present in the data be

assessed before any consideration is given to issue of confounding (Kleinbaum *et al.*, 1982). Pearce (1989) on the other hand suggested that the initial step in the analysis of interaction is to determine if there is an overall effect of exposure. To do that, it is necessary to control confounding effects and interactions among the confounders, and this should be done before the exposure interaction terms are included in the model. He suggested that confounding be evaluated purely on *a priori* considerations.

The modelling strategy adopted in the present thesis is the hierarchical method recommended by Kleinbaum *et al.* (1982). The steps recommended are listed and the factors and product terms applicable to the model building in the present thesis are given in parenthesis:

1) Specify the main effect variables for main exposures (Dust and Smoking) and confounding effects (Age).

2) Specify all the second and third order interaction terms which include the main exposure factors (Age x Dust, Age x Smoking, Dust x Smoking, Age x Dust x Smoking).

3) Determine any three factor interactions to be retained in the model.

4) Include in the further model any two factor products that are included in the three factor interactions.

5) Determine any other two factor interactions to be retained in the model.

6) Include in the further model any main factors that are included in the two factor interactions.

7) Assess the effect of the confounding factor (Age) if it is not included in any interaction term.

8) If possible eliminate from the model main effect terms that are not statistically significant and are not part of a significant two or three factor interaction.

The log likelihood statistic is recommended to identify significant terms to

be retained in the model.

As the Breslow and Storer general relative risk model includes the additive and multiplicative models as special cases, the above procedure will be followed for the additive and multiplicative model.

It is recommended that a general impression of the type of interaction present in the data be obtained by evaluating the stratum–specific effect measures resulting from stratified analyses, and the variables identified by the inspection as potential effect modifiers be included in the initial model.

3.8 Estimation of Attributable risk

The ultimate aim of epidemiological studies is to evaluate the proportion of cases with a disease that could be prevented by the elimination of a particular risk factors. In the case of chronic lung diseases associated with silica dust exposure the aim is to estimate the proportions of cases associated with dust and with tobacco smoking.

The method of quantifying the risk attributable to a specific risk factor in a population has received considerable attention in the epidemiological literature in recent years. It has been shown that an attributable risk (also called etiological fraction) can be estimated for multi–level multifactor types of associations for case–control study designs (Miettinen, 1974; Bruzzi *et al.*, 1985; Kjuus *et al.*, 1986). The method for calculation of attributable risk as developed by Miettinen (1974) and applied by Bruzzi *et al.* (1985) and Kjuus *et al.* (1986) was used for the present analysis and the term etiological fraction (EF) used by them is used in this thesis.

The etiological fraction for the ith level of an exposure factor (i=1,2,...,l) is

calculated as $EF_i = p_{ci} \cdot [(RR_i - 1)/ RR_i]$, where RR_i is the relative risk of disease for the exposed group relative to the non–exposed group standardized for the effect of the other exposure and any confounding effect, and p_{ci} is the proportion of cases found in the ith level. For the non–exposed, i.e. baseline, group $EF_0 = 0$. For the risk factor with k increasing levels of effect the summary etiological fraction is then calculated as $EF = \sum_{i=1}^{k} EF_i$, for $EF_i > 0$. The etiological fraction of two combined exposures of dust and smoking, EF_{ds}, is calculated by using the subjects with no exposure to any of the risk factors as the referent category and by summing over all two–way cross–classified cells. The 95% confidence intervals for the summary etiological fractions were estimated as in Denman and Schlesselman (1983).

Using the summary etiological fractions for the effect of individual risk factors and both risk factors together, the fraction of cases that could be prevented by elimination of a particular factor can be calculated as in Kjuus *et al.* (1986). Given exposure to dust and to smoking the etiological fractions that can be estimated from the data are EF_{dust}, $EF_{smoking}$, $EF_{dust+smoking}$. Applying these fractions, the proportion of cases that can be prevented by elimination of each exposure, preventable fractions, PF, can be can be estimated as follows:

$$PF_{dust} = EF_{dust+smoking} - EF_{smoking}$$

$$PF_{smoking} = EF_{dust+smoking} - EF_{dust}$$

$$PF_{dust+smoking} = EF_{dust+smoking} - PF_{dust} - PF_{smoking}$$

$$PF_{other} = 1 - EF_{dust+smoking}.$$

CHAPTER 4

SIMPLE CHRONIC BRONCHITIS

4.1 Objectives of chapter 4

The objective of this chapter is to examine whether tobacco smoking modifies the effect of underground gold mining dust exposure on the prevalence of simple chronic bronchitis. For the purpose of epidemiological studies the presence of simple chronic bronchitis is generally assessed by a questionnaire. It has been, however, suggested that miners have a tendency to exaggerate the response to a questionnaire on chronic bronchitis (Wiles and Faure, 1977). Thus, in order to validate the results obtained from the analysis of questionnaire response, the measurements of mucus production were analyzed also.

The data from the following two studies were re–analyzed:

a) The population study of chronic bronchitis done in Carltonville, South Africa (Sluis–Cremer *et al.*, 1967);

b) The cohort study of 2209 South African white gold miners and 483 controls by Wiles and Faure (1977).

The methods and results are presented in Sections 4.2 and 4.3, respectively.

4.2 THE CARLETONVILLE STUDY

4.2.1 Study subjects

A study of a mixed mining and non–mining population of Carltonville on the Witwatersrand was done in 1965 to determine whether there is an increased prevalence of simple chronic bronchitis among the gold miners. A detailed description of the study and the findings were published by Sluis–Cremer *et al.* (1967), and only a brief summary of relevant information is given here.

A sample of 827 white males was obtained by visiting every third house in the town and interviewing all the males over the age of 35. A person was classified as exposed to dust if he had worked, at any time in his life, in a scheduled dusty area on mines. The smoking data were collected in terms of tobacco consumption per day as at the time of the study, and for the purpose of the present analysis the subjects were categorized as smokers or non–smokers. Presence of chronic bronchitis was established by means of a standard questionnaire approved by the Committee on the Aetiology of Chronic Bronchitis of the Medical Research Council of Great Britain (1961).

Table 4.1 shows the observed prevalences for chronic bronchitis tabulated according to the presences of gold mining dust exposure and smoking habits, stratified by age categories 35–44, 45–54 and >55. In the original analysis of Table 4.1, the prevalences of the dust–exposed groups were compared to those in the non–exposed groups, in smokers only, and the differences were found to be significant at 5% probability level for the 35–44, and 45–54 age groups, but not for the >55 age group, where the sample size was too small. The conclusion drawn from the study was that gold mining exposure is a risk factor for chronic bronchitis, when age and smoking are standardized for.

4.2.2 Combined effect of mining dust and smoking on the presence of simple chronic bronchitis

Crude odds ratios

The null hypothesis of the absence of an interaction between tobacco smoking and the presence of gold mining dust exposure was tested as a departure from the additive and the multiplicative model in terms of the odds ratios as follows:

$OR_{11} > OR_{10} + OR_{01} - 1$, for departure from the additive model;

$OR_{11} > OR_{10} \cdot OR_{01}$, for departure from the multiplicative model.

The above tests were done for data stratified by age and for unstratified data in order to establish whether age confounds the effect of the combined exposures.

Table 4.1 shows the crude odds ratios relating the odds of simple chronic bronchitis in the three exposed categories to the odds in the unexposed category, within each age strata and for all the ages combined. As no subjects were found with simple chronic bronchitis in the non–miners age categories 35–44 and 55+, the value 0.5 was added to all the cell frequencies. This was done in order to establish whether there is a trend in the odds ratios.

The observed odds ratios for the joint exposures lie, on the whole, between the additive and multiplicative model.

Confounding by age

The effect of age is best seen by examining the marginal prevalences for the individual exposure categories. The prevalences for the smokers and nonsmokers, according to the age categories and ignoring the effect of dust, were (9.0, 12.9 and 17.5) and (2.5, 3.4 and 3.0), respectively. Thus age appears to modify the effect of smoking. The prevalences for miners and nonminers, according to the age categories and ignoring the effect of smoking, were 9.6, 10.7 and 14.8, and 1.6, 5.3 and 9.5, respectively. Here age does not appear to modify the effect of dust. The

trend with age appears to be similar in smoking nonminer (2.6, 5.6 and 13.3) and smoking miners (11.1, 16.2 and 21.2). Thus from the inspection of the unadjusted prevalences the interaction term for Dust x Smoking can be expected to be important in the multifactorial model. The marginal prevalences for age categories show a systematic trend (11.1, 16.2 and 21.2), indicating that age is a confounding factor in the association under question.

Before applying the general relative risk model to assess the scale on which the interaction effect is significant, the effect that age had on the above association was examined in a multifactorial analysis. The confounding and modifying effect of age was tested by the additive and the multiplicative relative risk models.

The indicator variables D and S, represented the presence of mining dust exposure as D=1 and the absence as D=0, and the smoking habits as S=1 if current smoker and as S=0 if current non–smoker. The three age categories were represented by two indicator variables AGE2 and AGE3, AGE2=1 if age=45–54, and AGE3=1 if age $\geq$ 55, with the variables equal to zero if otherwise.

The additive and the multiplicative relative risk models (Eq.'s 3.5 and 3.7, respectively) were fitted to the data with the following saturated model :

$$RR = \beta_0 + \beta_1 \cdot (AGE2) + \beta_2 \cdot (AGE3) + \beta_3 \cdot (D) + \beta_4 \cdot (S) + \beta_5 \cdot (S \cdot D) +$$
$$\beta_6 \cdot (AGE2 \cdot D) + \beta_7 \cdot (AGE3 \cdot D) + \beta_8 \cdot (AGE2 \cdot S) + \beta_9 \cdot (AGE3 \cdot S).$$

(For the additive relative risk model the value of unity was added.) The models with third order interactions were fitted also, but none of the models converged. The inclusion of third–order interactions would allow us to assess whether age modifies the interaction effect between dust and smoking. The second–order interactions allow us to test whether age modifies the individual effects of dust and smoking on the prevalence of chronic bronchitis. The log likelihood ratio statistic, X^2_{LR}, was applied to test the following hypotheses:

1) $H_0 : \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0$ versus

H_a : any of the β's $\neq 0$

2) $H_0 : \beta_5 = 0$ versus $H_a: \beta_5 > 0$

3) $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ versus

H_a : any of the >0.

The log likelihoods, $-2\log L$, and the X^2_{LR} obtained from the multiplicative models were as follows :

Effect tested	$-2\log L$	X^2_{LR}	df
2nd order interactions with age	443.56	2.4	4
Dust x Smoking	445.91	0.3	1
Main effects	446.20	33.4	4
Intercept	479.60		

The log likelihood ratio statistic, $X^2_{LR}=2.4$, df=4, $p>0.05$ does not support the hypothesis that age modifies the effect of dust or smoking, i.e. the alternative hypothesis 1) H_a. The hypothesis that the interaction between dust and smoking is significant in the multiplicative model is rejected as the value of the test statistic, X^2_{LR} (D · S|main effects)=0.29, falls in the rejection region, $p>0.50$. However, the main effects are statistically significant since the X^2_{LR} (main effects|intercept) = 33.4, with df = 4 and $p<0.001$.

The log likelihoods and the log likelihoods ratios obtained from the additive models were as follows:

Effect tested	$-2\log L$	X^2_{LR}	df
2nd order interactions with age	441.80	4.9	4
Dust x Smoking	446.72	5.2	1
Main effects	451.96	27.6	4
Intercept	479.60		

Again, the hypothesis that age interacts with dust or smoking is not supported as the value of the test statistic, X^2_{LR} (2nd order interactions|main effects, Dust·Smoking) = 4.92, df=4, falls in the rejection region, $p>0.20$. The hypothesis that the interaction between dust and smoking is significant on the

additive scale is, however, supported by the test, as X^2_{LR} (D · S|main effects)= 5.24, with df = 1 and $p<0.05$. Thus the results indicate that the interaction between the presence of gold mining dust exposure and tobacco smoking is significant in the additive model, and that, according to the additive model, age does not modify the effect of dust or smoking.

Estimations by the general relative risk model

To establish the scale on which the interaction is present the general relative risk model was fitted to the data as:

$$\text{logit}(p)=\alpha_0+\alpha_1\cdot(\text{AGE2})+\alpha_2\cdot(\text{AGE3})+\text{RR}(\text{dust},\text{smoking};\beta) \qquad (4.1)$$

where p is the probability of disease and RR(dust, smoking;β) is equal to:

1. $RR_a=\log(1+\beta_1\cdot(D)+\beta_2\cdot(S))$, for the additive model,
2. $RR_a=\log(1+\beta_1\cdot(D)+\beta_2\cdot(S)+\beta_3\cdot(D\cdot S))$, for the additive model with an interaction term,
3. $RR_m=\beta_1\cdot(D)+\beta_2\cdot(S)$, for the multiplicative model,
4. $RR_P=[(1+\beta_1\cdot(D)+\beta_2\cdot(S))^{\lambda}-1]/\lambda$, for the power model by Breslow and Storer,
5. $RR_t=\log\{[1+\beta_1\cdot(D)+\beta_2\cdot(S)]^{\lambda}[(1+\beta_1\cdot(D))\cdot(1+\beta_2\cdot(S))]^{1-\lambda}\}$, for the power model suggested by Thomas (1981).

The parameter estimates and the log likelihood values (–2logL) obtained from the five models are presented in Appendix A, Table A.1. According to the log likelihood statistics shown in Table A.1, the models 2 to 5 significantly improved the fit to the data above the main effect additive model. The additive model with the interaction term, model 2, and the multiplicative model, model 3, gave almost identical goodness–of–fit values (446.72, and 446.20, respectively). The two power models, model 4 and 5, showed slight improvement to the fit (445.91 and 445.96, respectively), and again the goodness–of–fit values were virtually identical.

The predicted and observed frequencies for chronic bronchitis are shown in Table 4.2. The two power models give virtually identical fitted values, which lie, most of the time, in between the predicted values obtained from the additive model with the interaction term and the multiplicative model.

The age adjusted odds ratios for the comparison of non–smoking nonminers with non–smoking miners, smoking nonminers and smoking miners were estimated:

a) by the additive model with interaction term (model 2) as (1+2.78)=3.78, (1+3.98)=4.98, and 1+2.78+3.78+7.23)=14.99, and

b) by the multiplicative model (model 3) as exp(1.02)=2.77 (1.4–5.5), exp(1.45)=4.26 (2.0–9.2), and exp(1.02+1.45)=11.82 (1.82–74), respectively.

The confidence interval for the cell with the combined exposure was estimated according to a formula given by Kleinbaum *et al.* (1982).

There is one observation that does not agree with the description of the power model as given by Breslow and Storer. The value of $\lambda = 2.2$ would suggest that the two exposure factors combine on a supra–multiplicative scale, and thus we would expect the interaction in the multiplicative model to be significant. Although there was some evidence of the presence of interaction on the multiplicative scale as given by the $OR_{11} = OR_{01} \cdot OR_{10}$ comparison, this evidence was not, supported by the goodness–of–fit statistic obtained from the multiplicative model with an interaction term; the interaction term estimated by the multiplicative model had a negative value.

4.2.3 Conclusions from the analysis

The multifactorial analysis of Table 4.1 has shown that the interaction between

the presence of gold mining dust exposure and tobacco smoking is statistically significant in the additive model, when age was standardized for. Thus, the statistical evidence supports the hypothesis that the effect of the combined exposure to tobacco smoking and gold mining dust on the presence of simple chronic bronchitis is significantly greater than the sum of the individual effects and of the background effect, i.e. that the two exposures act synergistically in the public health context.

The multiplicative model fitted the data significantly better than the additive model, and the age adjusted odds ratios for the comparison of non–smoking nonminers with non–smoking miners, smoking nonminers and smoking miners were estimated by the multiplicative model (model 3) as exp(1.02)=2.77 (1.4–5.5), exp(1.45)=4.26 (2.0–9.2), and exp(1.02+1.45)=11.82 (1.82–74), respectively.

The two power models applied to test the scale on which the interaction is present gave the same estimated values and goodness–of–fit statistics. However, both of the models fitted the data best when the value of the power parameter was approximately 2, indicating a supra–multiplicative scale. This may be due to the small sample and the zero values in some cells.

Table 4.1 Prevalence of chronic bronchitis, by presence of gold mining dust and smoking status, stratified by age

Age	Exposure		Chronic Bronchitis	Total	Chronic Bronchitis	Observed
	Dust	Smoking	No.	No.	(prevalence)	OR
35–44	No	No	0	52	0.0	1.00
		Yes	2	77	2.6	3.47
	Yes	No	3	68	4.4	5.60
		Yes	27	244	11.1	13.28
45–54	No	No	2	40	5.0	1.00
		Yes	3	54	5.6	1.12
	Yes	No	2	79	2.5	0.49
		Yes	19	117	16.2	3.68
≥55	No	No	0	12	0.0	1.00
		Yes	4	30	13.3	4.25
	Yes	No	1	21	4.8	1.83
		Yes	7	33	21.2	7.07
All ages	No	No	2	104	1.9	1.00
		Yes	9	161	5.6	3.02
	Yes	No	6	168	3.6	1.89
		Yes	53	394	13.5	7.93
TOTAL			70	827	8.46	

Table 4.2 Predicted No. with chronic bronchitis, estimated from different RR models, fitted to age stratified data

Age	Effects Dust	Effects Smoke	Total No.	Observed with Bronch. No.	Predicted No. with bronchitis by relative risk models 1	2	3	4	5
35–44	0	0	52	0	0.37	0.72	0.53	0.71	0.70
	0	1	77	2	5.10	3.02	3.23	3.03	3.03
	1	0	68	3	2.80	1.70	1.87	1.71	1.71
	1	1	244	27	23.50	26.55	26.38	26.55	26.55
45–54	0	0	40	2	0.42	0.87	0.63	0.86	0.84
	0	1	54	3	5.09	3.25	3.45	3.26	3.26
	1	0	79	2	4.69	3.06	3.33	3.07	3.07
	1	1	117	19	15.83	18.83	18.59	18.81	18.84
≥55	0	0	12	0	0.18	0.40	0.29	0.40	0.40
	0	1	30	4	3.95	2.73	2.88	2.74	2.75
	1	0	21	1	1.77	1.25	1.35	1.25	1.25
	1	1	33	7	6.13	7.59	7.49	7.61	7.62

4.3 THE WILES AND FAURE STUDY

4.3.1 Study subjects and methods

Study subjects

In accordance with the South African law, every miner is compelled to undergo an annual (so called periodical) medical examination when working in a dusty occupation. Many miners who have left mines continue to come to an occasional medical examination in order to keep their fitness certificate. By the age of 45 many miners request a special (so called benefit) medical examination to qualify for compensation for disability due to pneumoconiosis or chronic obstructive lung disease. A sample of 2209 gold miners was selected from men attending the MBOD during the four year period between 1968 and 1971 on the basis of the following criteria:

a) Bureau number between B9000 and C8999;

b) Age between 45 and 54;

c) Minimum 10 years of underground service on gold mines;

d) Resident in South Africa for not less than 20 years;

e) Service in mines other than gold, less than 2 years.

A control group of 483 non–miners was drawn from various municipal and government departments. A whole office or workshop was asked to volunteer (if aged between 45 and 54). This group also included 213 workers employed in gold mines who had always worked in non–dusty occupations, such as clerks, artisans etc.

Information collected on each man and variables derived from it

The investigations done on the men and relevant to this analysis were as follows:

a) **Bronchitis questionnaire.** An adapted version of the British Medical Council Bronchitis questionnaire (1961) was used. Chronic bronchitis was considered present if productive cough, that is cough and phlegm, occurred in the early morning and in addition at other times during the day for 3 months in a year.

b) **Clinical examination.** A full clinical examination including measurement of sputum production was done on each man. For the purpose of this analysis sputum production was considered abnormal when the man expectorated 2ml or more of mucus substance, in 5 minutes. Only mucotic substance was measured.

c) **Lung function tests.** The lung function parameters measured were: Vital Capacity (VC), Forced Expiratory Volume in one second (FEV_1) and Forced mid–Expiratory Flow (FEF_{25-75}).

d) **Mining service.** The mining service was estimated from the Chamber of Mines records which show the exact number of shifts worked in each occupation. The cumulative dust exposure was calculated as the sum of the products of number of years worked in each occupation and the dust concentration for the occupational group. The occupations were classified into seven occupational groups on the basis of dust concentration measured by Beadle (1971). The concentration was measured as respirable surface area (RSA), which is the total surface area of particles in the respirable fraction of the dust. Thus calculated cumulative dust exposure (RSA x years), was used in previous studies (Wiles and Faure, 1977; Irwig *et al.*, 1978). Several variables were used to analyze the effect of mining dust exposure. To indicate the presence/absence of gold mining dust exposure, a dichotomous (0, 1) variable (DUST) was used. To evaluate a disease dose–response relation with cumulative dust exposure, two variables were employed: a continuous variable (CDE) representing cumulative dust exposure and a categorical variable (CDEC) representing

categorized cumulative dust (RSA x years/1000) (0, 1–20, 21–30, 31–40, >40) where the actual values used in the analysis were the means of each range (Berry, 1980). The actual years of mining were also calculated as the total number of dusty shifts divided by 270 (the average number of shifts per year).

e) **Smoking habits.** Full details of smoking habits were obtained by means of a smoking questionnaire and the reliability of the answers was checked against smoking histories recorded in the medical files kept at MBOD. The data on the following questions were collected for cigarette and pipe smoking separately:

age when started smoking, how long ago stopped smoking, average cigarettes per day/ tobacco (oz. per week) consumed now, maximum cigarettes per day/ tobacco (oz. per week) ever consumed, and total temporary stoppages.

For the present analysis, the answers to these questions were used to derive the following variables for cigarettes, pipes and total tobacco:

i) Years of smoking, calculated as

(year of examination – year of birth – age started smoking – temporary stoppage – how long ago stopped).

ii) Pack years, calculated as follows:

Cigarettes=((cigarettes now + maximum)/(2·20)) · years smoked.

Pipes=((tobacco now + maximum)·28 /(2·7))· years smoked.

Cigarette equivalent=cigarettes+pipes.

Here 20 represents the number of cigarettes per packet, 28 the number of cigarettes per ounce of tobacco and 7 is the number of days per week.

iii) Smoking status, calculated as:

0 – never smoked

1 – stopped smoking tobacco 2 years or more

3 – still smoking or stopped less than 2 years ago.

iv) Number of cigarettes equivalent smoked per day

(cigarettes now + maximum)/2 +(tobacco now + maximum)·11/(2·7).

To select a smoking variable that was most strongly related to the presence of chronic bronchitis the stepwise logistic regression was applied. The smoking variables (i–iv), age, cumulative dust exposure and years of mining were included in the regression model. The levels of significance for the variables selected into the 'best' fitting model were as follows:

Effect	$\chi^2_{(1)}$	P–value
Smoking status	93.9	<0.0001
Cumulative dust exposure	21.2	<0.0001
Years of gold mining	13.9	<0.001
Age at the examination	7.1	<0.01
Cig. eq. pack–years	5.1	<0.025

The smoking status and cigarettes equivalent pack–years were the best predictors of the effect of smoking on the prevalence of chronic bronchitis in the data.

To establish whether the time when an ex–smoker stopped smoking explains some of the variation in the presence of chronic bronchitis, an analysis of the 527 ex–smokers was done. The variables selected into the above 'best' model were included together with the number of years since the subject stopped smoking tobacco. The only variable that was selected as being significant at 0.15 probability level was CDE (X^2=17.96, df=1, p<0.0001). This suggests that if an individual stops smoking for two or more years, then the number of years that he stopped and the previous smoking habits are not significantly related to the presence of chronic bronchitis.

To establish whether the effect caused by smoking can be effectively modelled by one smoking variable which would combine the information from the smoking status and cigarettes equivalent pack–years, several joint variables

combining the two variables were created and tested. The 'best' fitting model was obtained with a smoking variable which had the following categories: 0=non–smoker, 1=ex–smoker, 2=current smoker of <15 pack–years, 3=current smoker of 15–59 pack–years, and 4=current smoker of 60+ pack–years (SMOKEC). The levels of significance for the factors in the model were as follows:

Variable	$\chi^2_{(1)}$	p–value
SMOKEC	178.5	<0.0001
CDE	21.1	<0.0001
Years of gold mining	15.1	<0.001
Age at examination	7.7	<0.01

In further analyses the categorized smoking variable with values 0 to 4 (SMOKEC) and the actual values for cigarette equivalent pack–years (PACKYR) were used to estimate the effect of dose, and an indicator (0,1) variable (SMOKE) was used to compare smokers with non–smokers.

Comparison of miners according to type of examination

The study sample of miners consisted of men who came for a periodical medical examination and were still working on the mine (n=1669), who came for a periodical examination but were no longer working on the mine (n=99), who came for a benefit examination and were still working on the mine (n=312), and who came for a benefit examination, but were no longer working on the mine (n=129). The benefit examination is a more detailed examination for miners seeking compensation. To examine whether the miners who came for a benefit examination did not consist of men particularly susceptible to respiratory diseases, the four groups of men were compared for their dust exposure characteristics. Table 4.3

Table 4.3 Comparison of miners with different reasons for attending annual examination, for cumulative dust exposure and lung function tests

Dust/ Lung function	Reason for attending medical examination			
	Periodical examination		Benefit examination	
	Working in the mine		Working in the mine	
	Yes (n=1669)	no (n=99)	Yes (n=312)	no (n=129)
	mean±s.d.	mean±s.d.	mean±s.d.	mean±s.d
Cumulative dust exposure (RSA x years /1000) to				
1950	9.1±5.1	9.8±4.7	10.9±4.9	10.6±4.7
1960	19.6±8.3	19.9±7.2	23.4±8.2	21.8±7.6
1970	29.1±12.0	25.3±9.1	34.1±11.5	27.0±10.3
Lung function at 1968–71				
FEV1	3.20±0.72	3.15±0.69	2.82±0.73	2.86±0.92
VC	4.47±0.76	4.38±0.75	4.16±0.68	4.31±0.87
FEF_{25-75}	2.84±1.27	2.81±1.21	2.37±1.17	2.37±1.34

shows the cumulative dust exposure up to 1950, 1960 and to the examination in 1968–71. The men who came for a benefit examination had higher dust exposure in comparison with those who came for a periodical examination, and thus their disability was likely to be related to the high dust exposure rather than to higher susceptibility to respiratory diseases. Their lung function tests also reflect their disability. However, the type of medical examination that a men attends does not necessary reflects his disability. A miner may come for a benefit examination one year and for a periodical examination the next year, and a miner who was already compensated for disability will come for a periodical examination. More over, many miners who attend the benefit examination are not disabled and are not granted disability certificate.

Comparison of the miners and non–miners

Characteristics of the miners and non–miners are compared in Table 4.4. The differences that were important were those in the smoking habits. The non–miners had significantly higher cigarette consumption (cigarettes per/day, cigarette pack–years, and cigarette equivalent pack–years) than the miners.

Table 4.4 Characteristics of the miners and non–miners

Characteristic	miners (n=2209)	non–miners (n=483)	P–value
	mean±s.d.	mean±s.d.	
Age at examination	50.2±2.5	49.8±2.8	<0.01
Height (cm)	175.4±6.1	175.9±6.7	<0.01
Weight (kg)	81.4±12.3	77.7±13.3	<0.0001
Smoking habits			
Current cig./day	9.4±10.5	11.7±13.3	<0.001
Maximum cig./day	22.1±14.6	23.9±16.4	0.02
Average cig./day	15.8±10.6	17.8±12.9	0.001
Age at start	17.6±8.0	16.6±8.0	0.001
Years of smoking	21.9±11.3	21.4±11.9	0.40
Cigarette pack–years	20.5±16.2	23.5±21.2	0.002
Pipe eq. pack–years	3.9±6.9	3.6±8.0	0.50
Cig. eq. pack–years	24.4±17.3	27.1±21.2	0.008

Confounding and effect modification by age

Although the men in this study fall within a narrow age range (45–55), the confounding and the modifying effect of age on the association in question was examined. The multiplicative relative risk model (Eq. 3.7) fitted to the data in order to establish the effect of age, was as

$$RR_m = \alpha + \beta_1 \cdot (AGEC) + \beta_2 \cdot (CDEC) + \beta_3 \cdot (SMOKEC)$$
$$+ \beta_4 \cdot (CDEC \cdot SMOKEC) + \beta_5 \cdot (AGEC \cdot CDEC) + \beta_6 \cdot (AGEC \cdot SMOKEC)$$
$$+ \beta_7 \cdot (AGEC \cdot CDEC \cdot SMOKEC)$$

where AGEC=0 for age<50 and unity for age≥50.

The log likelihood ratio statistic was used to test the following hypotheses:

1: effect modification by age is absent

H_0: $\beta_5 = \beta_6 = \beta_7 = 0$ versus H_a: any of the β's $\neq 0$.

2: age is not a confounding factor

H_0: $\beta_1 = 0$ versus H_a: $\beta_1 \neq 0$.

The maximum log likelihood values and the X^2–test of significance estimated from the full and reduced models were as follows:

Effect tested	–2logL	X^2_{LR}	df	p–value
Interactions with age	2921.3	6.3	3	>0.05
AGEC	2927.6	6.9	1	<0.01
AGEC, CDEC, SMOKEC,CDEC·SMOKEC	2934.5	369	3	<0.0001
Intercept	3303.0			

The log likelihood ratios tests of the hypotheses 1 and 2 were:

1: X^2_{LR} = (age interactions | main effects, CDE · SMOKEC)

= 6.3; df=3; p>0.05

2: X^2_{LR} = (AGEC | CDE, SMOKEC, CDE · SMOKEC)

= 6.9; df=1; p<0.01.

Thus the null hypothesis 1) of the absence of effect modification by age cannot be rejected, but the null hypothesis 2), that age is not a confounding factor to the relationship in question, is, however, rejected.

4.3.2 Combined effect of mining dust and smoking on the presence of simple chronic bronchitis

The effect of combined exposure to gold mining dust and tobacco smoking on the

presence of simple chronic bronchitis was evaluated firstly for the presence/absence of mining dust and smoking using the indicator variables DUST and SMOKE, and secondly the effect of increasing dose was evaluated using the categorical variables CDEC, SMOKEC and the continuous variable PACKYR.

To test the null hypothesis of absence of interaction between dust and tobacco smoking, the following family of relative risk models was fitted to the data

$$\text{logit}\ (p) = \alpha_0 + \alpha_1\ (\text{AGEC}) + \text{RR}(\text{dust}, \text{smoking};\beta), \qquad (4.2)$$

where p is the probability of having the condition, α_0 is a nuisance parameter and RR(dust, smoke;β) represents

a) an additive model

$$\text{RR}_a = \log\{1 + \beta_1(\text{dust}) + \beta_2(\text{smoke}) + \beta_3(\text{dust}\cdot\text{smoke})\} \qquad (4.2a)$$

or b) a multiplicative model

$$\text{RR}_m = \beta_1(\text{dust}) + \beta_2(\text{smoke}) + \beta_3(\text{dust}\cdot\text{smoke}) \qquad (4.2b)$$

or c) a power model (Breslow and Storer)

$$\text{RR}_p = [(1 + \beta_1(\text{dust}) + \beta_2(\text{smoke}))^{\lambda} - 1] \,/\, \lambda. \qquad (4.2c)$$

The likelihood ratio test of H_0: $\beta_3 = 0$ against H_a: $\beta_3 \neq 0$ was applied to test the interaction terms in the models a) and b).

Dichotomous exposure variables

Table 4.5 shows the relations for chronic bronchitis, the presence/absence of dust exposure and smoking (yes/no). The crude odds ratio for the joint effect of the two exposures lies between the additive and multiplicative hypotheses, as

$$(10.0 + 5.2 - 1) < 41.9 < (10.0 \times 5.2).$$

The crude odds ratio for chronic bronchitis associated with the mining dust is almost twice as high as that due to smoking.

Table 4.5 Relations for chronic bronchitis, smoking (yes/no) and gold mining dust exposure (non–miner/miner)

SM[a]	DUST	BRONCHITIS NO	YES	OR (95%CI)	(adjusted) OR (95%CI)	(adjusted) Excess odds (95%CI)
No	No	67	1	1.0	1.0	0.0
	Yes	222	33	10.0(1.3;75)	9.6(1.3;72)	0.14 (0.06;0.22)
Yes	No	385	30	5.2(0.7;39)	5.2(0.7;39)	0.09 (0.01;0.16)
	Yes	1202	752	41.9(5.8;302)	41.5(5.7;299)	0.62 (0.53;0.70)
Total		1876	816			

[a] SM=SMOKE.

The parameter estimates and the standard errors from the additive, multiplicative and power models (Eqs. 4.2a, 4.2b, and 4.2c, respectively) are shown in Appendix A, Table A.2. For the additive model, the estimated coefficient for the interaction term (DUST·SMOKE) was significantly greater than zero (β_3=0.395, S.E.=0.049), and the likelihood ratio test of H_0:β_3=0 was highly significant (X^2_{LR}=45.96; df=1; p<0.0001). For the multiplicative model, the estimated interaction coefficient was negative (β_3=–0.187, S.E.=1.043), and the likelihood ratio test of H_0: β_3=0 was not significant (X^2_{LR}=0.03; df=1; p>0.20). As the inclusion of the interaction term in the multiplicative model increased the variances of the coefficients substantially, which is due mainly to the multicollinearity between the main effects and the interaction term, the regression coefficients were not as highly significant as those from the model with the main effects only (Greenland, 1989). Table 4.5 shows the age adjusted odds ratios estimated from the multiplicative model and the excess odds estimated from the additive model. The 95% confidence limits derived from the multiplicative model with the interaction term included (Kleinbaum *et al.*, 1982) (shown in Table 4.5) are in good agreement with those estimated from the normal approximation by the logit formula. The confidence intervals are large as they are determined by the

smallest frequency number. The likelihood ratio tests of H_0: $\beta_1=0$ and of H_0:$\beta_2=0$, for the additive model Eq.(4.2a). were highly significant (X^2_{LR}=117.8;p<0.0001 and X^2_{LR}=60.3;p<0.0001, respectively).

The power model, Eq.(4.2c), reached the lowest maximum likelihood value when λ=0.92, (−2logL=3013.9) indicating that the combined effect lies closer to being multiplicative rather than additive. Without an interaction term, the multiplicative model fitted the data better than the additive model, as the likelihood ratio test of H_0:(λ=0) vs H_a:(λ=1) was significant (X^2_{LR}=(3065.8–3014.0)=51.8; df=1; p<0.0001). However, the negative coefficient for the interaction term estimated by the multiplicative model appears to be an artefact related to the functional form fitted, rather than a result due to the underlying phenomena.

In terms of the observed probabilities of the disease, i.e. prevalences, the size of the interaction effect, estimated by the T index Eq.(3.3), using prevalences instead of risks is

$$T = P_{11} - P_{01} - P_{10} + P_{00} = 0.3849 - 0.0723 - 0.1294 + 0.0147 = 0.1979.$$

Here P_{11} denotes prevalence in the smoking miners, P_{01} in non–smoking miners, P_{10} in smoking non–miners and P_{00}in non–smoking non–miners. This interaction effect is statistically significant (Z=0.1979/0.0306=6.46; p<0.001, see Section 3.3.1.). According to the T index, approximately 20% of the smoking miners appear to suffer with chronic bronchitis which is due to the combined effect of tobacco smoking and underground gold mining exposure.

Dose–response relationship

The relations for chronic bronchitis and the ordinal categories of mining and smoking are shown in Tables 4.6 a) and b), respectively. The ORs adjusted for age and the other exposure were estimated by means of the multiplicative model in

which each category of exposure was included as an indicator variable (zero, 1). It is apparent that the adjusted odds ratios for dust have a higher value than the unadjusted ones. This effect is caused by an association between smoking and mining exposure, present in the data. The non–miners had a higher tobacco consumption than the miners, and there is also a decreasing trend in consumption with increasing levels of cumulative dust exposure among the miners. The average cigarette equivalent pack–years according to the categories of dust (CDEC) were 27.1(SD, 21.0), 25.2(SD, 18.2), 25.1(SD, 16.7), 23.2(SD, 17.2), 23.6(SD, 17.0).

The trend in the adjusted ORs and the highly significant chi–square tests for trend support the hypotheses that there is a significant increase in the prevalence of chronic bronchitis associated with the increasing level of dust exposure, and increasing level of smoking. The best functional form of the relative risk function was established by fitting the additive (RR_a), multiplicative (RR_m) and Breslow and Storer power model (RR_p) with each individual exposure factor. The functional form fitted to the smoking variable (SMOKEC) represents, however, the combined effect of smoking status and the pack–years in the current smokers. The variable SMOKEC was primarily used to explore the combined effect of dust and smoking taking the ex–smoker effect into account. The effect of age was controlled by means of a categorical variable (AGEC). However, as dust and smoking are correlated, inclusion of both variables may change the functional form or the estimated regression coefficients. The results from the models are presented in Table A.3 a) for dust and Table A.3 b) for smoking.

The relationship between chronic bronchitis and CDE was best fitted by the power model with λ=–0.24. The additive model fitted the data better than the multiplicative model (X^2_{LR}= 52.4; df=1; p<0.0001). The dose–response relationship between chronic bronchitis and the smoking variable SMOKEC was best fitted by the power model with λ=0.48 and the multiplicative model fitted

the data better than the additive model (X^2_{LR}=10.9:df=1;p<0.001).

The prevalence in the ex–smoker category was not significantly different from that in the nonsmoker category (see Table 4.6 b). The adjusted odds ratio for the ex–smoker category was 1.27 (0.8–2.0), which is not significantly different from unity, whereas the three current smoker categories had odds ratios significantly greater than unity. An alternative smoking variable was created by combining the never and ex–smokers into a category of current non–smokers and this category was then used as a baseline for estimating the odds ratio associated with the effect of cigarette equivalent pack–years in the current smokers. The power model with PACKYR fitted best when λ=0.31 indicating that the combined effect lies between the additive and multiplicative hypothesis. However, the additive model gave better fit to the data than the multiplicative model (X^2_{LR}=52.0; df=1; p<0.001). The linear approximation thus appears to be more appropriate for the smoking variable PACKYR.

Table 4.6 Dose–response relations for chronic bronchitis

a) and gold mining exposure

Bron–chi–tis	Cumulative dust exposure (RSA x years /1 000)					TOTAL
	0	15	25	35	48	
YES	31	137	265	206	177	816
NO	452	403	432	330	259	1876
TOTAL	483	540	697	536	438	2692
unadjusted OR	1.00	4.96	8.94	9.41	9.96	
adjusted for age and smoking OR[a]	1.00	5.66	9.61	10.16	11.10	
95%CI		(3.7–8.6)	(6.4–14)	(6.7–15)	(7.3–17)	

Test for trend: X^2=152.1; p<0.0001

b) and smoking habit

Bron–chi–tis	Smoking category (SMOKEC)					TOTAL
	never	ex–smk.	current smoker (pack–years)			
			<15	15–59	60+	
YES	34	70	82	586	44	816
NO	289	457	164	904	62	1876
TOTAL	323	527	246	1490	106	2692
unadjusted OR	1.00	1.30	4.25	5.51	6.03	
adjusted for age and mining exposure OR[a]	1.00	1.27	4.27	5.86	7.33	
95%CI		(0.8–2.0)	(2.7–6.7)	(4.0–8.6)	(4.2–12.6)	

Test for trend: X^2=184.0; p<0.0001

[a] Odds ratios were estimated by the maximum likelihood method.

Joint effect of dust and smoking at several levels of exposure

Table 4.7 a) shows a two–way frequency distribution of the response to the chronic bronchitis questionnaire according to dust and smoking categories. The effect of each combination of exposure levels was compared to the baseline group, who had no exposure, by means of the odds ratio statistic (see Table 4.7 b). The increasing trend in the size of the OR with the increasing level of dust is apparent for all levels of smoking. However, the trend is stronger for the current smoker than for the never or ex–smoker categories. Figure 4.1 shows the observed odds ratios plotted against dust for three smoking categories. The odds ratios for the current smokers with cumulative dust exposure above 20 000 indicate a higher increase than that expected from the sum of the individual effects, for example 69.2>1+16.8+7.2, suggesting the presence of synergism.

For the multifactorial analysis the relative risk models, Eq 4.2, were used. The smoking habits were analyzed using the variables SMOKEC and PACKYR, and the results are presented in Table A.4 a) for SMOKEC and Table A.4 b) for PACKYR. The log likelihood ratio tests of H_0:β_3=0 versus H_a:β_3≠0 for the additive model (Eq. 4.2a) were found to be statistically significant for both smoking variables SMOKEC and PACKYR,

(X^2_{LR}=(CDE·SMOKEC|main effects)=2912.2–2881.4=30.8;p<0.0001) and

(X^2_{LR}=(CDE·PACKYR|main effects)=2959.2–2876.0=63.2;p<0.0001).

The coefficients for the interaction terms in the multiplicative model were not

Table 4.7 Frequencies and odds ratios for chronic bronchitis, by smoking and cumulative dust exposure

a) frequencies

Mining (CDEC)	Bron–chitis	SMOKING HABITS (SMOKEC) never	ex–smk.	current smk.(pack–yrs) <15	15–59	60+	Total
0	yes	1	0	1	24	5	31
	no	67	101	33	222	29	452
15	yes	2	8	9	104	14	137
	no	68	112	32	182	9	403
25	yes	11	26	25	196	7	265
	no	60	93	42	226	11	432
35	yes	11	20	27	138	10	206
	no	58	71	38	154	9	330
48	yes	9	16	20	124	8	177
	no	36	80	19	120	4	259
Total	yes	34	70	82	586	44	816
	no	289	457	164	904	62	1876

b) the odds ratios

Mining (CDEC)	SMOKING HABITS (SMOKEC) never	ex–smk.	current smk.(pack–yrs) <15	15–59	60+	marginal adjusted
0	1.00	–	2.03	7.24	11.55	1.0
15	1.97	4.79	18.84	38.29	104.22	5.66
25	12.28	18.73	39.88	58.11	42.63	9.61
35	12.71	18.87	47.61	60.04	74.44	10.16
48	16.75	13.40	70.53	69.23	134.00	11.10
marginal adjusted	1.00	1.27	4.27	5.86	7.33	

CDEC= RSA x years / 1 000.

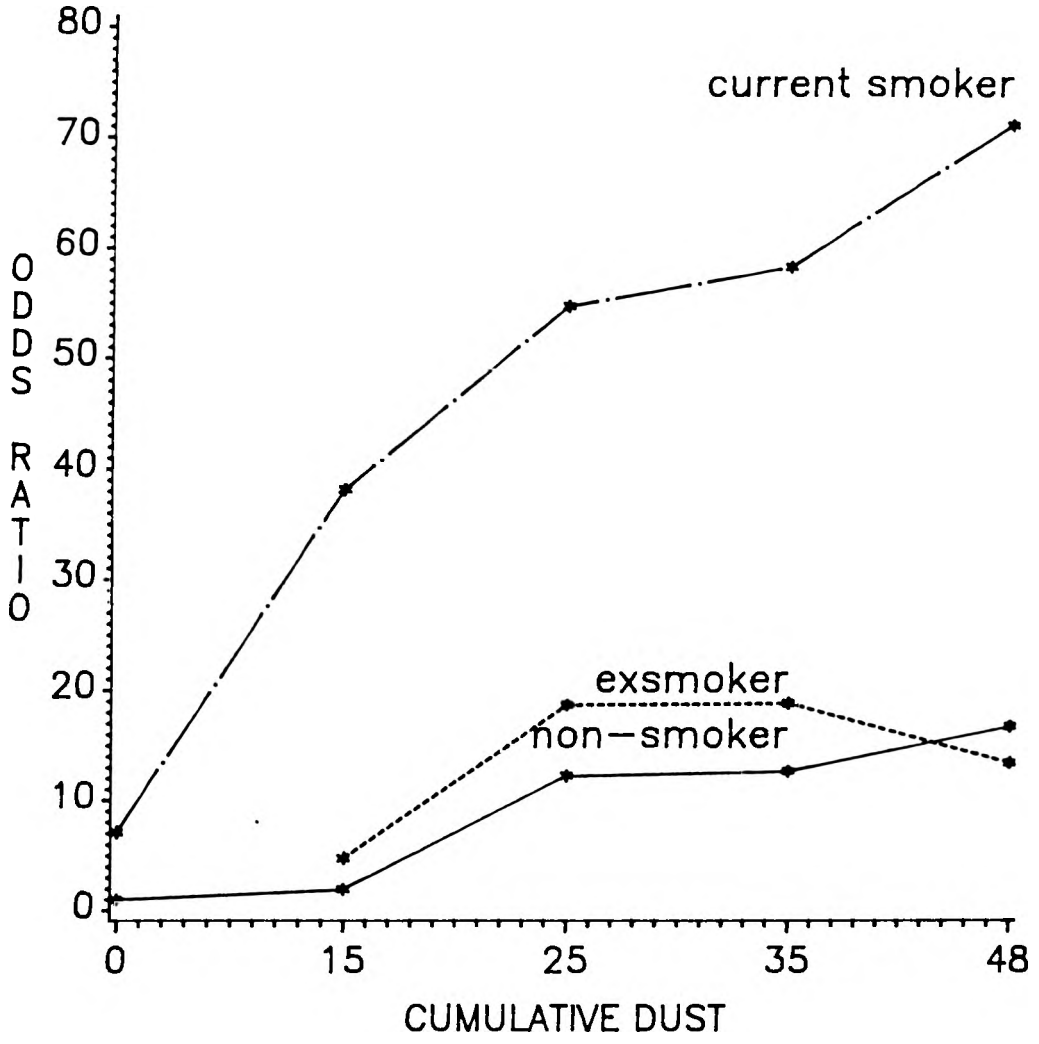


Fig 4.1 Observed odds ratios for chronic bronchitis and cumulative dust exposure, by smoking habits

statistically different from zero. The power model fitted best when λ=0.24 for SMOKEC and λ=0.25 for PACKYR, suggesting higher than additive effect.

As the best fitting individual dose–response curves were of a different functional form, the modified Thomas' model, RR_{MT}, was fitted to the data with variables CDE and SMOKEC. The suggestion of Lubin and Gaffey (1988), that the individual risk factors be fitted by their best fitting RR model in the RR_{MT} model, was followed. The RR_{MT} model with RR(CDE)=1+ β_1(CDE) and RR(SMOKEC)=exp(β_2·SMOKEC) was fitted to the data, and the results are shown in Table A.5 under model 4. When the risk function for SMOKEC was changed to the linear form RR(SMOKEC)=1+β_2(SMOKEC) and λ was restrained between .001 and 1.0, the best fit was found at λ=1.0 (see models 2 and 3). When λ was not restrained, the best fit was found at λ=6.92. The coefficient λ did not indicate correctly the scale on which interaction was present, when RR(SMOKEC) was entered as a linear function, but it did so when RR(SMOKEC) was entered as an exponential function (λ=0.40).

4.3.3 Combined effect of mining dust and smoking on abnormal sputum production

Every study subject was classified by the amount of mucus expectorated in 5 minutes as: abnormal when 2ml or more of mucus was produced, and normal otherwise (Wiles and Faure, 1977). The relationship between this outcome variable and dust and smoking was analyzed in the same fashion as for the response to the chronic bronchitis questionnaire, Section 4.3.2.

Dichotomous exposures

Table 4.8 shows the relations for abnormal sputum production, the presence/

absence of mining dust and smoking (yes/no) smoking. The crude odds ratio for the joint effect of the two risk factors lies between the additive and multiplicative models, as

$$(5.6 + 8.8 - 1) < 17.4 < (5.6 \times 8.8).$$

The crude and the age–adjusted ORs for dust were lower than those for smoking. The 95% confidence intervals excluded unity and were narrower than those for the questionnaire response, this is because the variance of each of the estimates depends on the smallest cell frequency, which is larger for sputum than for bronchitis.

The parameter estimates obtained from fitting the additive, multiplicative, and power models, Eq. (4.2), are shown in Table A.6. The likelihood ratio test of H_0: $\beta_3=0$ for the additive model was statistically significant (X^2_{LR}=3468.1–3461.1 =7.0; df=1; $p<0.01$), and the estimated coefficient was significantly greater than zero (β_3=0.213, SE=0.074; $X^2_{LR}(1)$=8.29; $p<0.01$). For the multiplicative model, the coefficient β_3 was negative, and not significantly different from zero. The log likelihood ratio test of H: ($\lambda=0$) vs ($\lambda\neq0$) was slightly significant (X^2_{LR}=(3468.1–3462.0) =5.2; $p<0.025$).

The results from this analysis support the hypothesis that the combined effect of dust and smoking on the presence of abnormal sputum production is more than additive. However, in terms of the observed prevalences of the abnormality, the size of the interaction effect measured by the index T is T=0.4451–0.2892–0.2039+ 0.0441 =–0.0039. This result indicates that the multifactorial analysis done on the odds of the abnormality may have overestimated the joint effect of the exposures. This is because the rare disease assumption was not satisfied.

Table 4.8 Relations for abnormal sputum production, smoking (yes/no) and gold mining dust exposure (non–miner/miner)

SM[a]	DUST	SPUTUM NO	SPUTUM YES	OR (95%CI)	(adjusted) OR (95%CI)	(adjusted) Excess odds (95%CI)
No	No	65	3	1.0	1.0	0.0
	Yes	203	52	5.6(1.7;19)	5.5(1.7;18)	0.17(0.07;0.27)
Yes	No	295	120	8.8(2.7;29)	8.9(2.7;29)	0.31(0.24;0.42)
	Yes	1076	863	17.4(5.5;56)	17.2(4.7;62)	0.71(0.61;0.81)
Total*		1639	1038			

[a] SM=SMOKE. *15 miners had missing information on sputum production.

Dose–response relationship

The relations for chronic bronchitis and the ordinal categories of dust and smoking are shown in Table 4.9 a) and b), respectively. The ORs adjusted for age and the other exposure factor, and the chi–square test for trend, support the hypothesis of a significant, increasing dose–response relation for both risk factors. Smoking appears to be a greater risk factor than mining. For dust the trend is not as strong as it was for the questionnaire response, while for smoking the trend is stronger and more consistent. The ex–smokers have a significantly elevated odds ratio (OR=1.58, 1.1–2.3), and there is a systematic increasing trend in the current smokers (3.0, 4.4 and 6.0). The best functional form of the relative risk function was established by fitting the power model, RR_p, to each exposure. The best fitted model for CDE was at λ=–0.78, which indicates a convex shape of the dose–response curves, and for SMOKEC the best fit was obtained at λ=0.53, which indicates a concave shape (see Table A.7 for the coefficients).

Table 4.9 Dose–response relations for abnormal sputum production
a) and gold mining dust exposure

SPUTUM	Cumulative dust exposure (RSA x years / 1 000)					TOTAL
	0	15	25	35	48	
YES	123	188	303	232	192	1038
NO	360	349	391	299	240	1639
TOTAL	483	537	694	531	432	2677
unadjusted OR	1.00	1.58	2.27	2.27	2.34	

Test for trend: X^2 =44.7; df=1; p<0.001

adjusted for age and smoking					
OR[a]	1.00	1.66	2.26	2.35	2.43
95%CI		(1.2–2.2)	(1.7–2.9)	(1.8–3.1)	(1.8–3.2)

b) and smoking habits

SPUTUM	Smoking category (SMOKEC)					TOTAL
	never	ex–smk.	cigarette eq. pack–yrs			
			<15	15–59	60+	
YES	55	129	93	705	56	1038
NO	268	395	148	779	49	1639
TOTAL	323	524	241	1484	105	2677
unadjusted OR	1.00	1.59	3.06	4.41	5.57	

Test for trend: X^2=184.0; df=1; p<0.0001

adjusted for age and mining exposure					
OR[a]	1.00	1.58	2.99	4.41	6.01
95%CI		(1.1–2.3)	(2.0–4.4)	(3.2–6.0)	(3.7–9.8)

[a] Odds ratios were estimated by the maximum likelihood method.

Joint effect of dust and smoking at several levels of exposure

Table 4.10 a) shows a two–way frequency distribution of abnormal sputum production according to dust and smoking categories. The odds ratios are presented in Table 4.10 b). The increasing trend in the size of the OR with the increasing level of dust is apparent for all the levels of smoking. Figure 4.2 shows the observed odds ratios plotted against dust, according to three smoking categories. The pattern in the odds ratios indicates some degree of synergism, but not as much as it did for the simple chronic bronchitis.

The confounding and effect modification by age was not significant at the 0.05 probability level. However, since the coefficient for AGEC was slightly significant ($0.05<p<0.10$), AGEC was included in the model.

The results obtained from fitting the models given in Eq. (4.2) are shown in Table A.8. The likelihood ratio test of $H:\beta_3=0$ for the additive model, Eq. (4.2a), (X^2_{LR}(CDE·SMOKEC|main effects)=9.2; df=1; $p<0.01$), suggests that a significant deviation from additivity is present. For the power model, Eq. (4.2c), the best fit was obtained at $\lambda=0.70$, which also suggests that the combined effect is synergistic. Surprisingly the λ coefficient is higher than for the questionnaire response.

To get an indication of whether the additive model overestimated the interaction effect, the size of the interaction effect was calculated using the T index with the observed prevalences. For the current smoker of 15–59 pack–years and cumulative dust exposure of 35 000, the T value was found to be

$$T=0.5257 - 0.2174 - 0.3374 + 0.0441 = 0.015,$$

which indicates a small interaction effect.

Table 4.10 Frequencies and odds ratios for abnormal sputum production, by smoking and cumulative dust exposure

a) frequencies

Mining	Abnormal Sputum	Smoking habits (SMOKEC)					Total
		never	ex–smk.	current smoker(p–yrs)			
(CDEC)				<15	15–59	60+	
0	yes	3	15	5	83	17	123
	no	65	86	29	163	17	360
15	yes	9	26	20	123	10	188
	no	61	94	21	161	12	349
25	yes	15	33	26	219	10	303
	no	56	85	41	201	8	391
35	yes	15	25	28	153	11	232
	no	54	65	34	138	8	299
48	yes	13	30	14	127	8	192
	no	32	65	23	116	4	240
Total	yes	55	129	93	705	56	1038
	no	268	395	148	779	49	1639

b) the odds ratios

Mining	Smoking habits (SMOKEC)					
	never	ex–smk.	current smk.(pack–yrs)			marginal
(CDEC)			<15	15–59	60+	adjusted
0	1.0	3.8	3.7	11.0	21.7	1.0
15	3.2	6.0	20.6	16.6	18.1	1.7
25	5.8	8.4	13.7	23.6	27.1	2.3
35	6.0	8.3	17.8	24.0	29.8	2.4
48	8.8	10.0	13.2	23.7	43.2	2.6
marginal adjusted	1.0	1.58	2.99	4.41	6.01	

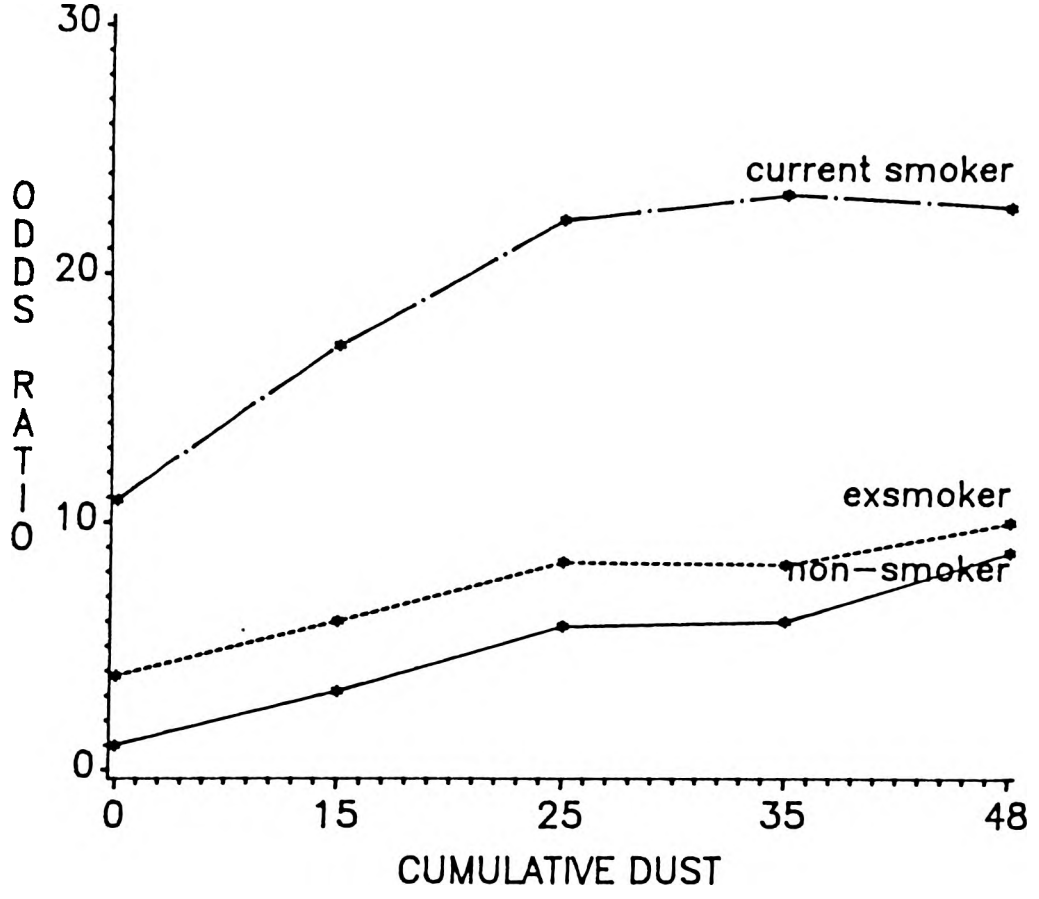


Fig 4.2 Observed odds ratios for abnormal sputum production and cumulative dust exposure, by smoking habits

4.4 Discussion

The results from the analysis of the two sets of data included in this chapter [Sluis–Cremer study (1967) and Wiles and Faure study (1977)] indicate that both of the risk factors, mining dust and tobacco smoking, increase the prevalence of simple chronic bronchitis as measured by a respiratory questionnaire. The age adjusted odds ratios obtained from the Carletonville study, for the comparison of non–smoking miners, smoking non–miners and smoking miners, to nonsmokers without mining exposure, were estimated by the multiplicative model as 2.8 (1.4–5.5), 5.0 (2.0–9.2) and 11.8 (1.8–74), respectively. The equivalent age adjusted odds ratios estimated from the Wiles and Faure study, were 9.6 (1.3–72), 5.2 (0.7–39) and 41.5 (5.7–299), respectively. For both data sets the interaction coefficient in the additive model was statistically significant, and the multiplicative model fitted better to the data than the additive model. Thus, the combined effect of the presence of dust and smoking on the prevelance of simple chronic bronchitis evaluated by a questionnaire was found to be synergistic.

For abnormal sputum production the equivalent age adjusted odds ratios, estimated from the Wiles and Faure study, were 5.5 (1.7–18), 8.9 (2.7–29) and 17.2 (4.7–62), which indicates a very small synergistic effect. The goodness–of–fit tests also confirmed a significant departure from the additive model. The test of synergism done with prevalences, using the T index, did not, however, show a significant interaction effect. This discrepancy may be due to the fact that the odds ratios for the combined exposures do not approximate the prevalence ratio.

The results from the analysis of the Wiles and Faure study, using the categorized mining and smoking exposure variables, have shown a consistent, increasing trend in the age adjusted odds ratios for simple chronic bronchitis with

increasing levels of cumulative dust exposure (1.0, 5.7, 9.6, 10.2, 11.1), but the equivalent odds ratios for abnormal sputum production were much lower (1.0, 1.7, 2.3, 2.4, 2.4). The odds ratios for simple chronic bronchitis and increasing levels of smoking (1.0, 1.3, 4.3, 5.9, 7.3) were in better agreement with those for abnormal sputum production and smoking (1.0, 1.6, 3.0, 4.4 and 6.0). In Table 4.7 b) and Figure 4.1 the odds ratios for simple chronic bronchitis indicate a departure from the additive hypothesis. This observation was confirmed by the multifactorial analysis which found a significant interaction term in the additive relative risk model. The results obtained from the analysis of abnormal sputum, Table 4.10 and Figure 4.2, display a similar picture, but the interaction effect is not as strong.

However, as the two conditions, simple chronic bronchitis and abnormal sputum production, are highly prevalent, especially in the higher exposure categories where the numerator for the exposure odds becomes greater than the denominator, the odds ratios for the higher exposure categories overestimate the actual prevalence ratios and make the effect appear stronger than it is. This applies to the individual effects as well as to the combined effect. Thus the general observation from the analysis of these data is that the additive relative risk model may overestimate the interaction effect for a highly prevalent condition such as chronic bronchitis, and may lead to biased conclusions. On the other hand, the additive model detects a synergistic effect which is quite strongly apparent, whereas the multiplicative model does not, as was the case with the questionnaire response.

CHAPTER 5
LUNG FUNCTION IMPAIRMENT

Our problem in terms of people who are sick or disabled and are no longer able to earn as much as they used to, is definitely in terms of obstructive airways disease and not in terms of silicosis; but what are the causes of this pulmonary disability in order that preventive measures can be applied? The available evidence we have is that obstructive airways disease occurs twice as often in gold miners as in a control population ...; but we have very little evidence at all that the incidence or the prevalence of this condition correlates with dust exposure. Dr G.K. Sluis–Cremer (Beadle *et al*, 1970).

5.1 Objectives of chapter 5

The objective of this chapter is to examine whether tobacco smoking modifies the effect of underground gold mining dust exposure on the prevalence of chronic respiratory impairment. The data from the study of 2209 white South African gold miners and of 483 non–miners, done by Wiles and Faure (1977), were reanalyzed in order to:

i) examine the individual lung function measurements, FEV_1, VC and $FEF_{25-75\%}$, for the dose–response trend with dust and for the presence of interaction between dust and smoking;

ii) examine the patterns and degree of ventilatory impairment as reflected in the spirogram, their relations to cumulative dust exposure and tobacco smoking and

the presence of interaction;

iii) estimate the attributable risk for tobacco smoking and gold mining dust exposure;

iv) examine how relevant is the pattern and degree of ventilatory impairment to the mortality from respiratory and cardiovascular diseases.

5.2 Study subjects and methods

Study subjects

The 2209 white gold miners and the 483 non–miners used in the analysis of simple chronic bronchitis were used for this analysis. Details of the original study design are given in Section 4.3.1 and only additional information relevant to the present chapter is given.

Mining exposure

The mining dust exposure was analyzed in terms of cumulative dust exposure (RSA x years / 1 000), cumulated to the date of the medical examination in 1968–71 (CDE), and as categorized cumulative dust exposure (RSA x years / 10 000) (0, 1–1.5, 1.6–2.5, 2.6–3.5, 3.6–4.5, 4.6+) (CDEC), where the means of the ranges were used to estimate the dose–response relationship (Berry, 1980). In a preliminary analysis, cumulative dust exposure was found to be a much stronger predictor of lung function tests than years of exposure.

Smoking habits

The information on smoking habits was utilized in the analysis in three ways:

1) The subjects were grouped into non–smokers, ex–smokers and current smokers and indicator (zero, one) variables SMK1 and SMK2 were created to represent the

grouping.

2) The information on tobacco consumption, in terms of pack–years, and the smoking status were combined. The subjects were divided into 5 categories as: 0=never–smoker, 2=ex–smoker, 3=current smoker <20 pack–years, 4=current smoker 20–39 pack–years and 5=current smoker ≥40 pack–years (SMOKEC).

3) Lifetime tobacco consumption was measured in terms of cigarette equivalent pack–years (CEPY).

Lung function tests

Three spirograms were measured on a Godart water sealed spirometer with the subjects seated and wearing a noseclip. Spirograms judged to be of poor technical quality were excluded from the original study by a lung physiologist and were not recorded on the computer file. The lung function parameters measured were: Vital Capacity (VC), Forced Expiratory Volume in one second (FEV_1), and Forced mid–Expiratory Flow (FEF_{25-75}).

For the analysis of the pattern of ventilatory impairments, the miners were grouped according to the presence of an obstructive or restrictive lung function profile, and those with the obstructive profile were further sub–divided according to the degree of obstruction. A method of categorization used by Harber *et al.* (1987) for asbestos–exposed workers was adapted by incorporating features that are used by the Pulmonary Unit of MBOD when categorizing the types of lung dysfunction. The observed lung function tests were considered low if they fell below a 95% or 99% confidence limit (CL). The confidence limits were calculated as (predicted value – SEE · z), where the predicted value was calculated from the equations to be given below, and SEE is the standard error of the estimate and z =1.645 or 2.324 for the 95% and 99% CL, respectively. The lung function tests were considered normal if they were above the lower 95% confidence limit.

The prediction equations were derived by the Lung Function Unit of the MBOD, and are routinely used for the assessment of lung function impairment in white miners who come to the MBOD for their medical examination. The equations were derived from 731 healthy miners consisting of 34% smokers, 41% ex–smokers and 25% non–smokers, who came to the MBOD during 1984. They were on average 51 years of age, 175 cm tall and weighed 81 kg. The lower 95% CLs derived from Crapo's (Crapo *et al.*, 1981) and MBOD's equations for a man 51 years of age, 174 cm tall and weighing 81 kg are: 3.64 and 3.76 for VC, 2.97 and 2.97 for FEV_1, 2.08 and 2.32 for FEF_{25-75}, and 71.85 and 70.51 for FEV_1/VC, respectively. The main difference in the two sets of prediction curves is the weight factor included in the MBOD prediction equations. Weight was a significant negative predictor of lung function in white South African miners, who are often overweight. The mid–expiratory flow time (MEFT) was calculated as $VC/(2 \cdot FEF_{25-75})$.

The prediction equations and standard errors of the estimate (SEE) were as follows:

$VC = -5.57964 - 0.02195(AGE) + 0.06822(HEIGHT) - 0.00777(WEIGHT)$

$SEE = 0.47227$

$FEV_1 = -3.00483 - 0.02491(AGE) + 0.04759(HEIGHT) - 0.00510(WEIGHT)$

$SEE = 0.37765$

$FEV_1/VC = 113.4064 - 0.23250(AGE) - 0.1470(HEIGHT) + 0.02649(WEIGHT)$

$SEE = 4.6250$

The subjects were grouped into six lung function profile categories as follows:

1.**Normal** (Reference category). Those with normal lung function tests, i.e. normal VC, FEV_1, FEV_1/VC and MEFT<=1 second.

2. **Minimal Obstructive**. Those who had normal VC and one or two features of

obstruction, but not all three present. The three features being low FEV_1, low FEV_1/VC and MEFT>1 second.

3. **Moderate Pure Obstructive.** Those with normal VC, FEV_1 between 95 and 99% lower confidence interval, low FEV_1/VC and MEFT>1 second.

4. **Marked Pure Obstructive.** Those with normal VC, FEV_1 below 99% confidence interval, low FEV_1/VC and MEFT>1 second.

5. **Marked Obstructive with Restriction.** As those with marked pure obstruction except that VC is below 99% CL. All the subjects with an obstructive and a restrictive element were found to have the FEV_1 value below 99% confidence limit.

5. **Pure Restrictive.** Those with low VC, normal FEV_1 and FEV_1/VC ratio, and MEFT≤1 second. Four men were found to have restrictive impairment with a minimal obstructive element and these were combined with the pure restrictive group.

There were 202 men whose lung function tests were excluded from the original study by a pulmonary physiologist, as they were considered to be sub–maximal efforts, and the lung function tests were not recorded on the data file. In addition to these, lung function tests of 21 men were rejected from this analysis on advice of the pulmonary physician, as they were found to have very unusual parameters and did not belong to any of the above categories.

The miners whose lung function tests were rejected from the original study (n=190, 12 of the rejected tests were from non–miners) were compared for exposure characteristics and the causes of death with those that were not rejected. Significant differences between the miners with rejected lung function tests and the rest of the miners were found in cumulative dust exposure, cumulated up to 1950 (10.3 compared to 9.4 ($p=0.03$)), in cumulative dust exposure up to year 1969, (32.0 compared to 29.3 ($p=0.001$)), in mortality risk from all causes (55% compared to 42% ($p<0.0001$)), and in mortality risk from COLD (8% compared to

3% ($p<0.0001$)).

Other medical and compensation data

The medical examination included assessment of simple chronic bronchitis by a questionnaire, rhonchi, sputum production, and radiological silicosis assessed for compensation purposes. The degree of compensation was also coded from the MBOD medical files. The MBOD medical files record dates and type of compensation awarded to the miner. Compensation for silicosis or COLD can be awarded if radiological silicosis or specified lung dysfunction is found, and the service on mines was at least 10 years in high dust or 20 years in low dust occupations.

Mortality data

The follow–up period for the collection of mortality data was from the date of the 1968–71 examination to 30 December 1986, by which time 954 of the 2209 miners died. Details of the mortality study are presented in Chapter 6. The causes of death were coded according to the 9th revision of the WHO International Classification of Diseases (1977) by an independent physician, using best available information derived from the death certificate, autopsy report and MBOD medical file. The causes of death of interest to this analysis are COLD (ICD 416, 491, 492 and 496) and ischaemic heart disease and IHD (ICD 410–414).

Statistical methods

Analysis of individual lung function measurements

Firstly, the relationship between the individual lung function measurements (FEV_1, VC and FEF_{25-75}) and dust and smoking was evaluated by means of a linear regression analysis in which the effect of age, body weight and height were

adjusted. Secondly, the individual lung function measurements were dichotomized into normal and abnormal if they fall above or below the 95% CL (or 99% CL) of the predicted value, respectively, and the general relative risk model was applied to assess the presence of interaction.

Analysis of lung function profiles

For the analysis of the lung function profiles, the cases with each of the five impaired profiles were compared to the reference category for differences in exposure prevalences, by means of the exposure odds ratio. The following statistical tests were applied:

a) Test for interaction. The null hypothesis of absence of an interaction effect was postulated in terms of the additivity of relative risks. To test the null hypothesis several types of exposure variables were employed in the following tests:

i) The dose–response trends in the odds ratios for each impairment and the individual dust levels were compared between non–smokers, ex–smokers and current smokers. The indicator smoking variables, SMK1 and SMK2, were employed to test whether the ex–smokers and current smokers differed from the non–smokers in their dose–response trend for the dust exposure (CDEC). The fully saturated additive RR model with those variables was of the form

$$\log[p/(1-p)] = \alpha + \log\{1+\beta_1\cdot(\mathrm{CDEC})+\beta_2\cdot(\mathrm{SMK1})+\beta_3\cdot(\mathrm{SMK2})+ \beta_4\cdot(\mathrm{SMK1}\cdot\mathrm{CDEC})+\beta_5\cdot(\mathrm{SMK2}\cdot\mathrm{CDEC})\}, \qquad (5.1)$$

where p is the probability of having a particular impairment profile, α is the nuisance parameter and $\{\cdot\}$ estimates the odds ratio for the exposed subjects relative to those whose exposure is zero. The β coefficients estimate the excess relative risk due to each effect. The coefficients β_2 and β_3 measure the change in the intercept value caused by the ex–smoker and current–smoker effect, and β_4 and β_5 measure the rate of change in the odds ratios with increasing dust level

for ex–smokers and current smokers, respectively. The coefficients β_4 and β_5 indicate whether tobacco smoking modifies the effect of mining dust exposure. The likelihood ratio test of the absence of interaction, i.e. a test of H_0: $\beta_4 = \beta_5 = 0$ against H_a:β_4 and/or $\beta_5 \neq 0$, is given by $X^2_{LR} = -2\log(L_1/L_2)$, where $\log L_1$ and $\log L_2$ are the maximized log likelihoods estimated from the model without interaction and with interaction, respectively. The multiplicative model with the same variables was fitted to the data for comparison.

b) A test of interaction using either both categorized or both continuous exposure variables was also performed. The additive RR model was applied as

$$RR(x_1,x_2; \beta) = 1 + \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \beta_3 \cdot x_1 \cdot x_2, \qquad (5.2)$$

where x_1 represents the dust variable (CDEC or CDE) and x_2 the smoking variable (SMOKEC or CEPY). The likelihood ratio statistic for testing the null hypothesis of no interaction, i.e. $H_0 : \beta_3 = 0$ against $H_a : \beta_3 \neq 0$, is given by $X^2_{LR} = -2\log(L_1/L_2)$ as above.

The computer program RISK (Thomas, 1980) was used to fit the relative risk Eq. (5.2) and the computer program MINUIT (James and Roos, 1975) was used to fit the Eq. (5.1). The program MINUIT is a general minimization program which allows much greater flexibility than the RISK package, e.g. fitting a grouped likelihood function. The maximum likelihood function for grouped data was used to fit Eq. (5.1) and where zero values occurred a value of 0.5 was added to each cell (McCullagh and Nelder, 1983). The results obtained using the two programs were compared and were found to be identical for the continuous and categorized variables.

b) Estimation of attributable risk

The etiological fraction for dust, EF_d, was estimated as

$$EF_d = \sum_{i=1}^{k} p_{ci} \cdot [(OR_i - 1)/OR_i]$$

where p_{ci} is the estimated proportion of cases in the ith exposure category (i=1,2,...,k) and $FE_i>0$, and OR_i are the odds ratios estimated by logistic regression in which the effect of age and smoking were adjusted. The etiological fractions for smoking, EF_s, and for the total effect of dust and smoking, $EF_{d\cdot s}$, were estimated in the same manner. The 95% CLs for the summary etiological fractions were estimated as in Denman and Schlesselman (1983). Using the EFs, the proportion of cases that could be prevented by elimination of each of the risk factors (PF) were calculated as: $PF_{dust}=EF_{d\cdot s}-EF_s$. $PF_{smoking}=EF_{d\cdot s}-EF_d$, $PF_{dust\cdot smoking}=EF_{d\cdot s}-EF_d-EF_s$, $PF_{other}=1.00-EF_{d\cdot s}$.

5.3 Results

5.3.1 Analysis of individual lung function measurements

Figure 5.1 shows the mean values for the observed FEV_1, VC, FEF_{25-75} and FEV_1/VC measurements, plotted according to categorized cumulative dust exposure for nonsmoker, ex–smoker, current smoker ≤30 pack–years, and current smoker >30 pack–years. The graphs display an interesting pattern. For FEF_{25-75} the values for the non–miners (DUST=0) are lower than those for the miners in the lowest exposure category (DUST=1.2). For all four variables there is a declining slope for the miners up to the dust level of 40 thousand. If one considers the data up to that point, then it can be observed that the decrease with dust is approximately parallel for all the smoking categories and that there in no indication that smoking modifies the effect of dust. The consistent increase in the lung function values in the highest dust exposure category is presumably due to some systematic effect, possibly the healthy worker effect. The rate of decrease associated with increasing dust level in the region up to 40 thousand is highest for

FEF_{25-75} and lowest for VC, but the changes in FEV_1 and VC are almost parallel.

The results from the linear regression done on the individual lung function tests have confirmed the above observation, namely that the interaction effect between dust and smoking is not statistically significant. For the dichotomized FEV_1 and VC variables (normal versus lower than 95% or 99% CL) the interaction term was also not statistically significant when tested by the additive model.

Table 5.1 shows the estimated regression coefficients from the "best' fitting, parsimonious model in which FEV_1, VC and FEF_{25-75} were each related to a set of predictor variables. The statistically significant predictor variables were age at 1968–71 examination, height (cm), cumulative dust exposure, ex–smoker (yes=1/no=0), current smoker (yes=1/no=0) and cigarette equivalent pack–years.

Table 5.1 Estimated linear regression coefficients for prediction of FEV_1, VC and FEF_{55-75}

Variable	FEV_1 $\beta \pm$ S.E.	VC $\beta \pm$ S.E.	FEF_{25-75} $\beta \pm$ S.E
Intercept	–1.296 ±0.538	–3.311 ±0.534	0.841±0.962
Age	–0.0339±0.0060	–0.0327±0.0059	–0.0386±0.0107
Height	0.0384±0.0024	0.0558±0.0012	0.0283±0.0044
CDE	–0.0062±0.0012	–0.0048±0.0012	–0.0109±0.0022
Ex–smoker	–0.1664±0.0569	–0.0920±0.0565	–0.3190±0.1018
Current–smoker	–0.4038±0.0580	–0.3040±0.0576	–0.7088±0.1037
Cig. Eq. p–yrs	–0.0048±0.0011	–0.0022±0.0011	–0.0100±0.0019

CDE= RSA x years / 1000.

The regression coefficients presented in Table 5.1 were applied to estimate the loss of lung function attributable to silica dust exposure and to tobacco smoking. The results are shown in Table 5.2.

Table 5.2 Estimated loss of lung function attributable to silica dust and to tobacco smoking, according to percentiles of each exposure variable

Percentile	Median value CDE/CEPY	Estimated loss of lung function		
		FEV_1 (ml)	VC (ml)	FEF (ml)
a) loss attributable to silica dust				
25	20	123	97	220
50	28	175	136	309
75	37	231	180	408
90	46	281	218	496
95	52	319	250	567
b) loss attributable to being a current smoker and to cigarette eq. pack–yrs				
25	20	500	348	909
50	30	548	370	1009
75	39	591	390	1099
90	50	644	414	1209
95	59	687	434	1299

CDE=RSA x years / 1000. CEPY= cigarette equivalent pack–years.

The above analysis has estimated that by 50 years of age the loss of FEV_1 attributable to dust ranges from about 123 ml to 320 ml, whereas the loss attributable to smoking can be as much as 687 ml. Even if we assume that the effects of dust and smoking are additive, the loss of lung function attributable to the two exposures can be quite substantial. The estimated effect of heavy dust exposure (95th percentile) on loss of lung function (FEV_1 and FEF_{25-75}) is approximately half of that of heavy tobacco smoking (95th percentile). The effect on VC is slightly lower than half.

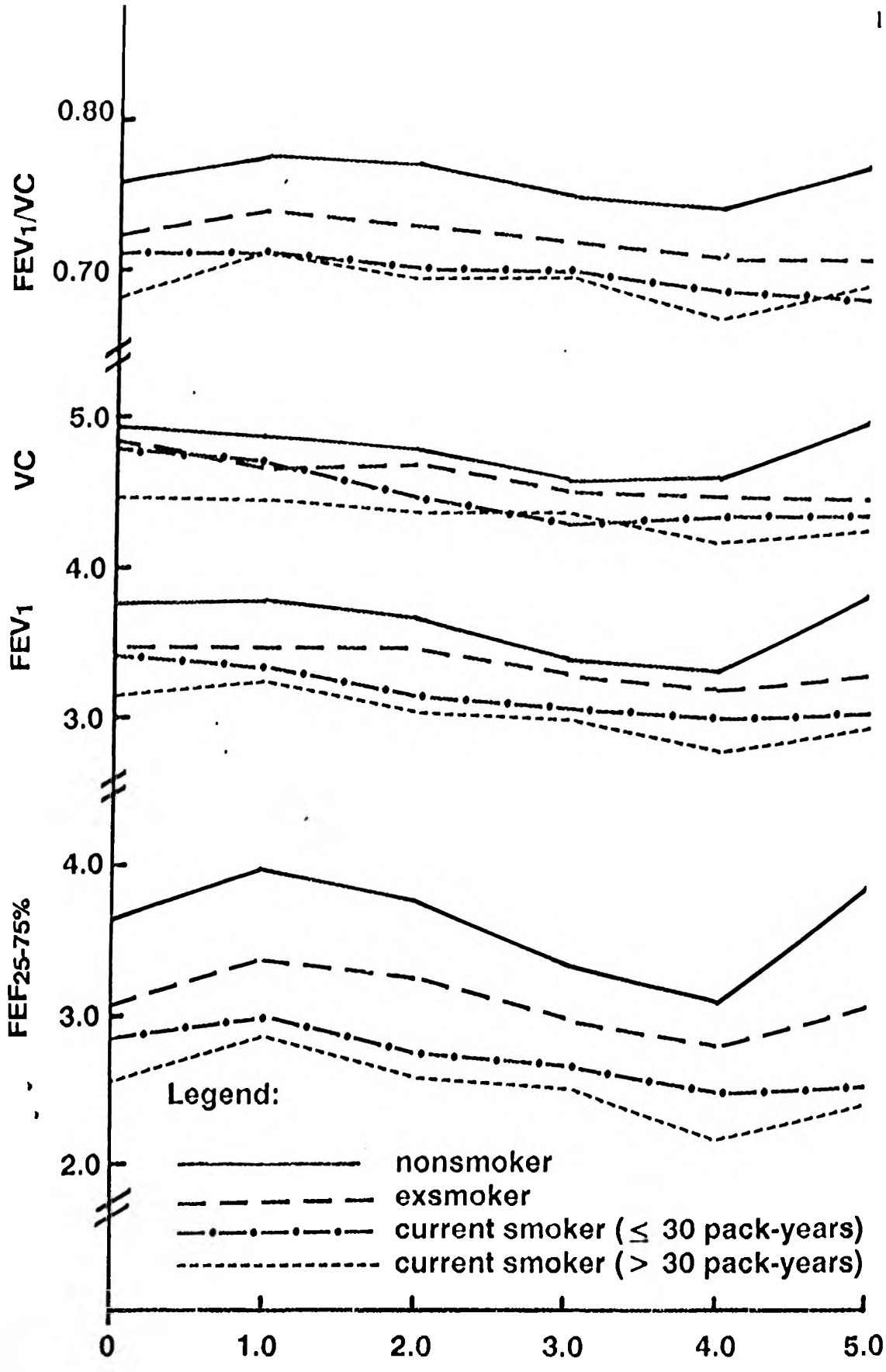


Fig 5.1 Individual lung function measurements, plotted against cumulative dust exposure (RSA x years/ 10 000), according to smoking habits

5.3.2 Analysis of lung function profiles

Table 5.3 shows the distribution of the miners into the six lung function profile groups, and their characteristics. The groups do not differ in age, but the two groups with the restrictive profile are considerably heavier in comparison to the other groups. The observed lung function measurements decrease, and the proportion of those having bronchitis and rhonchi increase, with increasing level of obstruction. Tobacco consumption displays an increasing trend with obstruction. The cumulative dust exposure increases up to moderate obstruction only and then declines slightly. This pattern is reflected by the proportion of silicotics which is highest in the moderate obstructive profile group. The proportions of those who came for a benefit examination (because they were seeking compensation) and of those already certified increase progressively with level of obstruction. The mortality from respiratory disease shows a strong association with increasing obstruction. The restrictive profile group had the highest percentage of death from IHD.

Table 5.3 Characteristics of the lung function profile groups[a]

Characteristic	Lung function profile group					
	'Normal'–reference group	Obstructive				Restrictive
		Minimal	Moderate	Marked	Marked+rest.	
	1 (n=1030)	2 (n=256)	3 (n=111)	4 (n=211)	5 (n=191)	6 (n=187)
	mean±s.d.	mean±s.d.	mean±s.d.	mean±s.d.	mean±s.d.	mean±s.d.
Age	50.1±2.6	50.0±2.5	50.6±2.4	50.3±2.5	50.5±2.5	50.4±2.6
Weight	82.5±12.0	80.3±11.5	78.8±12.0	78.1±12.3	79.3±12.2	85.0±12.9
Height	175.4±6.1	174.9±6.1	174.7±6.6	174.8±5.6	176.2±5.3	176.9±5.8
Body mass	26.8±3.3	26.2±3.2	25.8±3.4	25.5±3.6	25.5±3.5	27.1±3.8
Lung function measurements, symptoms and signs						
$FEV_{1.0}$	3.59±0.53	3.20±0.44	2.90±0.30	2.40±0.37	1.92±0.53	2.80±0.35
FEF_{25-75}	3.61±0.98	2.28±0.43	1.74±0.36	1.26±0.42	1.08±0.46	2.79±0.79
VC	4.68±0.65	4.78±0.67	4.55±0.57	4.31±0.55	3.48±0.48	3.64±0.46
Bronchitis %	23.5	30.9	35.1	48.3	60.7	44.9
Sputum %	32.6	42.2	41.4	52.6	60.0	50.0
Rhonchi %	14.7	20.7	21.6	36.0	44.5	26.7
Cigarette equivalent consumption						
Pack years	21.2±17.3	24.7±16.9	27.5±14.8	29.4±17.1	31.4±16.6	26.3±16.8
Years	19.2±12.1	22.5±10.7	25.3±8.4	26.1±9.4	27.0±7.6	23.2±10.1
Cumulative dust exposure and silicosis						
CDE	28.2±12.1	29.7±12.2	31.5±13.0	30.6±11.4	30.7±11.1	30.5±11.9
Silicosis %	4.6	6.6	18.0	8.1	8.4	7.0
Certification seeking or awarded						
Benefits %	12.9	17.2	17.1	22.3	35.6	21.4
Certified %	11.1	16.8	24.3	29.4	43.5	23.5
Mortality from all causes and from respiratory and cardiovascular diseases						
All causes %	32	45	40	55	64	50
COLD %	0.3	2.3	3.6	7.1	12.6	0
IHD %	14.2	10.9	12.6	21.3	20.9	27.3

[a] Only the miners are included in this table.
COLD=chronic obstructive lung disease, IHD=ischaemic heart disease,
CDE= RSA x years / 1000.

a) Test for interaction

Table 5.3 shows the distribution of the subjects into the six lung function profiles, according to smoking habits and dust exposure. The observed ORs indicate an increase in the prevalence of each impairment as the dust and smoking exposure increase, in comparison to the reference group. There was some difficulty in selecting the baseline groups for the calculation of the ORs. As the non–miners (DUST=0) had higher proportions of obstructive impairments in all three smoking categories than the miners in the lowest dust exposure category (DUST=1.2) they were excluded from further analyses and the miners in the lowest dust exposure (DUST=1.2) were used as baseline. The value of 0.5 was added to cell frequencies to enable the computation, this had the effect of increasing the prevalences in non–smokers, but should not change the dose–response trend in the smokers. While for most of the profile groups the non–smokers did not display an increasing trend with dust, the ex–smokers and current smokers did, and the trend increased with the degree of obstructive impairment.

Figure 5.2 shows the observed ORs and the ORs predicted by the additive and multiplicative models for each obstructive impairment. The ORs are plotted against cumulative dust exposure, according to the smoking status. The predicted odds ratios for the marked obstructive profiles were estimated from the additive model which included the interaction term. The multiplicative model did not include the interaction term as it was not significant.

Table 5.4 Lung function profile groups frequencies (N) and the observed ORs, according to smoking habits and cumulative dust exposure

Dust (CDEC)	Normal–refer. group 1	Lung function profile group: Obstructive Minimal 2	Moderate 3	Marked 4	Marked+rest. 5	Restrictive 6
	N	N OR	N OR	N OR	N OR	N OR
Nonsmokers						
0	56	6 –	1 –	2 –	1 –	0 –
1.2	24	4 1.0	0 1.0[a]	1 1.0[a]	0 1.0[a]	2 1.0
2.0	54	6 0.7	0 0.5	2 0.8	1 1.4	5 1.1
3.0	47	6 0.8	2 2.6	7 2.6	1 1.6	3 0.8
4.0	30	4 0.8	1 2.4	1 0.8	0 0.8	1 0.4
5.2	22	2 0.6	1 3.3	0 0.4	0 1.1	2 1.1
Exsmokers						
0	67	13 –	8 –	7 –	2 –	3 –
1.2	41	7 1.0	2 3.0	3 1.4	2 3.0	2 0.6
2.0	83	14 1.0	7 4.4	5 1.1	4 2.6	4 0.6
3.0	58	11 1.1	6 5.4	6 1.8	4 3.8	8 1.7
4.0	42	16 2.3	4 5.2	5 2.1	5 6.3	6 1.7
5.2	24	9 2.3	5 11.0	0 0.3	3 7.0	5 2.5
Current smoker						
0	142	39 –	29 –	46 –	26 –	17 –
1.2	62	16 1.6	9 7.5	10 2.7	9 7.5	9 1.7
2.0	194	52 1.6	21 5.4	54 4.6	52 13.2	45 2.8
3.0	188	56 1.8	24 6.4	54 4.7	46 12.1	56 3.6
4.0	104	36 2.1	18 8.7	41 6.5	50 23.7	22 2.5
5.2	57	17 1.8	11 9.8	22 6.4	14 12.4	17 3.6
Total	1295	314	149	266	220	207
Miners	1030	256	111	211	191	187

[a] A value of 0.5 was added to each cell frequency.
– Odds ratio was not calculated. CDEC= (RSA x years / 10000)

Table 5.5 shows the results of the analysis with the additive model, Eq. (5.1), fitted to the data for the marked obstructive profiles shown in Table 5.4, individually and combined. The results of the analysis from which the benefit examinations were excluded (n=248) are also shown. Table 5.5 shows the regression estimates, associated chi–square test of significance, goodness–of–fit test ($-2\log L_a$) and likelihood ratio test (X^2_{LR}) of the null hypothesis of no interaction, for the additive model. For comparison, the goodness–of–fit test ($-2\log L_m$) is presented for the multiplicative model fitted without an interaction term. Because of narrow range, age was not a significant predictor for any impairment category.

The regression coefficients estimate excess relative risk attributable to each exposure that is above the baseline risk, which is set to unity. Thus, for example, the estimated odds ratio for the marked obstruction with restriction impairment for the current smoker with a value of 5 for cumulative dust exposure (RSA x years / 10 000) is calculated as $1–0.025\cdot(5–1)+6.456+2.176\cdot(5–1) = 16.1$, and the observed value given in Table 5.4 was 12.4 (see Table 5.5 for the coefficients).

The interaction effect was found to be significant at 0.05 probability level for the combined marked obstructive profiles (4+5). As the interaction terms for the individual marked obstructive profiles were significant for the other variable types, it is likely that adding the 0.5 value to the cell frequencies and combining the current smokers into one category had the effect of decreasing the interaction effects ($0.05<p<0.10$). The parameter error estimates from the additive model are known to be unreliable for making statistical inference, and the log likelihood ratio test should be used as a test for absence of an interaction effect (Moolgavkar and Venzon, 1987).

As the miners with the lowest exposure were used as a baseline, the values of one was subtracted from the dust exposure in order to get the value of 1.0 for the OR in the lowest exposure category, this should not influence the strength of association.

Table 5.5 Results from the additive RR model, Eq. (5.1), fitted with indicator smoking variables and an ordinal dust variable to the marked obstructive profiles

Effect	β ± S.E	X^2_1	β ± S.E.	X^2_1
	Marked obstructive profiles			
	Without restriction		With restriction	
Intercept	−2.587 ± 0.289	80.1	−3.686 ± 0.461	63.9
CDEC	−0.143 ± 0.167	0.7	−0.025 ± 0.449	0.0
Exsmoker	0.107 ± 0.649	0.1	0.810 ± 1.320	0.3
Curr. smoker	1.631 ± 0.986	2.7	6.456 ± 3.883	2.8
CDEC·Exsmk	0.163 ± 0.247	0.4	0.878 ± 0.596	2.2
CDEC·Curr.smk	0.897 ± 0.322	7.8	2.176 ± 1.218	3.2
$-2\log L_a$	1099.2		1007.2	
X^2(LR),df=2	4.5		5.1	
$-2\log L_m$	1100.8		1008.0	
	Combined marked obstructive profiles (4+5)			
	Benefits included		Benefits excluded	
Intercept	−2.449 ± 0.266	84.8	−2.552 ± 0.282	82.0
CDEC	−0.021 ± 0.182	0.0	−0.082 ± 0.183	0.2
Exsmoker	0.337 ± 0.616	0.3	0.340 ± 0.627	0.3
Curr smoker	3.387 ± 1.421	5.7†	3.125 ± 1.472	4.5
CDEC·Exsmk	0.309 ± 0.234	1.7	0.242 ± 0.207	1.4
CDEC·Curr.smk	1.318 ± 0.534	6.1†	1.265 ± 0.298	18.0
$-2\log L_a$	1594.8		1239.9	
X^2(LR),df=2	7.3†		5.4	
$-2\log L_m$	1596.7		1241.1	

† $p<0.025$. CDEC= RSA x years / 10000.
$-2\log L_a$= −2 log likelihood, from additive model.
$-2\log L_m$= −2 log likelihood, from multiplicative model.

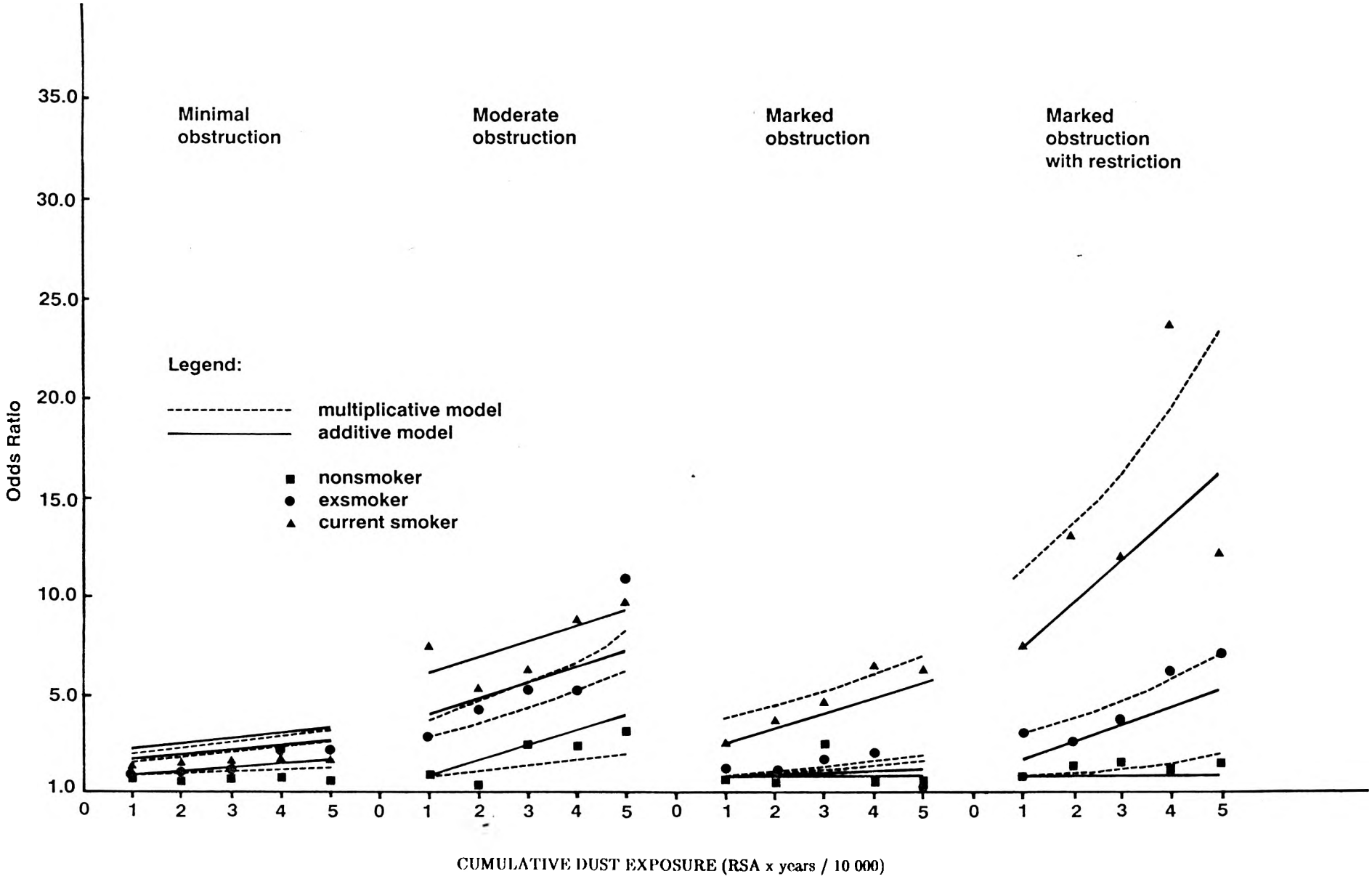


Fig 5.2 Observed and predicted odds ratios for individual lung function profiles, plotted against cumulative dust exposure (RSA x years / 10 000), according to smoking habits

Table 5.6 shows the results from the analysis with categorized exposure data, Eq. (5.2), as the results from the analysis with the continuous variables were very similar they are not presented. The X^2–test statistic shows that for the model fitted without the interaction term, the effect of dust and smoking increase with increasing level of obstruction. For the model fitted with the interaction term the interaction effect becomes significant for the marked obstructive profiles c) to e). Table 5.7 summarizes the results of the likelihood ratio tests of the absence of interaction effect for the obstructive and restrictive impairments, for the categorized and continuous exposure variables as included in Eqs (5.1) and (5.2). The interaction effect was not significant at the 0.05 probability level for the minimal and moderate obstructive profiles for any type of the exposure variables. For the marked obstructive profiles (without and with restriction), the interaction effect was significant for the continuous and categorized variables. For the restrictive profile the interaction effect between dust and smoking was not significant.

b) **Estimation of attributable risk**

Tables 5.8 to 5.12 show the two–way frequency distribution of the cases with the different impairment profiles and of the referents. Each table also shows the odds ratios for each two–way cell and the adjusted marginal odds ratios. The odds ratios for the two–way cells show that the highest smoking categories display increasingly stronger dose–response trend with dust as the degree of obstruction increases.

Each table also shows the etiological and preventable fractions. Thus for example, the etiological fractions for the cases with marked obstructive with restriction profile (Table 5.11) were estimated as EF_d=52%, 95% CL (41; 62) for dust, EF_s=94% (87; 100) for smoking, and $EF_{s.d}$=91% (84; 98) for the total effect

of dust and smoking. The preventable fraction indicated that approximately 39% of the cases could have been prevented by elimination of tobacco smoking, 0% by elimination of dust exposure only, and 40% by elimination of either dust or smoking.

Table 5.6 Results from the additive RR model fitted with the categorized exposure variables, for each obstructive profile

	Additive RR model			
	Without interaction		With interaction	
Effect	$\beta \pm$ S.E.	X^2_1	$\beta \pm$ S.E.	X^2_1
a) minimal obstructive v.s. reference group				
Intercept	-0.919 ± 0.044		-0.914 ± 0.063	
CDEC	$2.59\times10^{-3}\pm1.43\times10^{-3}$	3.3	$2.36\times10^{-3}\pm2.31\times10^{-3}$	1.0
CEPYC	$4.78\times10^{-2}\pm1.31\times10^{-2}$	13.3	$4.40\times10^{-2}\pm3.30\times10^{-2}$	1.8
Interact.			$0.15\times10^{-3}\pm1.21\times10^{-3}$	0.0
$-2\ln L$	1267.84		1267.82	0.02
b) moderate obstructive v.s. reference group				
Intercept	-1.025 ± 0.084		-1.028 ± 0.046	
CDEC	$2.09\times10^{-3}\pm0.68\times10^{-3}$	9.5	$2.23\times10^{-3}\pm1.57\times10^{-3}$	2.0
CEPYC	$3.74\times10^{-2}\pm0.67\times10^{-2}$	31.2	$4.01\times10^{-2}\pm2.47\times10^{-2}$	2.6
Interact.			$-1.11\times10^{-4}\pm8.21\times10^{-4}$	0.0
$-2\ln L$	703.34		703.32	0.02
c) marked obstructive v.s. reference group				
Intercept	-0.991 ± 0.031		-0.908 ± 0.060	
CDEC	$1.55\times10^{-3}\pm0.43\times10^{-3}$	13.0	$-1.57\times10^{-3}\pm1.94\times10^{-3}$	0.8
CEPYC	$7.62\times10^{-2}\pm0.77\times10^{-2}$	97.9	$1.10\times10^{-2}\pm3.02\times10^{-2}$	0.1
Interact.			$2.40\times10^{-3}\pm1.02\times10^{-3}$	5.5
$-2\ln L$	1068.80		1063.36	5.44
d) marked obstructive with restriction v.s. reference group				
Intercept	-1.010 ± 0.018		-0.980 ± 0.014	
CDEC	$8.52\times10^{-4}\pm1.04\times10^{-4}$	67.1	$-3.83\times10^{-4}\pm4.06\times10^{-4}$	0.9
CEPYC	$8.61\times10^{-2}\pm0.75\times10^{-2}$	132	$2.84\times10^{-2}\pm1.62\times10^{-2}$	3.1
Interact.			$2.13\times10^{-3}\pm0.55\times10^{-3}$	15.0
$-2\ln L$	965.46		958.34	7.12
e) marked obstructive (4+5) v.s. reference group – benefits excluded				
Intercept	-0.997 ± 0.035		-0.918 ± 0.061	
CDEC	$1.84\times10^{-3}\pm0.51\times10^{-3}$	13.0	$1.16\times10^{-3}\pm2.05\times10^{-3}$	0.3
CEPYC	$1.32\times10^{-1}\pm0.11\times10^{-1}$	144	$4.18\times10^{-2}\pm3.80\times10^{-2}$	1.2
Interact.			$3.34\times10^{-2}\pm1.39\times10^{-2}$	5.8
$-2\ln L$	1213.62		1207.18	6.44

Table 5.7 Summary table of the log likelihood ratio tests of no interaction

Lung function profile	Additive model		
	1	2	3
	X^2_2	X^2_1	X^2_1
Obstructive			
minimal	3.3	0.1	0.0
moderate	0.4	0.4	0.0
marked	4.5	6.3†	5.4†
marked+restr.	5.1	8.4†	7.1†
all marked	7.3†	12.8§	9.7‡
all marked[c]	5.3	9.3‡	6.4†
Restrictive	3.6	4.0	1.4

Model 1 used CDEC, SMK1, SMK2 (likelihood ratio test with 2 df)
Model 2 used CDE, CEPY (likelihood ratio test with 1 df)
Model 3 used CDEC, SMOKEC (likelihood ratio test with 1 df)
[c] marked obstructive without benefits
† $p<0.025$; ‡ $p<0.01$; § $p<0.001$

Table 5.8 Etiological fractions for minimal obstructive profile

Dust[a]	Case	Smoking category[b]					Total	OR (95%CL)
		0	1	2	3	4		
1.2	Yes	4	7	2	9	5	27	1.0
	No	24	41	19	30	13	127	
	(OR)	(1.0)	(1.0)	(0.6)	(1.8)	(2.3)		
2.0	Yes	6	14	8	28	16	72	1.0
	No	54	83	45	98	51	331	(0.6;1.6)
	(OR)	(0.7)	(1.0)	(1.1)	(1.7)	(1.9)		
3.0	Yes	6	11	18	28	10	73	1.1
	No	47	58	48	101	39	293	(0.7;1.9)
	(OR)	(0.8)	(1.1)	(2.3)	(1.7)	(1.5)		
4.0	Yes	4	16	12	19	5	56	1.5
	No	30	42	35	47	22	176	(0.9;2.5)
	(OR)	(0.8)	(2.3)	(2.1)	(2.4)	(1.4)		
5.2	Yes	2	9	3	12	2	28	1.3
	No	22	24	14	36	7	103	(0.7;2.4)
	(OR)	(0.5)	(2.3)	(1.3)	(2.0)	(1.7)		
Total	Yes	22	57	43	96	38	256	
	No	177	248	161	312	132	1030	
OR		1.0	1.9	2.1	2.5	2.4		
95%CL			(1.1;3.2)	(1.2;3.7)	(1.5;4.1)	(1.3;4.2)		

Summary etiological fractions:
EF_d=0.477, 95% CL(0.402;0.552)
EF_s=0.816, 95% CL(0.762;0.870)
$EF_{d\cdot s}$=0.896 95% CL(0.866;0.926)

Preventable fractions:
PF_{dust}=0.896–0.816=0.080
$PF_{smoking}$=0.896–0.477=0.419
$PF_{dust\cdot smoking}$=0.896–0.080–0.419=0.397
PF_{other}=1.00–0.896=0.104

[a] cumulative dust exposure (rsa x years / 10 000)
[b] smoking categories (never, exsmoker,current smoker:pack-years <20, 20-39, 40+)

Table 5.9 Etiological fractions for moderate obstructive profile

Dust[a]	Case	Smoking category[b]					Total	OR (95%CL)
		0	1	2	3	4		
1.2	Yes	0	2	1	4	4	11	1.0
	No	24	41	19	30	13	127	
	(OR)	(1.0)	(2.4)	(2.6)	(6.5)	(15.1)		
2.0	Yes	0	7	5	10	6	28	0.9
	No	54	83	45	98	51	331	(0.4;1.9)
	(OR)	(–)	(4.1)	(5.4)	(5.0)	(5.8)		
3.0	Yes	2	6	7	12	5	32	1.2
	No	47	58	48	101	39	293	(0.6;2.5)
	(OR)	(2.1)	(5.1)	(7.1)	(5.8)	(6.3)		
4.0	Yes	1	4	4	12	2	23	1.5
	No	30	42	35	47	22	176	(0.7;3.3)
	(OR)	(1.6)	(4.7)	(5.6)	(13)	(4.5)		
5.2	Yes	1	5	2	7	2	17	2.0
	No	22	24	14	36	7	103	(0.9;4.5)
	(OR)	(2.2)	(10)	(7.0)	(9.5)	(14)		
Total	Yes	4	24	19	45	19	111	
	No	177	248	161	312	132	1030	
OR		1.0	4.4	5.7	6.7	6.7		
95%CL			(1.5;13)	(1.9;17)	(2.4;19)	(2.2;20)		

Summary etiological fractions:
EF_d=0.199, 95% CL(0.044;0.356)
EF_s=0.798, 95% CL(0.697;0.899)
$EF_{d\cdot s}$=0.839, 95% CL(0.794;0.884)

Preventable fractions:
PF_{dust}=0.839–0.798=0.041
$PF_{smoking}$=0.839–0.199=0.640
$PF_{dust\cdot smoking}$=0.839–0.041–0.640=0.158
PF_{other}=1.00–0.839=0.161

[a] cumulative dust exposure (rsa x years / 10 000)
[b] smoking categories (never, exsmoker,current smoker:pack-years <20, 20-39, 40+)

Table 5.10 Etiological fractions for marked obstructive profile

Dust[a]	Case	Smoking category[b]					Total	OR (95%CL)
		0	1	2	3	4		
1.2	Yes	1	3	2	6	2	14	1.0
	No	24	41	19	30	13	127	
	(OR)	(1.0)	(1.8)	(2.5)	(4.8)	(3.7)		
2.0	Yes	2	5	11	22	21	61	1.5
	No	54	83	45	98	51	331	(0.8;2.8)
	(OR)	(0.9)	(1.5)	(5.9)	(5.4)	(9.9)		
3.0	Yes	7	6	11	30	13	67	2.0
	No	47	58	48	101	39	293	(1.1;3.7)
	(OR)	(3.6)	(2.5)	(5.5)	(7.1)	(8.0)		
4.0	Yes	1	5	7	23	11	47	2.4
	No	30	42	35	47	22	176	(1.2;4.5)
	(OR)	(0.8)	(2.9)	(4.8)	(12)	(12)		
5.2	Yes	0	0	5	12	5	22	2.0
	No	22	24	14	36	7	103	(1.0;4.2)
	(OR)	(–)	(–)	(8.6)	(7.8)	(17)		
Total	Yes	11	19	36	93	52	211	
	No	177	248	161	312	132	1030	
OR		1.0	1.3	3.6	4.8	6.5		
95%CL			(0.6;2.5)	(2.0;7.3)	(2.5;9.1)	(3.2;13)		

Summary etiological fractions:
EF_d=0.431, 95% CL(0.327;0.535)
EF_s=0.691, 95% CL(0.611;0.771)
$EF_{d\cdot s}$=0.808, 95% CL(0.764;0.852)

Preventable fractions:
PF_{dust}=0.808–0.691=0.117
$PF_{smoking}$=0.808–0.431=0.377
$PF_{dust\cdot smoking}$=0.808–0.117–0.377=0.314
PF_{other}=1.00–0.808=0.192

[a] cumulative dust exposure (rsa x years /10 000)
[b] smoking categories (never, exsmoker,current smoker:pack-years <20, 20-39, 40+)

Table 5.11 Etiological fractions for marked obstructive with restriction profile

Dust[a]	Case	Smoking category[b]					Total	OR (95%CL)
		0	1	2	3	4		
1.2	Yes	0	2	1	4	4	11	1.0
	No	24	41	19	30	13	127	
	(OR)	(1.0)	(2.7)	(2.9)	(7.5)	(17.2)		
2.0	Yes	1	4	8	33	11	57	1.8
	No	54	83	45	98	51	331	(0.9;3.7)
	(OR)	(1.0)	(2.7)	(9.9)	(19)	(12.1)		
3.0	Yes	1	4	10	23	13	51	1.9
	No	47	58	48	101	39	293	(1.0;3.9)
	(OR)	(1.2)	(3.9)	(12)	(13)	(18.6)		
4.0	Yes	0	5	8	28	14	55	3.7
	No	30	42	35	47	22	176	(1.8;7.4)
	(OR)	(–)	(6.7)	(13)	(33)	(35.6)		
5.2	Yes	0	3	1	8	5	17	2.0
	No	22	24	14	36	7	103	(0.9;4.5
	(OR)	(–)	(7.0)	(4.0)	(12)	(39.9)		
Total	Yes	2	18	28	96	47	191	
	No	177	248	161	312	132	1030	
OR		1.0	6.6	16.6	28.3	32.1		
95%CL			(1.6;28)	(4.0;70)	(6.8;115)	(7.6;132)		

Summary etiological fractions:
EF_d=0.517, 95% CL(0.414;0.620)
EF_s=0.941, 95% CL(0.872;1.000)
$EF_{d\cdot s}$=0.906, 95% CL(0.836;0.975)

Preventable fractions:
PF_{dust}=0.906–0.941=0.0
$PF_{smoking}$=0.906–0.517=0.389
$PF_{dust\cdot smoking}$=0.906–0.0–0.389=0.517
PF_{other}=1.00–0.906=0.094

[a] cumulative dust exposure (rsa x years /10 000)
[b] smoking categories (never, exsmoker·current smoker·pack-years <20, 20-39, 40+)

Table 5.12 Etiological fractions for marked obstructive (4+5), benefit examinations were excluded

Dust[a]	Case	Smoking category[b]					Total	OR (95%CL)
		0	1	2	3	4		
1.2	Yes	1	3	3	9	5	21	1.0
	No	22	36	19	27	13	117	
	(OR)	(1.0)	(1.8)	(3.5)	(7.3)	(8.5)		
2.0	Yes	2	8	13	40	25	88	1.5
	No	52	75	39	91	48	305	(0.9;2.5)
	(OR)	(0.9)	(2.4)	(7.3)	(9.7)	(11.5)		
3.0	Yes	6	7	16	37	18	84	1.9
	No	37	48	36	87	33	241	(1.1;3.3)
	(OR)	(3.6)	(3.2)	(9.8)	(9.4)	(12.0)		
4.0	Yes	1	5	9	35	17	67	2.5
	No	28	38	24	40	19	149	(1.4;4.5)
	(OR)	(0.8)	(2.9)	(8.3)	(19)	(19.7)		
5.2	Yes	0	2	5	12	8	27	1.9
	No	18	20	12	29	6	85	(1.0;3.6)
	(OR)	(−)	(2.2)	(9.2)	(9.1)	(29.3)		
Total	Yes	10	25	46	133	73	287	
	No	157	217	130	274	119	897	
OR		1.0	1.8	5.6	7.6	9.8		
95%CL			(0.9;3.9)	(2.7;12)	(3.9;15)	(5.0;19)		

Summary etiological fractions:
EF_d=0.422, 95% CL(0.331;0.513)
EF_s=0.801, 95% CL(0.737;0.865)
$EF_{d\cdot s}$=0.863 95% CL(0.827;0.899)

Preventable fractions:
PF_{dust}=0.863–0.801=0.062
$PF_{smoking}$=0.863–0.422=0.441
$PF_{dust\cdot smoking}$=0.863–0.062–0.441=0.360
PF_{other}=1.00–0.863=0.137

[a] cumulative dust exposure (rsa x years /10 000)
[b] smoking categories (never, exsmoker,current smoker:pack-years <20, 20-39, 40+)

Table 5.13 summarizes the etiological and preventable fractions. The size of etiological fraction is a reflection of the size of the marginal ORs for dust. The etiological fractions for dust increase consistently with increasing obstruction, but the preventable fractions indicate that this is mostly due to the combined effect of dust and smoking, which increases with increasing obstruction as 14%, 16%, 32% and 52%. For the two marked obstructive profiles the etiological fraction for dust are more than twice as high as that for the moderate obstructive profile which had the highest dust exposure and highest proportion of silicotics. Thus the results show that heavy tobacco smoking potentiates the effect of dust on lung function impairment. The preventable fractions show that as the severity of the obstructive impairment increases, smoking on its own and the combined effects of smoking and dust become more important.

Table 5.13 Summary of the etiological and preventable fractions

Lung function profile	Risk factor		
	Dust	Smoking	Dust and Smoking
	% (95% CL)	% (95% CL)	% (95% CL)
a) etiological fractions			
Obstructive			
minimal	13.6(2.3;24.9)	50.0(41.8;58.2)	36.3(30.9;41.7)
moderate	20.0(4.4;35.6)	79.8(69.7;89.9)	83.9(79.4;88.4)
marked	43.1(32.7;53.5)	69.1(61.1;77.1)	80.8(76.4;85.2)
marked+rest.	51.7(41.4;62.0)	94.1(87.2;100.)	90.6(83.6;97.5)
all marked	47.7(40.2;55.2)	81.6(76.2;87.0)	89.6(86.6;92.6)
all marked[a]	42.2(33.1;51.3)	80.1(73.7;86.5)	86.3(82.7;89.9)
Restrictive	41.6(30.3;52.9)	59.8(50.9;68.7)	58.0(32.7;83.2)
b) preventable fractions			
Obstructive			
minimal	0	23	14
moderate	4	64	16
marked	11	38	32
marked+rest.	0	39	52
all marked	8	42	40

[a] Benefit examinations were excluded.

The etiological fractions for the marked obstructive profiles from which those that came for a benefit examination were excluded were only slightly lower than those obtained from all the data (Table 5.12). Figure 5.3 shows the percentages of all the marked obstructive cases (n=402) that could have been prevented by the elimination of which ever exposure.

5.4 Discussion

The results of the present analysis indicate that for the individual lung function measurements (FEV_1, VC and FEF_{25-75}) and the ratio FEV_1/VC, the dose–response trends with dust are similar for nonsmokers, ex–smokers and current smokers (Figure 5.1). The interaction effect of dust and smoking was not statistically significant in the linear or logistic regression model. The estimated loss of lung function (FEV_1) attributable to silica dust exposure ranged from 123 ml to 319 ml. In current smokers the estimated loss of FEV_1 ranged from 500 to 687 ml (Table 5.2). Thus the effect of dust exposure was estimated to be about half of that of smoking. The absence of statistically significant interaction between dust and smoking on the individual lung function measurements is consistent with the observations from studies done on coal miners (as reviewed by Marine *et al.* (1988). Marine *et al.* (1988) stated that the results from the analysis of individual lung function tests, namely FEV_1, do not support the idea that cigarette smoking potentiates the effect of dust exposure on the chance of developing bronchitis symptoms or lung function decrements.

However, in our study, when the miners were grouped according to their lung function impairment profiles, a synergistic effect between dust and smoking

became apparent. The miners with the marked obstructive profiles were found to have the highest tobacco consumption and strongest dose–response trend with dust, yet they did not have the highest dust exposure or the highest proportion of silicotics (Table 5.3). The dose response trend with dust was found to be stronger for the marginal odds ratios adjusted for smoking (see Tables 5.8 to 5.12), and for the current smokers (see Table 5.4 and Figure 5.2) which was mainly for those with tobacco consumption was over 20 pack–years (see Tables 5.10 and 5.11). It can be seen that the adjusted odds ratios for dust were much lower for the moderate obstructive profile than for the two marked obstructive profiles. Figure 5.2 shows that with increasing obstruction the number of impaired lung functions in the nonsmokers decreases and the dose–response trend with dust becomes progressively steeper for the current smokers.

The log likelihood ratio tests obtained from fitting the additive relative risk model to different types of exposure variables confirmed that the joint effect of dust and tobacco smoking on the prevalence of marked obstructive impairment was higher than it would be expected if the effects were additive (Tables 5.5 to 5.7). No interaction term was significant for the multiplicative model.

The etiological fractions, estimated for each impairment profile in comparison with the reference group, also display an increasing trend with the degree of obstruction (Table 5.13). Reducing the respirable dust to which the miners are exposed on South African gold mines to a non–harmful level would theoretically prevent the occurrence of 48% of the marked obstructive impairments (n=402). Elimination of tobacco smoking would prevent 82% of the marked obstructive cases and 94%, 95%CL (87;100), of the marked obstructive cases with restriction (n=191) who had the highest mortality from COLD (Table 5.3).

The results from this analysis are consistent with those obtained from the mortality study reported in Chapter 6. Those results showed a similar pattern,

namely that the death from COLD occurred in tobacco smokers only, and that the dose–response trend with dust increased with increasing years of smoking. The combined effect of the two exposures was close to being multiplicative. As the miners who died from COLD were mostly those who suffered from marked obstruction, the exposure patterns should be similar.

The association between the restrictive profile and mortality from IHD was observed in several other studies (Kannel *et al.*, 1983), and in this study it may be associated with obesity.

Possible biases in this study may be related to the selection of the non–mining controls, who were heavier smokers (the average cigarette equivalent pack–years were 27.1, SE 1.0, and 24.3, SE 0.4, $p<0.01$ for non–miners and miners, respectively) and had a higher prevalence of the obstructive impairments. However, the nonsmoking non–miners also had a higher prevalence of the obstructive impairments than the nonsmoking miners with the lowest dust level, and this could indicate that either the miners were fitter when selected originally, or that a healthy worker effect was present. The use of the miners in the lowest dust category as baseline groups could have underestimated the effect of dust, but less so than if the non–miners were used as baselines.

A selection bias, whereby the fitter workers accumulate higher dust exposures, was probably involved also in this study, and may have caused underestimation of the effect of dust and of the synergistic effect. Furthermore, a negative association between tobacco consumption and dust exposure was observed among the miners. The mean pack–years, according to the categories of cumulative dust, were 27.1, 25.2, 25.2, 23.2 and 23.6. This indicates that the miners who smoked less were fitter to continue to work in dusty occupations. Zwi (1959) has also observed that gold miners doing the heaviest work with the greatest dust exposure had the best lung function, and more recently this

observation was confirmed for the black gold miners (Cowie, 1987). This effect may have been responsible for the change in dose–response trend in the highest dust exposure category, as shown in Fig. 5.1.

The miners whose original spirograms were rejected because of inconsistencies, and consequently their lung function data were not recorded on the computer file, were compared with the miners who were included in the study. They were found to have higher cumulative dust up to the 1970 examination ($p=0.001$) but the difference in cigarette equivalent pack–years was not statistically significant. They had higher mortality from all causes, 55% versus 46% ($p<0.0001$), and from COLD, 8% versus 3% ($p<0.0001$). Thus, the exclusion of these men from the original study and from this analysis has probably resulted in an underestimation of the effect of dust.

The miners selected for this study could have come to the MBOD for a regular annual (so called periodical) examination or a specialized (so called benefit) examination. A miner may attend a benefit examination during one year and a periodical examination during the next year. By the age of 45 years many miners have already accumulated sufficient service required for compensation, and thus request a benefit examination to establish whether they qualify for compensation. Many miners who attend the benefit examination are not found to be disabled and are not granted disability certification which entails compensation. Table 4.3 shows the miners categorized according to the reason for attending the medical examination. The mean dust exposures, cumulated to the individual decades, were higher and lung function measurements were lower for the benefit examinations than for the periodical examinations. No important differences in age, body weight or height, body mass index or tobacco consumption were observed between the four groups. The possibility that the fact that the men who came for their benefit examination could have biased the results of this study was

considered. Tables 5.5 to 5.7 report on analyses from which those who had attended the benefit examinations were excluded. The pattern of dose response trends and the regression results were very similar to those for the whole data. Table 5.12 shows the two–way frequency table and the etiological fractions for the marked obstructive profiles (4+5) from which those attending the benefit examinations were excluded. The etiological fractions changed only slightly and the change was most noticeable for dust, but not so much for smoking, which is in agreement with the results from the comparison of those attending the benefit examination with the rest of the cohort. Generally, the overall pattern was the same as that obtained from all the data.

It can be seen (Table 5.3) that the disability certification which leads to a financial compensation was related to the amount of tobacco consumption. The moderate obstructive group which had higher dust exposure, higher rate of silicosis and lower tobacco consumption, in comparison to the two marked obstructive groups, had lower rate of certification for COLD or pneumoconiosis than the latter two groups. Thus it appears that the heavily smoking miners had an advantage for compensation over the more heavily dust exposed miners who smoked less. Since the effect of smoking on loss of lung function is almost twice as high than the effect of dust this is to be expected.

The results from the analysis of the individual lung function measurements and from the analysis of the lung function profiles indicate that the miners who are most likely to become disabled because of low lung function are the heavy smokers, and the smokers with heavy dust exposure. Thus, tobacco smoking, and the combined effect of dust and tobacco smoking, are the main risk factors for respiratory disability in the white South African gold miners. Figures 5.4 to 5.6 show the survival of the cohort during the follow–up period, according to lung function profiles, smoking habits and quartiles of dust exposure. The relative

importance of dust and smoking on loss of lung function and subsequent mortality can be confirmed from these three graphs.

There are several possible reasons for the different results obtained from the analysis of the individual tests and from the analysis of the profile groups, with respect to the interaction effect. Firstly, categorizing the men on the basis of the spirogram performance may have helped to eliminate misclassification errors for the disability. Secondly, obstructive impairment is a function of FEV_1 and VC (Morgan, 1987), and if exposure to dust causes parallel reduction of FEV_1 and VC (Soutar *et al.*, 1986) than the ratio FEV_1/VC, often used as a measure of obstruction, may not change with increasing dust level. In our study the parallel reduction in airflow as well as in lung volume was considered. Also the observed mid–expiratory flow time, used in the grouping, was independent of the predicted values. Thirdly, the impaired miners were compared to miners who had 'normal' lung function tests, and in relation to these miners the synergistic effect of dust and smoking became apparent. Finally, by grouping the men into the lung function profiles, the groups became more homogeneous with respect to their exposure pattern and thus the individual and combined effects could be more easily identified.

It has been shown that the logistic model, often used for the analysis of dichotomous outcome variables, assumes multiplicativity of relative risks and thus may not detect synergistic effects that are important from the public health point of view (Rothman, 1974). It was proposed. that when the multiplicative model fits the data substantially better than the additive model without an interaction term, then a synergistic effect of public health importance could be assumed (Blot and Day, 1979). The results from this study support the proposition. Figure 5.2 and the log likelihoods in Table 5.5 show that the additive model with interaction and the multiplicative model without interaction fitted almost equally well to the data,

and better than the additive model without the interaction term. It has been shown that the parameter error estimates obtained from the additive model are unreliable for inference purposes and that the log likelihood ratio statistic for *absence* of effect can be used to test the hypotheses of additivity or multiplicativity of relative risks (Moolgavkar and Venzon, 1987). It was also observed in the present study that the inference based on the parameter error estimates and on the log likelihood ratio statistic did not always lead to the same conclusion. Thus the log likelihood ratio test was used for inference purposes.

In conclusion, it can be said that this study supports the hypothesis that tobacco smoking potentiates the effect of the underground gold mining dust exposure to increase the chances of developing a severe obstructive lung function impairment which leads to increased mortality from COLD. The main risk factors for marked respiratory disability are tobacco smoking on its own, and the combined effect of tobacco smoking and silica dust exposure. Silica dust exposure on its own can rarely cause a serious repiratory disability.

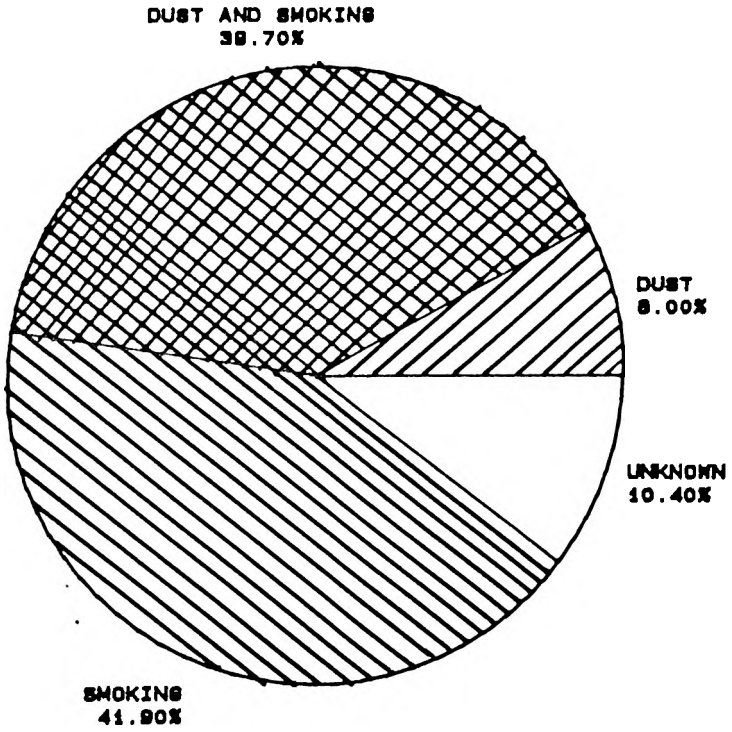


Fig 5.3 The risk approportionment for the cases with marked obstructive impairment profile (n=402)

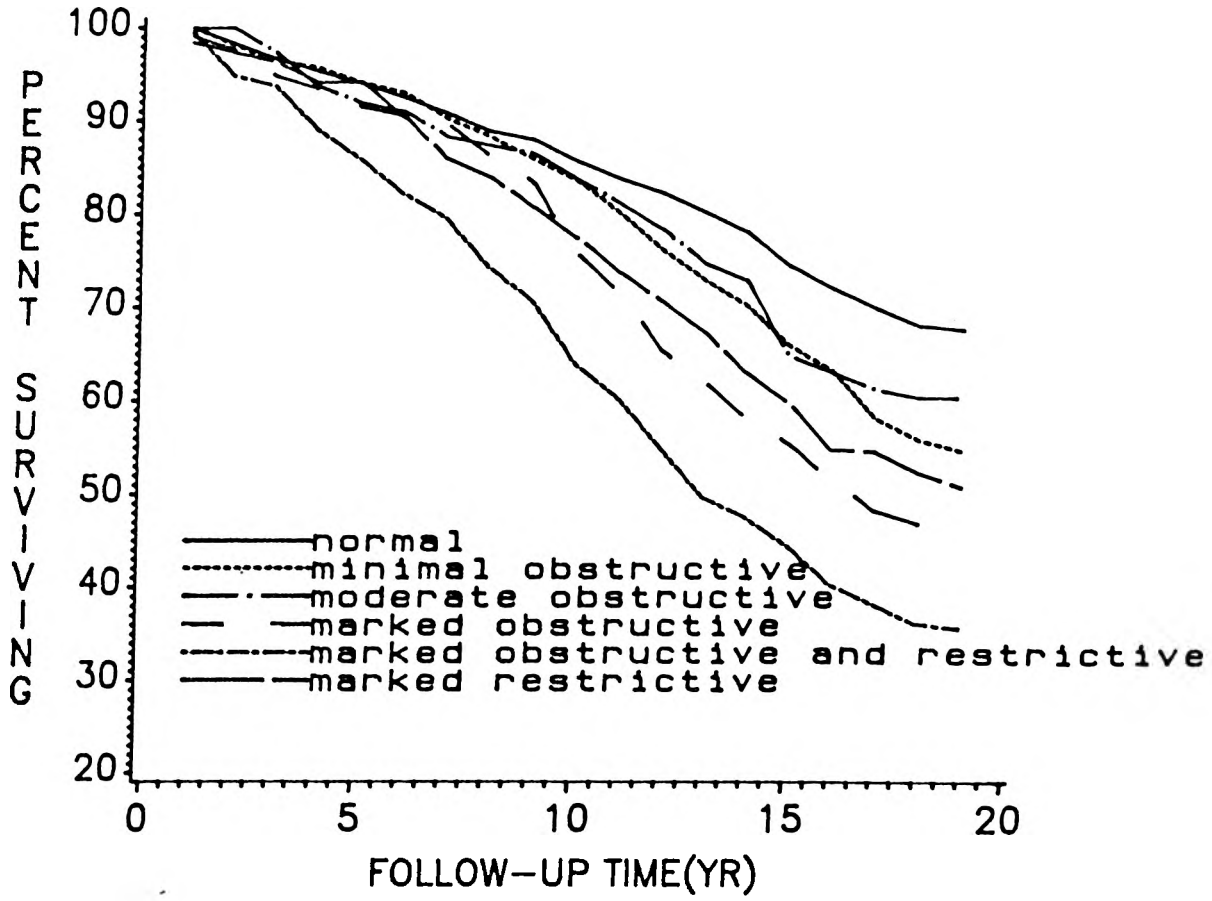


Fig 5.4 Percent surviving during the follow-up period, according to lung function impairment profile

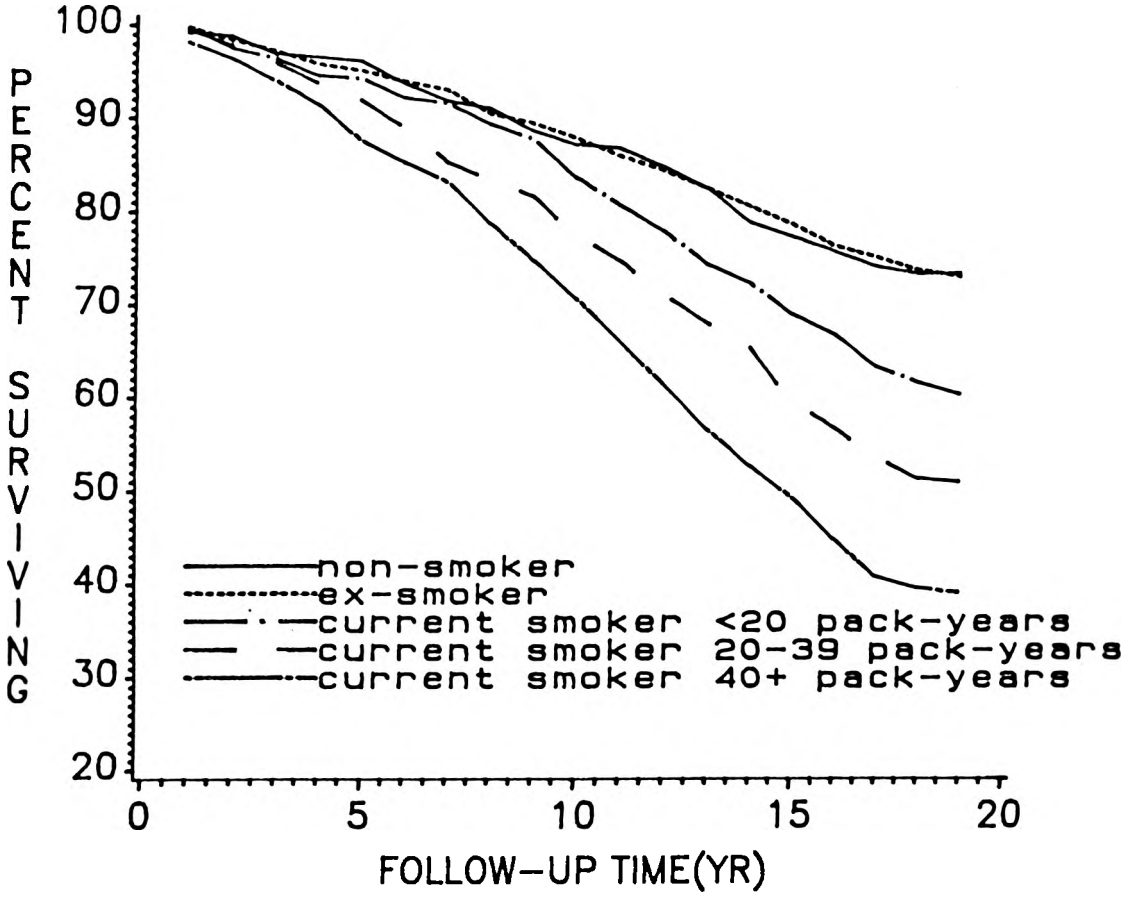


Fig 5.5 Percent surviving during the follow-up period, according to smoking habits

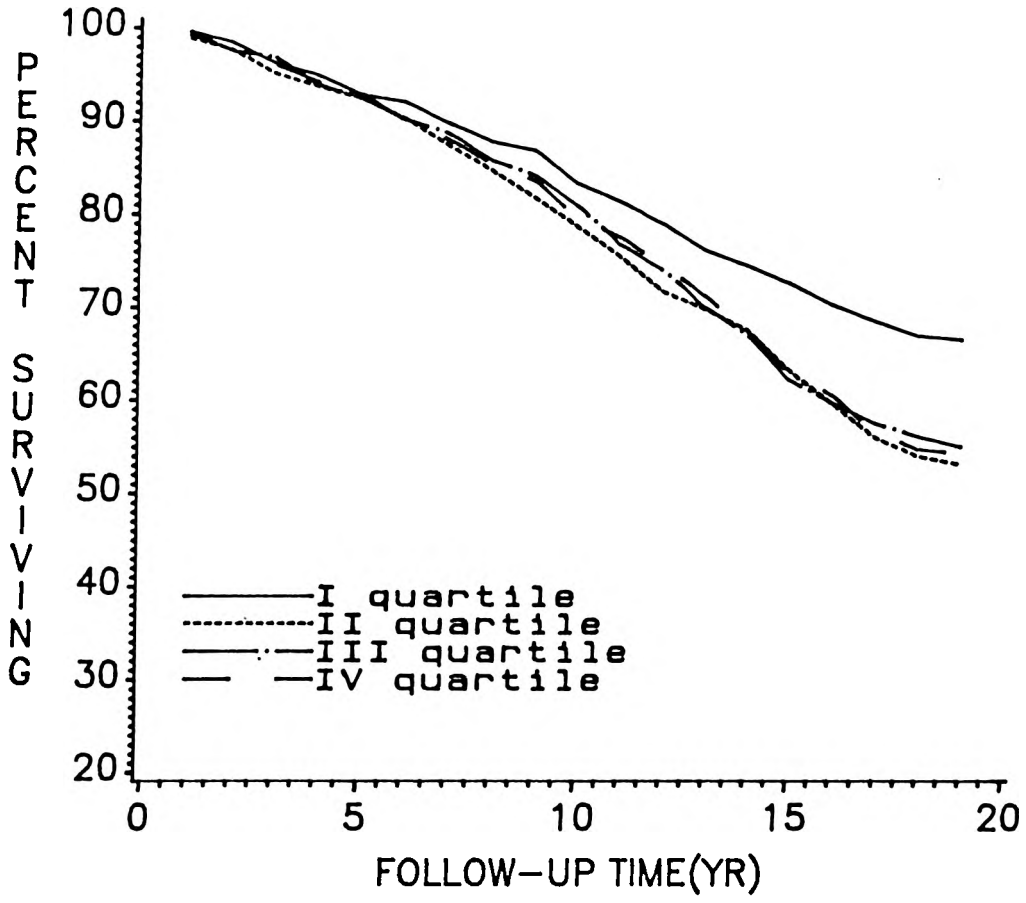


Fig 5.6 Percent surviving during the follow-up period, according to cumulative dust exposure quartiles

CHAPTER 6

MORTALITY FROM CHRONIC OBSTRUCTIVE LUNG DISEASE

6.1 Objectives of chapter 6

In Chapter 5, it was reported that the cases with the more severe respiratory impairment had significantly higher tobacco consumption and exposure to gold mining dust than the controls who had normal lung function tests, and that the interaction between these two risk factors was statistically significant in the additive relative risk model. Further, it was shown, that the cases with respiratory impairment had significantly higher than expected mortality from diseases related to the respiratory and cardio–vascular organs.

The objective of this chapter is to examine the effect of gold mining dust exposure and tobacco smoking on mortality from chronic obstructive lung disease (COLD), and to answer the following questions:

(1) Are silica dust and tobacco smoking risk factors for mortality from COLD ?

(2) What are the shapes of the dose response curves ?

(3) Is the combined effect of silica dust and tobacco smoking additive or synergistic ?

(4) What is the relative contribution of silica dust and tobacco smoking to the risk of death from COLD ?

The mortality data were obtained on the 2209 white gold miners who were studied in 1968–71 for COLD. The miners who were 45–54 years of age in 1968–71

were followed up to 31 December 1986, when their vital status was established. There were 945 deaths, of which 66 (7.0%) were evidently from COLD, according to the best available information. The prime purpose of collecting the mortality data was to establish whether death from COLD is related to mucus hypersecretion after adjustment for chronic obstructive lung disease (Wiles and Hnizdo, 1991).

The dose–response relation between age–specific mortality risks and rates for COLD and the exposures factors were evaluated by means of the general relative risk model (see Chapter 3), using:

1) a nested age matched case–referent study design and the conditional maximum likelihood method of analysis, and

2) the cohort data and the Poisson regression.

Both methods of analysis are supposed to give the same results of measures of association (Pearce *et al.*, 1988). In the present study, the age matched case–referent study was executed first in order to select the exposure variables most strongly associated with the outcome (Breslow and Day, 1987), and then the measures of association were calculated applying both methods, in order to check that the application of the two methods was correct.

6.2 **Study subjects and methods**

All white gold miners who were included in the 1968–71 study of COLD were followed up to 30 December 1986 when their vital status was established. The selection criteria for the 1968–71 study and description of the subjects was given in Chapter 4, and here only the additional information relevant to the mortality study are given.

The vital status of the cohort was established from the records of the

Provident Fund (a gold miners pension fund), from the medical files and from the Department of Interior. A list of names and dates of birth of all the 2209 miners was sent to the Provident Fund and to the Department of Interior, where the miners vital status was established. The Department of Interior keeps record of all deaths that have occurred in the Republic of South Africa. The copies of the death registration certificates were obtained from the Department of Interior. Of the 945 miners who died during the follow–up period, 794 (84.0%) had an autopsy examination performed at the National Center for Occupational Health (NCOH) for medico–legal purposes. Medical files were available on all of the dead miners. The cause of death was established from the best available information by an experienced specialist physician who was not aware of the purpose of the study, and did not consider the exposure to dust or smoking habits when making the diagnosis. The causes of death were encoded in accordance with the 9th revision of the WHO International Classification of Disease (ICD) (1977). COLD was coded as present when the person evidently died of bronchitis (ICD=491, n=1), emphysema (ICD=492, n=15) or chronic airways obstruction (ICD=496, n=50).

Mining exposure

The detailed description of the method for estimation of the gold mining dust exposure was given in Section 4.3.1. In a preliminary analysis of the case–referent data, it was observed that the cumulative dust exposure to the beginning of the mortality follow–up (1968–71) was higher for the cases, and that thereafter the dust exposure for the controls became higher. Since the original coding of the dust exposure was done in such a way that the exposure for each decade – namely, the 1940s, 1950s, 1960s and 1970s – could be calculated, the cumulative dust exposures for the individual decades as well as the total cumulative dust exposure up to the start of follow–up were considered in this analysis. .

Smoking habits

Tobacco consumption was evaluated in terms of number of cigarettes smoked per day, years of cigarette smoking (CYSM), cigarette pack–years (CPYR), cigarette equivalent pack–years (CEPY), and as categorized smoking variable SMOKEC with categories 0 to 4 for non–smoker, ex–smoker, and current smoker (<20, 20–39 and ≥40 pack–years), see Section 4.3.1 for details.

Lung function tests and respiratory symptoms and signs

The respiratory measurements collected during the 1968–71 and used in the present study were: Vital Capacity (VC), Forced Expiratory Volume in one second (FEV_1), and Forced mid–Expiratory Flow (FEF_{25-75}), assessment of the presence of simple chronic bronchitis by means of a questionnaire and measured mucus, and the presence of rhonchi.

Relative risk function applied

The general relative risk model proposed by Breslow and Storer, Eq. 3.9, was fitted to the mortality data as

$$\log(RR) = \{[1+\beta_1\cdot(\text{dust})+\beta_2\cdot(\text{smoking})]^{\lambda}-1\}/\lambda, \quad \text{if } \lambda\neq 0 \qquad (6.1)$$

$$= \log\{1+\beta_1\cdot(\text{dust})+\beta_2\cdot(\text{smoking})\}, \quad \text{if } \lambda=0$$

6.3 Nested case–referent study

The purpose of the nested case–referent study was to examine the relations between the age–specific mortality risks from COLD and the individual exposures, dust and smoking, and their combined effect. For each case of a miner who died of COLD, a set of six referents was selected at random from those miners who were born closest to the date of birth of the case (within one year) and had survived the case in age (Liddell *et al.*, 1977; Whittemore and McMillan, 1983).

Characteristics of the cases and referents

Table 6.1 shows the case–referent comparison for personal characteristics, exposure to dust, smoking habits, lung function tests, and respiratory signs and symptoms, as measured during the 1968–71 examination. The case–referent differences were compared by the T–test (the application is justified by using the results for exploratory purposes only; the results from the multifactorial analyses are used for making statistical inferences). The cases had lower body weight at the 1968–71 examination than the referents. The average body weights at the initial examination at around 20 years of age were, however, similar for the two groups (± 70 kg), and subsequently the referents increased their weight at a higher rate than the cases. The cases had lower lung function tests (FEV_1, FVC, FEF_{25-75}) and higher prevalence of respiratory symptoms and signs (simple chronic bronchitis, measured mucus≥2ml, and rhonchi) than the referents at the time of the 1968–71 examination.

Table 6.1 Comparison of the COLD cases and referents

Characteristic	Cases	Referents	P–value
	mean ± S.D. (n=66)	mean ± S.D. (n=396)	
Year of birth	1918.3 ± 2.34	1918.3 ± 2.34	0.99
Start of follow–up	1969.2 ± 1.12	1969.3 ± 1.07	0.99
Body weight (kg)	74.3 ± 11.3	81.4 ± 11.7	0.0001
Body height (cm)	175.4 ± 6.1	175.5 ± 5.7	0.97
Age at death	61.8 ± 5.0		
Gold mining exposure			
(cumulative dust exposure (RSA x years))			
to 1949	11250 ± 4983	9735 ± 5043	0.024
1950–1959	11611 ± 4134	10545 ± 4542	0.075
1960–1969	9566 ± 4192	8891 ± 5165	0.315
1969–start[a]	259 ± 543	280 ± 562	0.770
start–1986	2630 ± 4043	4015 ± 6637	0.022
(years)			
to 1949	8.1 ± 3.1	7.6 ± 3.0	0.21
1950–1959	8.7 ± 1.9	8.6 ± 1.9	0.84
1960–1969	7.5 ± 2.4	7.4 ± 2.7	0.80
1969–start	0.2 ± 0.5	0.2 ± 0.4	0.97
start–1986	2.2 ± 3.0	3.2 ± 2.9	0.012
Smoking habits			
Cigarette(pack/day)	1.7 ± 1.4	1.6 ± 1.5	0.80
Cigarette(years)	28.5 ± 6.7	21.6 ± 11.7	0.0001
Cigarette(pack–yrs)	26.5 ± 15.4	19.8 ± 17.3	0.0004
Cig.equiv(pack–yrs)	31.3 ± 16.3	23.6 ± 18.2	0.002
Never smoked No.(%)	0 (0.0)	53 (13.4)	
Ex–smoker No.(%)	11 (16.7)	89 (22.5)	
Current smk. No.(%)	55 (83.3)	254 (64.1)	0.0001[b]
Lung function tests			
FEV1	1.98 ± 0.83	3.10 ± 0.72	0.0001
VC	3.98 ± 0.85	4.40 ± 0.72	0.0001
FEF_{25-75}	1.09 ± 0.77	2.75 ± 1.30	0.0001
Respiratory symptoms and signs, No. (%)			
With rhonchi	31 (47.0)	89 (22.5)	0.0001
With bronchitis	38 (57.6)	147 (31.8)	0.01
Sputum >=2ml	34 (51.5)	152 (38.6)[c]	0.04

[a] Start of the follow–up study.
[b] Mantel–Haenszel χ^2 test for trend (Mantel, 1963).
[c] 2 subjects did not have sputum measured

The dust exposure pattern indicates that the cases had higher exposure to dust up to the end of the 1960s and thereafter the exposure to dust for the referents became higher. The actual years of exposure were not significantly different.

Selection of best predictor of COLD from dust variables

In order to establish the cumulative dust exposure pattern that was most strongly associated with the mortality from COLD, the cumulative dust exposures (RSA x years / 1000) from the start of the gold mining exposure to the end of 1949 (CD40), to 1959 (CD50), to 1969 (CD60) and the total to the start of follow–up (1968–71) (CDTS) were examined individually in the conditional logistic regression model in which the effect of smoking was adjusted for using cigarette years of smoking. Table 6.2 shows the goodness–of–fit statistic (–2logL) for the models with the individual dust exposure variables and CYSM. The best predictor of dust was CD40 (X^2_{LR}=5.9; df=1; p=0.015) and this variable was used to assess the combined effect of dust and smoking. It was noted that the likelihood ratio for dust did not change by inclusion of smoking into the model, and vice versa.

Selection of best predictor of COLD from smoking variables

In a preliminary analysis of the case–referent data, all the variables available on smoking were related to death from COLD by means of the stepwise logistic model. The smoking variables found to be most strongly related to death from COLD were years of cigarette smoking, the categorized smoking variable (SMOKEC), cigarette equivalent pack–years (this includes pipe smoking), and cigarette pack–years, in that order. The above smoking variables were then included individually in the conditional logistic model in which the effect of dust was adjusted, using CD40, and the best predictor for smoking was selected

according to the goodness–of–fit criteria. Table 6.3 shows the goodness–of–fit statistic for the models with the individual smoking variables and CD40. As the model with years of cigarette smoking had the highest log likelihood ratio, 27.6, this variable is used in the analysis of the combined effect. The years of smoking and the categorized smoking variable (SMOKEC) take into account whether a person stopped smoking or not, and this may be the reason why these variables were more strongly related to death from COLD than cigarette or cigarette equivalent pack–years.

Table 6.2 Goodness–of–fit statistic for the multiplicative RR model, RR_m, fitted with different cumulative dust variables, adjusted for smoking (CYSM)

Variables fitted	–2logL	$X^2_{LR(1)}$	P–value
CYSM	229.3	27.5	<0.001
CYSM, CD40	223.4	5.9	0.015
CYSM, CD50	224.2	5.2	0.023
CYSM, CD60	226.2	3.1	0.078
CYSM, CDTS	226.7	2.8	0.096

Table 6.3 Goodness–of–fit statistic for the multiplicative RR model, RR_m, fitted with different smoking variables, adjusted for dust (CD40)

Variables fitted	–2logL	$X^2_{LR(1)}$	P–value
CD40	251.0	5.9	0.015
CD40, CYSM	223.4	27.6	<0.001
CD40, SMOKEC	236.0	15.0	<0.001
CD40, CEPY	241.1	9.9	0.002
CD40, CPY	242.7	8.2	0.004

Dose–response trend for individual exposures

Table 6.4 shows the two way frequency distribution of the cases and controls for categorized cumulative dust exposure (RSA x years / 1000) to 1949 (CD40C) (<=5, >5–10, >10–15, 15+), and categorized years of cigarette smoking (CYSMC) (<10, 10–19, 20–29, 30+). The mean values, given in Table 6.4 for each range, were used in the analyses to estimate the shape of the dose–response curve (Berry, 1980). The marginal observed odds ratios (ORs) display a significant increasing trend for cumulative dust ($X^2_{(1)}$=8.9, p=0.003) and for years of cigarette smoking ($X^2_{(1)}$=22.2, p<0.0001). The age adjusted marginal ORs, estimated from the conditional logistic regression by means of (0,1) indicator variables, also confirmed a significant dose–response relation between death from COLD and the individual exposure variables.

The relative risk model, Eq 6.1, was applied to assess the shape of the dose–response curves for dust and smoking individually (Moolgavkar *et al.*, 1987). Table 6.5 shows the goodness–of–fit statistic and the parameter estimates for the models with changing value of λ, for the categorized and continuous dust variables. For both dust variables CD40C and CD40 the maximum log–likelihood displayed an insignificant change when λ varied from 0 to 1.0. The chi–square test of significance for the β coefficients, $X^2_{(1)}$, is also shown in Table 6.5, and it is interesting to note that as $\lambda \rightarrow 0$, the chi–square value decreases. Table 6.4 shows that the estimated ORs have a large error attached, and consequently all the models fitted equally well to the data.

It should be pointed out that in order to obtain the relative risk of unity at the lowest dust exposure, the value of 3.991 was subtracted from the fitted exposures (CD40C), as otherwise the additive relative risk model would estimate the relative risk values incorrectly.

Table 6.4 Two–way frequency table for death from COLD, by categorized cumulative dust to 1949 and years of cigarette smoking

Dust[a] (CD40C)	Death COLD	Years of cigarette smoking (CYSM) 2	16	26	34	Total n	OR[b] (95%CI)
3.99	yes	0	0	3	2	5	1.0
	no	19	9	41	11	80	
	(OR)[c]	(1.0)	–	(3.3)	(8.5)	(1.0)	
7.44	yes	1	2	9	8	20	2.5
	no	21	15	58	37	131	(0.9;6.9)
	(OR)[c]	(2.7)	(6.3)	(6.3)	(8.8)	(2.4)	
12.35	yes	0	2	12	9	23	3.3
	no	30	18	47	26	121	(1.2;9.4)
	(OR)[c]	–	(5.3)	(10)	(14)	(3.0)	
17.58	yes	0	1	7	10	18	5.3
	no	14	8	23	19	64	(1.8;15.9)
	(OR)[c]	–	(6.9)	(13)	(21)	(4.5)	
Total	Yes	1	5	31	29	66	
	no	84	50	169	93	396	
	(OR)[c]	(1.0)	(8.4)	(15.4)	(26.2)		
OR[b] (95%CI)		1.0	8.0 (0.9;69.7)	15.8 (2.1;118.0)	32.3 (4.2;248.2)		

[a] Mean cumulative dust exposure (RSA x years /1000).
[b] ORs adjusted for age.
[c] Unadjusted OR; value 0.5 was added to all frequencies.

Table 6.5 Goodness–of–fit statistic, parameter estimate, and X^2–test of significance, estimated from RR model.RR_p, fitted with changing λ, and categorized and continuous dust variables

λ	–2logL	β ± S.E	$X^2_{(1)}$
a) for categorized cumulative dust (CD40C)			
1.0	247.2	0.0976 ± 0.0322	9.2
0.5	246.8	0.1396 ± 0.0581	5.8
0.25	246.6	0.1822 ± 0.0914	4.0
0.0	246.3	0.2672 ± 0.1770	2.3
–0.23±.46	246.2	0.4331 ± 0.5710	0.6
b) for continuous cumulative dust (CD40)			
1.0	251.96	0.0699 ± 0.0295	5.6
0.75	250.94	0.0813 ± 0.0385	4.5
0.5	250.98	0.0972 ± 0.0537	3.3
0.25	251.08	0.1192 ± 0.0798	2.2
0.0	251.34	0.1440 ± 0.1260	1.3

Table 6.6 Goodness–of–fit statistic, parameter estimate, and X^2–test of significance, estimated from RR model, RR_p, fitted with changing λ, and categorized and continuous smoking variables

λ	–2logL	β ± S.E	$X^2_{(1)}$
a) for categorized years smoked (CYSMC)			
1.0	226.94	0.0945 ± 0.0215	19.3
0.75	226.70	0.1270 ± 0.0363	12.2
0.5	226.59	0.1940 ± 0.0729	7.1
0.5±0.2	226.58	0.1941 ± 0.0742	6.9
0,25	226.94	0.3650 ± 0.2361	2.4
0.0	228.68	0.7861 ± 0.8062	1.0
b) for continuous years smoked (CYSM)			
1.0	229.34	0.084 ± 0.019	18.8
0.75	229.05	0.113 ± 0.032	12.7
0.5	228.92	0.173 ± 0.067	6.7
0.25	229.34	0.333 ± 0.212	2.5
0.0	231.6	0.633 ± 0.668	0.9

The β coefficient for dust estimated from the univariate multiplicative model (λ=1) shows that a person with cumulative dust exposure of 10 000 has an increased relative risk for COLD of exp(0.0976·10)=2.65 when estimated with CD40C, and exp(0.0699·10) =2.01 when estimated with CD40, see Table 6.5.

Table 6.6 shows the estimated regression coefficients for the categorized years of cigarette smoking (CYSMC), and the continuous variable (CYSM). The best fits were obtained with λ=0.50 and λ=0.51, respectively. For both variable types the model with λ=0.5 gave the best fit, but the likelihood ratio test of λ=0 v.s. λ=1 was not statistically significant.

Combined effect of tobacco smoking and gold mining dust

Table 6.4 shows the risk of dying from COLD for those exposed to various combinations of dust and smoking, relative to those exposed to the lowest dust and smoking level, as estimated by the OR. The calculation was complicated by the zero value in the lowest exposure cell, and thus 0.5 was added to each frequency number. Only one subject died of COLD among those who smoked cigarettes for less than 10 years, and thus there is no observable trend with dust for that smoking category, but an increasing trend with dust occurred for all the other smoking categories, especially for those miners who smoked for 20 or more years. On inspection of the medical file, it was found that the miner who died of COLD and was coded as smoking for less than 10 years was in fact a pipe smoker for 32 years and consumed on average four ounces of tobacco a week; according to the medical file and questionnaire on smoking he did not smoke cigarettes.

The trend in the OR indicates a synergistic effect. The ORs in the two highest categories for dust and smoking had higher values than would be expected from merely adding the individual effects– for example, 21>8.5+1.3–1 and

12.5>3.3+1.3–1. The synergistic effect appears to be present for cumulative dust exposure of more than 10 000 and for years of smoking of 20 years or more.

The combined effect was assessed by fitting equation 6.1 to the data. Table 6.7 shows the log–likelihood value (–2logL), the β coefficients, standard errors and the chi–square test of significance estimated from the RR_P model with changing λ. For the categorized variables best fit was obtained from the model with λ=0.72 (–2logL=217.2), and for the continuous variables the model with λ=0.51 fitted best. The test of H_0:λ=0 versus H_a:λ=1 was significant (X^2_{LR}=229.6–217.4=12.2; df=1; p<0.001 for the categorized data, and X^2_{LR}=230.5–223.4=7.1; p<0.01 for the continuous data), indicating that the combined effect of dust and smoking is synergistic.

Table 6.7 Goodness–of–fit statistic, parameter estimates and X^2–test of significance, estimated from RR model, RR_P, fitted with different λ, and categorized and continuous exposure variables

		Dust		Smoking	
λ	–2logL	$\beta \pm$ S.E	X^2_1	$\beta \pm$ S.E.	X^2_1
a) with categorized exposure variables (CD40C and CYSMC)					
1.0	217.36	0.100 ± 0.033	9.2	0.096 ± 0.022	19.2
0.75	217.22	0.155 ± 0.055	7.9	0.140 ± 0.039	13.2
0.72±0.7	217.22	0.165 ± 0.236	0.5	0.148 ± 0.192	0.6
0.5	217.38	0.313 ± 0.136	5.3	0.264 ± 0.099	7.1
0,25	218.24	1.729 ± 1.463	1.4	1.298 ± 1.062	1.5
0.0	229.64	1.849 ± 1.956	0.9	2.208 ± 2.398	0.8
b) with continuous exposure variables (CD40 and CYSM)					
1.0	223.4	0.071 ± 0.030	5.8	0.084 ± 0.019	18.7
0.75	223.04	0.111 ± 0.050	5.0	0.125 ± 0.035	12.6
0.51	222.78	0.233 ± 0.101	5.3	0.241 ± 0.075	10.3
0.5	222.80	0.230 ± 0.127	3.3	0.244 ± 0.102	5.7
0.25	222.8	1.492 ± 1.578	0.9	1.453 ± 1.185	1.5
0.0	230.5	1.802 ± 2.202	0.8	6.508 ± 6.982	0.8

The additive model with an interaction term.

$$\log(RR)=\log\{1+\beta_1(CD40C)+\beta_2(CYSMC)+\beta_3(CD40C\cdot CYSMC)\}$$

was also fitted to the data. The test of $H_0:\beta_3=0$ versus $H_a:\beta_3\neq 0$ was significant ($X^2_{LR}=207.5-201.3=6.2$; df=1; $p<0.025$), which also supports the hypothesis that the combined effect of the two exposures is synergistic.

6.4 Cohort data analysis

Method of analysis

In the cohort data analysis the effect measure is the mortality rate and the Poisson regression can be used for making statistical inferences about mortality rates (Breslow and Day, 1987). The method of analysis of cohort data used here for the assessment of the individual and combined effects was described by Lubin and Gaffey (1988), Breslow and Day (1988) and in Section 3.1.3 a).

The data were cross classified by five age categories (45–49, 50–54, 55–59, 60–64, and 65–69), five calendar years at risk categories (<1970, 1970–4, 1975–9, 1980–4, and 1984 onwards), four years of smoking categories (CYSMC) (<10, 10–19, 20–29, ≥30), and four categories for cumulative dust exposure (RSA x years /•1000) to the end of 1949 (CD40C) (<5, 5–9, 10–14, ≥15). Using the notation of Breslow and Day, there are I categories for dust, J for smoking, K categories for the nuisance effect of age and L categories for the year at risk effect. For dust i, smoking j, age k and year l, the data in the (i,j,k,l)th cell consist of the observed number of deaths d_{ijkl} from COLD, person–years N_{ijkl}, the mortality rates $h_{ijkl}=d_{ijkl}/N_{ijkl}$, and the mean values for dust and smoking denoted as $\underline{X}_{ijkl}=(X_{1ijkl}, X_{2ijkl})$, where X_{1ijkl} represents the dust variable and X_{2ijkl}

represents the smoking variable.

The expected mortality rate within each ijkl cell, h_{ijkl}, can be estimated using the proportional hazards model as

$$h_{ijkl} = h_{0kl} \cdot RR(\underline{X}_{ijkl}; \beta), \qquad (6.2)$$

where h_{0kl} denotes the age–and–year specific disease rate at RR(0)=1, and $RR(\underline{X}_{ijkl})$ is the relative risk given by a specific risk function, in this case equation 6.1. The equation 6.1 is estimated by the maximum log–likelihood function (Breslow and Day, 1987)

$$L(h) = \sum_i \sum_j \sum_k \sum_l d_{ijkl} \cdot \log h_{ijkl} - N_{ijkl} \cdot h_{ijkl}. \qquad (6.3)$$

The computer program GLIM (Baker and Nelder, 1978), and the general minimization program MINUIT (James and Roos, 1975), were used to fit the models.

The person–years in each (ijkl)th cell and the age and year at risk adjusted SMRs given in Table 6.9 have been calculated according to a method described by Frome and Checkoway (1985), and Pearce and Checkoway (1987).

Results of the analysis

Table 6.8 shows the number of deaths and person–years summarized by the age and year at risk categories. Table 6.9 shows the summary data according to years of cigarette smoking and cumulative dust to the end of 1949. Table 6.9 also shows the number of deaths, person–years, SMRs, and mean values for CYSM and CD40. This table is collapsed over age and year at risk categories; in the final analysis there were 127 cells with non–zero values for person–years. The age and year at risk adjusted SMRs shown in Table 6.9 were calculated as a ratio between the observed and predicted number of deaths estimated from a model which included only the age and year at risk effects. Table 6.9 also presents the marginal RRs

adjusted for age, year at risk and each other exposure, and the 95%CI, as estimated from the Poisson multiplicative model (λ=1). Figure 6.1 shows the SMRs plotted against cumulative dust, according to each category of smoking. The increase in the dose response trend for the two highest smoking categories is apparent.

Table 6.8 Number of deaths from COLD (d) and person–years (P–Y), by age and year at risk

Age at risk		Year at risk					Total
		<70	70–74	75–79	80–84	84+	
<50	d	0	0	0	0	0	0
	P–Y	654	1361	0	0	0	2015
50–54	d	0	6	3	0	0	9
	P_Y	577	5823	1615	0	0	8015
55–59	d	0	4	9	3	0	16
	P–Y	50	2545	5619	1444	0	9658
60–64	d	0	0	3	13	4	20
	P–Y	0	47	2155	4705	917	7824
65+	d	0	0	0	11	10	21
	P–Y	0	0	38	1701	1741	3480
Total	d	0	10	15	27	14	66
	P_Y	1281	9776	9427	7850	2658	30992

Table 6.9 Number of deaths from COLD (d) and person–years (P–Y), by cumulative dust exposure and years of cigarette smoking

Dust (CD40C)		Years of cigarette smoking(CYSM) <10	10–19	20–29	30+	Total (RR)[d] (95%CI)
<5	d	0	0	3	2	5
	P–Y	1528	779	3515	1042	6864
	SMR[a]	0.00	0.00	0.46	0.88	0.37
	CD40[b]	3122	3582	2925	3264	(1.0)
	CYSM[c]	1.1	16.0	26.4	33.1	
5–9	d	1	2	9	8	20
	P–Y	1935	1487	4712	2070	10204
	SMR	0.23	0.68	0.97	1.72	0.94
	CD40	7454	7554	7141	7866	(2.49)
	CYSM	1.3	16.7	26.2	33.4	(0.9–6.7)
10–14	d	0	2	12	9	23
	P–Y	2024	1387	4048	2117	9576
	SMR	0.00	0.67	1.51	1.92	1.14
	CD40	12242	12447	12219	12234	(3.12)
	CYSM	2.1	16.5	26.5	32.9	(1.2–8.3)
>14	d	0	1	7	10	18
	P–Y	971	712	1573	1092	4348
	SMR	0.00	0.59	2.04	4.25	1.82
	CD40	17379	18886	17555	17944	(5.30)
	CYSM	2.7	15.7	26.0	33.6	(2.0–14)
Total	d	1	5	31	29	66
	P–Y	6458	4365	13848	6321	30992
	SMR	0.07	0.54	1.14	2.08	
(RR)[d]		(1.0)	(7.29)	(16.15)	(29.18)	
(95%CI)			(0.9–62)	(2.3–118)	(4.0–214)	

[a] SMR adjusted for the effect of age and year at risk.
[b] Mean values for cumulative dust exposure (RSA x years/1000) before 1950.
[c] Mean values for years of cigarette smoking to 1970 examination.
[d] RR adjusted for age, year at risk, and the other exposure.

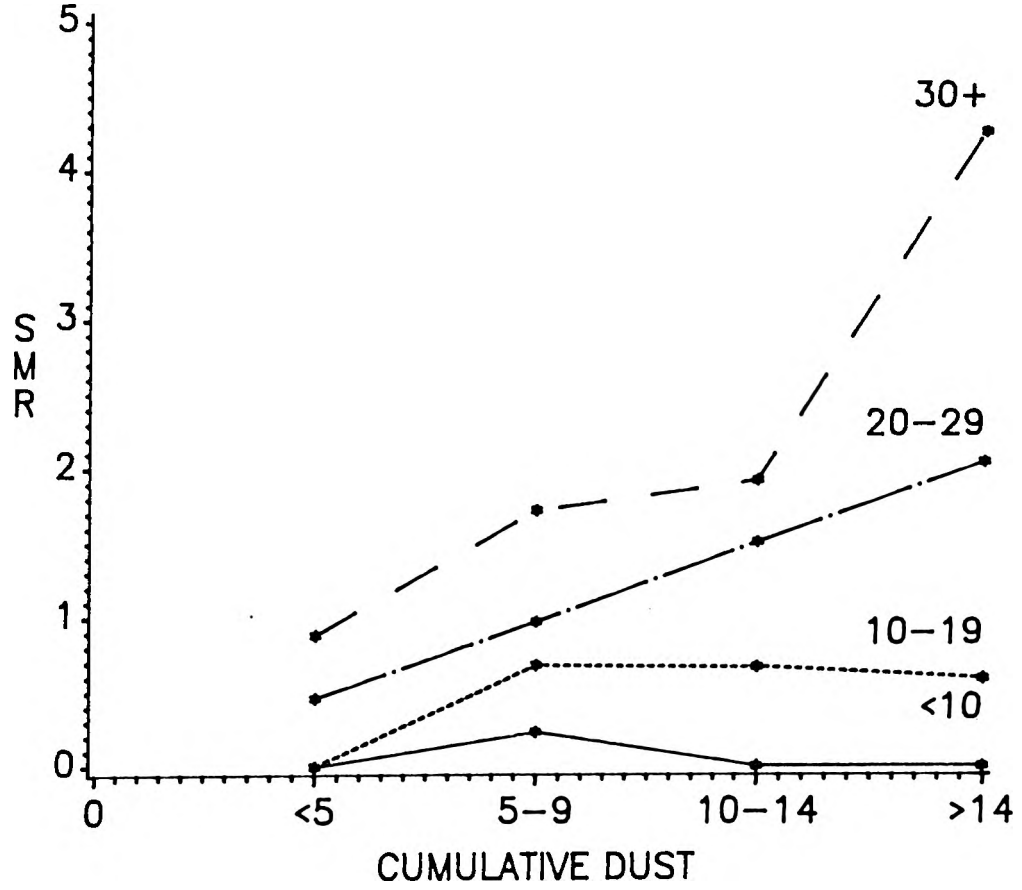


Fig 6.1 SMRs for death from COLD by cumulative dust exposure, and by years of smoking

Table 6.10 shows for the individual dust and smoking categories the estimated regression coefficient, the associated standard error, relative risk and 95% confidence limit, adjusted for the age and year effect, as estimated from the multiplicative model with the program GLIM. It can be noted that the relative risks estimated by the Poisson regression are in a good agreement with the results from the nested case–referent study.

Table 6.10 Estimated regression coefficients for dust and smoking, adjusted for age and year, estimated from the multiplicative RR model, by Poisson regression, and using the program GLIM

Factor	level	Regression coeff. ±Standard Error	Relative (95%CI) Risk
CD40C	<5		
	5–9	0.91 ± 0.50	2.49 (0.9; 6.7)
	10–14	1.14 ± 0.49	3.12 (1.2; 8.3)
	15+	1.67 ± 0.51	5.30 (2.0;14.4)
CYSMC	<10		
	10–19	1.99 ± 1.10	7.29 (0.9; 62.4)
	20–29	2.78 ± 1.02	16.15 (2.2;118.3)
	30+	3.37 ± 1.02	29.18 (4.0;214.1)

In the next step the program MINUIT was used to fit the general relative model (Eq. 6.2) using the maximum log–likelihood function (Eq. 6.3). The parameter estimates obtained from GLIM and MINUIT, for the multiplicative model, i.e. λ=1, are shown in Table 6.11. These estimates are almost identical, but the errors obtained for the intercept from MINUIT are smaller.

Table 6.11 Comparison of parameter estimates from programs GLIM and MINUIT obtained from fitting the mortality data

Factor	Poisson regression estimates form	
	GLIM	MINUIT
	$\beta \pm$ S.E.	$\beta \pm$ S.E.
Intercept	18.68 ± 2.03	18.70 ± 0.85
AGEDC	0.022 ± 0.034	0.022 ± 0.018
YEARDC	0.106 ± 0.037	0.106 ± 0.027
Dust	$9.257 \times 10^{-5} \pm 2.591 \times 10^{-5}$	$9.264 \times 10^{-5} \pm 2.543 \times 10^{-5}$
CYSMC	$9.473 \times 10^{-2} \pm 1.971 \times 10^{-2}$	$9.503 \times 10^{-2} \pm 1.218 \times 10^{-2}$
Deviance (df=122)	37.09	37.09

Dust= cumulative dust exposure before 1950 (RSA x years).

Table 6.12 shows the maximum log–likelihood (–2logL), deviance (Breslow and Day, 1987), and the estimates for CD40C and CYSMC for a given value of λ. The log–likelihoods indicate that for λ=1 to λ=0.6 the fit to the data was equally good, but from λ=0.5 to λ=0 the goodness–of–fit became gradually worse. The test of H_0:λ=0 versus H_a:λ=1 was significant (X^2_{LR}=883.82–867.24=16.6; df=1; p<0.001), indicating that the multiplicative model fits the data significantly better than the additive model, and that there is a synergistic effect between dust and smoking on mortality from COLD.

The RR's estimated by the Poisson regression are in good agreement with the results from the case–referent study. The standardized RR for dust estimated from the multiplicative model (λ=1) was estimated as exp(0.0926, SE 0.0254)=1.097 (95% CI,1.044–1.153) for a unit of cumulative dust exposure of 1000. Thus a person with cumulative dust exposure of 10 000 during the 1940s has a 2.5 times higher risk of dying from COLD than the person with the lowest dust exposure; for a person with cumulative dust exposure of 20 000 the risk is 6.4 times

higher. This probably applies to smokers only, however, as all the cases were smokers for 10 years or more.

Using the data from Table 6.9 the attributable risk due to each exposure was calculated according to the method described in Section 3.8. The attributable risks were estimated as 64% for dust, 93% for smoking and 98% for dust and smoking combined. Thus the approximate estimate of the proportion of cases caused by dust only is 5% (98%–93%), by smoking only 34% (98%–64%), and by the combined exposure 59% (98%–34%–5%).

Table 6.12 Parameter estimates and goodness–of–fit statistic estimated from the RR model, RR_p, fitted with different λ

λ	Cumulative dust[a]	Years of smoking	−2logL Dev.
	β ± S.E	β ± S.E.	
1.0	$9.26\times10^{-5}\pm2.54\times10^{-5}$	$9.50\times10^{-2}\pm1.22\times10^{-2}$	867.2 37.1
0.9	$1.02\times10^{-4}\pm2.44\times10^{-5}$	$1.04\times10^{-1}\pm1.14\times10^{-2}$	867.4 37.2
0.8	$1.20\times10^{-4}\pm2.56\times10^{-5}$	$1.22\times10^{-1}\pm1.46\times10^{-2}$	867.3 37.1
0.7	$1.17\times10^{-4}\pm4.96\times10^{-5}$	$1.63\times10^{-1}\pm3.87\times10^{-1}$	867.4 37.2
0.6	$2.28\times10^{-4}\pm7.77\times10^{-5}$	$2.19\times10^{-1}\pm6.42\times10^{-2}$	867.4 37.2
0.5	$3.28\times10^{-4}\pm1.10\times10^{-4}$	$3.03\times10^{-1}\pm8.65\times10^{-2}$	867.7 37.5
0.4	$5.84\times10^{-4}\pm2.85\times10^{-4}$	$5.38\times10^{-1}\pm2.30\times10^{-1}$	867.7 37.5
0.3	$1.54\times10^{-3}\pm9.75\times10^{-4}$	1.378±0.910	867.9 37.8
0.2	$7.09\times10^{-3}\pm1.62\times10^{-3}$	6.125±2.594	868.6 38.4
0.1	0.691±0.216	633.7±334.4	869.1 38.9
0.0	$5.70\times10^{-4}\pm1.54\times10^{-3}$	4.010±3.869	883.8 53.7

[a] Cumulative dust exposure (RSA x years)

Table 6.13 shows the observed and predicted death rates per 1000 for COLD, for for different value of λ. It can be observed that the predicted death rates obtained from the multiplicative model are closer to the observed values than those obtained from the additive relative risk model. Figure 6.2 shows the observed death rates according to cumulative dust exposure and years of cigarette smoking. Figures 6.3 to 6.6 show the predicted values estimated for different values of λ.

Table 6.13 Observed death rates per 1000, h, and deaths rates predicted from RR model, RR_p, for given λ

Yrs smoked	Dust[a]	P–Y[b]	Obs. h	Estimated death rates for given λ				
				0.0	0.2	0.5	0.6	1.0
<10	<5	1528	0.0	0.21	0.02	0.05	0.06	0.12
	5–9	1935	0.5	0.29	0.09	0.12	0.12	0.18
	10–14	2024	0.0	0.40	0.26	0.26	0.25	0.29
	15+	971	0.0	0.58	0.68	0.59	0.54	0.53
10–19	<5	779	0.0	1.53	0.40	0.42	0.41	0.45
	5–9	1487	1.3	1.61	0.74	0.71	0.68	0.69
	10–14	1387	1.4	1.77	1.37	1.24	1.18	1.11
	15+	712	1.4	1.90	2.37	2.11	2.01	1.83
20–29	<5	3515	0.9	2.42	1.18	1.19	1.16	1.16
	5–9	4712	1.9	2.52	1.85	1.81	1.76	1.71
	10–14	4048	3.0	2.59	2.93	2.86	2.80	2.69
	15+	1573	4.5	2.75	4.50	4.46	4.43	4.28
30+	<5	1042	1.9	3.25	2.36	2.35	2.36	2.36
	5–9	2070	3.9	3.36	3.43	3.43	3.47	3.47
	10–14	2117	4.3	3.37	4.83	4.97	5.06	5.15
	15+	1092	9.2	3.54	7.43	8.14	8.48	9.07
Total		30992	2.13					
Deviance				53.68	38.42	37.52	37.20	37.09[c]

[a] Cumulative dust to the end of 1949 (RSA x years / 1000) (CD40C).
[b] Person–years.
[c] df=122.

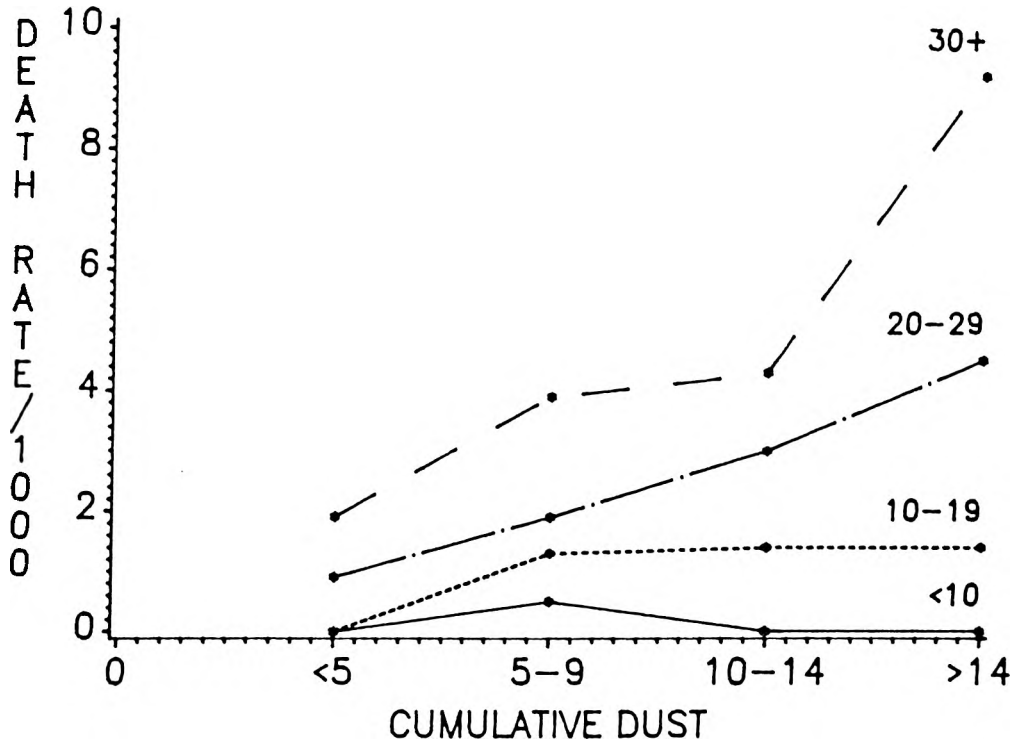


Fig 6.2 Observed death rates per 1000 person–years for COLD, by cumulative dust exposure and years of cigarette smoking

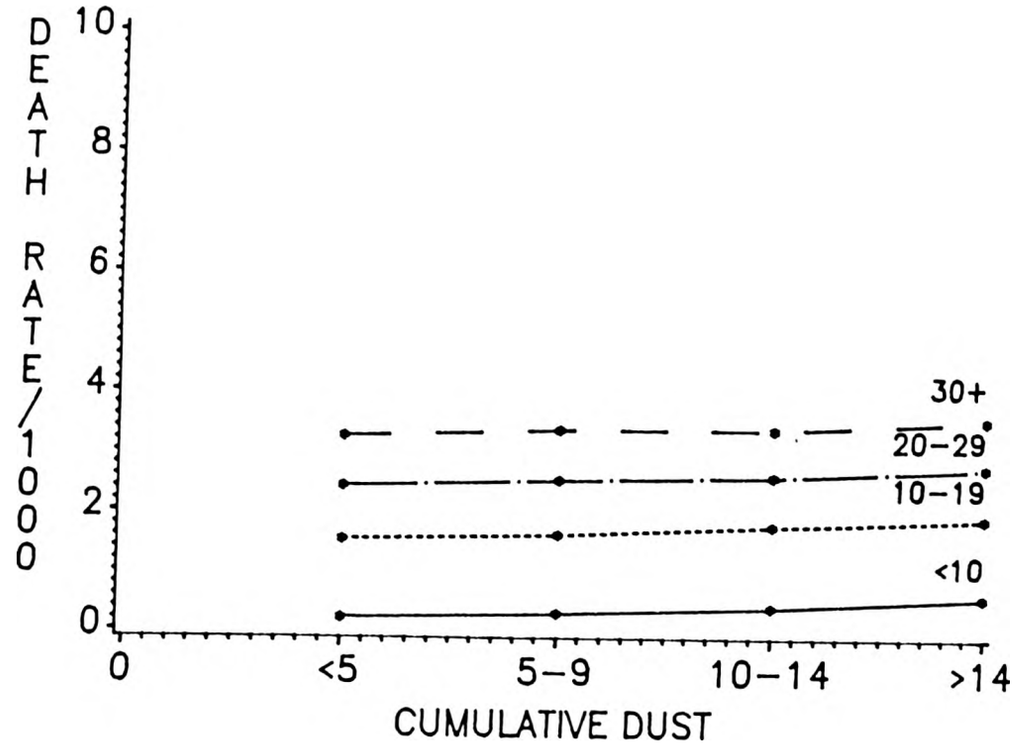


Fig 6.3 Predicted death rates per 1000 person–years for COLD, estimated from the GRRM with $\lambda=0$

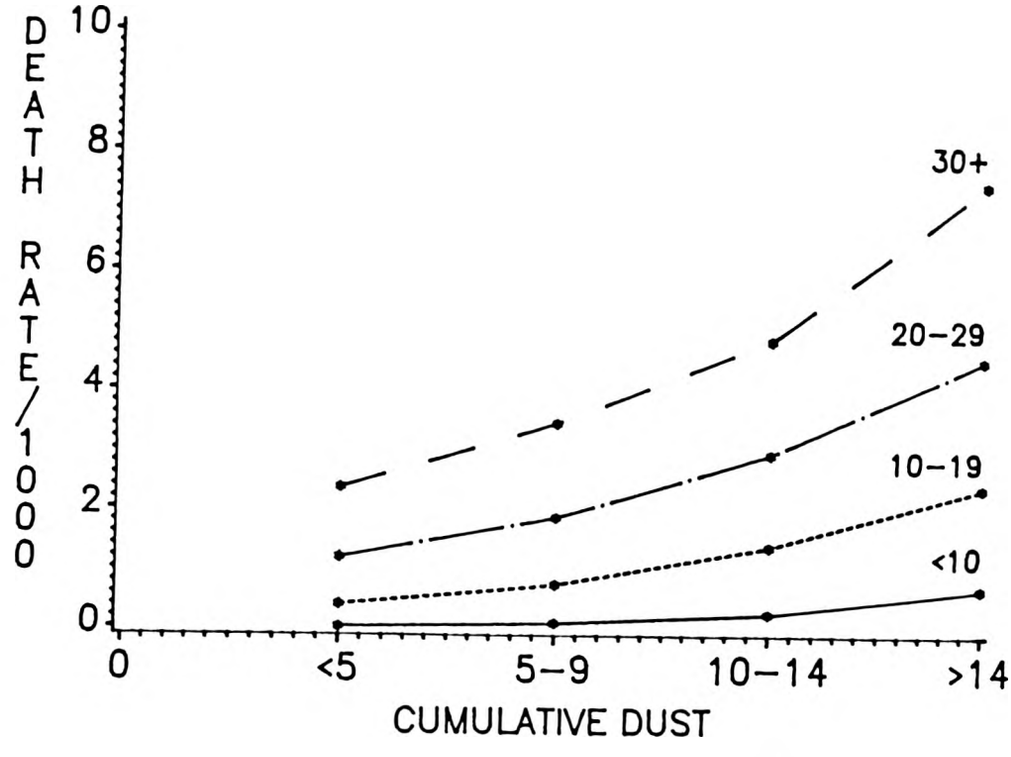


Fig 6.4 Predicted death rates per 1000 person–years for COLD, estimated from the GRRM with $\lambda=0.2$

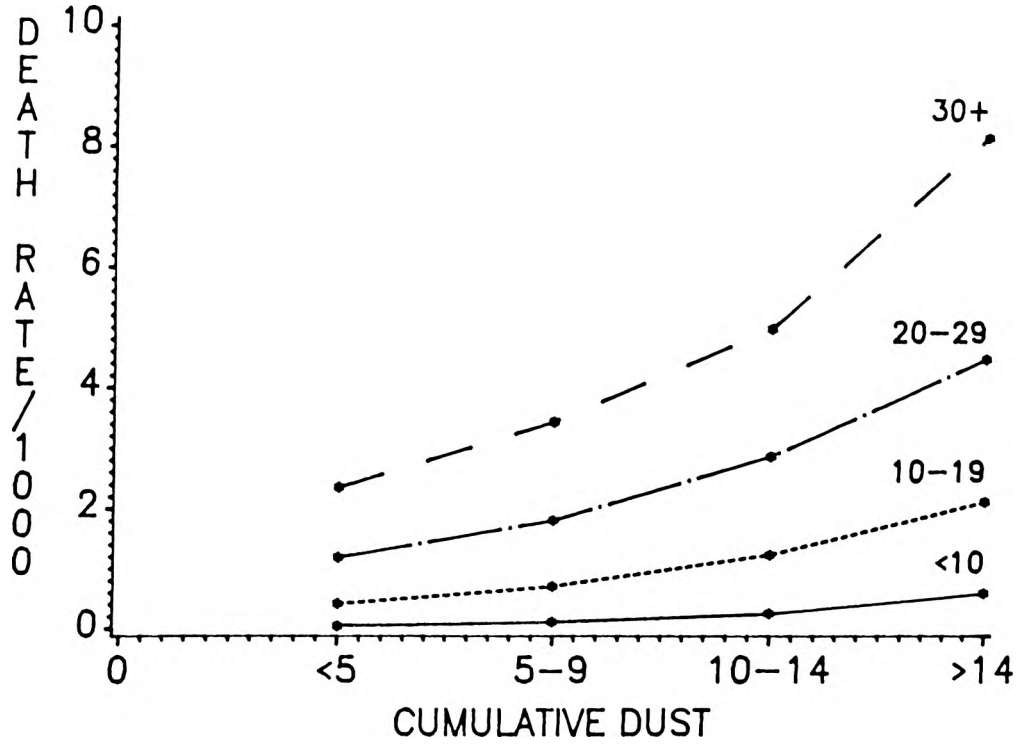


Fig 6.5 Predicted death rates per 1000 person–years for COLD, estimated from the GRRM with λ=0.5

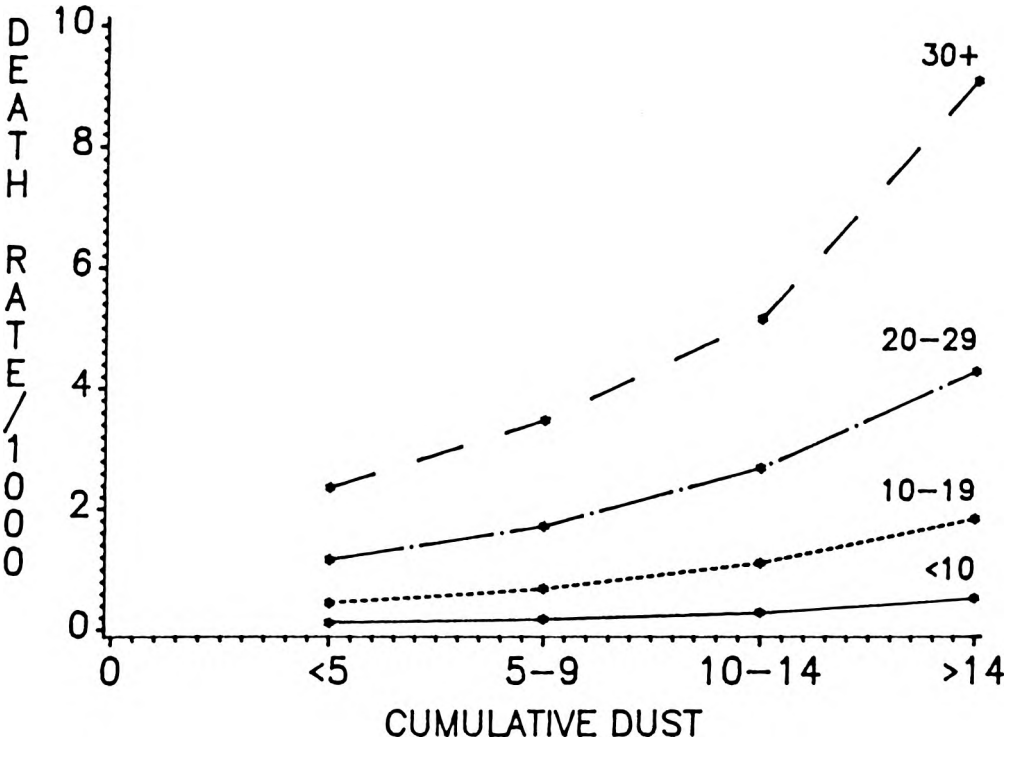


Fig 6.6 Predicted death rates per 1000 person–years for COLD, estimated from the GRRM with λ=1

6.5 Discussion

The epidemiological evidence on the effect of gold mining dust exposure on mortality from COLD is rather inconclusive. Studies on mortality of cohorts exposed to silica have consistently shown increased mortality rates for non–malignant respiratory disease, largely accounted for by silicosis (US DHHS, 1984). Davis *et al.* (1983) did not find an excess mortality from bronchitis in Vermont Granite workers. Finkelstein *et al.* (1987) found non–malignant respiratory disease (8th revision ICD 460–519) to be the most frequent cause of death, with SMR over 600, among Ontario miners receiving compensation for silicosis. In a mortality study of Western Australian gold miners, the overall death rate from non–malignant respiratory disease was slightly less than expected, but the death rate from silicosis was high (Armstrong *et al.*, 1979). In a previous study of mortality in middle–aged South African gold miners (Wyndham *et al.*, 1986), a cohort of 3971 white miners, born between 1 January 1916 and 31 December 1930, who were alive on 1 January 1970 and currently working on the gold mines, was followed up to December 1978. Standardized mortality ratios were calculated using as a reference population the total white male population of the Republic of South Africa. Increased SMR for lung cancer (161.2), chronic respiratory diseases (165.6), and acute and chronic nephritis (381.0) were reported. Of the 530 deaths found, 29 (5.5%) miners had died of chronic respiratory diseases, which included bronchitis, emphysema, and pulmonary heart disease, coded according to best available information drawn from the death certificates, MBOD medical files and autopsy findings. A relative risk for chronic respiratory diseases estimated from a nested case–referent study was 2.48 per unit of 10 000 particle–years ($p<0.03$).

In the present mortality study, 945 (43%) of the gold miners have died during the follow–up period from 1968–71 to 30 December 1986, at an average age

of 60.1 years. The miners who died had significantly higher tobacco consumption measured by cigarette equivalent pack–years ($p<0.0001$), and lower lung function tests (FEV_1, FVC, FEF_{25-75}; $p<0.001$ for all three tests) at the start of the follow–up than the miners who survived. Those who died of COLD, however, had, in comparison with other main groups of causes of death, the lowest average values for the three lung function measurements. This suggests that for the other causes of death the lower lung function may have been a contributory factor, whereas for the COLD deaths, loss of lung function was a main cause of death. Cumulative dust and tobacco smoking were also significantly related to mortality from all causes; thus the effect of combined exposure on mortality from all causes may also be an important factor to consider. Nine miners died of pulmonary tuberculosis and they were found to have the highest exposure to dust in comparison to other groups created according to the main categories of the ICD; no miner died of pneumoconiosis .

The results from the present analysis show a statistically significant dose–response relationship between death from COLD and cumulative dust, after adjustment for smoking. The exposure to dust during the 1940s was most strongly related to death from COLD. These results are in agreement with the findings from the case–referent study of emphysema identified by necropsy, done on South African white gold miners (Becklake *et al.*, 1987). In that study, the number of shifts worked in high dust before 1950 were most strongly related to the presence of emphysema at necropsy in men 50–70 years old who died during 1981. The authors suggested that the heavy exposure in early adulthood may be most important to the subsequent development of emphysema. The same phenomenon was observed in asbestos workers – namely that asbestos exposure received early in a man's working career was a more important determinant of subsequent air–flow limitation than later exposure (Copes *et al.*, 1985). An alternative

explanation is that the dust concentrations in the 1940s were still relatively high in comparison with those found later, even though the concentrations in the 1940s were reportedly equal to those in the 1960s (Beadle *et al.*, 1971; Page–Shipp and Harris, 1972). Becklake *et al.* (1987) have also observed that those who had emphysema at death were slightly taller, but weighed less when they entered the mining industry than those who did not develop emphysema. This finding was not confirmed by our study; where the COLD cases and referents were approximately of the same height and weight at 20 years of age (70 kg). The subsequent 5 yearly increases in weight were, however, systematically higher in the referents than in the cases.

In an attempt to find out whether the cases were of a slimmer build than the referents, the radiograms taken in 1960 and 1970 were measured blindly by an independent person. The findings indicated statistically significant differences in all the measurements of lung height (cm). The cases had significantly longer lungs in 1960 (24.5 cm standard deviation (SD), 1.9), and in 1970 (25.2 cm SD, 2.1) in comparison with the referents whose respective values were (22.6 cm SD, 2.4 and 22.6 cm SD, 2.3). The height of the lung increased significantly for the cases but remained unchanged for the referents. The left and right lungs gave the same results. This result indicates that by 1960 the lungs of the cases had already started to expand as a consequence of emphysema, and continued to do so in the 1970s; this corresponded with their increased disability and retirement from dusty occupations during the late 1960s and 1970s.

The relative risk for cumulative dust before 1950, standardized for years of cigarette smoking and age, estimated from the case–referent data by the logistic model, was exp(0.1003)=1.106 with 95%CI(1.104;1.179) for one unit of cumulative dust of 1000. The estimate obtained from the cohort study was exp(0.0926)=1.097 with 95%CI(1.044;1.153). Thus, for a man exposed to cumulative dust of 17 500

during the 1940s, the expected risk that he died of COLD when he was over 50 years of age was 5.06 timeshigher than for a man who was exposed to cumulative dust of less than 5 000 during the same period. As all the miners who died of COLD were smokers, however, the relative risk for dust most likely applies to smokers only.

For both study designs the data have indicated that the combined effect of dust and smoking is more than additive. The general relative risk model indicated that the combined effect of the two exposures may be closer to the multiplicative model than the additive model. For the case–referent data, the mixture parameter λ for the best fitting model was estimated to be 0.72 for the categorized data, and 0.51 for the continuous data. For the cohort data the parameter λ ranged from 0.6 to 1.0 for the best fitting models.

The results from the general relative risk model are in line with the results presented in Tables 6.4 and 6.9, which show that the ORs and SMR's for those who smoked over 20 years, and had cumulative dust exposure of more than 10 000, were substantially higher than would be expected from adding the individual effects. Figure 6.1 shows that the increasing trend with dust for the age and year–at–risk adjusted SMRs was higher for the smoking categories 20–29 and $\geq$30. Only one miner died of COLD who did not smoke cigarettes. This miner was, however, a pipe smoker of four ounces of tobacco a week over 32 years.

The relative importance of each exposure factor is given by the estimates of the attributable risk due to each exposure. According to those estimates, about 5% of the cases were due to the effect of dust only, 34% due to smoking only, and 59% due to the combined effect of dust and smoking. Thus tobacco smoking is more dangerous for workers exposed to silica dust than for unexposed people, as silica dust and tobacco smoking act synergistically in causation of obstructive lung disease.

Possible biases involved in this study may be related to the original selection criteria. One of the selection criteria for this study was a minimum of 10 years of underground gold mining exposure. Thus miners who may have been potentially sensitive to dust, to the extent that they could die from COLD owing to the effect of dust only, may have withdrawn from the dusty underground occupations earlier. Therefore, the healthy worker effect may be the reason why there were no deaths from COLD in those who smoked for less than ten years. The possibility that the cases of COLD were subjects who came for a benefit examination in 1970 was also considered. Of the 66 cases, 59% came for a periodical examination and were still working in the mine at the time of the 1970 examination, 3% came for a periodical examination but were no longer working in the mine, 27% came for a benefit examination (for compensation purposes) but were still working in the mine, and 11% came for a benefit examination and were no longer working in the mine. Thus the distribution of the number of deaths from COLD is approximately proportionally distributed according to the reasons for attending the medical examination.

In conclusion, it appears that dust on its own does not cause disability that results in death from COLD, but smoking on its own and the combined effect of dust and smoking are the main environmental risk factors for death from COLD in the present data. Tobacco smoking is more dangerous for miners exposed to silica dust than for other men, as the two exposures act synergistically in causing chronic obstructive lung disease.

CHAPTER 7
MORTALITY FROM LUNG CANCER

7.1 Objectives of chapter 7

The objective of this chapter is to examine the effect of gold mining dust exposure and tobacco smoking on mortality from lung cancer, and to answer the following questions:

1) Are silica dust and tobacco smoking related to mortality from lung cancer?

2) Is the combined effect of silica dust exposure and tobacco smoking additive or synergistic?

3) What is the relative contribution of silica dust exposure and tobacco smoking to the risk of death from lung cancer?

4) Is there an association between the diagnosis of silicosis and diagnosis of lung cancer at necropsy?

The mortality data on the cohort of 2209 gold miners who were studied in 1968–71 for respiratory impairments and were subsequently followed up to 30 December 1986 are used.

7.2 Study subjects and methods

Study subjects

The details of the follow–up study were given in Section 6.2 and thus only

information relevant to this chapter is given here.

The cohort of 2209 white gold miners was selected in 1968–71, when the miners were 45–54 years old. for a study of respiratory impairment (Wiles and Faure, 1977), and the mortality follow–up was up to 30 December 1986 (Wiles and Hnizdo, 1991). The miners started their exposure in the 1940s and worked on average 23.5 (S.D. 4.9) years in gold mines with minimal exposure in other mines.

During the 17–year follow–up period, 945 (43%) of the miners died. The specialist physician who identified the causes of death, according to the best available information from death certificates, medical files and postmortem reports, reported 81 deaths from lung cancer (ICD=162, 9th revision). These 81 lung cancer deaths were then further scrutinized by the director of the Epidemiology Research Unit, Dr G.K. Sluis–Cremer, and by an experienced NCOH pathologist, who has done extensive research on lung cancer in gold miners, Dr S. Klempman. Four cases were rejected because of insufficient evidence that the cancer originated in the lung. Of the 77 accepted cases, 66 were confirmed by a necropsy, in addition five were proven by a biopsy, and six were identified by a death certificate.

The occupational records of the lung cancer cases were searched to find out if any of the miners had worked as a boilermaker (known to have exposure to asbestos), or had another type of asbestos exposure. Of the 77 miners, only one miner worked for 1.7 years in a platinum mine, the rest worked in gold mines only, and none was found to have worked as a boilermaker. The occupational records indicate such exposures reliably.

Mining exposure

The cumulative dust and the actual years of exposure were calculated for each individual decade, namely the 1940s, 1950s and 1960s, and to the start and to the

end of the follow–up period, see Section 4.3.1.

Smoking habits

Tobacco consumption was evaluated in terms of number of cigarettes/day, years of smoking (cigarettes or/and pipe), cigarette pack–years, and cigarette equivalent pack–years.

Lung function tests and respiratory symptoms and signs

The information collected during the 1968–71 examination and considered in the present analysis included: Forced Vital Capacity (FVC), Forced Expiratory Volume in 1 second (FEV1) and Forced mid–Expiratory Flow ($FEF_{25-75\%}$), assessment of the presence of simple chronic bronchitis by means of a questionnaire and measured mucus volume, and of the presence of rhonchi.

Method of analysis

In Chapter 6, the nested case–referent study was used to select the risk factors most strongly related to the mortality from COLD, and the combined effect of dust and smoking was evaluated by the conditional maximum likelihood method and the Poisson regression. The results from the cohort data analysis were in a good agreement with those obtained from the case–referent study, indicating that both methods of analysis were executed correctly. In the present chapter the following alternative method of analysis was adopted:

1) The Cox's proportional hazards model was applied to select the smoking and gold mining dust exposure variables most strongly related to the death from lung cancer (Cox, 1972). This method of analysis takes into account the individual period of follow–up, which was from the start of the follow–up to the time when the subject either died or the follow–up period ended. SAS computer program

PROC PHGLM (SAS, 1986) was used. The selection procedure involved fitting a set of hierarchical equations to the data. Firstly, the predictive value of each smoking variable was evaluated in a set of models which included age at the time of the start of the follow–up. Next, the predictive value of each dust exposure variable was evaluated in a set of models which included age and the best predictor of smoking. The predictive value of each variable was evaluated by means of the log–likelihood ratio test (X^2_{LR}).

2) Then the Poisson regression approach and the general relative risk model was applied to evaluate the combined effect of dust and smoking on mortality from lung cancer, the analysis was done in the same way as in Chapter 6, see Sections 6.4 and 3.1.3 for the description of the method.

The data were crossclassified by five age categories (45–49, 50–54, 55–59, 60–64, and 65–69), five calendar–year–at–risk categories (<70, 70–74, 75–79, 80–84, and 84+), four cigarette–equivalent pack–years categories (≤10, 11–25, 26–35, 36+), and four categories for cumulative dust exposure to the start of follow up (RSA x years/ 1000) (≤15, 16–30, 31–40, 41+). The death rates, h, within each cell were related to the average dust and smoking level for each cell as described in Section 6.4.

Association between lung cancer and silicosis

The association between death from lung cancer and the presence of silicosis found at necropsy was examined using the 745 miners who had an necropsy examination done by pathologists at NCOH. The necropsy examinations are done for medico–legal purposes, and for every miner the silicosis of the parenchyma, pleura and hilar glands are coded as none, slight, moderate or marked, according to the findings from a macroscopic examination. The diagnosis is then confirmed by a microscopic examination. The histological examination includes tissue from the

upper, middle and lower zones of both lungs and hilar glands. As the silicotic lesions are unique in appearance, the diagnosis is generally considered to be reliable.

The associations between death from lung cancer and the degree of silicosis of the parenchyma, pleura and hilar glands were examined by crude odds ratios and by adjusted odds ratios estimated from the logistic regression in which the effects of age, cumulative dust exposure and smoking were adjusted.

7.3 Selection of best predictors of death from lung cancer

Table 7.1 summarizes data for the cases and the remainder of the cohort separately. The cases had higher cumulative dust exposure for all the time intervals examined, but the differences in the actual years of exposure are not as consistent. As expected, tobacco consumption of the cases was substantially higher; this was indicated by all the smoking variables examined. The cases had lower lung function measurements (FEV_1, FVC, FEF_{25-75}), and a higher prevalence of respiratory symptoms and signs (simple bronchitis and rhonchi) at the time of the 1968–71 examination than the rest of the cohort.

Table 7.2 shows the sequential results of the hierarchical analyses by the Cox's proportional hazards model applied to select best predictors for smoking and dust. Of the smoking factors examined, cigarette pack–years were most strongly related to death from lung cancer. When the lung cancer cases who had zero value for cigarette pack–years were, however, examined, it was observed that some of them had very high cigarette equivalent pack–years due to pipe smoking, and thus the cigarette equivalent pack–years were selected to be used in the Poisson regression analysis. The difference in the log likelihood ratios for the two variables

was, however, very small.

Table 7.1 Comparison of lung cancer cases with rest of the cohort

Characteristic	Cases mean ± S.D. (n=77)	Cohort mean ± S.D. (n=2132)
Age category at the start of follow–up No. (%)		
45–50	33 (42.9%)	1157 (54.3%)
51–55	44 (57.1%)	975 (45.7%)
Gold mining exposure		
(cumulative dust exposure (RSA x years))		
to 1949	11238 ± 5471	9436 ± 5035
1950–1959	11304 ± 4870	10778 ± 4760
1960–1969	10324 ± 5642	8881 ± 4984
1940–start	33228 ±13004	29421 ±11871
1940–1986	38138 ±16066	33365 ±17044
(years)		
to 1949	8.0 ± 3.0	7.3 ± 2.9
1950–1959	8.5 ± 2.2	8.6 ± 2.0
1960–1969	7.5 ± 2.7	7.4 ± 2.7
1940–start	24.4 ± 5.1	23.5 ± 4.9
1940–1986	27.6 ± 7.2	26.4 ± 6.2
Smoking habits		
Cigarette (No./day)	20.6 ±11.3	15.2 ± 10.3
Years of smoking	27.6 ± 9.1	21.7 ± 11.3
Cigarette (pack–years)	30.1 ± 18.4	20.1 ± 16.0
Cig. eq.(pack–years)	34.5 ± 18.2	24.0 ± 17.2
Never smoked No. (%)	3 (3.9)	252 (11.8)
Ex–smoker No. (%)	7 (9.1)	419 (19.7)
Current smoker No. (%)	67 (87.0)	1461 (68.5)
Lung function tests		
FEV1	2.89 ± 0.64	3.13 ± 0.74
VC	4.21 ± 0.70	4.42 ± 0.76
FEF_{25-75}	2.36 ± 1.09	2.76 ± 1.26
Respiratory signs symptoms		
With rhonchi No. (%)	27 (35.1)	472 (22.1)
With bronchitis No. (%)	39 (50.7)	746 (35.0)
Measured sputum >=2ml	41 (53.3)	874* (41.1)

* 3 subjects did not have sputum measured

Table 7.2 Results from the Cox's proportional hazards model applied to select best predictors of lung cancer

Variable entered into the model	Effects included in the model			−2logL	X^2_1
	Dust $\beta \pm$ S.E.	Smoking $\beta \pm$ S.E.	Age $\beta \pm$ S.E.		
Intercept				1142.5	
Age			0.150±0.046	1131.5	11.0‡
Smoking variables[a]					
Cigarette/day		0.046±0.010	0.163±0.046	1111.5	20.0‡
Yrs smoking		0.064±0.015	0.115±0.047	1104.6	26.9‡
Cig.eq. p–yrs		0.032±0.006	0.125±0.046	1101.9	29.6‡
Cig. p–years		0.033±0.006	0.134±0.046	1101.7	29.8‡
Gold mining dust variables[b]					
CDE to 1949	0.062±0.023	0.033±0.006	0.092±0.048	1094.4	7.5†
CDE to 1959	0.031±0.013	0.033±0.006	0.103±0.047	1096.4	5.5*
CDE to start	0.023±0.009	0.033±0.006	0.104±0.046	1095.2	6.7†
CDE to 1986	0.010±0.003	0.033±0.006	0.122±0.046	1096.8	5.1*
Yrs of mining	0.023±0.002	0.032±0.006	0.115±0.047	1101.0	0.9

[a] Adjusted for age at the start of the follow–up, and fitted independently to the data.
[b] Adjusted for age and cigarette equivalent pack–years, and fitted independently to the data.
CDE= cumulative dust exposure (RSA x years / 1000).
‡ $P<0.001$
† $P<0.01$
* $P<0.025$

The effect of gold mining dust exposure, adjusted for age at the start of follow–up and cigarette equivalent pack–years, was examined using cumulative dust exposure to the end of 1949, 1959, to the start of the follow–up, and to the end of follow–up, as well as the actual years of exposure. The cumulative dust exposure for all the periods was significantly related to death from lung cancer, indicating that the lung cancer cases had higher exposure over the whole period of employment, unlike the cases who died from COLD whose dust exposure declined after the 1960s. The regression coefficient for cigarette equivalent pack–years and the contribution in terms of the chi–square value ($\simeq$ 30) changed very little by adding the individual dust variables to the model. The cumulative dust exposure

to the start of the follow–up is used to evaluate the combined effect. By the time of the start of the follow–up, the miners already had on average 24 (SD ± 5.0) years of underground exposure, and the average exposure after the 1968–71 examination was only 3.3 (SD ± 1.8) years. The actual length of the mining exposure was not significant. The correlation coefficient between cumulative dust exposure and cigarette equivalent pack–years estimated on the 2209 miners was –0.036.

According to the results of the Cox's proportional hazards model the adjusted relative risk for lung cancer associated with one unit of cumulative dust exposure (RSA x years / 1000), cumulated to the start of the follow–up period, was 1.023 with 95%CI(1.005; 1.042).

7.4 Assessment of the combined effect of dust and smoking

Table 7.3 shows the cohort data crosstabulated by the categories of cumulative dust exposure and cigarette equivalent pack–years. The table shows the number of deaths from lung cancer, person–years, the mean values for cumulative dust exposure and cigarette equivalent pack–years, and the observed death rates per 1 000 person–years. This table is summarized over the age and year at risk categories, see Table 7.4 for the two way table by age and year at risk. The mean values for dust and smoking calculated for each cell were used to estimate the dose–response trend. In the final analysis there were 272 cells with a non–zero value for person–years.

The marginal adjusted relative risks and 95%CIs indicate an increasing trend for dust and for smoking (Table 7.3). The estimates were adjusted for age, year at risk, and the other exposure. The individual dose–response curves for dust

and smoking were best fitted by the additive model. The differences in the log–likelihood for the additive and the multiplicative model were, however, very small and insignificant.

Table 7.3 Number of deaths (d) from lung cancer and person–years (P–Y), by cumulative dust exposure and cigarette equivalent pack–years

Dust[a]		Cigarette equivalent pack–years				Total	RR[e] (95%CI)
		≤10	11–25	26–35	36+		
≤15	d	0	0	2	2	4	1.00
	P–Y	838	1043	727	614	3222	
	dust[b]	12.20	11.74	12.99	12.11		
	cig.[c]	1.3	18.7	29.6	51.4		
	SMR	0.0	0.0	1.19	1.41		
16–30	d	0	11	9	10	30	1.54
	P–Y	3182	4569	3133	3439	14323	(0.6;4.3)
	dust	22.31	22.77	23.54	21.65		
	cig.	2.6	17.9	30.3	47.1		
	SMR	0.0	1.00	1.20	1.23		
31–40	d	2	3	3	12	20	2.07
	P–Y	2118	2212	1482	1571	7383	(0.7;6.0)
	dust	35.77	34.72	34.74	34.72		
	cig.	2.6	17.7	30.2	47.7		
	SMR	0.37	0.54	0.82	2.96		
41+	d	3	5	4	11	23	2.92
	P–Y	1708	1844	1333	1179	6064	(1.02;8.4)
	dust	46.57	47.69	50.05	48.68		
	cig.	3.2	17.5	29.7	50.0		
	SMR	0.64	1.05	1.19	3.68		
Total	d	5	19	18	35	77	
	P–Y	7846	9668	6675	6803	30992	
RR[e] (95%CI)		1.00	3.33 (1.3;8.9)	4.60 (1.7;12.4)	8.91 (3.5;22.7)		

[a] Cumulative dust to the start of follow–up (RSA x years/1 000).
[b] Mean value for cumulative dust to start of follow–up/1000.
[c] Mean value for cigarette equivalent pack–years.
[d] Mortality rates per 1 000 person–years.
[e] Adjusted RR estimated from the multiplicative model.
P–Y=person–years of follow–up.

Table 7.4 Number of deaths (d) from lung cancer and person–years (P–Y), by age and year at risk

Age at risk		Year at risk					Total
		≤70	70–74	75–79	80–84	84+	
≤50	d	1	1	0	0	0	2
	P–Y	654	1361	0	0	0	2015
50–54	d	0	7	0	0	0	7
	P–Y	577	5823	1615	0	0	8015
55–59	d	0	4	9	3	0	16
	P–Y	50	2545	5619	1444	0	9658
60–64	d	0	0	9	15	4	28
	P–Y	0	47	2155	4705	917	7824
65+	d	0	0	0	15	9	24
	P–Y	0	0	38	1701	1741	3480
Total	d	1	12	18	33	13	77
	P–Y	1281	9776	9427	7850	2658	30992

The SMRs given in Table 7.3 indicate an increasing trend with smoking, which is consistent for all dust exposure levels. The increasing trend with dust is apparent in the lowest and the highest smoking categories, see Figure 7.1. The steep increase in the highest smoking category suggests that dust and smoking act synergistically.

To assess the combined effect of dust and smoking the Breslow and Storer model, Eq. (6.1), was fitted to the data. The best fit was obtained when $\lambda=0.25$, indicating that the combined effect is slightly higher than if the individual effects were additive. The test of $H_0:\lambda=0$ versus $H_a:\lambda\neq0$, i.e. the additive versus non–additive model, was found to be statistically significant for $\lambda=0.25$ ($X^2_1=3.9$, $p<0.05$). Table 7.5 shows the parameter estimates and goodness–of–fit statistic estimated from the general relative risk model, Eq. (6.1) with the changing value of λ. Table 7.6 shows the predicted death rates per 1 000 person years

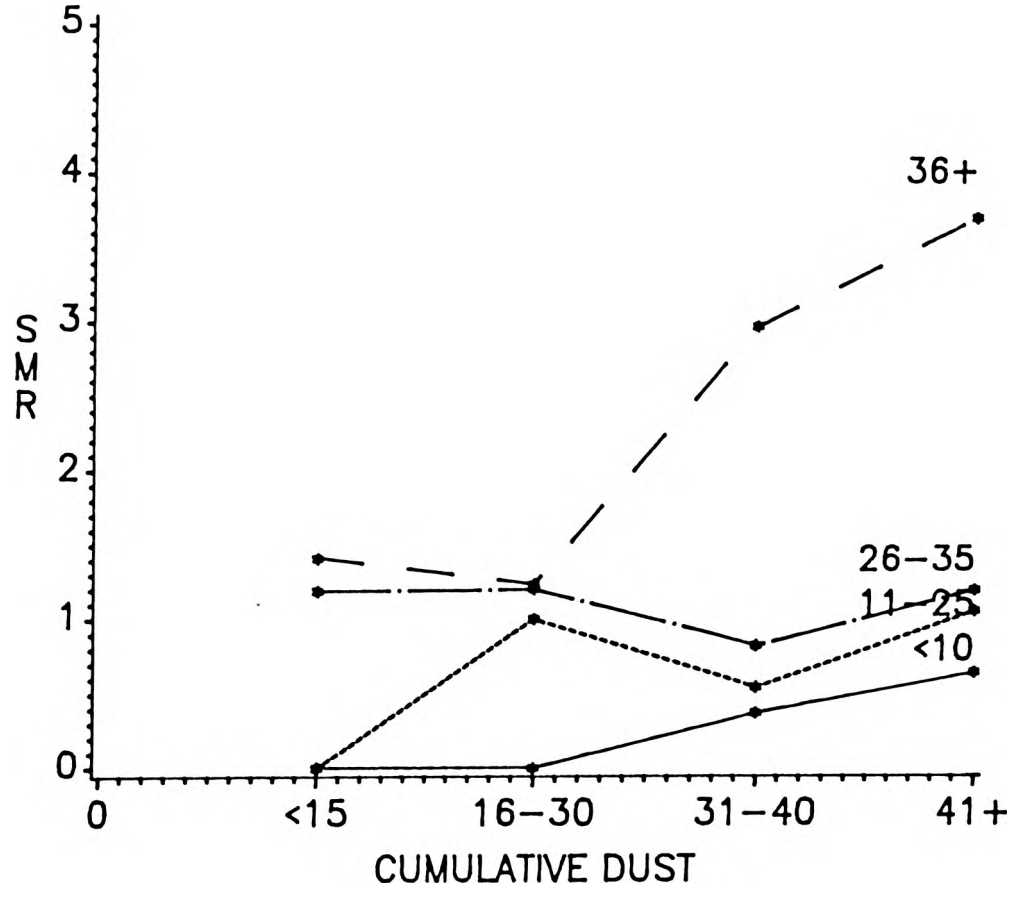


Fig 7.1 SMRs for death from lung cancer by cumulative dust exposure, and by cigarette equivalent pack–years

estimated from the models with different values of λ, and the goodness–of–fit statistic (X^2_{LR}). It can be seen that the estimated values change very gradually between λ=1.0 and λ=0.25, but for λ=0.0, i.e. the additive model, the estimated death rates for the two highest dust exposure categories are underestimated, indicating that the combined effect of dust and smoking is more than additive. As the goodness–of–fit statistic does not indicate a significant difference between the models with λ=0.25 and λ=1.0, we can assume that the combined effect is multiplicative and that the proportionality of hazards applies. Figure 7.2 shows the observed death rates and Figures 7.3 to 7.6 show the predicted death rates for changing λ, plotted against cumulative dust exposure, according to the smoking categories. It can be noted that there is a very gradual increase in the predicted slopes with increasing λ. Only the difference between the predicted lines from the models with λ=0 and λ=0.25 is apparent. In comparison with the observed and predicted lines for the COLD death rates (Fig. 6.2), the interaction effect for lung cancer is less significant.

Table 7.5 Results from the RR model, Eq. 6.1, for changing λ.
Parameter estimates for dust and smoking are adjusted for age and year at risk.

		DUST[a]		SMOKING[a]	
λ	$-2\hat{\log}L$	$\hat{\beta} \pm \hat{S.E}$	X^2_1	$\hat{\beta} \pm \hat{S.E.}$	X^2_1
1.0	1003.1	0.028 ± 0.009	9.2	0.042 ± 0.007	32.3
0.75	1002.4	0.038 ± 0.018	4.5	0.058 ± 0.019	9.3
0.5	1002.0	0.066 ± 0.032	4.3	0.101 ± 0.029	12.1
0.25	1001.5	0.242 ± 0.082	8.8	0.377 ± 0.091	17.2
0.0	1005.4	1.011 ± 1.023	1.0	5.597 ± 6.163	0.8

[a] The mean values shown in Table 7.3 were used.

Table 7.6 Observed and predicted death rates/1000. h. estimated form the general relative risk models with different value of λ

CEPY	Dust[a]	P–Y	Obs. h	Estimated death rates for given λ				
				0.0	0.25	0.50	0.75	1.0
≤10	≤15	838	0	0.36	0.23	0.35	0.44	0.50
	16–30	3182	0	0.62	0.48	0.58	0.66	0.70
	31–40	2118	0.94	0.84	0.87	0.92	0.96	0.98
	41+	1708	1.76	1.15	1.53	1.53	1.53	1.52
11–25	≤15	1043	0	1.66	0.83	0.85	0.89	0.92
	16–30	4569	2.41	1.90	1.29	1.28	1.28	1.28
	31–40	2212	1.36	2.12	1.90	1.84	1.80	1.77
	41+	1844	2.71	2.44	2.87	2.78	2.71	3.22
26–35	≤15	727	2.75	2.66	1.64	1.58	1.55	1.54
	16–30	3133	2.87	2.92	2.31	2.22	2.15	2.11
	31–40	1482	2.02	3.19	3.21	3.10	3.02	2.97
	41+	1333	3.00	3.46	4.48	4.45	4.39	4.41
36+	≤15	614	3.26	4.30	3.75	3.63	3.55	3.52
	16–30	3439	2.91	4.30	4.33	4.27	4.22	4.16
	31–40	1571	7.64	4.81	5.89	5.97	6.00	6.03
	41+	1179	9.33	4.97	7.60	8.05	8.37	8.74
Total		30992						
X^2_{LR}†				141.9	138.1	138.6	139.0	139.7

† Degree of freedom is 267.
CEPY= cigarette equivalent pack–years.
Dust= (RSA x years /1000)

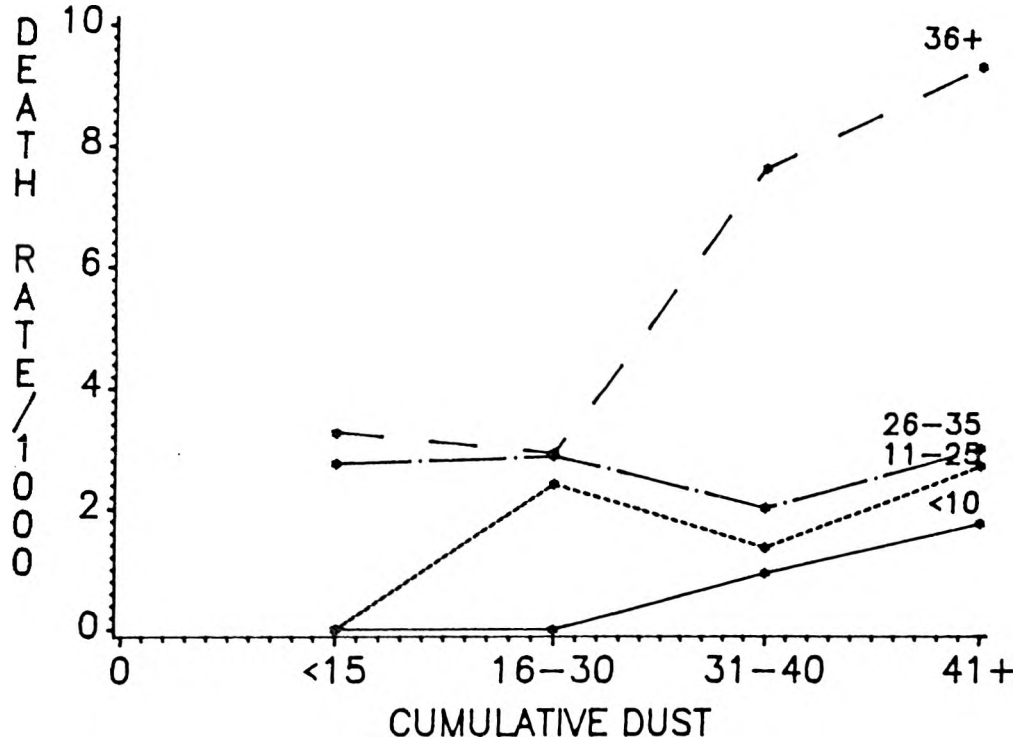


Fig 7.2 Observed death rates per 1000 person–years for lung cancer, by cumulative dust exposure and cigarette equivalent pack years

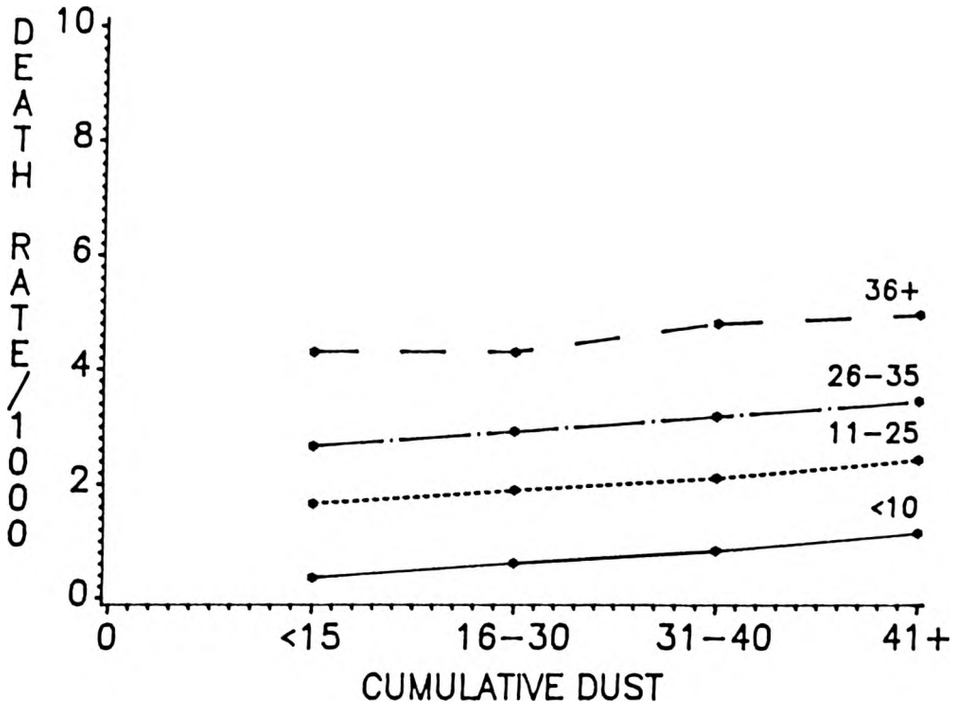


Fig 7.3 Predicted death rates per 1000 person years for lung cancer, estimated from the GRRM with $\lambda=0$

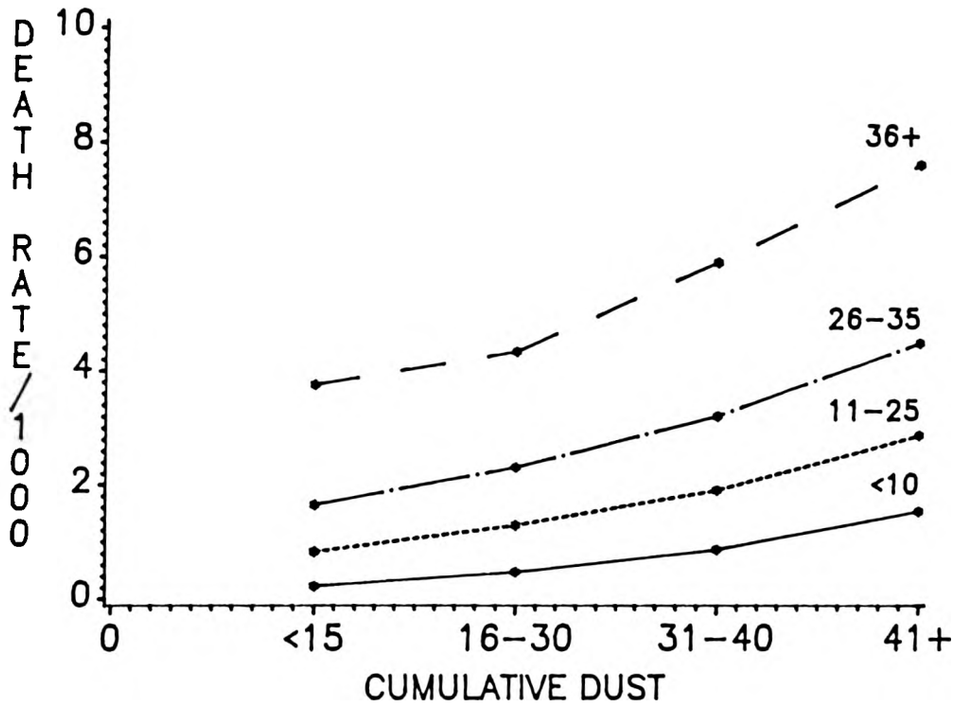


Fig 7.4 Predicted death rates per 1000 person years for lung cancer, estimated from the GRRM with $\lambda=0.2$

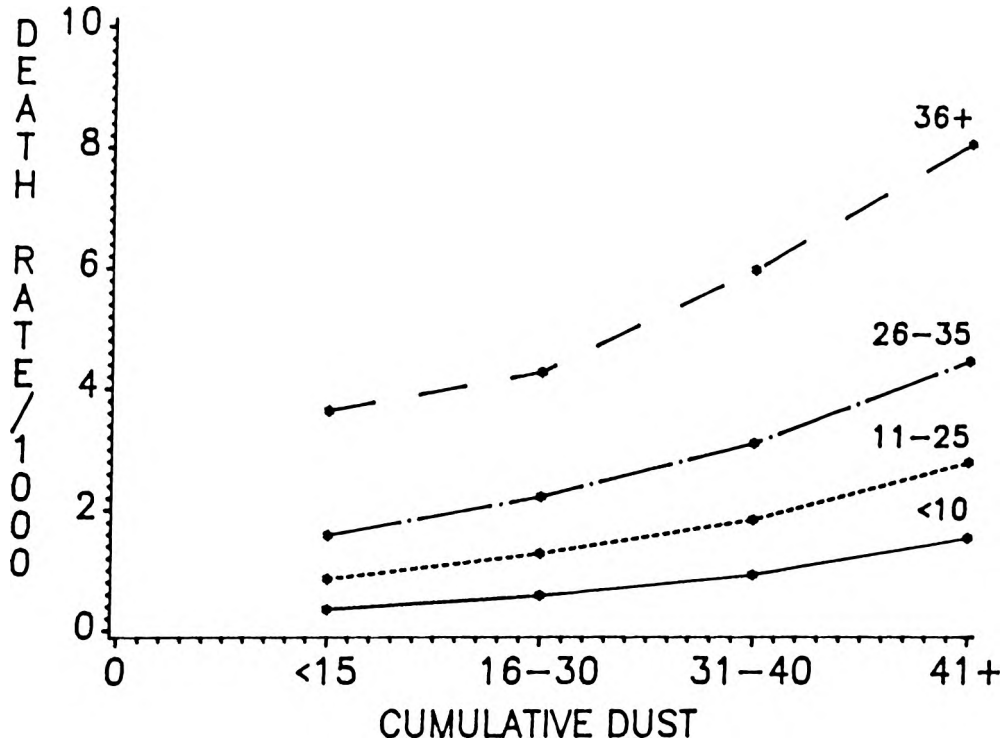


Fig 7.5 Predicted death rates per 1000 person years for lung cancer, estimated from the GRRM with λ=0.5

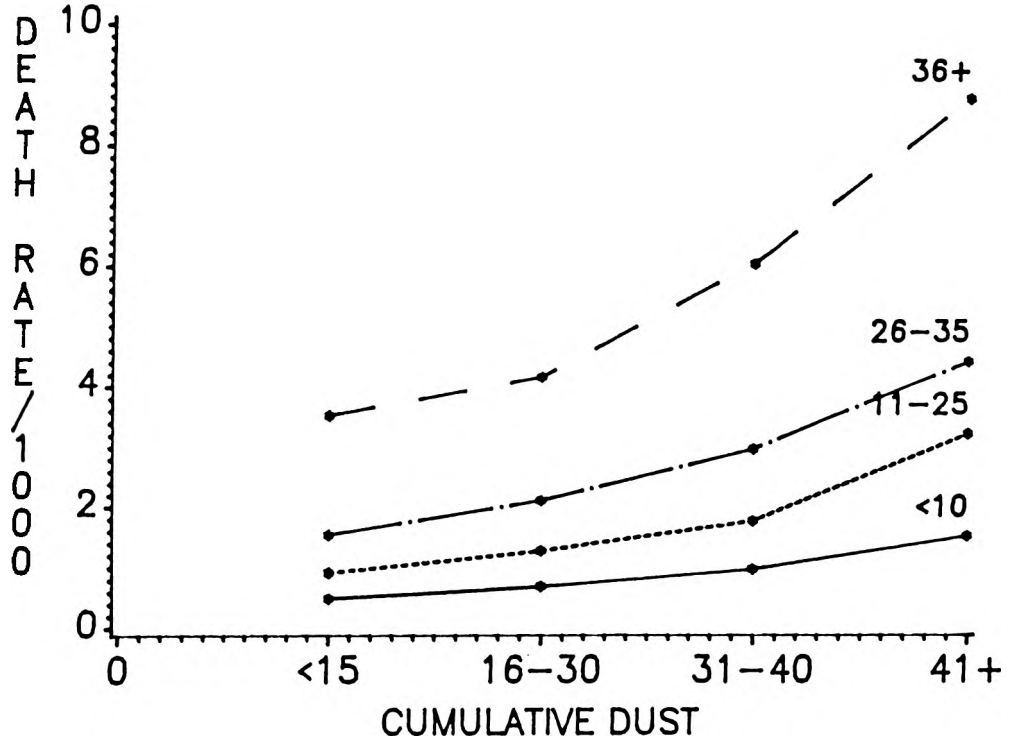


Fig 7.6 Predicted death rates per 1000 person years for lung cancer, estimated from the GRRM with λ=1

The five cases in the lowest smoking category (Table 7.3) were re–checked for their smoking habits as recorded in the medical files. Three men were found to be life–long nonsmokers and two were light pipe smokers. Their actual length of exposure on the gold mines was high, though, ranging from 24 to 32 years.

The adjusted relative risk for dust estimated by the multiplicative model (λ=1.0) was 1.028 95%CI(1.009;1.048) for one unit of cumulative dust exposure (RSA x years/ 1 000), which is only slightly higher than the estimate obtained from the Cox's proportional hazards model. The risk approportionment estimated from the data in Table 7.3 indicated that the proportion of cases associated with dust only was 22%, smoking only 45%, the combined effect of dust and smoking 30%, and 3% was associated with an unknown effect (0.5 was added to all cell frequencies for the estimation of the proportion due to dust and smoking).

7.5 Association between lung cancer and silicosis

The crude odds ratios and odds ratios adjusted (OR_a) for the effect of age at death, cumulative dust exposure to the start of the follow–up and cigarette equivalent pack–years were calculated for the association between lung cancer and silicosis, for all the 745 miners who had an necropsy examination. No association was observed with the presence of silicosis of the lung (OR_a=0.9, 95%CI 0.5;1.6), or silicosis of the pleura (OR_a=1.2, 95%CI 0.7;2.0), but there was a positive association with silicosis of the hilar glands (OR_a=3.9, 95%CI 1.2;12.7) (see Table 7.7).

Table 7.7 Association between death from lung cancer and silicosis of hilar glands

Death from lung cancer	Silicosis of the hilar glands				Total
	none	slight	moderate	marked	
Yes	3	10	39	17	69
No	109	132	309	174	724
Total	112	142	348	191	793†
$\chi^2_{\text{for trend}}$=4.5, p=0.035					
OR	1.0	2.8	4.6	3.6	
OR adjusted	1.0	2.9	4.2	3.1	
95% CI		(0.8;10.7)	(1.3;14.0)	(0.9;11.2)	

† One subject had a missing value for the hilar gland variable.

7.6 Association between silicosis and tobacco smoking

For all the 2209 gold miners the association between cigarette equivalent pack–years and cumulative dust exposure to start of follow–up was negative, with the Pearson correlation coefficient r=–0.036, p<0.09. For the 794 miners who underwent necropsy the correlation coefficients between cigarette equivalent pack–years and the ordinal categories of silicosis were (–0.076, p=0.03) for parenchyma, (–0.1334, p<0.001) for pleura, and (–0.044, p=0.22) for hilar glands. When adjusted for dust and age at death, by means of the multiple linear regression, only the association between silicosis of the pleura and smoking remained significant (the regression coefficient β=–0.006, p=0.0005), suggesting a negative association between tobacco smoking and silicosis of the pleura.

7.7 Discussion

The results from this study show a significant dose–response relation between death from lung cancer and exposure to cumulative dust and cigarette equivalent pack–years. The relative risk for lung cancer associated with the cumulative dust exposure (RSA x years/1000), standardized for smoking and age, estimated by the proportional hazards model was exp(0.023)=1.023, 95%CI(1.006;1.041). Thus the expected risk for the men in the highest exposure category with an average cumulative dust exposure of about 50, relative to those with less than 15 but with more than ten years of gold mining, is 3.18, 95%CI(1.34; 7.45). The cumulative dust to the start of the follow–up ranged from 7 to 80 (RSA x years/1000).

The adjusted relative risk for lung cancer estimated from the previous mortality study of 3971 white South African gold miners was 1.77 95%CI (0.94–3.31) per one unit of cumulative dust exposure of 1000 particle–years (p=0.06) (Wyndham *et al.*,1987). The results from the two follow–up studies are in disagreement with the two case–referent studies done on deceased white South African gold miners (Hessel *et al.*, 1986; Hessel *et al.*, 1990), in which no associations between lung cancer and dust exposure or silicosis were found. As dust was found to be a significant predictor of mortality from all causes it is possible that by selecting dead controls of the same age at death as the cases the case–referent studies might have been over matched for dust exposure. The other possibility is, that the selection procedure employed for the two case–referent studies may have missed miners with high dust exposure who were awarded maximum compensation in life and for whom the next of kin did not request an autopsy examination. Because of the inconsistencies in the results found by the different studies, further studies of the association between gold mining dust exposure and lung cancer are planned.

The combined effect of dust and smoking appears to be higher than would be expected from merely adding the effects of individual exposures. The power model fitted best at λ=0.25, and the discrepancies between the expected number of deaths as estimated by the additive model and the observed number of deaths (see Table 7.6) indicate some degree of synergism for the highest smoking and dust exposure categories. The goodness–of–fit statistic indicates that the hypothesis of multiplicativity of the combined effects fits the data better than does additivity. It appears that there is a synergistic effect between the two exposures for those with more than 35 pack–years of smoking and cumulative dust exposure of more than 30 (RSA x years/ 1000). These findings are consistent with other studies in which the combined effect of gold mining dust exposure and smoking was evaluated (Goldsmith *et al.*,1986).

The risk approportionment estimated from the data in Table 7.3 indicated that the proportion of the cases associated with dust only was 22%, smoking only 45%, the combined effect of dust and smoking 30%, and 3% were associated with an unknown effect. The proportion of cases associated with dust only may be slightly overestimated, since the baseline category for smoking was ≤10 years of smoking. however, three out of the five lung cancer cases in that category were life long non–smokers and two were light pipe smokers. Their actual length of gold mining exposure was high, though, ranging from 24 to 32 years.

No association between silicosis of the parenchyma or pleura and death from lung cancer was observed. The cases with lung cancer were, however, found to have increased prevalence of silicosis in the hilar glands in comparison to those who did not die from lung cancer (Table 7.7). The possibility that the increase in the hilar glands silicosis is due to the fact that the cases were heavier smokers and also had heavier dust exposure, but had no causal association with the development of lung cancer, was also considered. The correlation coefficients for

the association of lung cancer with smoking, with dust and with degree of hilar glands silicosis were found positive and statistically significant. The correlation between hilar glands silicosis and smoking was not significant, whereas that with dust was highly significant. The odds ratios for lung cancer and silicosis of the hilar glands adjusted for dust remained significant, and so did the effect of dust. As dust and silicosis of the hilar glands are, however, inter–correlated, it is difficult to draw a conclusion about their association with lung cancer.

Several possible mechanisms for the effect of silica dust on lung cancer have been proposed. Goldsmith and Guidotti (1986) proposed three inter–related hypotheses to explain the excess of lung cancer in silica exposed workers: 1) silica directly induces lung cancer, 2) silica causes silicosis, which may be an intermediate pathological state leading to lung cancer, and 3) silica combined with polycyclic aromatic hydrocarbons from cigarette smoke impairs lung clearance, thus increasing the effective dose in smokers. Cohen *et al.* (1978) suggested that cigarette smoke in addition to initiating tumors may also cause pulmonary insult which increases the individual susceptibility to a variety of malignant and non–malignant conditions. Guidotti and Coley *et al.* (1986) suggested that lymphatic obstruction by silica containing macrophages and by silica may decrease lymphatic flow, reduce lymphatic dust clearance and thus increase residence time of active carcinogens. Archer *et al.* (1986) argued, however, that silicosis among a group of miners usually reflects the presence of both a siliceous type of ore and poor ventilation. With poor ventilation, the miners are also exposed to elevated radon daughter levels which, they suggest, is responsible for the increased mortality from lung cancer in silica exposed miners.

This study confirms that an association between gold mining dust exposure with high level of silica dust and lung cancer is present. The results do not support the hypothesis that silicosis of the parenchyma or pleura is an intermediate

pathological state leading to lung cancer. As the intensity of dust exposure was found to be related to lung cancer, and not to the actual duration of the underground exposure, the hypothesis that silica directly induces lung cancer cannot be rejected. There is, however, a possibility that the association between lung cancer and exposure to silica dust in South African gold mines may be confounded by the effect of radon daughters, as the increased intensity of dust exposure may be related to increased levels of radon daughters (Archer *et al.*, 1986). Further research of this issue is required.

The following factors that could have contributed to the presence of positive association have been examined: asbestos exposure of the cases, positive misclassification of the cases (the cases were rechecked for positive necropsy and biopsy findings by three medical practitioners), and the calculation and coding of the dust exposures were re–checked. The lung cancer cases were proportionally distributed according to the reason for attending the 1968–71 medical examination (69% came for a periodical examination and were working in the mine, 3% from periodical examinations not working in the mine, 16% from benefit examinations working in the mine, and 12% from benefit examinations not working in the mine). The dose response relationship for lung cancer remained present even after those who attended the benefit examination and were not working on the mine were excluded from the analysis. The possibility that another unknown factor biased the results could not, however, be excluded.

In conclusion, the relative risk of lung cancer associated with one unit of cumulative dust exposure (RSA x years/ 1000) was estimated as 1.023 95%CI (1.005;1.042), the cumulative dust exposure ranged from 7 to 80 for the whole cohort. The combined effect of dust and smoking was adequately fitted by the multiplicative model, indicating the presence of synergism. The odds ratio for the association between lung cancer and silicosis of the hilar gland was 3.9 95%CI (1.2;

12.7). No association was observed with silicosis of the parenchyma or pleura.

CHAPTER 8

GENERAL CONCLUSION

In this chapter a summary is given of the findings from the present thesis and of their significance in the context of gold miners health. The application of the general relative risk model to investigate the presence of synergism of public health significance is discussed also.

It should be mentioned that cumulative dust exposure was estimated in terms of surface area of respirable dust (Beadle *et al*, 1970). Since this kind of measurements is no longer used in South Africa, or in other countries, an attempt was made to convert the surface area of respirable dust into measurements comparable with gravimetric sampling. The cumulative dust exposure, estimated in terms of number of hours in each occupation multiplied by the mass of respirable dust (g/m^3), i.e. gram hours/m^3 (gh/m^3), ranges approximately from 6 to 22 gh/m^3 for the 2209 white gold miners.

8.1 The changing pattern of lung diseases in the gold miners

In the early years of gold mining the high levels of respirable dust in the air caused a high incidence of disabling silicosis which was often complicated by tuberculosis of the lung. This resulted in a high mortality at an early age from these two causes. Introduction of stringent dust control measures effectively

reduced dust exposure levels by the 1930s to the exposure levels of the present day. As a consequence the incidence and severity of silicosis was reduced to the extent that death from silicosis is very rarely reported as a primary cause of death in white gold miners (Wyndham *et al.*, 1986). However, with the decline in the number of miners affected by disabling silicosis the disability due to chronic obstructive lung disease became more prominent. This may also be as a consequence of increased tobacco consumption which started on large scale during the World War I, and increased progressively during the subsequent decades (McCusker, 1988).

Chronic obstructive lung disease has been a compensable disease since 1952. Since then epidemiological and experimental studies have demonstrated an association between silica dust exposure and chronic obstructive lung disease (Wiles and Faure, 1977; Wright *et al.*, 1988), and emphysema diagnosed at autopsy (Becklake *et al*, 1987; Hnizdo *et al*, 1991; Churg *et al*, 1989). Two theories about the silica dust involvement in the pathogenic process are suspected. Firstly, silica particles may cause inflammation and narrowing of the small airways thus increasing the airways obstruction. Secondly, silica particles may be digested by the alveolar macrophages which then bring about increased production of proteolytic enzymes which can cause destruction of the alveolar walls thus causing emphysema.

More recent evidence shows that chronic obstructive lung disease is more important in terms of morbidity and mortality among white gold miners than is silicosis (Irwig *et al.*; 1978; Wyndham *et al.*, 1986). Beadle *et al.*(1970) observed that 61% of the white gold miners included in his study were certified for disability without radiological evidence of silicosis and that the certification rate of men who did not have radiological signs of silicosis was not related purely to their dust exposure. A question was raised "What are the causes of the

pulmonary disability in order that preventive measures can be applied?"

In black miners the pattern of lung diseases is probably different from that in white miners. The black miners who work mostly as laborers were, and are, generally exposed to higher personal lung burdens of dust than the white miners who work mostly in supervisory positions (Cowie, 1989; Wyndham, 1974). Most of the black miners are, however, migrant workers employed on short contracts with periods away from the mines, which may help to reduce the risk of silica induced lung diseases. Generally, the black miners also smoke less than the white miners. The main causes of lung disability are still related to silicosis and tuberculosis of the lung, both of which are highly prevalent in black miners (Cowie, 1987). It is believed that exposure to silica dust and the presence of silicosis in the lung depresses the immune response to pulmonary infections and thus increases the risk of tuberculosis (Bruce, 1968). In addition, the risk of tuberculosis in the black miners is considered to be related to the incidence of tuberculosis in the general black population of South Africa and other African countries from which these miners are recruited (Cowie, 1989). Because of lack of reliable medical data, exposure data, and mortality statistics, in addition to the selection process during the re–recruitment of the miners, the effect of gold mining dust exposure on chronic obstructive lung impairment in black miners is difficult to investigate.

8.2 Findings of this thesis in the context of gold miners health

8.2.1 Simple chronic bronchitis

This condition is highly prevalent in white South African gold miners. The present thesis (Chapter 4) has shown that the prevalence of chronic bronchitis

increases with increasing levels of cumulative dust exposure. The increasing trend with dust was observed in non–smokers, ex–smokers and current smokers (Table 4.7 and Figure 4.1). However, in the current smokers the trend was significantly increased in comparison with the non–smokers and ex–smokers. When tested by the additive relative risk model the combined effect of dust and smoking was found to be synergistic. The effect of dust was found to be much stronger than the effect of smoking. For measured sputum the pattern was similar, however, the effect of smoking was much stronger than the effect of dust and the synergistic effect was less apparent (see Table 4.10 and Figure 4.2).

Several studies have suggested that simple chronic bronchitis and airways obstruction are two separate diseases of the lung (Fletcher *et al.*, 1976). In a study done on the 2209 gold miners used in the present thesis, we have found that the presence of chronic bronchitis, when lung function impairment was adjusted for by the FEV1/height^3, was not related to increased mortality from chronic respiratory diseases (Wiles *et al*, 1991). This result was confirmed by several other studies (Peto *et al.*, 1983), and leads to a conclusion that the presence of chronic bronchitis without lung function impairment is not a risk factor for mortality from chronic obstructive lung disease. Simple chronic bronchitis may, however, predispose for lung function impairment.

8.2.2 Lung function impairment

Loss of lung function attributable to dust and smoking

Cumulative dust exposure was shown to be associated with decreasing lung function measurements (FEV_1, VC, FEF_{25-75}). The mid–expiratory flow, which is considered to be an early indicator of small airways obstruction, has shown greater decrease with increasing cumulative dust exposure than the other lung function measures (see Figure 5.1 and Tables 5.1 and 5.2). Smoking did not

appear to modify significantly the dose–response trend of the individual lung function measurements with dust (Figure 5.1). This observation was confirmed by a multiple linear regression and logistic regression. The estimated effect of cumulative dust exposure on loss of lung function (FEV_1) ranged from 123 ml to 319 ml. The estimated loss of FEV_1 in current smokers ranged from 500 ml to 687 ml (Table 5.2). Thus the effect of high dust exposure on the loss of lung function measurements appears to be about half of that of the effect of heavy smoking.

Assuming that the effects of dust and smoking on the loss of lung function (FEV_1) are additive, then the expected loss of FEV_1 for a 50 years old miner who is a current smoker with a smoking history of 60 cigarette equivalent pack–years, and who worked in high dust occupations is approximately 319+687= 1 006 ml. This result indicates that the combined effect of silica dust and tobacco smoking has a potential to decrease the lung function to such an extent that the smoking miners are more likely to become disabled than the smoking non–miners.

Contribution of dust and smoking to seriousness of respiratory disability

In a further study the 2209 miners were grouped according to their lung function impairment profiles as "normal", minimal obstructive, moderate obstructive, marked obstructive, marked obstructive with restriction, and restricitve. The groups with the impaired profiles were then compared to the normal group for differences in dust exposure and smoking habits.

The miners with the marked obstructive impairment profile (n=402) had the highest tobacco consumption and the strongest dose–response trend with dust, yet in comparison with the moderate obstructive profile they had lower dust exposure and lower proportion of silicotics (see Table 5.3). The

dose–response trend with dust was increased for the marginal odds ratios and for those with tobacco consumption over 20 pack–years (see Tables 5.10 and 5.11). The synergistic effect between dust and smoking was found to be statistically significant in the additive relative risk model for the comparison of the marked obstructive profile with the reference group.

The estimated etiological fractions for dust increased with increasing obstructive impairment (14%, 20%, 43% and 52%). However, the preventable fractions indicated that the proportion of cases associated with dust only was small, whereas the proportion associated with smoking and dust together increased (14%, 16%, 32% and 52%). Of the cases with the marked obstructive profile (n=402), approximately 8% were associated with dust only, 42% with smoking only, 40% with dust and smoking, and 10% with an unknown cause. Of the 191 miners with the obstructive and restrictive profile (who had the highest mortality from COLD and from all causes) 0% were associated with dust only, 39% were associated with smoking alone, and 52% were associated with the combined effect of dust and smoking. This indicates that elimination of tobacco smoking would prevent most of the cases with the serious obstructive and restrictive impairment.

The miners with the marked obstructive profile had the highest mortality from COLD and from all causes (see Figure 5.3); 239 (60%) of these miners died during the follow–up period. The miners with the restrictive impairment profile had high mortality from ischaemic heart disease which may be associated with the higher body mass.

In conclusion, the analyses indicated that the main risk factors for the serious respiratory disability in white South African gold miners are tobacco smoking, and the combined effect of tobacco smoking and silica dust. Figure 5.1 also indicates that the miners most likely to become disabled are the tobacco

smokers with heavy dust exposure.

8.2.3 Mortality from COLD

The risk factors most strongly associated with death from COLD were years of cigarette smoking and cumulative dust exposure before 1950. The miners who smoked 20 or more years had significantly increased dose–response trend with increasing dust exposure. The synergistic effect between dust and smoking was found to be statistically significant in the additive relative risk model. All the miners who died of COLD were tobacco smokers. The risk approportionment indicated that 5% of the COLD deaths were associated with dust only, 34 with smoking only and 59% with dust and smoking. The 5% is mainly due to one miner who did not smoke cigarettes but was a pipe smoker over 30 years.

These results are in agreement with those obtained from the study of lung function impairment.

8.2.4 Mortality from lung cancer

There are several suspected risk factors for lung cancer in the gold mines. These include silica dust, radon daughters and diesel fumes. The epidemiological evidence on the association between silica dust and lung cancer obtained on South African gold miners is conflicting. Two case–control studies did not find any association between silicosis or silica dust exposure (Hessel *et al.*, 1986; Hessel *et al.*, 1990), and a mortality study found an association with cumulative dust exposure but not with years of exposure (Wyndham *et al.*, 1986). The findings of this thesis are that a) death from lung cancer is associated with cumulative dust exposure, b) the combined effect of dust and smoking is synergistic in the highest dust and smoking levels, c) the association between silicosis of the hilar gland and lung cancer is statistically significant, and d) no

association between silicosis of the lung parenchyma and lung cancer is present. As the evidence from other epidemiologic studies is still considered inconclusive, further studies of the association are required to resolve the issue.

8.2.5 Other evidence from mortality studies

Of the 2209 gold miners followed from 1968–71 up to 31 December 1986 945 have died. Table 8.1 shows the 945 deaths grouped according to the main categories of ICD and their exposure to dust and smoking. According to these results 49% of the miners died from cardio–vascular disease, 21% from respiratory diseases and 8% from lung cancer. Applying the Cox's proportional hazards model it was found that cumulative dust exposure was related to death from pulmonary tuberculosis, lung cancer, COLD and nephritis. This result is very much in line with the results from the first mortality study done on 3971 middle aged white gold miners (Wyndham *et al.*, 1986). In comparison to the general white South African male population the 3971 gold miners had increased mortality from all causes SMR=117.6%. This was attributed to excess mortality from lung cancer SMR=161.2% (115–220), from COLD SMR=166 (108–243) and from nephritis SMR=381 (164–750). No death that could be attributed directly to silicosis or tuberculosis was found. It was suggested that certain aspects of life style which are known to be destructive to health, such as dietary excesses, smoking and excessive alcohol intake were responsible for some of the excess deaths observed in the study (Wyndham *et al.*, 1986). In another study, Wyndham *et al.* (1986) also reported that the main risk factors related to death from cardiovascular disease were hypertension, smoking, hypercholesterolaemia, adiposity and, very marginally, length of underground exposure.

In the mortality study of 2209 gold miners we have found also that lung function and tobacco smoking were related to all the other main categories of

death, except tuberculosis, and neoplasm other than lung cancer (see Table 8.1). Since silica dust exposure reduces lung function, the combined exposure to silica dust and tobacco smoking may be reflected in mortality from several other causes of death. The presence of radiological silicosis was not found to be a predictor of death from all causes.

Table 8.1 Exposure profile of the categorized causes of deaths

Cause of death (ICD code)	N	Age death	Years of mining	CDE	Cigarette pack–yrs	FEV1 (l)
Alive	1264		23.5±4.9	28.8±12.1	21.0±16.3	3.28±0.68
Pulmonary TB (011 and 012)	10	64	25.7±4.8	40.5±14.0	24.3±17.2	3.19±0.70
Neoplasm (141–161,163–208)	94	61	23.4±5.3	29.8±11.4	24.5±17.1	3.17±0.71
Lung cancer (162)	77	62	24.4±5.1	33.2±13.0	34.5±18.2	2.89±0.64
IHD (410–412)	367	60	23.5±4.6	29.6±10.5	28.6±16.9	2.95±0.69
Cardiovascular (394–404,415–453)	100	61	23.3±4.7	30.4±12.1	29.3±20.3	2.93±0.75
COLD (490–496)	67	62	24.5±4.9	32.5±9.9	31.2±16.2	1.99±0.82
Other respir. (460–487,500–519)	48	62	22.1±4.1	27.9±12.3	32.5±17.6	2.85±12.0
Digestive (520–609)	50	60	23.5±5.2	28.6±12.0	30.4±16.0	2.84±0.70
Nephritis (582–588)	12	59	25.2±3.3	33.9±7.2	33.9±15.2	2.73±0.90
Others	116	58	23.2±5.3	31.5±12.7	25.8±18.0	3.06±0.76
Total	2205[a]					

[a] 4 cases classified as lung cancer were rejected.
CDE=RSA x years / 1000.

8.3 Implications for preventive measures

It is apparent that dust exposure levels to which the South African gold miners were exposed during the period from 1940 to 1970 should be considered harmful in view of the associations with lung function impairment, death from COLD, lung cancer and tuberculosis. However, the epidemiological evidence suggest that compensable chronic obstructive lung disease is mainly an occupational lung disease of smoking miners, as very few non–smokers have very serious lung function impairment (or a significant degree of emphysema found at autopsy (Hnizdo, Emphysema in nonsmoking gold miners, in preparation).

Thus, since most of the cases with compensable COLD are caused by smoking and the combined effect of dust and smoking, the prevention of this disease in terms of dust reduction is difficult to determine. A solution is to improve dust conditions to levels that minimize the incidence of radiological silicosis. Beadle (1970) estimated that in miners with 30 years of underground exposure the approximate risk of radiological silicosis increases according to the levels of respirable dust as follows: 10% of silicotics at an average level of 0.24 mg/m^3, 25% at 0.29 mg/m^3, and 40% at 0.34 mg/m^3. Using data from the 2209 white gold miners who died by 1990 and had an autopsy examination (n=969) we have observed a very similar dose–response trends for the moderate and marked silicosis diagnosed at autopsy, as was reported by Beadle (unpublished results). From 1991 the Epidemiological Research Unit will have access to computerized records of dust exposures done by the Government Mining Inspectors on individual mines. This data should be utilized to evaluate the risk of silicosis for white and black gold miners in order to establish the acceptable level of dust exposure in terms of risk of silicosis. In addition, the changes in the prevalence of COLD should be monitored in black and white miners on a regular

basis.

The other factor that may play some role in the miners health is the compensation scheme applied at present. Any miner who worked 20 years in low dusty occupations or 10 years in high dusty occupations and shows a certain degree of respiratory disability is compensated. The miners who worked in more dusty occupations such as developers, shaft sinkers and stoppers are also directly responsible for mine production. They are usually more intelligent and fitter miners who cary greater responsibility, have higher job pressure and have less time to smoke (see the negative association between cumulative dust exposure and cigarette equivalent pack–years). It may be that these miners are discriminated against by the present compensation scheme which seems to favor the heavy smokers (see Table 5.3). Thus to a great extent heavy smoking by the miners is being rewarded by the present compensation scheme. However, as COLD is related to silica dust it has to be compensable. Thus, to avoid compensation for smoking the miners should be convinced that smoking is bad for them as it causes disability and premature deaths from many causes.

The other aspect of the miners health which have strongest impact on white miners disability and mortality are life style related factors. Tobacco smoking is associated with mortality from many diseases especially with death from cardio–vascular diseases which constitute the major cause of death amongst the white miners. The other factor is obesity which is prevalent amongst the white miners (Wyndham *et al.*, 1986). In our study, the 187 miners with the restrictive impairment (Table 5.3) had increased average body mass index (weight/height^2), 27%, and large proportion of deaths from IHD.

The overwhelming evidence suggests that life style related intervention programmes should be introduced in the mines and epidemiological studies should be designed to assess the effect of the intervention programmes. The

financial expenditure involved would be reflected in the reduction of future compensation for chronic obstructive lung disease and an increase in productive years among the miners. In addition, such a campaign is also important in view of the increasing tobacco consumption of black miners. Experience has shown that any intervention program aimed at change of life style works best when supported by community based campaigns. This usually involves schools, media, non–smoking public places and workplaces, and a ban on advertising of products that are known to be harmful to health such as tobacco.

It must be pointed out that the present discussion did not include evidence on trauma, mine accidents, hearing loss and skin diseases which also play an important part the health of miners.

8.4 Observations made from using the GRRM

The logistic regression model, often applied in the analysis of epidemiological data, does not detect a synergistic effect that is often apparent and important from the public health point of view. However, application of the additive relative risk is often problematic. Firstly, the range of the β coefficients is restricted by the requirements that $\beta x > -1$ for all variables, since negative risk would otherwise result. Consequently, the log–likelihood function is skewed and the estimated standard errors of the estimates are unreliable for inferences purposes. Secondly, there are often problem with convergence unless the starting values of the parameters are close to the expected values.

The power relative risk model, namely the Breslow and Storer model, is easier to fit to the data if the following procedure suggested by Breslow and Day (1987) is followed. The multiplicative model with $\lambda=1$ is fitted to the data first, as it almost always converges. The β coefficients obtained at convergence of the multiplicative model are then used as starting values for the power model with

λ=0.9. This procedure is followed for λ=0.8 and so on, until the additive relative risk model with λ=0, is fitted. However, in some cases it is difficult to fit models with very small λ values, even though the model with λ=0 converges.

Since the standard errors of the parameter estimates are unreliable for inferences purposes the log likelihood ratio statistics should be used to test the statistical significance of the interaction effect.

The above method worked well with large study samples and when the disease was sufficiently rare, for example, in the mortality studies presented in Chapters 6 and 7. However, the results obtained from the analysis of simple chronic bronchitis may be unreliable because the rare disease assumption did not apply.

REFERENCES

American Thoracic Society. Medical Section of the American Lung Association. Standardization of spirometry – 1987 update. Am Rev Respir Dis 1987; 136:1285–1298.

Annesi I, Moreau T, Lellough J. Efficiency of the logistic regression and Cox proportional hazards models in longitudinal studies. Statistics in Medicine 1989; 8:1515–1521.

Archer VE, Roscoe JR, and Brown D. Is silica or radon daughters the important factor in the excess lung cancer among underground miners? In: Silica, Silicosis, and Cancer: Controversy in Occupational Medicine (Editors: Goldsmith DF, Winn DM, Shy CM). Cancer Research Monographs, Vol 2. Publ. Praeger, New York, 1986, pp. 375–384.

Armitage P, Doll R. Stochastic models for carcinogenesis. In: Neyman J. (ed.). Contributions to Biology and Problems in Medicine. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol 4. Berkeley, University of California Press, 1961, pp. 19–38.

Armstrong BK, McNulty JC, Levitt LJ, Williams KA, Hobbs MST. Mortality in gold and coal miners in Western Australia with special reference to lung cancer. Br J Ind Med 1979; 36:199–205.

Auerbach O, Stout AP, Hammond EC, Garfinkel L. Smoking habits and age in relation to pulmonary changes. N Eng J Med 1963; 269:1045–1054.

Attfield MD. Longitudinal decline in FEV1 in United States coalminers. Thorax 1985; 40:132–137.

Baker RJ and Nelder JA. The GLIM system: Release 3. Oxford. Numerical Algorithms Group. 1978.

Barlow WE. General relative risk models in stratified epidemiological studies. Appl Statist 1985;34:246–57.

Battista G. Belli S, Carboncini F. Comba P *et al.* Mortality among pyrite miners with low–level exposure to radon daughters. Scan J Work Environ Health 1988; 14:280–285.

Beadle DG, Bradley AA. The composition of airborn dust in South African Gold mines in pneumoconiosis. In: Shapiro, HA, Ed.,Pneumoconiosis. Proceedings of the International Conference, Johannesburg 1969. Cape Town, Oxford University Press,1970. pp.462–466.

Beadle DG. The relationship between the amount of dust breathed and the development of radiological signs of silicosis: an epidemiological study of South African gold miners. In: Walton WH. (ed.). Inhaled Particles III. Oxford: Pergamon Press. 1971;953–964.

Beaty TH, Cohen BH, Newill CA, Menkes HE, Diamond EL, and Chen CJ. Impaired pulmonary function as a risk factor for mortality. Am J Epidemiol 1982;116:102–113.

Becklake MR, Du Preez L, Lutz W. Lung function in silicosis of the Witwatersrand gold miner. Am Rev Tuberc 1958; 77:400–412.

Becklake MR. Chronic airflow limitation: its relationship to work in dusty occupations. Chest 1985; 88:608–614.

Becklake MR, Irwig L, Kielkowski D, Webster I, De Beer M, Landau S. The predictors of emphysema in South African gold miners. Am Rev Respir Dis 1987;135:1234–41.

Berry G, Newhouse M, Turok M. Combined effect of asbestos exposure and smoking on mortality from lung cancer in factory workers. Lancet 1972;

2:476–479.

Berry G. Dose–response in case–control studies. J Epidemiol Comm Health 1980; 34:217–222.

Berry G. The analysis of mortality by the subject–years method. Biometrics 1983;39:173–184.

Bignon J. Methods of quantitative pathology with special reference to emphysema and chronic bronchitis. Bull Europ Physiopath Resp 1976;12:675–692.

Blot WJ, Day NE. Synergism and interaction: are they equivalent? Am J Epidemiol 1979;110:99–100.

Breslow NE, Day NE, Halvorsen KT, Prentice RL, Sabai C. Estimation of multiple relative risk functions in matched case–control studies. Am J Epidemiol 1978; 108:299–307.

Breslow NE and Day NE. Statistical methods in cancer research, Vol 1– The analysis of case–control studies. IARC, Lyon, 1980.

Breslow NE. Elementary method of cohort analysis. International Journal of Epidemiology 1984; 13:112–115.

Breslow NE and Day NE. Statistical methods in cancer research, Vol 2– The design and analysis of cohort studies. IARC, Lyon, 1987.

Breslow NE, Storer BE. General relative risk functions for case–control studies. Am J Epidemiol 1985;122:149–162.

Brown DP. Kaplan SD, Zumwalde RD. Kaplowitz M and Archer VE. Retrospective cohort mortality study of underground gold mine workers. In: Silica, Silicosis, and Cancer: Controversy in Occupational Medicine (Editors: Goldsmith DF, Winn DM, Shy CM). Cancer Research Monographs, Vol 2. Publ. Praeger, New York, 1986, pp. 335–350.

Bruce T. Silicotuberculosis. Scand. J Respir Diseases 1968; 65:139–146.

Bruzzi P, Green SB, Byar DP, Brinton LA, Schairer C. Estimating population

attributable risk for multiple risk factors using case–control data. Am J Epidemiol 1985; 122:904–914.

Burns DM, Froines JR, Murray EJ. Biological interactions between smoking and occupational exposures. Am J Ind Med 1988:13:169–79.

Burrows B. An overview of obstructive lung diseases. Medical Clinics of North America 1981;65:455–71.

Churg A, Hobson J, Wright J. Functional and morphological comparison of silica – and elastase induced airflow obstruction. Experimental lung research 1989; 15:813–822.

Cohen BH. Is pulmonary dysfunction the common denominator for the multiple effect of cigarette smoking? Lancet 1978; 2:1024–1027.

Cohen D, Ara SF, Brain JD. Smoking impairs long–term dust clearance from the lung. Science 1979; 204:514–517.

Copes RL, Thomas DC, Becklake MR. Temporal patterns of exposure and nonmalignant pulmonary abnormality in Quebec chrysotile workers. Arch Environ Health 1985; 40:80–86.

Corn JK, Starr J. Historical perspective on asbestos: Policies and protective measures in World War II shipbuilding. Am J Ind Med 1987: 11:359–373.

Cornfield J. A method of estimating comparative rates from clinical data. Application to cancer of the lung, breast, and cervix. J Natl. Cancer Inst 1951; 11: 1269–1275.

Cornfield J. A method of estimating comparative rates from clinical data. Application to cancer of the lung, breast and cervix. J Natl Cancer Inst 1975; 11:1269–1275.

Cotes JE. Lung function. Assessment and application in medicine. 3 ed. Blackwell Scientific Publication 1975.

Cowie RL. Silicosis, pulmonary dysfunction and respiratory symptoms in South

African gold miners. Ph.D. Thesis, McGill University, Montreal, Canada, 1987.

Cowie RL. Pulmonary tuberculosis in South African gold miners. Am Rev Respir Dis 1989; 139:1086–1089.

Cowie RL. The five stages of pulmonary tuberculosis and the South African goldminer. South African Medical Journal 1989; 76:566–567.

Cox DR. Regression models and life tables (with discussion). J R Stat Soc B 1972; 34:187–220.

Crapo RO, Morris AH, Gardner RM. Reference spirometric values using techniques and equipment that meet ATS recommendations. Am Rev Respir Dis 1981;123:659–664.

Davis LK, Wegman DH, Monson RR, Froines J. Mortality experience of Vermont granite workers. Am J Ind Med 1983; 4:705–723.

Day NE, Brown C. Multistage models and primary prevention of cancer. J Nat Cancer Inst 1980;64:977–989.

Denman DW, Schlesselman JJ. Interval estimation of the attributable risk for multiple exposure levels in case–control studies. Biometrics 1983;39:185–192.

Doll R. Mortality from lung cancer in asbestos workers. Br J Ind Med 1955; 12:81–86.

duToit RSJ. Period of service before certification for pneumoconiosis on South African gold mines. In: Shapiro, HA, Ed.,Pneumoconiosis. Proceedings of the International Conference, Johannesburg 1969. Cape Town, Oxford University Press,1970, pp.247–249.

Elmes PC. Relative importance of cigarette smoking in occupational lung disease. Br J Ind Med 1981;38:1–13.

Finkelstein M, Liss GM, Krammer F, Kusiak RA. Mortality among workers

receiving compensation awards for silicosis in Ontario 1940–85. Br J Ind Med 1987;44:588–594.

Fleiss JL. The Mantel–Haenszel estimator in case–control studies with varying number of controls matched to each case. Am J Epidemiol 1984; 12:1–3.

Fletcher CM, Peto R, Tinker CM, Speizer FE. The natural history of chronic bronchitis and emphysema. Oxford: Oxford University Press, 1976.

Foxman B, Higgins ITT, and Oh MS. The effect of occupation and smoking on respiratory disease mortality. Am Rev Respir Dis 1986;134: 649–652.

Frome EL. The analysis of rates using Poisson regression models. Biometrics 1983; 39:665–674.

Frome EL, Checkoway H. Use of poisson regression models in estimating incidence rates and ratios 1985;121:309–323.

Gillam JD, Dement JM, Lemen RA, Wagoner JK, Archer VE, and Blejer HP. Mortality patterns among hard rock gold miners exposed to an asbestiform mineral. NY Acad Sci; 271: 336–344.

Goldsmith DF and Guidotti TL. Combined silica exposure and cigarette smoking: a likely synergistic effect. In: Silica, Silicosis, and Cancer: Controversy in Occupational Medicine (Editors: Goldsmith DF, Winn DM, Shy CM). Cancer Research Monographs, Vol 2. Publ. Praeger, New York, 1986, pp. 451–459.

Guidotti TL, Coley BD, Goldsmith DF. Silica exposure and intrathoracic lymphatic changes. In: Silica, Silicosis, and Cancer: Controversy in Occupational Medicine (Editors: Goldsmith DF, Winn DM, Shy CM). Cancer Research Monographs, Vol 2. Publ. Praeger, New York, 1986, pp. 147–155.

Goldsmith DF, Winn DM, Shy CM, (eds.). Silica, Silicosis, Controversy in Occupational Medicine. Cancer Research Monographs, Vol 2. Publ.

Praeger, New York. 1986.

Gordon D. Dust and history. Med J Aust 1954; 2:161–166.

Greenland S. Limitations of the logistic analysis of epidemiological data. Am J Epidemiol 1979;110:693–698.

Greenland S, Thomas D. On the need for the rare disease assumption in case–control studies. Am J Epidemiol 1982;116:547–553.

Greenland S, Poole C. Invariants and noninvariants in the concept of interdependent effects. Scand J Work Environ Health 1988; 14:125–129.

Greenland S. Modeling and variable selection in epidemiologic analysis. Am J Public Health 1989; 79:340–349.

Guerrero VM, Johnson RA. Use of the Box–Cox transformation with binary response models. Biometrika 1982;69:309–314.

Guidotti TL, Coley BD, Goldsmith DF. Silica exposure and intrathoracic lymphatic changes. In: Silica, Silicosis and Cancer. Contraversy in occupational medicine. (Editors: Goldsmith DF, Winn Dm, Shy CM). Cancer Research Monographs, Vol 2. Publ. Praeger, New York, 1986. pp. 147–155.

Hammond E, Selikoff I, Seidman H. Asbestos exposure, cigarette smoking, and death rates. Ann NY Acad Sci 1979: 33:473–489.

Harber P, Tashkin DP, Lew M, Simmons M. Physiological Categorization of asbestos–exposed workers.Chest 1987;92:494–99.

Hessel PA, Sluis–Cremer GK, Hnizdo E. Case–control study of silicosis, Silica Exposure, and lung cancer in white South African gold miners. Am J Ind Med 1986;10:57–62.

Hessel PA, Hnizdo E. Pathological findings in mine workers:II Quality of the PATHAUT data. Am J Ind Med 1987; 12:81–89.

Hessel PA, Sluis–Cremer GK. X–ray findings, lung function, and respiratory

symptoms in Black South African vermiculite workers. Am J Ind Med 1989; 15:21–29.

Hessel PA, Sluis–Cremer GK, Hnizdo E. Silica exposure, silicosis and lung cancer: an autopsy study. Br J Ind Med 1990:47:4–9.

Higgins ITT. The epidemiology of chronic respiratory disease. Prev Med 1973; 2:14–33.

Hnizdo E, Sluis–Cremer GK. Effect of tobacco smoking on the presence of asbestosis at postmortem and on the reading of irregular opacities on roentgenograms in asbestos–exposed workers. Am Rev Respir Dis 1988; 138:1207–1212.

Hnizdo E and Sluis–Cremer GK. Response to effect of tobacco smoking on the presence of asbestosis at postmortem and on the reading of irregular opacities on roentgenograms in asbestos–exposed workers. American Review of Respiratory Disease 1989; 139:1568–1569.

Hnizdo E, Sluis–Cremer GK, Abramowitz JA. Emphysema type in relation to silica dust exposure, in South African gold miners. Am Rev Respir Dis (in press).

Hogan MD, Kupper LL, Most BM, Haseman JK. Alternatives to Rothman's approach for assessing synergism (or antagonism) in cohort studies. Am J Epidemiol 1978; 108:60–67.

Holford TR, White C, Kelsey JL. Multivariate analysis for matched case–control studies. Am J Epidemiol 1978; 107:245–256.

Holford TR. The analysis of rates and survivorship using log–linear models. Biometrics 1980; 36:229–306.

Holman CDJ, Psaila–Savona P, Roberts M, McNulty JC. Determinants of chronic bronchitis and lung dysfunction in Western Australian gold miners. Br J Ind Med 1987;44:810–18.

International Agency for Research on Cancer. IARC monographs on the evaluation of the carcinogenic risk of chemicals to humans. Silica and some silicates. Vol 42. IARC, Lyon, France, 1987.

Irvine LG, Simson FW, Southerland Strachan A. The clinical pathology, radiology and symptomatology of silicosis. International Silicosis Conference 13–27 August 1930, No. 6. Geneva:ILO, 1930: 251–293.

Irwig LM, Rocks P. Lung function and respiratory symptoms in silicotic and nonsilicotic gold miners. Am Rev Respir Dis 1978; 117:429–435.

Irwig LM, du Toit RSJ, Sluis–Cremer GK, *et al.* Risk of asbestosis in crocidolite and amosite mines in South Africa. Ann NY Acad Sci 1979; 330:35–52.

James F, Roos M. "MINUIT"– A system for function minimization and analysis of the parameter errors and correlations. Comput Phys Commun 1975; 10:343–67.

Kannel WB, Hubert H, Lew EA. Vital capacity as a predictor of cardiovascular disease: The Framingham Study. Am Heart J 1983;105:311–315.

Kathren RL, Petersen GR. Units and terminology of radiation measurements: a primer for the epidemiologist. Am J Epidemiol 1989; 130:1076–1087.

Kilburn KH. Cigarette smoking does not produce or enhance the radiological appearance of pulmonary fibrosis. Am J Ind Med 1981; 2:305–308.

Kilburn KH, Lilis R, Anderson HA, Miller A, Warshaw RH. Interaction of asbestos, age, and cigarette smoking in producing radiological evidence of diffuse pulmonary fibrosis. Am J Med 1986; 80:377–81.

Kjuus H, Langard S, Skjaerven R. A case–control study of lung cancer, occupational exposures and smoking. III Etiological fraction of occupational exposures. Scand J Work Environ Health 1986;12:210–15.

Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic Research: Principles and Quantitative Methods. Belmont, California, Lifetime Learning

Publications, 1982.

Koopman JS. Interaction between discrete causes. Am J Epidemiol 1981: 113:716–724.

Kreiss K, Greenberg LM, Kogut SJH, Legotte DC, Irvin CG, Cherniack RM. Hard rock mining exposure affect smokers and nonsmokers differently. Am Rev Respir Dis 1987; 136:1487–1493.

Kupper LL, Hogan MD. Interaction in epidemiologic studies. Am J Epidemiol 1978;108:447–453.

Kurppa K,Gudbergsson H, Hannunkari I, Koskinen H. Hernberg S, Koskela R S, Ahlman K. Lung cancer among silicotics in Finland. In: Silica, Silicosis, and Cancer: Controversy in Occupational Medicine (Editors: Goldsmith DF, Winn DM, Shy CM). Cancer Research Monographs, Vol 2. Publ. Praeger, New York, 1986, pp. 311–119.

Leigh J, Wiles An, Glick M. Total population study of factors affecting chronic bronchitis prevalence in the coal mining industry of New South Wales, Australia. Br J Ind Med 1986;43:263–71.

Liddell FDK, McDonald JC, Thomas DC. Methods of cohort analysis: Appraisal by application to asbestos mining. J Roy Statist Soc (A) 1977;140:469–491.

Liddell FDK, Thomas DC, Gibbs GW, McDonald JC. Fibre exposure and mortality from pneumoconiosis, respiratory, and abdominal malignancies in chrysotile production in Quebec. 1926–1975. Singapore Ann Acad Med 1984; 13 (suppl):339–344.

Love RG, Miller BG. Longitudinal study of lung function in coal miners. Thorax; 37:193–197.

Lubin JH, Gaffey W. Relative risk models for assessing the joint effects of multiple factors. Am J Ind Med 1988; 13:149–167.

Manfreda J, Sidwall G, Maini K, West P, Cherniack RM. Respiratory

abnormalities in employees of the hard rock mining industry. Am Rev Respir Dis 1982; 126:626–634.

Mantel N and Haenszel W. Statistical aspects of analysis of data from retrospective studies of disease. J Nat Cancer Inst 1959; 22(4):719–748.

Mantel N. Chi–square tests with one degree of freedom: extentions of the Mantel–Haenszel procedure. J Am Stat Assoc 1963;58:690–700.

Marine WM, Gurr D, Jacobsen M. Clinically important respiratory effects of dust exposure and smoking in British coal miners. Am Rev Respir Dis 1988;137:106–12.

McCullagh P, Nelder JA. Generalized linear models. London: Chapman and Hall 1983:72–101.

McCusker K. Landmarks of tobacco use in the United States. Chest 1988; 93: (Suppl Feb) 34S–35S.

McDonald JC, Gibbs GW, Liddell FDK, and McDonald AD. Mortality after long exposure to Cummingtonite–grunerite. Am Rev Respir Dis 19788;

McDonald JC. Editorial. Silica, silicosis and lung cancer. Br J Ind Med 1989; 46:289–291.

McNulty JC. The prevalence of respiratory symptoms in Western Australian gold miners compared with coal miners. In: Proceedings of the first Australian Pneumoconiosis Conference, 1968. Sydney: University of Sydney 1968:411–427.

Miettinen OS. Proportion of disease caused or prevented by a given exposure, trait or intervention. Am J Epidemiol 1974;99:325–31.

Miettinen OS. Estimability and estimation in case–referent studies. Am J Epidemiol 1976; 103:226–235.

Miettinen OS. Causal and preventive interdependence. Scan J Work Environ Health 1982; 8:159–168.

Miettinen OS. Theoretical epidemiology. John Wiiley & Sons. New York. NY 1985.

Moolgavkar SH, Venzon DJ. General relative risk regression models for matched case–control studies. Am J Epidemiol 1987;126:949–961.

Morgan WKC. On dust, disability, and death. Am Rev Respir Dis 1986;134:639–41.

Moolgavkar SH, Venzon DJ. General relative risk regression models for epidemiological studies. Am J Epidemiol 1987; 126:949–961.

Morgan WKC. Industrial bronchitis. Br J Ind Med 1978; 35:285–291.

Morgan WKC. Industrial bronchitis and other nonspecific conditions affecting the airways. In: Morgan WKC, Seaton A. Occupational lung disease, 2nd ed. Philadelphia: WB Saunders CO, 1984:521–540.

Morgan WKC. On dust, disability, and death. Am Rev Respir Dis 1986; 134:639–641.

Morgan WKC. Update on lung disease in coal miners. (Correspondence) Br J Ind Med 1987;44:716–720.

Morgan WKC, Lapp NL. Clinically important respiratory effects of dust exposure and smoking in British coal miners. Am Rev Respir Dis 1988;138:1643–1644.

Mossman BT, Craighead JE. Mechanism of asbestos carcinogenesis. Environ Res 1981; 25:260–280.

Mullen JBM, Wright JL, Wiggs BR, Pare PD, Hogg JC. Structure of central airways in current smokers and ex–smokers with and without mucus hypersecretion: relationship to lung function. Thorax 1987;42:843–848.

Muller J, Wheeler W, Gentleman J, Suranyi G, and Kusiak R. Study of mortality of Ontario miners: 1955–1977, Part I. Ontario Ministry of Labour, Toronto 1983.

Olofson J, Skoogh B–E, Bake B, Svardsudd K. Mortality related to smoking habits, respiratory symptoms and lung function. Eur J Respir Dis 1987; 71:69–76.

Pearce NE, Checkoway H. A simple computer program for generating person–time data in cohort studies involving time–related data factors. Am J Epidemiol 1987; 125:1085–91.

Pearce NE, Checkoway H, Dement K. Exponential models for analyses of time related factors, illustrated with asbestos textile workers mortality data. Journal of Occupational Medicine 1988; 30:517–522.

Pearce NE. Incidence density matching with a simple SAS computer program. Int J Epidemiol 1989; 18:981–984.

Pepys J. Occupational asthma: an overview. J Occ Med 1982; 24:534–36.

Peto R. Epidemiology, multistage models, and short–term mutagenicity tests. In Hiatt HH, Watson JD, Winsten JA (Eds). Origins of Human Cancer. Cold Spring Harbor: Cold Spring Harbor Laboratory,1977, pp. 1403–1428.

Peto R, Pike MC, Armitage P, Breslow NE, Cox DR. Howard SV, Mantel N, McPherson K, Peto J, Smith PG. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. Br J Cancer 1977;35:1–39.

Peto R, Speizer FE, Cochrane AL, Moore F, Fletcher CM, Tinker CM, Higgins ITT, Gray RG, *et al.*. The relevance in adults of air–flow obstruction, but not mucus hypersecretion, to mortality from chronic lung disease. Am Rev Respir Dis 1983;128:491–500.

Prentice RL, Self SG. Asymptotic distribution theory for Cox–type regression models with general relative risk form. Ann Stat 1983;11:804–813.

Prentice RL and Breslow NE. Retrospective studies and failure time models. Biometrica 1978; 65:153–158.

Pride N. Definitions of emphysema, chronic bronchitis, asthma and airflow obstruction: 25 years on from Ciba symposium. Thorax 1984; 39:81–85.

Prowse CM. Aspects of pulmonary function in silicosis in South African gold miners. In: Shapiro HA. (ed.). Pneumoconiosis: Proceedings of the International Conference, Johannesburg 1969. Cape Town, Oxford University Press, 1970, pp. 508–511.

Robins JM, Gail MH, Lubin JH. More on 'biased selection of controls for case–control analyses of cohort studies'. Biometrics 1986; 42:293–299.

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 coal miners. Br J Ind Med 1973; 30:217–226.

Rothman KJ. Synergy and antagonism in cause–effect relationships. Am J Epidemiol 1974;99:385–88.

Rothman KJ. The estimation of synergy or antagonism. Am J Epidemiol 1976;103:506–511.

Rothman KJ. Occam's razor pares the choice among statistical models. Am J Epidemiol 1978;108:347–349.

Rothman KJ, Greenland S, Walker AM. Concepts of interaction. Am J Epidemiol 1980; 112:467–470.

Saffiotti U. The pathology induced by silica in relation to fibrogenesis and carcinogenesis. In: Silica, Silicosis, and Cancer: Controversy in Occupational Medicine (Editors: Goldsmith DF, Winn DM, Shy CM). Cancer Research Monographs, Vol 2. Publ. Praeger, New York, 1986, pp. 375–384.

Saracci R. Asbestos and lung cancer: An analysis of the epidemiological evidence of the asbestos–smoking interaction. Int J Cancer 1977;20:323–331.

Saracci R. Interaction and synergism. Am J Epidemiol 1980;112:465–466.

SAS Institute Inc. Sugi Supplemental Library User's Guide, Version 5 Edition. Cary, NC:SAS Institute Inc.,1986. 575 pp.

Schuler G and Ruttner JR. Silicosis and lung cancer in Switzerland. In: Silica, Silicosis, and Cancer: Controversy in Occupational Medicine (Editors: Goldsmith DF, Winn DM, Shy CM). Cancer Research Monographs, Vol 2. Publ. Praeger, New York, 1986, pp. 357–366.

Selikoff IJ, Lee DHK. Asbestos and Disease. New York, Academic Press, 1978, 549 pp.

Selikoff I, Sedman H, Hammond E. Mortality effects of cigarette smoking among amosite asbestos factory workers. JNCI 1980; 65:507–513.

Siemiatycki J, Thomas DC. Biological models and statistical interactions: An example from multistage carcinogenesis. Int J Epidemiol 1981;10:383–387.

Sluis–Cremer GK, Walters LG, Sichel HS. Chronic bronchitis in miners and non–miners: An epidemiological survey of a community in the gold–mining area in the Transvaal. Br J Ind Med 1967; 24:1–12.

Sluis–Cremer GK, Walters LG, Sichel HS. Ventilatory function in relation to mining experience and smoking in a random sample of miners and non–miners. Br J Ind Med 1967; 24:13–25.

Sluis–Cremer GK. Silica exposure and silicosis in Witwatersrand gold miners in South Africa. In: Silica, Silicosis, and Cancer: Controversy in Occupational Medicine (Editors: Goldsmith DF, Winn DM, Shy CM). Cancer Research Monographs, Vol 2. Publ. Praeger, New York, 1986, pp. 67–70.

Sluis–Cremer GK, Hessel PA, Hnizdo E. Factors influencing the reading of small irregular opacities in a radiological survey of asbestos miners in South Africa. Arch Environ Health 1989; xxx–xxx.

Snider GL. Chronic obstructive pulmonary disease– a continuing challenge. Am Rev Respir Dis 1986; 133:942–944.

Soutar CA, Hurley JF. Relation between dust exposure and lung function in miners and ex–miners. Br J Ind Med 1986; 43:307–320.

Steenland K, Thun M. Interaction between tobacco smoking and occupational exposures in the causation of lung cancer. J Occup Med 1986; 28:110–118.

Thomas DC. Program "RISK" User's Manual (1980). Montreal:McGill University, 1980.

Thomas DC. General relative–risk models for survival time and matched case–control analysis. Biometrics 1981;37:673–686.

Thomas DC, Whittemore A. Interaction between tobacco use and workplace exposure in determining disease risk. Am J Ind Med 1988;13:131–148.

Thurlbeck WM, Henderson JA, Fraser RG, Bates DV. Chronic obstructive lung disease. A comparison between clinical, roentgenologic, functional and morphologic criteria in chronic bronchitis, emphysema, asthma, and bronchietasis. Medicine 1970; 49:81–1945.

U.S. Department of Health and Human Services. The Health Consequence of Smoking: Cancer. A Report of the Surgeon General. U.S. Department of Health and Human Services, Public Health Service, Office on Smoking and Health. DHHS Pub. No. (PHS) 82–50179, 1982, 322 pp.

U.S. Department of Health and Human Services. The Health Consecuence of Smoking: Chronic Obstructive Lung Disease. A Report of the Surgeon General. U.S. Department of Health and Human Services, Public Health Service, Office on Smoking and Health. DHHS Pub. No. (PHS) 84–50205. 1984, 568 pp.

U.S. Department of Health and Human Services. The Health Consequence of Smoking: Cancer and Chronic Lung Disease in the Workplace. A Report of the Surgeon General. U.S. Department of Health and Human Services, Public Health Service, Office on Smoking and Health. DHHS Pub. No.

(PHS) 84–50205, 1985, 542 pp.

Walter SD, Holford TD. Additive, multiplicative and other models of disease risks. Am J Epidemiol 1978;108:341–346.

Walker AM. Proportion of disease attributable to the combined effect of two factors. Int J Epidemiol 1981; 10:81–85.

Walker AM, Rothman KJ. Model of varying parametric form in case–referent studies. Am J Epidemiol 1982;115:129–137.

Walter SD, Holford TR. Additive, multiplicative, and other models for disease risks. Am J Epidemiol 1978; 108:341–346.

Weed DL, Selmon M, Sinks T. Links between categories of interaction. Am J Epidemiol 1988; 127:17–27.

Weiss W. Cigarette smoking and diffuse pulmonary fibrosis. Am Rev Respir Dis 1969; 99:67–72.

Weiss W. Cigarette smoke, asbestos, and small irregular opacities. Am Rev Respir Dis 1984; 130:293–301.

Wiles FJ, Faure MH. Chronic obstructive lung disease in gold miners. In: Walton, WH (ed.). Inhaled Particles IV, Part 2. Oxford, Pergamon Press, 1977. pp. 727–735.

Wiles FJ, Hnizdo E. Relevance of airflow obstruction and mucus hypersecretion to mortality. Respiratory Medicine 1991; 85:27–35.

Wiles FJ, Baskind E. Hessel PA. Hnizdo E. Lung function in silicosis. Ann Occup Hyg (submitted).

Whittemore AS. Estimating attributable risk from case–control studies. Am J Epidemiol 1983; 117:76–85.

Whittemore AS, McMillan A. Lung cancer mortality among US uranium miners: A reappraisal. J Natl Cancer Inst 1983;71:489–499.

Wright JL, Harrison N, Churg A. Quartz but not iron oxide cause airflow

obstruction, emphysema, and small airway lesions in the rat. Poster Session. Am Rev Respir Dis 1988;137:49.

Wyndham CH. 1973 Yant Memorial Lecture. Research in the human sciences in the gold mining industry. Am Indust Hygiene Assoc J 1974; 35:113–136.

Wyndham CH, Heyns AJ, Keeley WH *et al.* Lung cancer rates among white South African miners. Johannesburg: Chamber of Mines, 1970. (Internal COM research report No 23/70).

Wyndham CH, Greenacre MJ, Wiles FJ. Effects of personal, environmental and occupational factors on ischaemic heart disease in white miners in South Africa. International Journal of Epidemiology 1986; 15:507–512.

Wyndham CH, Bezuidenhout BN, Greenacre MJ, Sluis–Cremer GK. Mortality in middle aged South African gold miners. Br J Ind Med 1986; 43:677–684.

Zambon P, Simonato L, Mastrangelo G, Winkelmann R, Saia B, Crepet M. Mortality of workers compensated for silicosis during the period 1959–1963 in the Veneto region of Italy. Scan J Work Environ Health 1987; 13:118–123.

Zwi S. Lung function of Witwatersrand gold miners. In: Orenstein AJ, ed. Proceedings of the pneumoconiosis conference Johannesburg 1959. London: Churchill, 1960; 522–530.

APPENDIX A

This appendix contains tables showing the parameter estimates and the standard errors obtained from the general relative risk model fitted to the dichotomous outcome variables for simple chronic bronchitis and abnormal sputum production, as described in Chapter 4.

Table A.1. Simple chronic bronchitis – Carletonville study.
Parameter estimates (E) and standard errors (± SE) estimated from the GRRM, Eq.(4.1)

Effect	P	Relative Risk Model[a]				
		1	2	3	4	5
Int.	α	−4.92±0.12	−4.85±0.13	−4.58±0.50	−4.27±0.41	−4.30±0.29
Age2	α_1	0.38±0.23	0.45±0.21	0.44±0.28	0.45±0.28	0.45±0.28
Age3	α_2	0.76±0.33	0.99±0.31	0.88±0.37	0.90±0.37	0.90±0.37
Dust	β_1	4.90±0.17	2.78±1.29	1.02±0.35	0.47±0.14	0.90±0.52
Smk	β_2	8.76±0.28	3.98±1.36	1.45±0.39	0.73±0.19	2.02±0.80
D·S	β_3		7.23±2.22			
	λ				2.25±0.15	2.19±0.90
−2lnL		451.96	446.72	446.2	445.91	445.96

[a] 1= additive model, 2= additive model with interaction, 3= multiplicative model, 4 and 5= power models

Table A.2 Simple chronic bronchitis – Wiles and Faure study.
Parameter estimates (E) and standard errors (± SE) estimated from the GRRM, Eq.(4.2), with dichotomous exposure variables

	Relative risk model		
	RR_a	RR_m	RR_p
Effect	E ± S.E.	E ± S.E.	E ± S.E.
Intercept	−1.029±0.015	−4.698±1.018	−4.303±0.347
AGEC	0.058±0.025	0.318±0.090	0.318±0.090
DUST	0.139±0.033	2.261±1.025	2.305±0.483
SMOKE	0.085±0.030	1.650±1.025	1.645±0.500
DUST·SMOKE	0.395±0.049	−0.187±1.043	λ =0.92
−2lnL	3019.8	3014.0	3013.9

Table A.3 Simple chronic bronchitis – Wiles and Faure study.
Parameter estimates (E) and standard errors (± SE) estimated from the GRRM, Eq. (4.2), with individual exposure variables

a) continuous dust exposure variable (RSA x years) (CDE)

	Relative risk model		
	RR_a	RR_m	RR_p
Effect	E ± S.E.	E ± S.E.	E ± S.E.
Intercept	–0.943±0.015	–2.065±0.163	–2.940±0.044
AGEC	0.017±0.001	0.196±0.089	0.420±0.055
CDE	0.043±0.025	0.036±0.003	0.649±0.064
λ			–0.244±0.016
–2lnL	3082.6	3135.0	3075.8

Hypotheses tested:
H: (λ=0) vs (λ=1) X^2_{LR}=52.4; p<0.0001
H: (λ=0) vs ($\lambda \neq 0$) X^2_{LR}= 6.8; p<0.01

b) smoking habits (SMOKEC)

	Relative risk model		
	RR_a	RR_m	RR_p
Effect	E ± S.E.	E ± S.E.	E ± S.E.
Intercept	–1.049±0.037	–2.766±0.190	–2.580±0.141
AGEC	0.091±0.026	0.355±0.089	0.363±0.087
SMOKEC	0.169±0.011	0.581±0.045	0.960±0.241
λ			0.484±0.208
–2lnL	3095.9	3085.0	3082.2

Hypotheses tested:
H: (λ=0) vs (λ=1) X^2_{LR}= 10.9; p<0.001
H: (λ=0) vs ($\lambda \neq 0$) X^2_{LR}= 13.7; p<0.001

Table A.4 Simple chronic bronchitis– Wiles and Faure study.
Parameter estimates (E) and standard errors (± SE) estimated from the GRRM, Eq. (4.2), with categorized or continuous exposure variables

a) for gold mining variable (CDEC) and smoking variable (SMOKEC)

	Relative risk model			
	RR_a	RR_a	RR_m	RR_p
Effect	E ± S.E.	E ± S.E.	E ± S.E.	E ± S.E.
Inter.	−1.320±0.026	−1.130±0.001	−3.790±0.315	−6.525±0.045
AGEC	0.107±0.028	0.055±0.001	0.234±0.092	0.447±0.058
CDEC	0.016±0.001	0.007±0.001	0.043±0.008	0.568±0.023
SMOKEC	0.166±0.012	0.076±0.001	0.667±0.100	7.919±0.266
CDEC·SMOKEC		0.005±0.001	−0.002±0.003	λ=0.235
−2lnL	2912.2	2881.4	2919.3	2906.5

b) for gold mining variable (CDEC) and smoking variable (PACKYR)

	Relative risk model			
	RR_a	RR_a	RR_m	RR_p
Effect	E ± S.E.	E ± S.E.	E ± S.E.	E ± S.E.
Inter.	−1.256±0.025	−1.023±0.018	−3.435±0.157	−3.907±0.045
AGEC	0.074±0.023	0.038±0.019	0.156±0.092	0.146±0.057
CDEC	0.018±0.001	0.007±0.001	0.035±0.005	0.218±0.009
PACKYR	0.012±0.001	0.004±0.001	0.021±0.004	0.138±0.009
CDEC·PACKYR		0.001±0.0001	0.0002±0.0001	λ=0.250
−2lnL	2959.2	2876.0	2993.0	2952.2

CDEC=RSA x years/1000.

Table A.5 Simple chronic bronchitis– Wiles and Faure study.
Parameter estimates (E) and standard errors (± SE) estimated from the Thomas' modified model, R_{MT}

	Relative risk model			
	1	2	3	4
Effect	E ± S.E.	E ± S.E.	E ± S.E.	E ± S.E.
Inter.	−1.146±0.017	−1.263±0.020	−1.288±0.021	−1.784±0.009
AGEC	0.065±0.012	0.092±0.024	0.098±0.025	0.066±0.012
CDEC	0.008±0.001	0.013±0.001	0.014±0.005	0.014±0.001
SMOKEC	0.081±0.007	0.138±0.009	0.150±0.009	0.118±0.004
λ	6.92	1.0	0.5	0.40
−2lnL	2881.4	2896.2	2902.4	2902.1

CDEC=RSA x years/1000.

Table A.6 Abnormal sputum production– Wiles and Faure study.
Parameter estimates (E) and standard errors (± SE) from the GRRM, Eq.(4.2), with dichotomous exposure variables

	Relative risk model		
	RR_a	RR_m	RR_p
Effect	E ± S.E.	E ± S.E.	E ± S.E.
Intercept	–0.963±0.026	–3.169±0.604	–2.447±0.041
AGEC	0.070±0.043	0.160±0.082	–0.037±0.054
DUST	0.170±0.049	1.703±0.611	1.300±0.068
SMOKE	0.331±0.048	2.184±0.600	1.669±0.065
DUST·SMOKE	0.213±0.074	–1.045±0.622	λ=0.667
–2lnL	3461.1	3460.0	3462.0

Table A.7 Abnormal sputum production – Wiles and Faure study.
Parameter estimates (E) and standard errors (± SE) from the GRRM, Eq.(4.2), with individual exposure variables

	Exposure variable	
	RR_p	RR_p
Effect	E ± S.E.	E ± S.E.
Intercept	−1.140±0.112	−1.766±0.041
AGEC	0.107±0.081	0.192±0.054
exposure	0.072±0.034[a]	0.692±0.026[b]
λ	−0.778±0.282	0.527±0.050
−2lnL	3533.8	3411.8

[a] CDE=RSA x years, [b] SMOKEC

Table A.8 Abnormal sputum production – Wiles and Faure study.
Parameter estimates (E) and standard errors (± SE) from the GRRM, Eq.(4.2), with categorized exposure variables

	Relative risk model			
	RR_a	RR_a	RR_m	RR_p
Effect	E ± S.E.	E ± S.E.	E ± S.E.	E ± S.E.
Inter.	−0.978±0.015	−0.969±0.020	−2.321±0.208	−2.258±0.041
AGEC	0.023±0.032	0.047±0.036	0.128±0.084	0.122±0.054
CDEC	0.008±0.001	0.005±0.001	0.027±0.007	0.026±0.002
SMOKEC	0.209±0.016	0.160±0.021	0.582±0.076	0.659±0.023
CDEC·SMOKEC		0.003±0.001	−0.004±0.003	λ=0.702
−2lnL	3357.8	3348.6	3350.4	3349.6