

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Lucy Katherine Smith BSc, MSc (Leicester) MA (Warwick)
Department of Ophthalmology
University of Leicester

November 1997

UMI Number: U116396

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U116396

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS

I would like to thank my supervisor, John Thompson, for all of his help and guidance throughout my studies. I am grateful to the University of Leicester Department of Ophthalmology particularly Geoffrey Woodruff, Fiona Hiscox and John Thompson for their hard work in carrying out the 1983 multi-centre amblyopia study and allowing me access to the data for my thesis. I appreciate the work of Diane Oliver, who made diagnoses for all of the children in the 1992 amblyopia study, and Bharti Patel, who helped me collect the data for this study and conducted all of the interviews. I gratefully acknowledge the Anne Allerton fund for funding Bharti Patel. I would also like to thank Elizabeth Draper and Micheal Clarke of the Department of Epidemiology and Public Health for allowing me access to data from the perinatal mortality study. Finally, I am grateful to Paul Lambert for his support throughout. .

Material from Crown-copyright records was made available through the Post Office and the ESRC data archive.

TABLE OF CONTENTS

CHAPTER 1 : INTRODUCTION

1.1 AIMS OF THE THESIS	1
1.2 INEQUALITIES IN HEALTH	1
1.3 DEPRIVATION MEASUREMENT	2
1.4 AMBLYOPIA AND VISION SCREENING	2
1.5 THE EFFECT OF DEPRIVATION ON AGE AT PRESENTATION OF AMBLYOPIA	3
1.6 EVALUATION OF CHANGES MADE TO A VISION SCREENING PROGRAMME	3
1.7 COMPARING AREA AND INDIVIDUAL DEPRIVATION MEASURES	3
1.8 ERROR IN THE MEASUREMENT OF DEPRIVATION	4
1.9 CONCLUSIONS	4

CHAPTER 2 : INEQUALITIES IN HEALTH

2.1 AIMS OF THE CHAPTER	6
2.2 INTRODUCTION	6
2.3 THE CONCEPTS	7
2.4 THE BLACK REPORT	9
2.4.1 THE BACKGROUND TO THE BLACK REPORT	9
2.4.2 THE MESSAGE OF THE BLACK REPORT	9
2.4.3 METHODOLOGICAL PROBLEMS IN THE BLACK REPORT	10
2.5 EXPLANATIONS	11
2.5.1 ARTEFACT EXPLANATIONS	11
2.5.2 NATURAL OR SOCIAL SELECTION EXPLANATIONS	12
2.5.3 CULTURAL OR BEHAVIOURAL EXPLANATIONS	12
2.5.4 MATERIALIST OR STRUCTURALIST EXPLANATIONS	13
2.6 FURTHER EVIDENCE	14
2.6.1 ETHNICITY	15
2.7 INEQUALITIES IN CHILD HEALTH AND PREVENTIVE SERVICES	16
2.8 SOLUTIONS AND GOVERNMENTAL POLICY	18
2.8.1 GENERAL RECOMMENDATIONS	18
2.8.2 CHILD HEALTH AND PREVENTIVE SERVICES	20
2.9 FUTURE RESEARCH AND INFORMATION NEEDS	21
2.10 SUMMARY AND CONCLUSIONS	22

CHAPTER 3 : MEASUREMENT OF DEPRIVATION

3.1 AIMS OF THE CHAPTER	24
3.2 INTRODUCTION	24
3.3 SOCIAL CLASS	25
3.4 CONSTRUCTING ALTERNATIVE MEASURES	27
3.4.1 THE CENSUS	27
3.4.2 CONCEPTUALISING POVERTY	28
3.4.3 DEVELOPING A MEASURE OF DEPRIVATION	28
3.4.4 COMBINING INDICATORS	30
3.4.5 THE PROBLEM OF ECOLOGICAL FALLACY	31
3.4.6 THE PROBLEM OF UNDERENUMERATION IN THE 1991 CENSUS	33

3.5 THE JARMAN SCORE.....	34
3.5.1 AN OVERVIEW OF THE JARMAN SCORE.....	34
3.5.2 THE RESPONSE TO THE JARMAN SCORE.....	34
3.5.3 FURTHER WORK.....	35
3.6 UNIDIMENSIONAL INDICATORS	35
3.6.1 A COMPARISON OF THE SCORES	35
3.6.2 CONCLUSIONS ON CHOOSING A UNIDIMENSIONAL SCORE.....	36
3.7 THE TOWNSEND SCORE.....	37
3.7.1 INTRODUCTION	37
3.7.2 CALCULATION OF THE TOWNSEND SCORE.....	38
3.7.3 PROBLEMS WITH USING THE TOWNSEND SCORE	39
3.8 MULTI-DIMENSIONAL INDICATORS.....	39
3.9 SUMMARY AND CONCLUSIONS	42

CHAPTER 4: PATTERNS OF DEPRIVATION IN LEICESTERSHIRE

4.1 AIMS OF THE CHAPTER.....	44
4.2 POPULATION OF LEICESTERSHIRE.....	44
4.2.1 POPULATION DENSITY	44
4.2.2 ETHNICITY.....	46
4.3 DEPRIVATION IN LEICESTERSHIRE.....	46
4.3.1 MATERIAL DEPRIVATION INDICATORS IN LEICESTERSHIRE	47
4.3.2 COMPARING DEPRIVATION SCORES.....	48
4.4 SUMMARY AND CONCLUSIONS	58

CHAPTER 5 : AMBLYOPIA AND VISION SCREENING

5.1 AIMS OF THE CHAPTER.....	59
5.2 WHAT IS AMBLYOPIA ?.....	59
5.2.1 AMBLYOPIA AND VISUAL DEVELOPMENT.....	59
5.3 PRESENTATION AND DIAGNOSIS OF AMBLYOPIA.....	61
5.4 SCREENING FOR AMBLYOPIA	62
5.4.1 PUBLIC HEALTH IMPORTANCE	63
5.4.2 AVAILABILITY OF TREATMENT	63
5.4.3 DIAGNOSIS AND TREATMENT EFFICACY	64
5.4.4 LATENT OR EARLY SYMPTOMATIC STAGE OF THE CONDITION	65
5.4.5 TESTS FOR DETECTING THE CONDITION.....	65
5.4.6 NATURAL HISTORY OF THE CONDITION.....	67
5.4.7 DEFINITION OF A CASE	68
5.4.8 EARLY TREATMENT SHOULD IMPROVE OUTCOME.....	68
5.4.9 COST OF SCREENING	68
5.4.10 SCREENING SHOULD BE A CONTINUOUS PROCESS	68
5.4.11 SUMMARY OF SCREENING REQUIREMENTS.	69
5.4.12 CURRENT RECOMMENDATION FOR AMBLYOPIA SCREENING	69
5.4.13 VISION SCREENING IN LEICESTERSHIRE	70
5.5 SUMMARY AND CONCLUSIONS	71

CHAPTER 6 : AGE AT PRESENTATION IN AMBLYOPIA AND DEPRIVATION: ANALYSIS OF A MULTI-CENTRE STUDY

6.1 AIMS OF THE CHAPTER.....	72
6.2 PRESENTATION OF AMBLYOPIA.....	72
6.3 BACKGROUND TO THE MULTI-CENTRE STUDY	73
6.3.1 REASONS FOR THE STUDY.....	73
6.3.2 STUDY DESIGN.....	74
6.4 PRELIMINARY ANALYSIS	75
6.4.1 CLASSIFICATION OF AMBLYOPIA	75
6.4.2 CLINICS	76
6.4.3 TOWNSEND DEPRIVATION SCORE.....	76
6.4.4 OTHER EXPLANATORY VARIABLES	80
6.5 REGRESSION ANALYSIS	82
6.5.1 MODEL INTERPRETATION	88
6.5.2 ALTERNATIVE MODELS	90
6.6 DISCUSSION OF RESULTS.....	91
6.7 IMPLICATIONS OF THE FINDINGS.....	93
6.8 OTHER ASPECTS OF AMBLYOPIA	94
6.9 SUMMARY AND CONCLUSIONS	95

CHAPTER 7 : DESIGN OF A STUDY TO INVESTIGATE THE EFFECT OF CHANGES TO VISION SCREENING IN LEICESTERSHIRE

7.1 AIMS OF THIS CHAPTER.....	97
7.2 STUDY OBJECTIVES	97
7.2.1 AGE AT PRESENTATION AND DEPRIVATION	97
7.2.2 AREA AND INDIVIDUAL DEPRIVATION.....	99
7.2.3 FURTHER ASPECTS.....	100
7.3 STUDY DESIGN	100
7.3.1 OVERVIEW OF STUDY DESIGN.....	100
7.3.2 SAMPLE SIZE AND POWER CALCULATIONS	101
7.3.3 ETHICAL CONSIDERATIONS	104
7.3.4 QUESTIONNAIRE DESIGN.....	105
7.3.5 METHODS AND TIMETABLE	106
7.4 SUMMARY AND CONCLUSIONS	108

CHAPTER 8 : ASSESSING THE EFFECT OF SCREENING CHANGES ON AGE AT PRESENTATION OF AMBLYOPIA

8.1 AIMS OF THE CHAPTER.....	109
8.2 HYPOTHESIS	110
8.3 STUDY DATA.....	110
8.3.1 DEPRIVATION.....	111
8.4 PRELIMINARY RESULTS	112
8.5 REGRESSION ANALYSES	113
8.5.1 MICROTROPIA OR NO STRABISMUS.....	113
8.5.2 LARGE ANGLE STRABISMUS	120
8.6 DISCUSSION OF REGRESSION ANALYSIS	122
8.7 SUMMARY AND CONCLUSIONS	127

CHAPTER 9 : ASSESSING THE PROBLEM OF ECOLOGICAL FALLACY

9.1 AIMS OF THE CHAPTER.....	129
9.2 SAMPLES OF ANONYMISED CENSUS RECORDS.....	130
9.2.1 DESCRIPTION OF THE DATA.....	132
9.2.2 COMPARISON OF INDIVIDUAL AND AREA DATA	133
9.2.3 DEPRIVATION AND LIMITING LONG TERM ILLNESS	135
9.2.4 REGRESSION ANALYSES	139
9.2.5 SENSITIVITY ANALYSES	143
9.2.6 INTERPRETATION OF REGRESSION ANALYSIS	144
9.2.7 ETHNICITY AND DEPRIVATION	147
9.2.8 CONCLUSIONS OF SAR ANALYSES.....	149
9.3 PERINATAL MORTALITY DATA.....	152
9.3.1 DESCRIPTION OF THE DATA.....	152
9.3.2 ASSESSING AGREEMENT	153
9.3.3 DEPRIVATION AND MEASURES OF CHILD HEALTH	155
9.3.4 SENSITIVITY ANALYSIS	158
9.3.5 CONCLUSIONS OF ANALYSES OF PERINATAL MORTALITY DATA	160
9.4 SUMMARY AND CONCLUSIONS	161

CHAPTER 10 : COMPARING THE EFFECT OF DEPRIVATION AT INDIVIDUAL AND AREA LEVEL ON THE AMBLYOPIA STUDY RESULTS

10.1 AIMS OF THE CHAPTER.....	163
10.2 DATA COLLECTED	164
10.3 INFORMATION ON INDIVIDUAL DEPRIVATION	164
10.4 COMPARING INDIVIDUAL AND AREA LEVEL DATA	164
10.5 ANALYSIS OF AGE AT PRESENTATION AND DEPRIVATION.....	167
10.6 DISCUSSION OF THE FINDINGS	168
10.7 SUMMARY AND CONCLUSIONS	169

CHAPTER 11 : ADJUSTING THE EFFECT OF DEPRIVATION FOR MEASUREMENT ERROR

11.1 AIMS OF THE CHAPTER.....	171
11.2 MEASUREMENT ERROR	172
11.3 THE EFFECTS OF MISCLASSIFICATION.....	173
11.3.1 NON-DIFFERENTIAL MISCLASSIFICATION	173
11.3.2 DIFFERENTIAL MISCLASSIFICATION.....	174
11.3.3 ADJUSTING FOR MISCLASSIFICATION.....	175
11.3.4 USE OF RELATIVE RISK AND ODDS RATIO	175
11.3.5 TERMINOLOGY	176
11.4 FLEGAL ET AL'S METHOD.....	177
11.5 ESPELAND AND HUI'S METHOD.....	180
11.5.1 THE ESPELAND AND HUI MODEL APPLIED TO SAR DATA.....	182

11.6 EKHOLM'S METHOD	184
11.6.1 PARAMETERISATION OF MODEL.....	184
11.6.2 DESCRIPTION OF EXPONENTIAL FAMILY NON-LINEAR MODELS	186
11.6.3 FITTING THE MODEL IN GLIM.....	187
11.7 SUMMARY OF RESULTS FROM THE THREE METHODS	189
11.8 NONDIFFERENTIAL MISCLASSIFICATION ASSUMPTION	190
11.9 DISCUSSION	196
11.10 SUMMARY AND CONCLUSIONS	198

CHAPTER 12 : APPLICATION OF MISCLASSIFICATION ADJUSTMENT METHODS TO THE AMBLYOPIA DATA

12.1 AIMS OF THE CHAPTER.....	199
12.2 ANALYSIS OF AMBLYOPIA DATA	200
12.3 RESULTS OF ADJUSTING FOR MISCLASSIFICATION	202
12.4 DISCUSSION OF RESULTS	204
12.4.1 DIFFERENCES IN MISCLASSIFICATION BETWEEN TYPES OF AMBLYOPIA.....	204
12.4.2 DIFFERENCES IN MISCLASSIFICATION AND AGE AT PRESENTATION	206
12.4.3 DIFFERENCES IN MISCLASSIFICATION BETWEEN COHORTS.....	207
12.5 SUMMARY AND CONCLUSIONS	208

CHAPTER 13 : SUMMARY AND CONCLUSIONS

13.1 AIMS OF THE CHAPTER.....	209
13.2 USING A MEASURE OF DEPRIVATION TO INVESTIGATE INEQUALITIES IN THE PRESENTATION OF AMBLYOPIA	209
13.3 EVALUATING CHANGES MADE TO VISION SCREENING	212
13.4 THE ISSUE OF ECOLOGICAL FALLACY	214
13.5 ADJUSTING FOR MISCLASSIFICATION OF DEPRIVATION	216
13.6 FURTHER APPLICATIONS	217

APPENDIX A	219
APPENDIX B	220
APPENDIX C	224
APPENDIX D	228
APPENDIX E	233
APPENDIX F	237
APPENDIX G	242
APPENDIX H	243
APPENDIX I	247
APPENDIX J	248
BIBLIOGRAPHY	250

CHAPTER 1

INTRODUCTION

1.1 AIMS OF THE THESIS

This thesis describes an investigation of inequalities in health in paediatric ophthalmology from a health services research perspective and examines some associated methodological issues. I undertake an exploration of inequalities in the presentation of amblyopia, a common childhood eye disease. An area based deprivation measure is used to investigate the relationship between age at presentation of amblyopia and deprivation of children presenting to seven orthoptic clinics. I extend this work by evaluating the effect of a screening intervention on existing inequalities at one particular clinic.

I then explore some methodological issues related to this work. The analyses of amblyopia data are based on an ecological measure of deprivation. Since there are problems in assuming that exposure effects at the area level apply at the individual level, I therefore investigate the relationship between individual and area level deprivation measures. I assess whether information about this relationship can be used to estimate of the effect of deprivation at the individual level in studies where only area level information is available.

1.2 INEQUALITIES IN HEALTH

To aid the understanding of the development and use of area deprivation measures, in chapter 2 I discuss the main concepts of inequalities in health together with possible explanations and proposed solutions in a general epidemiological context. This discussion concentrates on the Black report (DHSS, 1980) and highlights the important aspects of the wealth of work prompted by its publication with particular reference to child health and preventive services. This research shows health to vary with geography, ethnicity and gender. However in this thesis I concentrate on the variation in health with respect to socio-economic status.

Various explanations for inequalities in health are discussed, focusing on those relating to material deprivation since they explain most of the variation in health. A discussion of research needs and possible solutions to reduce inequality is made with particular reference to preventive services that have relevance to vision screening for amblyopia.

1.3 DEPRIVATION MEASUREMENT

To assess inequalities in the presentation of amblyopia, a measure of socio-economic status is needed. Epidemiological studies in the UK have traditionally used occupational social class as a proxy for wealth. In chapter 3 I discuss the problems of its use in terms of classification and interpretation.

With the availability of computerised census data there has been a deluge of work constructing various indices that are used to measure deprivation in epidemiological studies. Although these indices are commonly used to measure material deprivation, few were originally designed for this purpose. I investigate these measures and discuss the theory behind their construction and make recommendations for their use in studies of amblyopia. The Townsend score (Townsend et al, 1988), a material deprivation score on a continuous scale based on area level census data, is selected for use in subsequent analyses. A discussion of the practicalities of its use is made with respect to the analysis of data on amblyopia. The issue of ecological fallacy is highlighted where causal effects are assumed to exist at the individual level based on area level observations. This is investigated in more depth later in the thesis. Mapping techniques are performed in chapter 4 to demonstrate the pattern of deprivation in Leicestershire, the setting for some of the later research work. Various deprivation scores are mapped to show whether the ranking of areas is seriously affected by the choice of measure.

1.4 AMBLYOPIA AND VISION SCREENING

In chapter 5 I discuss the presentation of amblyopia. Amblyopia, often known as ‘lazy eye’, is the most common visual disability in children, and is a loss of vision in an otherwise healthy eye. I discuss the social significance of amblyopia and highlight issues relating to diagnosis and presentation of the disease, particularly the effectiveness of screening. Research on factors

affecting the age at presentation of amblyopia is reviewed and the need for work on inequalities in health in this area is indicated.

1.5 THE EFFECT OF DEPRIVATION ON AGE AT PRESENTATION OF AMBLYOPIA

Amblyopia due to visible misalignment of the eye in a child is usually noticed by their parents and when they seek medical help any associated amblyopia is likely to be detected. In contrast asymptomatic amblyopia offers no obvious outward signs and is usually only detected by a vision screening test. Research has suggested that many children's screening services are inequitable since they fail to serve deprived areas. This may lead to inequalities in the age at presentation of conditions that rely on screening for their detection. In chapter 6 I analyse data from a large multi-centre survey of children with amblyopia to investigate the effect of deprivation on age at presentation of both asymptomatic amblyopia and amblyopia due to a visible abnormality. The implications of the findings are then discussed.

1.6 EVALUATION OF CHANGES MADE TO A VISION SCREENING PROGRAMME

Although it is important to demonstrate inequalities in health, it is also important that, where possible, steps are taken to try and reduce them. In chapter 7, I describe changes made to the vision screening programme in Leicestershire and the study I carried out to look at the effect of deprivation before and after these changes to screening. I then present my analysis of the relationship between deprivation and age at presentation before and after the introduction of the new screening programme and discuss the implications of the results in chapter 8.

1.7 COMPARING AREA AND INDIVIDUAL DEPRIVATION MEASURES

Having demonstrated the use of deprivation scores in assessing the extent of inequalities in health and also the effect of changes to a vision screening programme on these inequalities it is

important to see whether using ecological data may be affecting the analyses. Many studies have used area level data to show relationships between health and deprivation based on scores calculated using various area level census data. However there are problems with assuming that effects observed at the area level also exist at the individual level.

In chapter 9 I use census data to explore the relationship between area and individual level deprivation to investigate the issue of ecological fallacy. The analyses use individual anonymised records from the census to assess the relationship between limiting longterm illness and ecological and individual measures of deprivation. I then extend this research by looking at the data from a study of perinatal mortality in Leicestershire that has information on individual characteristics of deprivation and, in addition, area measures of deprivation based on smaller areas than those recorded by the individual census data. In chapter 10 I then investigate the effect of using individual level data on the relationship between deprivation and age at presentation for the data from the 1992 amblyopia study.

1.8 ERROR IN THE MEASUREMENT OF DEPRIVATION

The use of ecological data on deprivation evidently leads to measurement error. Consequently the area deprivation score assigned to an individual may misclassify the deprivation seen at the individual level. The effects of this are demonstrated in chapter 9. Exposure misclassification has been shown to lead to bias in the estimation of the exposure effect. Methods for adjusting for exposure misclassification have been developed for use in other research areas. In chapter 11 I adapt several of these methods in order to apply them to my work. I look at whether it is possible to adjust relative risk estimates for deprivation obtained using area data to estimate the relative risk of deprivation at the individual level in other studies using census data. This work on adjustment for misclassification of deprivation is then applied to the data from the amblyopia studies and the results are discussed.

1.9 CONCLUSIONS

In this thesis I investigate the practical use of deprivation measures in undertaking a piece of health services research to look at inequalities in the presentation of amblyopia. Although I

describe the application of a deprivation measure in a particular setting, the issues are applicable to many other areas of epidemiology. The analysis of the amblyopia multi-centre study shows how area measures can be used to investigate the extent of inequalities in health and how these methods could be used to explore other ophthalmic conditions and more general epidemiological hypotheses.

The exploration of the changes made to vision screening in Leicestershire investigates a previously unevaluated service. It provides insight as to the usefulness of such interventions and whether their implementation in other regions could lead to reductions in inequalities in health elsewhere.

The relationship between area and individual level deprivation is relatively under-researched. If deprivation effects at the individual level are different to those shown by area level measures it is possible this work could inform on health policy. Resources may need to be targeted at deprived people rather than deprived areas. Further if the use of a method to correct for misclassification of deprivation is effective then it would be possible to use the data investigated in this work to estimate the effect of deprivation at the individual level in studies that only have access to area level information.

Finally, I discuss the use of deprivation measures in future amblyopia research and suggest possible uses in future epidemiological studies.

CHAPTER 2

INEQUALITIES IN HEALTH

2.1 AIMS OF THE CHAPTER

In chapter 1 I described how this thesis investigates inequalities in the presentation of amblyopia through the use of an area based deprivation measure. In this chapter I review research into inequalities in health that has prompted the development of deprivation measures to investigate and monitor these inequalities further.

To aid understanding of this research area, I discuss the concepts of inequalities in health. This chapter concentrates on the Black report (DHSS, 1980), since this work inspired a wealth of publications in this area. More recent evidence is also reviewed, focusing on evidence of inequalities in child health and the use of child health services which is of particular relevance to the research undertaken in this thesis. Explanations and proposed solutions for inequalities in health are discussed with reference to child health and screening.

2.2 INTRODUCTION

Research into inequalities in health has increased dramatically since the publication of the Black report. This is illustrated by the reviews of Feinstein (1993), Davey Smith, Bartley and Blane (1990) and Morris (1990). Over the last century 'health' has generally improved in Britain with a dramatic drop in mortality rates and a significant increase in life expectancy. However these improvements have not benefited the population alike. Although it is evident that there are naturally occurring differences in the health experiences of different areas or cultures, it is also apparent that unfair or unacceptable levels of 'inequality' exist with vast differences in the health experiences of those of contrasting levels of socio-economic status (Morris, 1990).

If equality was to exist in Britain then health experiences should not depend on social or environmental factors such as financial resources, social position, ethnic origin, geography, or gender. Furthermore, the health system should serve all parts of the community distributing resources according to need. The wealth of evidence in this area shows that this is not the case with major differences between the health experiences of the socially advantaged and disadvantaged. Increasing differences in mortality rates between affluent and deprived individuals over the last twenty years show the gap to be widening (DHSS, 1980; Davey Smith et al, 1990). This may be perpetuated by the fact that the more disadvantaged not only have higher mortality and morbidity rates but are also less likely to receive good health care.

Although inequalities in health are found within every country in varying magnitudes (Feinstein, 1993), this thesis concentrates on the UK. Inequalities in health can refer to differences in health experience related to gender, ethnicity, and socio-economic status and these issues are often inseparable. However in this thesis, socio-economic status will be the principal issue of interest.

2.3 THE CONCEPTS

A definition of the meanings of 'health' and 'inequalities' is necessary to be able to study this area. The concept of health is not simple. The Black report shows the disparity in various definitions of 'health', where, in historical terms, it can be thought of as the result of freeing a person from disease and disorder, while, in a more modern sociological sense, it can be thought of as a person's 'vigorous, creative and even joyous involvement in environment and community, of which presence and absence of disease are only a part' (DHSS, 1980).

With these differing definitions of health it follows that there are numerous ways of measuring health that will lead to different implications in attempts to improve it. Health outcomes are multi-dimensional, often qualitative and measurement can be affected by inaccuracy, bias, confounding and chance (Orchard 1994). Inaccuracy can occur when routine patient data are poorly coded, recorded and measured. Biases can arise in observational data where clinician's treatment and referral decisions may be affected by non-medical factors as well as disease severity, comorbidity and anticipated effectiveness and acceptability of the intervention.

Confounding, 'the presence of additional factors which render apparent associations spurious' (Orchard, 1994), can occur in observational studies. For example, Orchard indicates that coronary artery bypass grafting has shown apparently better outcomes than angioplasty because it was confounded by the severity of the patient's condition. Finally, small numbers can lead to differences due to random variation, i.e. chance. Mortality rates are the most common measure used in research, but there is now further work looking at the wider definition of health using measures such as morbidity rates, restricting illness rates, health service utilisation, and long-term illness rates. An increase in life expectancy has led to research into the quality of life.

The terms inequality and inequity are used almost interchangeably but need to be differentiated (Carr-Hill, 1994). Equality means equal distribution while equity means fairness. In this thesis I use the term inequality with reference to unequal health experiences. I use the term inequity with reference to unfair allocation of health care resources where resources are not distributed according to need. The World Health Organisation recognised the need to reduce unfair variation in health and health care in their 'Health for all' targets (WHO 1985), defining levels of health and health care that should be attained (Whitehead, 1988):

In health terms:

'ideally everyone should have a fair opportunity to attain their full health potential and, more pragmatically, none should be disadvantaged from achieving this potential if it can be avoided'

In health care:

The principle of equity 'leads to equal access to available care for equal need, equal utilization for equal need and equal quality of care for all.' (Whitehead 1988)

When measuring socio-economic inequalities in health the aim is to divide the population according to their resources and ways of living. These divisions are extremely subjective. Although work has now extended into looking at measures such as unemployment, income and housing (Townsend et al, 1988; Carstairs and Morris, 1989a), traditionally occupation based social class has been used. The most common scale is the Registrar General's scale where classifications are based on the occupation of the head of the household and are ranked from professional employment down to unskilled manual workers. This measure is frequently used as it is a convenient and well-recognised measure. As the measurement of social factors is

subjective, there are many ongoing issues relating to its measurement which are discussed in Chapter 3.

2.4 THE BLACK REPORT

2.4.1 THE BACKGROUND TO THE BLACK REPORT

Prior to 1980, although research had shown the existence of differing health experiences between socio-economic groups, there had been no major attempt by the government to assess their current extent. With evidence of increasing inequalities, a research working group chaired by Sir Douglas Black was appointed by the Labour Government to investigate inequalities in health and propose future policy recommendations. This report issued in 1980 is commonly known as 'the Black report' (DHSS, 1980). The new Conservative Government did not welcome the results, printing only 200 copies and including a foreword by the Secretary of State for Health dissociating the government from the conclusions drawn by the working group. Despite this, the report went on to generate great interest in inequalities in health both in Britain and other countries.

2.4.2 THE MESSAGE OF THE BLACK REPORT.

For ease of interpretation and comparison the Black report investigates inequalities in health by using occupational mortality data from the 1970-1972 decennial supplement (OPCS, 1978). The principle and dramatic message of the report is that there exist distinct differences in mortality rates between different levels of occupational social class. This relationship persists in all age groups and both sexes. Class gradients are steeper at younger ages with differences in mortality rates particularly evident in neonatal deaths (deaths at birth or within the first month of life) where the rate among children of parents in social class V (unskilled manual workers) is double that among children of parents from social class I (professional workers). This class gradient is reduced in older children, but increases again in early adult life before falling in middle and old age. This pattern is seen for most causes of death and long-term illness. There is further evidence of inequalities in the utilization of health services particularly preventive services such as screening.

Although overall mortality rates have improved, the Black report demonstrates that since the 1930's there seems to have been a deterioration in the health experiences of men in social class V compared to those in social class I which remain even after adjusting for changes in occupational classification (Table 2.1).

Table 2.1: Mortality of men by occupational class (1930's-1970's) (standardised mortality ratios)

Occupational class		MEN AGED 15-64					
		1930-32	1949-53	1959-63	1959-63 *	1970-72	1970-72 *
				unadjusted	adjusted	unadjusted	adjusted
I	Professional	90	86	76	75	77	75
II	Managerial	94	92	81	-	81	-
III	Skilled manual and non-manual	97	101	100	-	104	-
IV	Partly skilled	102	104	103	-	114	-
V	Unskilled	111	118	143	127	137	121

* Occupations in 1959-63 and 1970-72 have been adjusted according to the 1950 classification

Source: Black Report Table 7, (DHSS 1980),

A comparison of mortality rates with other industrial countries shows perinatal and infant mortality rates to be distinctly higher in Britain than in Scandinavia and the Netherlands.

Although the improvement in perinatal mortality is similar to other countries, the improvement in infant mortality in Britain is less than in all other comparable countries. The infant mortality rate is thought to reflect social conditions (Grant, 1994) and so this further highlights evidence of inequalities in Britain.

The Black report considered various explanations for these inequalities in health and proposed various policy recommendations which are discussed in 2.5 and 2.81 respectively.

2.4.3 METHODOLOGICAL PROBLEMS IN THE BLACK REPORT

Although the message conveyed by the Black report is dramatic, several methodological problems need to be noted. Feinstein's review of the report (1993) sets out five main issues.

Firstly mortality rates are calculated using the occupation on the death draft as the numerator and occupation as recorded by the census as the denominator. These are not necessarily the same since occupation may change over time. Secondly, the sample may not be representative since the information is limited to people under 65 which may overestimate the extent of inequalities in Britain as mortality differentials tend to decrease with older age. Thirdly, there is a possible problem of reverse causality in the relationship between occupational class and mortality as people of poor health have been shown to descend the occupational scale over time. Fourthly, there have been changes in the distribution of occupational social class over the period 1930 to 1970 with a drift towards increasing numbers in the higher social classes compared to the lower ones. Evidence of widening inequalities could be due to changes in occupation definition and improvements in classifications. Lastly, there is a lack of adjustment for other explanatory factors apart from gender such as diet, smoking prevalence and access to health care. These could affect mortality differentials but they are not considered in a multivariate way.

However, it is important to remember that although all of these points are significant, work following the Black report addressed many of these issues, as will now be discussed, and yet still came to the same overall conclusions on the levels of inequalities in health.

2.5 EXPLANATIONS

There is no single, simple explanation of inequalities in health. Explanations are of a multicausal nature. The Black report divides the most widely discussed explanations into four groups: artefact, natural or social selection, cultural or behavioural, and materialist or structuralist.

2.5.1 ARTEFACT EXPLANATIONS

These explanations are linked to the methodological problems of the Black report. Artefact could explain the relationships found by the Black report since mortality rates were calculated using numerators and denominators from different data sources as discussed in section 2.4.3. An analysis of data from the OPCS longitudinal study addressed this problem (Fox et al, 1985). The study looked at a 1% sample of the population identified in 1971, studying mortality over time and recording data at an individual level. It used social class in 1971 to

classify individuals at death and this analysis resulted in the same conclusions on the association between social class and health.

Patterns of widening inequalities in health are also attributed to the drift up the social class scale over time with decreasing numbers in social class V. This small extreme group could artefactually show widening inequalities. Whitehead (1988) illustrates this problem using data from the longitudinal study. The SMR for all cause mortality was 125 in social class V men yet they only contributed 5% of all deaths. However this study investigated alternative measures of socio-economic status and found that men without access to a car and living in rented accommodation had an SMR of 123 (similar to social class V) but contributed 21% of all deaths.

Further only people under 65 are studied in the Black report making the sample unrepresentative. The OPCS study also addressed this and showed that there were slightly reduced but still substantial mortality differentials in those over sixty-five (Fox et al, 1985).

2.5.2 NATURAL OR SOCIAL SELECTION EXPLANATIONS

Health state is thought to determine socio-economic position. People who are unhealthy are thought to be downwardly mobile particularly in the years before death. This could lead to a higher concentration of people with a higher risk of dying in the lower classes accounting for the mortality differentials. The longitudinal study (Goldblatt, 1989) used social class data from 1971 and 1981 to investigate this relationship between mortality and social mobility. They found that there was some evidence of downward mobility, but it did not account for the mortality differentials.

2.5.3 CULTURAL OR BEHAVIOURAL EXPLANATIONS

Health related behaviour refers to differences in the way individuals from different social groups choose to lead their lives. Feinstein (1993) divides these behavioural explanations into two different sources of inequality, those relating to the lifespan and those relating to access and use of the health care system. For example, in terms of lifespan, characteristics such as smoking, diet, exercise patterns, alcohol consumption, drug use and driving habits can affect health experiences. The Whitehall study, a major study of the mortality of British Civil servants using individual data, showed employment grade differences in health risk behaviour

including smoking, diet and exercise (Marmot et al, 1991). Smoking is more prevalent among lower social classes and contributes to poor health experiences in these groups. In terms of diet, there are little differences in fat consumption but among the lower social classes there is evidence of reduced consumption of vitamin C, carotene, fibre and an increased sodium to potassium ratio which could have a detrimental effect on health.

There are also other behaviour related explanations with respect to access to and use of health care services. Attitudes towards health, particularly with respect to ideas about cure and prevention of disease vary greatly between social groups. It has been shown that people of low socio-economic status are less likely to use children's immunisation and screening programmes (see 2.7).

2.5.4 MATERIALIST OR STRUCTURALIST EXPLANATIONS

Materialist or structuralist explanations relate to the role of the external environment. They demonstrate the hazards within society that individuals have to endure given their current state of income and available opportunity. Feinstein (1993) again divides materialistic explanations into those related to the life-span and those related to access to the health care system. For example in terms of lifespan, people of lower socio-economic status experience a more unhealthy environment with more hazardous working conditions and poorer housing. In terms of access to the health care system, they generally have fewer resources available for using health services and obtaining necessities for health. Unemployment has been shown to be associated with increased mortality (Morris et al, 1994), increased psychiatric admissions (Kammerling and O'Connor, 1993) and increased mortality (Morris et al, 1994).

There has been interest in whether deprivation in early childhood or later life is more important to health (Ben-Shlomo and Davey Smith, 1991). Studies have inferred that some diseases have their origins in poor conditions during development. However, as poor conditions in childhood are evidently related to poor conditions in adulthood, it is extremely difficult to separate out these factors.

It appears that health inequalities cannot be explained as artefact and although natural and social selection do play a part, this is not sufficient to explain the variation in health experiences between people of differing socio-economic status. Behavioural and material

explanations are extremely important. People of low socio-economic status have fewer resources (both financial and psychological) with which to respond to major life events. Whitehead (1988) indicated that adverse social conditions can limit the choice of lifestyle, making it extremely difficult to separate these material and behavioural ideas.

2.6 FURTHER EVIDENCE

Several reviews have looked at work following the Black report and demonstrate the wealth of research in this area (Whitehead, 1988; Davey Smith et al, 1990; Feinstein, 1993; Morris, 1990). This work shows that the social inequalities highlighted by the Black report still persist and are possibly worsening, with increasing differences in mortality rates between the affluent and deprived over the last twenty years (DHSS, 1980; Davey Smith et al, 1990). McCarron et al (1994) have shown increasing mortality differentials between 1980 and 1992 in Glasgow. These mortality differentials may be perpetuated by the fact that the more disadvantaged not only have higher mortality and morbidity rates but are also less likely to receive good health care. Whichever method is chosen to measure socio-economic status, whether it is based on occupation as in the Whitehall study, or material resources as in the OPCS Longitudinal study, the conclusions are similar. Throughout life those at the top of the social ladder have better health experiences and lower death rates. Although the Black report concentrated on inequalities in mortality there have now been more studies undertaken to look at morbidity. Data from the Health and Lifestyle Survey has shown striking differences between social classes in self-defined health status, the reported incidence of illness and measured physiological fitness (Blaxter, 1990).

Further research has confirmed regional differences in mortality shown by the Black report and indicates a distinct north-south divide (Townsend et al, 1988; Carstairs and Morris, 1989b, Drever and Whitehead, 1995). Studies have shown that inequalities exist in neighbouring communities and these areas of poverty seem to be more common in the North (Townsend et al, 1988). However more recent evidence has indicated that although the north-south divide may still exist in older age groups, it is disappearing in younger groups (Illsley and Le Grand, 1993). Whitehead (1988) illustrates the differences in health experience and gender with women having a lower death rate but a higher sickness rate than men.

2.6.1 ETHNICITY

Although this thesis concentrates on socio-economic inequalities in health, ethnicity is an inseparable issue. The confounding effects of deprivation and ethnicity are difficult to untangle. Ethnicity is shown to be related to health with higher rates of accidents, worse health experiences and poorer access to health care among ethnic minorities (Whitehead, 1988). Atri et al (1996) found that women from minority ethnic groups were less likely than white women to have evidence of mammography or cervical smears recorded in their general practitioner's medical records illustrating differential uptake of preventive services. It is important to note that, as many people from minority ethnic groups endure both adverse housing and working conditions, the experience of these people may be extremely poor. A recent review of child health and poverty (Spencer, 1996) has indicated that although in some cases genetic differences in ethnic origin influence child health, it is socio-economic and environmental factors which explain the majority of the variations in health both between and within ethnic groups. A recent study in Tower Hamlets has shown that minority ethnic groups were considerably more disadvantaged than white people and were five times more likely to live in overcrowded accommodation, three times less likely to own their own home, twice as likely to be in social classes IV and V and less likely to be employed (Atri et al, 1996).

Smaje (1995) illustrates a variety of studies indicating higher rates of morbidity and mortality among minority ethnic groups. He argues that these differences cannot be explained by differences in socio-economic status alone. Racism and discrimination appear to be distinct problems needing further investigation (Benzeval et al 1995). Until recently ethnicity has not been assessed in the census, and most work has been based on country of birth. This has made it extremely difficult to use national data to look at the health experiences of British-born children of minority ethnic groups. With the introduction of a question on ethnicity to the 1991 census, research in this area should improve. Although this thesis concentrates on socio-economic inequalities in health, the confounding effect of ethnicity is investigated.

2.7 INEQUALITIES IN CHILD HEALTH AND PREVENTIVE SERVICES

In terms of child health the class gradients are at their steepest. As discussed in 2.3.2 the rate of neonatal deaths in social class V is double that seen in social class I. Judge and Benzeval (1993) show that these steep gradients may even be underestimated with the exclusion from conventional class based analysis of child mortality of children whose parents are classified as 'unoccupied' who are largely economically inactive single mothers.

Inequalities are not only evident when looking at mortality and Wadsworth (1988) has emphasised the need to look at more positive measures of child health. The National Child Development Study (NCDS) is a national study of people born in one week in March 1958. These children were studied at birth and then followed up over time. Whitehead (1988) reports results from this study that show differences in health experiences at ages 7 years and 23 years between children living in local authority accommodation and those living in owner occupied accommodation in terms of height, 'malaise', self-reported health, hospital admissions and psychiatric morbidity. Children living in local authority accommodation experienced consistently poorer health. Essen and Wedge (1982) report results from the same study illustrating clear differences in height between children classed as disadvantaged and those with no disadvantaged characteristics. Although these results are somewhat dated, recent evidence has confirmed these patterns. Reading, Jarvis and Openshaw (1993) showed that in Northumberland the proportion of birth weights less than 2800g was much higher in the most deprived areas than the least deprived areas and that mean height of children between 5 and 8 years was lower for those from the most deprived areas than those from the least deprived areas. Confirming the patterns of differential hospital admissions, Spencer et al (1993) have shown that infants from deprived areas of Sheffield were overrepresented in the multiple hospital admissions groups in both 1980 and 1985.

Further evidence of inequalities in a wide range of areas including mortality, growth, physical morbidity and accidents has been illustrated by recent reviews of child health and poverty (Spencer, 1996; Reading 1997). However the emphasis of this thesis is on inequalities in access to child health screening services which may lead to inequalities in health in later life. Inequalities in access to health care among children are very evident, particularly for

preventive services. Essen and Wedge (1982) used data from the NCDS to show that disadvantaged children received less health care than those with no disadvantaged characteristics with lower rates of reported immunisations against diphtheria, polio and smallpox by age seven and immunisations against tuberculosis, rubella and smallpox by the age of sixteen. Blaxter (1981) also quotes results from the NCDS to show differences in the measles vaccine and the triple vaccine with a lower proportion of disadvantaged children being vaccinated. Blaxter uses data from the study of Child Health and Education in the Seventies (CHES), a longitudinal study of children born in one week in 1970, to show similar differences in the uptake of the triple vaccination with children from manual classes having a lower uptake than those from non-manual classes.

More recent work again shows that these patterns are still clearly evident. Reading, Jarvis and Openshaw showed social inequalities in the uptake of childhood immunisation and screening programmes (1993). Marsh and Channing (1986) have also illustrated lower uptake of childhood immunisation in a deprived neighbourhood compared with a more endowed neighbourhood. Similarly, Lynch (1995) found that practices serving populations in socially deprived areas and with poorer health were less likely to achieve high targets for childhood immunisations. Of particular reference to this thesis is the work of Bowman et al (1996) who showed that in a study of children referred by their general practitioner with suspected amblyopia or strabismus, patients from deprived areas were less likely to attend their first ophthalmology hospital outpatient appointment than those from less deprived areas. Similar results have been found by a Glasgow study of vision screening (Williamson et al 1995) with higher default rates in areas of lower socio-economic status.

The consequences of inequalities in access to preventive services are highlighted by Macintyre (1989) since these inequalities:

'could influence the distribution of disease or death across social groups at each of three levels of prevention - primary, secondary or tertiary. Preventive procedures, such as immunisation, could influence the incidence of disease among different social groups if differentially available or used; screening or treatment procedures could influence cure or survival if differentially available, used or effective; and rehabilitation or after-care service could influence the consequences of disease again if differentially available, used and effective'

Hence this evidence of inequalities in access to and use of preventive services may have serious impact on the existence of inequalities in health. For example, Bowman et al (1996) found that more than half of patients who failed to attend their first appointment at ophthalmology outpatient clinics never reached the service and missed out on the opportunity of treatment for amblyopia. Those who have not had the condition treated will carry it on into later life.

2.8 SOLUTIONS AND GOVERNMENTAL POLICY

2.8.1 GENERAL RECOMMENDATIONS

Several policy recommendations were made by the Black report. An important message of the Black report is that inequalities in health are not just a problem for the NHS to tackle. A wider strategy is needed to tackle these issues, with participation from research, health and social services, and the Government departments of Education, Home Office, Environment, Trade and Transport. The three main objectives of the recommendations are to give children a better start, to use preventive and educational action to encourage good health, and to extend and improve the quality of life of disabled people. Improvements in social conditions to abolish child poverty were also proposed by the working party with a government policy to oversee this task. Researchers in child health and poverty have also emphasised this need for an approach not purely from the health service and call for well-evaluated multidisciplinary interventions (Spencer, 1996; Reading, 1997; Roberts, 1997).

The World Health Organisation (1985) compiled a strategy for health in Europe aiming for 'Health for all by the year 2000' and the plan was accepted by the UK. It recognised that essential social conditions would have to be met before this could be achieved such as equal opportunities for all, satisfaction of the basic needs for food, basic education, clean water and sanitation, decent housing, secure work and a useful role in society, and political will and public support to launch the necessary action. These targets aim to reduce inequalities in health both between and within countries by at least 25% and was deemed achievable if the goals attacking poverty were met.

These targets prompted the Government to publish 'The Health of the Nation' green paper (DoH, 1991) as a consultative document which stimulated extensive public debate. This was followed by the publication of a white paper (DoH, 1992) with a strategy for health selecting five key areas for action, setting national objectives and targets in these key areas, indicating action needed to achieve the targets, outlining initiatives to implement the strategy and setting the framework for monitoring, development and review. Although prompted by the WHO targets, social inequalities in health were not a priority.

The strategy concentrates on diseases and interventions rather than social, economic and environmental changes that might produce more long-lasting and equitable improvements in health (Thunhurst and MacFarlane, 1992). This lack of emphasis on inequalities in health has caused lengthy debate (Radical Statistics Health Group, 1991; Delamothe, 1991; Townsend, 1993). The strategy acknowledges the effects of 'social circumstances' and 'physical and social environment'. When asked about inequalities in health the Secretary of State for Health's claimed *'the divisions were so fundamental, complicated, long-lasting and recalcitrant that they were not a suitable government target. Targets need to be specific, measurable, and most of all achievable otherwise the whole thing comes into disrepute'* (Delamothe, 1991). This reasoning is challenged by many other countries where relatively successful policies have been used to tackle inequalities (Mackenbach, 1994). Benzeval et al (1995) also challenge this opinion indicating 'that there is no room for reasonable doubt that observed social inequalities in health are amenable to purposeful policy interventions'. In their book they set out a practical agenda for action to improve the current situation.

A problem with making inequalities in health a key area indicated by Delamothe (1991) is that the solutions lie beyond the Department of Health. He quotes Wilkinson (Quick and Wilkinson, 1991) who claims that 'the Chancellor has a much greater impact on health than the Secretary of State for Health, a thought that may well not cross the minds of either'. He argues that standards of health depend more on the overall distribution of income rather than the average level. Hence making incomes more equal could improve health. Japan for example, has the most evenly distributed income in the world and the longest life expectancy. In Britain, in 1979, 35% of income was distributed among the top fifth of the population

compared with 41% in 1990-1991 (OPCS, 1994). This increase in income inequity coincides with the pattern of widening inequalities in health.

Inequalities in health are of a multi-causal nature and need monitoring and changing in a multi-disciplinary way. More recently research has turned towards tackling these inequalities. A systematic review of interventions to reduce 'variations in health' has shown that it is possible to move towards greater equality in health experiences (NHS Centre for Reviews and Dissemination, 1995). However it emphasises the need for rigorous evaluations of interventions and the need to estimate the degree to which these interventions will reduce health variations. These mainly regional policies can help on a small scale level, but research shows the need for major government strategy involving the collaboration of various departments if inequalities are to be significantly reduced. Benzeval et al (1995) emphasise the need for commitment to reducing inequalities in health at the highest levels of government. Whitehead and Dahlgren (1991) indicate the greatest gains could be achieved by national and European policies. The incoming Labour government have indicated a concern for a change in the NHS (Labour party, 1995). They have appointed a Minister of Public Health and pledged to measure and monitor poverty and establish targets associated with inequality. An inquiry to update the Black report is to be launched (Laurance, 1997). However this is still at an early stage and the results may be a long way away.

2.8.2 CHILD HEALTH AND PREVENTIVE SERVICES

As discussed, the Black report recommended preventive and educational action to encourage good health since greater equality of care is thought to depend on a high national standard of knowledge about self-care, care of children and other dependants, and the pursuit of activities conducive to health. The aims were wide ranging and included expanding health education and selective screening.

People from lower socio-economic status have been shown to make less use of preventive services for themselves and their dependants. There are significantly different levels of uptake between socio-economic groups for screening and health promotion services, (Waller et al, 1990), and attendance at screening and immunisation clinics for children as previously discussed. A variety of factors affect screening attendance including under-provision, costs of attendance (financial and psychological), and a lifestyle which prevents rational actions

towards future good health. These factors need investigation in order to improve screening effectiveness.

Health education is thought to be related to the uptake of screening services. The Black report recommended further government funding of health education. 'The Health Divide' (Whitehead, 1988) commissioned by the Health Education Council reinforced this idea. However between commission and publication of this work, the government disbanded the independent Health Education Council and formed the Health Education Authority within the NHS. This led to health education being more closely controlled by the government and some feel it may not attack the issues of health-related behaviour that could affect the government both financially and politically.

New NHS reforms could also affect services such as screening. Budgets being held by general practices for certain types of health care may lead to a reduction in the number of general practitioners holding screening services at their own practices, as they may require the money for activities they deem to be more needy. Paton (1992) believes that general practitioner contracts will create 'weak' and 'strong' practices due to competition for patients, with weak practices being more common in inner-city areas where resources are poor. Weaker consumers (both economically and sociologically) would tend to suffer in this situation. Screening uptake in lower socio-economic groups may be unlikely to improve with these reforms. A recent study of screening programmes in general practice showed that only 57.3% of nonfundholding group practices offered child health surveillance programmes compared with 83.3% of fundholding practices indicating inequalities in the provision of services between practices (Li and Logan, 1996).

2.9 FUTURE RESEARCH AND INFORMATION NEEDS

Research has emphasised the need for monitoring inequality in Britain. However, the current routine recording of statistics requires improvement. Current data is inadequate but there is no consensus of opinion on exactly what statistics are needed (Thunhurst and MacFarlane, 1992). Whitehead (1988) indicated the need to increase recording of socio-economic factors in routinely collected health statistics and to include health factors in social and economic

statistics. Better measures of health and social factors are needed urgently in order to implement this. The development of health policy indicators, social and economic indicators related to health, indicators of the provision of health care and health status indicators could all be used for monitoring health.

In terms of health, measures need to be improved to look at the impact of various factors on health rather than just using traditional measures such as morbidity and mortality. One example of this is the recent introduction of questions on long-term illness to the 1991 census.

Since the publication of the Black report, there has been emphasis on compound indicators of socio-economic status (Carstairs and Morris, 1989a; Townsend et al, 1988; Jarman, 1983) rather than using occupational social class. Whitehead (1988) has called for more sensitive resource allocation formulae and numerous resource allocation formulae for general practice payments have been developed (e.g. Jarman, 1983; Hopton et al, 1992).

2.10 SUMMARY AND CONCLUSIONS

In this chapter I have discussed current research on inequalities in health that has prompted the work in this thesis. Scientific evidence has demonstrated the dramatic differences in health and socio-economic status. This cannot be explained by artefact or natural selection but by behaviour and material deprivation. Of particular relevance to the work in this thesis is the evidence showing inequalities in the use of childhood preventive services such as immunisation and screening. Macintyre (1989) has indicated that these inequalities in the use of preventive services may be leading to further inequalities in health in later life. This work has prompted my investigation of inequalities in the presentation of amblyopia.

Research has called for measuring and monitoring of health inequalities. In chapter 3 I present a review of various measures of deprivation that have been used for these purposes. I then use a measure of deprivation later in the thesis to assess inequalities in the age of presentation of amblyopia and discuss the possibility of using deprivation measures to monitor inequalities in this area in the future.

To reduce inequalities in health, research reviewed in this chapter has indicated that action is needed to reduce poverty and deprivation at a national level and that a multidisciplinary approach is required to achieve this. Preventive services and health education are a suggested method of reducing inequalities. Since national level interventions have not recently been forthcoming in this thesis I look at the effect of a local small-scale screening intervention on inequalities in the presentation of amblyopia and discuss the value of interventions at a local level.

CHAPTER 3

MEASUREMENT OF DEPRIVATION

3.1 AIMS OF THE CHAPTER

As discussed in chapter 1 this thesis looks at inequalities in the presentation of amblyopia. In order to estimate inequalities in health some measure of socio-economic status is needed. A wide variety of measures have been developed and implemented. In this chapter I discuss the problems of using social class in measuring inequalities in health and investigate several recently developed deprivation scores based on area census data. I discuss my selection of an area based score to look at inequalities in the age at presentation of amblyopia.

3.2 INTRODUCTION

One of the conclusions of the work on inequalities in health reviewed in chapter 2 was the need for health purchasers and providers to monitor health inequalities at both a local and national level. It was suggested that by identifying and targeting areas of special need, inequalities in health could be reduced. Previously occupational social class had been used to look at health inequalities but reports such as the Black report (DHSS, 1980) and the Health Divide (Whitehead, 1988) called for the development of more up-to-date and sensitive measures of material and social resources. Routine data on socio-economic factors are now widely available and with the increasing power and accessibility of computers that are able to cope with large datasets and complex statistical analyses there has been a large amount of work developing 'deprivation scores' for small areas. Similar types of scores have been concurrently developed for use in health care planning, health needs assessment and resource allocation. The Government and, at a local level, health authorities are employing deprivation scores in order to locate areas of need and improve resource allocation.

3.3 SOCIAL CLASS

In discussing social class it is important to clearly distinguish between the theoretical concept of social class and the practical measures of social class. Although the theoretical concept of social class may still be relevant in discussing inequalities in health, this thesis is concerned with the practical measurement of inequalities and this section is concerned with the usefulness of social class measures in measuring these inequalities. For many years occupational social class measures have been the principal proxy measure of income and lifestyle used in most research on inequalities in health. Their popularity is partly due to them being well-recognised, simple measures. Since occupation is recorded on UK death certificates, it makes this type of measure more accessible. Liberatos et al (1988) found 40% of studies on chronic diseases in the American Journal of Epidemiology between 1982 and 1985 incorporated a measure of social class. The way social class is used in research can affect conclusions drawn from a study, and poor classification could lead to misclassification and a reduction in any observed associations.

Although sociologists tend to think of the concept of social class as a multi-dimensional measure, in Britain epidemiologists attempting to operationalise this concept have traditionally used a scale just based on occupation. The most frequently used measure is that of the Registrar General which was first drawn up in 1911. The head of the household is allocated to one of five social classes:

- I. Professional (e.g. accountant, doctor, University lecturer)
- II. Intermediate (e.g. nurse, schoolteacher, manager)
- IIIN. Skilled non-manual (e.g. clerical worker, secretary, shop assistant)
- IIIM. Skilled manual (e.g. bus driver, butcher, carpenter)
- IV. Partly skilled (e.g. agricultural worker, bus conductor, postman)
- V. Unskilled (e.g. cleaner, labourer)

This classification, although a single indicator, was thought to also reflect education and culture.

There are many problems with this single indicator. Due to changes in skills and status of different occupations the scale is revised every ten years. Table 3.1 shows changes in the proportion of the population in each social class to lead to a shift up the scale over time.

Table 3.1: :Percentage of males by social class for 1931, 1951 and 1971 for England and Wales. (Source: OPCS (1977))

	Males aged 16-44			Males aged 45-64		
Occupational class (Registrar General)	1931	1951	1971	1931	1951	1971
I	1.8%	3.1%	5.0%	3.2%	3.3%	4.1%
II	10.0%	11.9%	14.3%	18.3%	17.5%	19.5%
III	51.1%	56.3%	46.4%	43.5%	46.2%	44.9%
IV	18.4%	15.0%	13.8%	16.6%	16.3%	18.2%
V	16.3%	9.8%	6.1%	17.1%	15.3%	8.1%
All	100%	100%	100%	100%	100%	100%
Population (000s)	8660.6	8910.3	9525.6	4068.4	4902.4	5722.4

* The *All* category includes students, the unoccupied and persons whose occupations were inadequately described

As discussed in chapter 2 there is now a smaller proportion in social class V but a larger proportion in social classes I and II. Unemployment is also reducing the proportion in social class V. It is therefore difficult to make comparisons over time. Variations in income, bonuses and benefits, and living standards within individual occupations can be great and there are also differences in income between occupations in the same social class. This has led to wide variations in mortality experiences within social classes. Where social class is available it is often poorly recorded, leading to potential misclassification. Women, children, the retired and the increasing number of 'never-worked' are extremely difficult to classify. With the increase in the proportion of working women there are problems as to who to define as the head of household. Carstairs and Morris (1989a) showed that in the census 42% of women aged 16 to 64 were not assigned to an occupation on their own account compared with 12% of men in that age group. They also showed that 68% of death certificates of women aged 20 to 59 failed to have an occupation recorded compared with 3% of men.

Another problem is the lack of theoretical basis in the way the scale was drawn up. Jones and Cameron (1984) claim that the ranks were made to 'conform to the prejudices of narrow-minded professionals' with some of the rankings being inappropriate. Also the scale was primarily developed to investigate infant mortality rates (Leese and Fox, 1977). In 1921 the

scale was graded so that there was an increase in infant mortality down the scale. Hence investigating mortality and social class should by definition show an association.

Alternatives to the Registrar General's scale are necessary. Jones and Cameron (1984) propose that the classification should be abandoned and not replaced. Both Jones and Cameron, and Liberatos (1988) have prompted the investigation of new measures of material and social resources.

3.4 CONSTRUCTING ALTERNATIVE MEASURES

There are many different aspects of social stratification which can be investigated besides occupational social class such as education, housing or income. Some of the work undertaken in this thesis was based on data collected from patient records which did not provide any access to personal levels of socio-economic status. The recent increased availability of census data has given access to detailed national data on socio-economic circumstances, making small area census data the most appropriate way of assessing this in my work.

3.4.1 THE CENSUS

The results of the censuses (started in 1801) have provided an invaluable source of information for social researchers, demographers and town planners as they give a picture of demographic characteristics of an area and an idea of changes over time. The census covers both the household (e.g. size, type, amenities, tenure, car availability) and its occupants (e.g. age, sex, occupation, ethnic group since 1991 and country of birth). Data are available on Britain as a whole comprising about 20 million households down to enumeration districts of about 150 to 200 households. Information on smaller areas is not possible since the census must preserve individual confidentiality.

In 1981 census data were made available via a mainframe computer system and became far more accessible than before. The increased accessibility of census data has broadened the possibilities of estimating socio-economic status. Previously, studies which aimed to measure social class involved interviewing patients or detailed investigation in order to obtain occupational descriptions. With access to census data it is possible to use area data as a proxy for individual data. It has been possible to match postcodes to census enumeration districts

(ED) since 1991 and wards since 1981. Reading and Openshaw (1993) found that prior to the 1991 census computerised matching of postcodes to EDs was highly inaccurate but this had little effect on the allocation of socio-economic status. Socio-economic inequalities were blunted but not eradicated by this mismatching.

3.4.2 CONCEPTUALISING POVERTY

Using census data can limit the areas of socio-economic status that can be investigated. For example, there is no information on the income of the household in the 1981 or 1991 census. The main recent work on inequalities in health based on census data has concentrated on the concept of poverty and the use of deprivation scores as proxy measures for poverty. These methods have the underlying assumption that poverty can be measured by deprivation scores.

There are two main definitions of poverty, absolute poverty and relative poverty (Blane, 1991). Absolute poverty or subsistence poverty is discussed in terms of a minimum standard that can be applied to all societies, below which individuals are assumed to be in poverty. This concept has been criticised since it assumes that there is a minimum basic level of needs that applies to all people in all societies. Needs vary greatly between and within societies and so the concept of relative poverty has been used more frequently where the definition of poverty must be related to the standards of a particular society at a particular time. Townsend (1979, cited in Blane, 1991) defines relative poverty as where ‘the resources (of those in relative poverty) are so seriously below those commanded by the average individual or family that they are, in effect, excluded from ordinary living patterns, customs and activities’. Hence the ‘poverty line’ which divides those who are and those who are not in poverty will vary with the affluence of the society but must be constructed with the concept of absolute poverty in mind.

3.4.3 DEVELOPING A MEASURE OF DEPRIVATION

The operationalisation of the concept of relative poverty however is not simple since it is difficult to decide what are the ‘normal’ or ‘acceptable’ standards of living. Generally measures of relative deprivation are used as proxies for poverty. There are two main approaches that may be taken, the measurement of material deprivation or the measurement of multiple deprivation. Material deprivation refers to a lack of material resources to maintain an acceptable standard of living and this is often measured using information on income. Alternatively, there is the idea of multiple deprivation as a proxy for poverty which refers to a

much wider concept including inadequate educational opportunities, unpleasant working conditions or powerlessness. These factors are not directly related to poverty and imply that wider changes than the redistribution of income are necessary to eliminate poverty.

A problem indicated by Thunhurst (1985) is that deprivation is a multi-dimensional problem and reducing it to a single score loses this essential multi-dimensionality.

Scores have been constructed in order to measure a variety of forms of deprivation. There are scores to measure material deprivation (Townsend et al, 1988; Carstairs and Morris, 1989a, 1989b), to identify underprivileged areas in terms of G.P. services (Jarman, 1983; Jarman, 1984; Irving, 1983) and to look at needs assessment (DOE, 1983; Duguid and Grant, 1983). There is also work trying to identify which measure is best in terms of correlation with mortality statistics, with little thought going into the construction of the scores. For example, Carr-Hill et al (1992) demonstrate this by quoting Jarman (1990) who at the beginning of a paper stated

'the term 'social deprivation' in health care is used to refer to the social characteristics of individuals or populations which are most associated with above average levels of morbidity or mortality or with need for health care provision (or any combination of these)'

Jarman then ends the paper saying that one of the conclusions from his analysis is that

'there is strong evidence of a relationship between social deprivation and the need for health care provision'.

This type of tautology is common and gives rise to high but uninformative correlations.

The principal problem encountered when measuring deprivation is that of choosing a sensible combination of indicators which clearly and accurately define deprivation. There is a frequent tendency to dredge for possible indicators with no overall justification for their inclusion.

Deprivation indicators can be divided into direct and indirect indicators. Direct indicators can be thought of as those which represent conditions or states of deprivation, while indirect measures reflect the persons experiencing that deprivation. Townsend et al (1988) illustrate the problem of using indirect indicators:

'From a sociological perspective it is important to distinguish between the measurement of deprivation in different areas and the kind of people experiencing that deprivation.

Otherwise there is a danger of treating social categories like age, ethnicity and single parenthood as causes of the phenomenon under study.'

The aim is to find how many of these minorities are deprived rather than to assume that they all are. This issue of indirect and direct indicators is also related to whether the aim is to measure social or material deprivation. In general the direct indicators reflect material deprivation while indirect measures reflect racism, sexism and ageism that can lead to social deprivation.

Thunhurst (1985) goes on to define interpretative indicators which are not measures of deprivation but aid the geographical analysis of the distribution of direct and indirect indicators e.g. the amount of furnished and unfurnished rented accommodation or the number of in-migrants in the last year.

The final choice of indicators is an extremely political one needing careful thought and justification. Bartley and Blane (1994) show it is important to evaluate and understand deprivation scores in terms of 'the purpose for which they have been used and the validity of the assumptions about social and economic life that they embody'. Although there are many problems with using scores, Carr-Hill et al (1992) state that if they are used in a way which is 'theoretically informed and contextually considered and methodologically appropriate' they can be used as part of the decision making process.

3.4.4 COMBINING INDICATORS

Most of the deprivation scores developed are composite. It is not only important to make informed decisions on how to choose the indicators themselves but also how to and whether to combine them into an overall score.

Thunhurst (1985) points out the problems of combining indicators into a single score. Multicollinearity is a major problem with many of the indicators being highly correlated. Thunhurst shows an example where percentage of households lacking basic amenities, percentage of households with more than one person per room, percentage of pensioners living alone and percentage of New Commonwealth immigrants all showed high correlation with the standardised mortality ratio and an indicator of need for acute services. Yet, when Thunhurst investigated the partial correlations, only the two direct indicators were important showing that correlations can be misleading and spurious associations may arise.

Thunhurst shows the problems of combining skewed variables. Transforming values to remove skewness is important (Gilthorpe, 1995). However Thunhurst indicates that it does not remove all

skewness, does not equalise skewness between variables and performed indiscriminately will have the effect of introducing an unknown weighting factor.

The technique used to combine indicators into a single score can also drastically affect the results. A common method used is the standardised z score where each indicator is converted into a standardised value based on the mean and standard deviation of the population. Summing the resulting z-scores results in a score where each indicator is weighted equally. If standardisation is not performed then those indicators having longer scales will be more heavily weighted. For example, two frequently used indicators are the percentage of households not owning a car, which has a mean of approximately 27% in Britain, and the percentage of households which are overcrowded, which has a mean of 2.4%. If these indicators were just summed then car ownership would dominate the measure. Standardisation is not always used as in some cases indicators are deliberately weighted unequally. This is particularly seen in resource allocation scores where indicators are often weighted according to their demand on the G.P.'s time. For example 'lone pensioners' is weighted very highly in the Jarman score because of their greater needs.

3.4.5 THE PROBLEM OF ECOLOGICAL FALLACY

A further issue in deprivation measurement based on area level census data is whether area level deprivation (e.g. ward or enumeration district level) accurately reflects individual deprivation. Studies where the unit of analysis is a group are known as ecologic studies (Morganstern, 1982). Typically, as in this case, the group is a geographical area. One of the major problems with ecologic studies is the potential for bias in effect estimation. This problem, known as the 'ecological fallacy', arises from making a causal inference about individuals based on the observations of groups (Morganstern, 1982). Robinson (1950) was one of the first to discuss the problem of 'ecological fallacy' where relationships which exist at one level of analysis (e.g. areas) are then extrapolated to another level (e.g. individuals) and lead to distortions and even reversal of effects. Piantidosi et al (1988) have alerted epidemiologists to 'the problem of serious errors resulting from inferences based on ecological analyses'.

Despite these problems, ecological studies are still widely undertaken for several reasons (Piantidosi, 1994). Firstly, they are relatively simple to undertake and can provide an opportunity to investigate relationships that would otherwise be impossible to look at.

Secondly, large disparities between ecologic and individual exposures are unintuitive. Thirdly, in some circumstances analyses are valid because ecologic effects are of primary interest. Fourthly, ecologic analyses can be useful for generating hypotheses. Rhind and Tannenbaum (1983) argue that although there is no solution to the problem, when there is no alternative to using aggregate data the effects can be minimised by using the most disaggregated data available. For data on the 1981 census it was only possible to relate postcodes accurately to ward level data. Only 6% of wards have a population of less than 1000 and 25% have a population of less than 2000 and the wards vary greatly in size (up to over 15000). Also ward boundaries may be defined to encapsulate a certain number of voters as opposed to identifying a natural community. A 'natural' community can often be split into several wards and different communities can be combined into one ward. Therefore it is difficult to produce a complete analysis of deprivation using data at ward level. Enumeration district level data (based on the households covered by each census enumerator) may be more appropriate (about 150 to 200 households) minimising problems of dividing communities but this is not always possible. However Carr-Hill and Rice (1995) have indicated that ward data may be as good at reflecting the aggregate characteristics of an area as ED level data.

If individual and area level deprivation are highly correlated then ecological bias should not occur although the observed effect of deprivation on the outcome of interest may be blunted. If they are poorly correlated then this can lead to ecological bias and possible reversal of effects. In the case of deprivation, MacRae (1994) argues that there is evidence such as that shown by Sloggett and Joshi (1994) using individual level data to support the argument that ecological correlations between deprivation and health can be seen to arise from associations at the individual level and that sceptics cannot legitimately use the issue of ecological fallacy to dismiss the findings of ecological studies. Hence, it is more probable that analyses based on area level data will lead to underestimation of the effects of deprivation on health rather than overestimation or reversal of effects.

Whenever using aggregate data it is important to appreciate this serious problem. In many ecologic studies where ecological bias is a problem both the outcome and the exposure are measured at an aggregate level. However, in the studies of amblyopia described in this thesis it is only the exposure, deprivation, which is being measured at an aggregate level and information on the outcome, age at presentation, is available at the individual level. This

reduces the problem somewhat. In this thesis I investigate the relationship between area level and individual level deprivation to see whether area measures are distorting the effect of deprivation found at an individual level.

3.4.6 THE PROBLEM OF UNDERENUMERATION IN THE 1991 CENSUS

Unlike most surveys the Census has an extremely high rate of coverage at 97.8%. However, although the rate of enumeration is high, the 2.2% who were not enumerated do not appear to be randomly distributed in terms of their demographic characteristics. The rates of underenumeration appear to be highest among men aged 20-29 years in inner city areas (Majeed et al 1995) with the rates of underenumeration being 9% for men aged 20-29 nationally and nearly 20% in inner London. Glover (1993) has shown that these rates of underenumeration are even higher for young male Black Caribbeans and that this may lead to overestimating rates of disease or use of services in affected areas. Carr-Hill (1993) illustrates that it is not only Black Caribbeans but other minority ethnic groups who are underenumerated. High levels of underenumeration in areas with deprived, mobile populations may have an impact on the deprivation payment system where a deprivation measure is used to allocate payments to GPs (Majeed et al 1996). OPCS have produced factors for adjusting for underenumeration (Thompson 1995) but these are based on the whole of the UK. Since the work in this thesis was performed, gold standard population estimates (Estimating with Confidence) have been produced by the Census Microdata Unit at Manchester University based on age, sex and other sociodemographic characteristics such as ethnicity (Leese et al, 1995) and these are recommended for future use. Leese et al illustrate the size of the problem since population adjustment factors of 1.4 were used for men aged 25-29 for a particular ward in inner city London meaning that nearly 30% of the population are estimated to be not enumerated by the census.

Since the work in this thesis concentrates on the use of deprivation scores in Leicestershire, where there is a high population of Asians living in the inner city area, underenumeration is a potential problem. It could lead to serious errors in deprivation scores calculated and the size of the underlying populations in enumeration districts. However since this work concentrates on children the effect of underenumeration is thought to be less of a problem. Glover (1993) indicates that one of the main reasons for underenumeration in young male Black Caribbeans is that a far higher proportion of Caribbean than White households contain only one adult. He believes that those households containing only one woman are more likely to include children and are consequently more home based and easier to contact than single male households. Since this work is looking at

children and deprivation scores are to be based on households with at least one dependent child under 15 then these households are less likely to be underenumerated. Although underenumeration is a problem, it is less of a problem than is first apparent in this case and the effects are likely to be marginal.

3.5 THE JARMAN SCORE

3.5.1 AN OVERVIEW OF THE JARMAN SCORE

Jarman (1983) published a paper describing a method for allocation of resources based on factors affecting workload identified in a survey of general practitioners. It was developed using questionnaires sent to a one in ten sample of G.P.'s from a commercial mailing list. G.P.'s were asked to score 13 factors according to the degree to which each one *'increases workload or contributes to the pressure of work when it is present'*. Ten of these variables were then used to construct the score, transforming each one by an arcsin transformation and weighting them according to the average score received in the survey. Jarman found the score to have a high correlation with measures of illness in different areas, and claimed it was a method which could be used for 'identifying underprivileged areas'. This claim led to its use in many studies as a 'deprivation score' rather than a method of resource allocation. It was also adapted to include just 8 variables, the 'underprivileged area score' (UPA8) and was subsequently used for resource allocation.

3.5.2 THE RESPONSE TO THE JARMAN SCORE

As discussed previously, the Jarman score has been frequently adopted to measure material deprivation in epidemiological studies. However, there was a great deal of criticism about its use in this context. The majority of the indicators in the UPA8 score are indirect measures of deprivation e.g. ethnic minorities and lone pensioners. Although these people often suffer extreme deprivation it is wrong to assume that they all do. Also, two of the variables with the highest weights reflect the age distribution of the population ('under age five' and 'lone pensioner'). This leads to particularly high scores (i.e. deprived) in popular retirement areas which are not necessarily areas of material deprivation. This reflects the way the measure was developed as a resource allocation formula. Carr-Hill et al (1992) criticise the score as it ignores the inter-correlation between the variables. Some variables may be adding little extra information to the score. Furthermore, the correlations between the transformed variables are even higher.

The inclusion of variables such as 'ethnic minorities' led to high scores in the south, particularly areas of London, while areas in the North had generally lower scores. Seven of the health authorities with the highest scores were in London. This goes against the trend of increasing mortality rates in the North compared with the South 'flying in the face of most observation and experience' (Townsend et al, 1988). I believe a further problem is that comparison of deprivation scores between ethnic groups is invalidated because of the inclusion of ethnic group in the score.

3.5.3 FURTHER WORK

One of the principal problems with using the Jarman score is that it is a measure of G.P. workload which has been applied to a wider context. This has led to the development of further scores of deprivation for use in studies of inequalities in health. The main composite scores of material deprivation to have arisen from this work are the Townsend score (Townsend et al, 1988) and the Carstairs score (Carstairs and Morris, 1989a,b). There are further scores such as that developed by Thunhurst (1985) which attempts to retain the multi-dimensionality of deprivation. Finally there are a variety of scores prepared for resource allocation but these are not suitable for use in looking at deprivation and health owing to the context of their construction.

3.6 UNIDIMENSIONAL INDICATORS

3.6.1 A COMPARISON OF THE SCORES

Carstairs and Morris (1989a,b) developed their measure of material deprivation to overcome the problems of social class. It consists completely of direct indicators of deprivation (table 3.2). Independently, Townsend et al (1988) developed a measure of material deprivation, again based on direct measures of deprivation. This work was performed in an attempt to show the areas of extreme deprivation in the North which had not been identified by the Jarman score. The Department of the Environment (1983) also developed a measure in relation to urban policies. Table 3.2 demonstrates the indicators and the weights attached to these scores and the UPA8 score (Jarman 1983) which have all been frequently used for epidemiological studies of deprivation and health. This table shows the different methodological variations, particularly the number, type and weights of the indicators. The Jarman and the Department of the Environment measures are based predominantly on unequally weighted combinations of indirect indicators while the Carstairs and Townsend scores are based on equally weighted combinations of direct indicators.

Table 3.2: Table of indicators and associated weights for four deprivation scores

	INDICATOR	CARSTAIRS	UPA8	TOWNSEND	DOE
UNEMPLOYMENT	DIRECT	1	3.34	1	2
NO CAR		1		1	
LOW SOCIAL CLASS		1			
UNSKILLED			3.74		
OVERCROWDING		1	2.88	1	1
NOT OWNER OCCUPIER				1	
LACKING AMENITIES					1
SINGLE PARENT	INDIRECT		3.01		1
UNDER AGE 5			4.64		
LONE PENSIONERS			6.62		1
1-YEAR IMMIGRANTS			2.68		
ETHNIC MINORITIES			2.50		1

Morris and Carstairs (1991) investigated these measures (with a variation on the DOE measure) and each individual variable in relation to their performance in explaining variation in a range of health measures based on postcode sector data in Scotland. They found that Carstairs, Townsend and DOE scores correlated most highly with measures of death and sickness, but the Jarman and DOE scores correlated more highly with hospital use variables. This relationship was explained by the inclusion of variables relating to elderly people living alone since they make more use of these services. Removing the 'lone pensioner' variable from the Jarman score led to a significant reduction in correlation with health measures. They also found that 'No car' performed as well as the composite indicators and was appealing in its simplicity. However, they concluded that composite scores take more account of the different aspects of deprivation and are less susceptible to instability caused by a rapid change in one indicator. Campbell, Radford and Burton (1991) and Haynes et al (1996) have performed similar analyses recommending the use of unemployment as a simple indicator and emphasise its usefulness in intercensal years. However although simple, this has the same problem as using 'No car'.

3.6.2 CONCLUSIONS ON CHOOSING A UNIDIMENSIONAL SCORE

The Jarman score appears to correlate less strongly with health measures than the other deprivation scores considered. To use it in analyses of deprivation and health, despite its previous wide usage, is inadvisable since it was developed for resource allocation. The other measures are becoming more widely known and are much simpler to calculate, with the

Townsend score and the Carstairs score both comprising just four indicators weighted equally. They also had a much higher correlation with health measures. Unemployment rates are useful as they also correlate highly with health measures but these rates can change rapidly for various areas and so a composite score seems more stable. The ideology behind the construction of the Townsend score led to a more direct measure of material deprivation than any of the other scores. Its use in research relating to England is perhaps more suitable than the Carstairs score which is very similar but constructed using Scottish health data. Also the Carstairs score includes social class which Carstairs and Morris had previously argued against using. Therefore, the use of the Townsend score in health research has greater potential.

3.7 THE TOWNSEND SCORE

3.7.1 INTRODUCTION

The aim of Townsend et al (1988) was to map health differences and to see to what extent they were matched by differences in material and social conditions. They concentrated on the Northern region as the Jarman score had generally given low UPA scores (i.e. not underprivileged) to areas in the North that were deprived. The score is based on four indicators:

Unemployment:	% of unemployed economically active residents aged 16-59/64.
Car ownership:	% of private households who do not possess a car.
Home ownership:	% of private households not owner occupied.
Overcrowding:	% of private households with more than one person per room.

These indicators were chosen to directly reflect material deprivation and *not* the type of people who generally suffer material deprivation. Townsend et al give the following reasons for their choice of variables. Unemployment was chosen as it indicates a lack of access to earned income and the facilities of employment. It also shows a general lack of material assets and insecurities than can arise from this. Car ownership is a good surrogate income measure in the short term as they have to be bought, replaced, repaired and licences, insurance and MOT's must be paid for. Non-owner occupancy is a good long-term surrogate measure of lack of wealth as well as income. Finally, overcrowding is a good guide to overall housing conditions.

Previously households without sole use of basic amenities have been used as a measure of deprivation (i.e. no bath or no indoor toilet) but this is no longer a useful measure. This is particularly evident in council estates where all basic facilities are available but the actual structural conditions are poor and the quality of the basic amenities is very low.

3.7.2 CALCULATION OF THE TOWNSEND SCORE

The score is a combination of the four indicator variables discussed: unemployment, car ownership, house ownership and overcrowding. The percentage prevalence of the four indicators in each area are used to construct the score. For illustrative purposes, the mean and standard deviation of these percentages for each of the indicators was calculated by Townsend et al for wards in the Northern region concentrated on by the study and these can be seen in table 3.3.

Table 3.3: Summary statistics of the variables in the Townsend score for the Northern region

	Untransformed variables		Transformed variables	
	Mean	Standard Deviation	Mean	Standard Deviation
Unemployment	11.72	6.56	2.42	0.51
Lacking car	42.01	17.58		
Not owner occupied	50.10	22.77		
Overcrowding	3.31	2.01	1.36	0.45

Both unemployment and overcrowding were found to be non-normally distributed and so were transformed by a natural log transformation ($y=\ln(x+1)$). The score appears to ignore the fact that all percentage scales are bounded between 0 and 100 and so none of the indicators would be normally distributed. This was not a problem in 1981 but subsequent reductions in the prevalence of these indicators may cause some concern since the aim when composing the score was to weight each variable equally and if the scores are not normally distributed then this may lead to unequal weighting. The score is then constructed by standardising the observed indicator values to form z-scores by using the mean and standard deviation and then summing the four values to give a score. The mean and standard deviation of the whole area under study should be used to form the z-scores. Deprivation scores then indicate the level of deprivation relative to the average deprivation in the area under study.

3.7.3 PROBLEMS WITH USING THE TOWNSEND SCORE

Since the Townsend score and other measures were first constructed, the situation in Britain has changed dramatically. House ownership has increased substantially inferring a decrease in Townsend score for some areas. However, large mortgages and high interest rates, the subsequent slump in house prices and rising unemployment means that people in these areas could be experiencing far greater material deprivation than before. Also, car ownership is increasing so this variable is only detecting an extremely small deprived group. Unemployment rates are constantly changing making census data quickly out-of-date. It is important to look carefully at the census data being used to construct the deprivation score and take into account the changes that have occurred since the collection of the data. Mass unemployment due to local industry changes can dramatically and quickly change the material deprivation suffered in an area yet census data could be up to ten years out of date and hence would not take these factors into account.

Townsend performed a log-transformation on two of the variables, unemployment and overcrowding, since the data was extremely skewed in 1981. With the changes in house and car ownership noted above all four of the variables are now skewed and perhaps the possibility of improving the measure could be investigated.

3.8 MULTI-DIMENSIONAL INDICATORS

Thunhurst (1985) argued that reducing deprivation to a single value on a scale ignored the multifaceted nature of the problem both for individuals and between areas. This is particularly important when looking at urban and rural areas which are inherently different. Reading, Openshaw and Jarvis (1994) indicate that 'a unidimensional view may hide more than it reveals, given the complexity of interactions between socio-economic, environmental and behavioural influences on health'.

Thunhurst (1985) and Folwell (1995) have encouraged the use of multidimensional area classifications to reflect the diversity of areas when investigating deprivation and health. These classifications are usually derived using cluster analyses looking at various census variables to achieve 'a classification of residential neighbourhoods' of homogeneous areas, with similar

patterns of socio-economic, housing, demographic and environmental factors known as 'profiles' or 'lifestyles'. Generally, geographers have attempted to develop these area identifiers while the unidimensional scores have been developed by sociologists, general practitioners and health services researchers.

Much of the early work in multi-dimensional measures was done by market research groups to distinguish between different types of consumer behaviour for use as a marketing tool. The ACORN lifestyles were developed in this way but have also been used to investigate variations in health (Morgan and Chinn, 1983). The 1971 ACORN classification was created from 40 variables relating to age structure, employment status, family structure, type of housing, social status and car ownership. Cluster analysis techniques were used to identify 37 ACORN types identifying different types of neighbourhoods and these were collapsed to form eleven ACORN groups. These groups are not ranked and are given short titles to reflect the population such as 'low income areas with immigrants' or 'modern family housing higher income'. Morgan and Chinn (1983) found that these groups explained as much of the variation in health as social class but they were concerned that underlying regional variations may have confounded the relationship.

More recently new classifications of ED's have been developed for the 1991 census such as the GB profiles developed by the School of Geography at Leeds University (Openshaw and Blake, 1995). A neural Network Classification Procedure uses 80 different variables from the census small area statistics for each enumeration district to produce 10, 49, 64 or 100 clusters for the whole of Great Britain. Data on demography, ethnicity, housing, household composition, socio-economic status, migration, health and work travel are used to produce these clusters. The program then gives a classification of the area. These classifications comprise of three parts, an overall description of the area (e.g. struggling, aspiring, established, climbing, prospering), a description of the residents (e.g. age, ethnicity, family status, occupation) and a description of the physical amenities (e.g. housing type and tenure). For example an enumeration district may be described as 'Aspiring; young married suburbia - young well-off blue collar couples and families - mixed tenure terraces' or alternatively 'Struggling; multi-ethnic area - pensioners and single parents - high unemployment - LA rented flats'. There is no explicit discrimination between rural and urban areas. Evidently the labels are extremely subjective and there is no explicit ranking of areas.

There are many advantages of using multi-dimensional area classifications. They maintain the multi-dimensional view of deprivation. Areas with a particular combination of socio-economic and environmental factors which have a great influence on health can be identified. These measures also remove the arbitrary ranking of areas which can be inaccurate particularly in comparing two different regions. They can also be more reliably used for health service provision (Thunhurst 1985).

There are also drawbacks in using this type of measure. For example, a frequent use of deprivation scores is to look for a trend with increasing deprivation. In using this type of multi-dimensional measure there is no rank order to the lifestyles and so this cannot be done. This also reflects the fact that it is difficult to understand the hierarchical structure of the classifications. If health variations are found to be related to multi-dimensional classifications it is impossible to identify the contribution of individual variables that may be related to the health outcome.

In areas where using different measures with different results could lead to major practical problems such as the allocation of GP payments for deprived areas where the choice between two measures or types of deprivation measures could lead to very real monetary differences for certain practices, the choice of indicator is crucial. Multi-dimensional classifications may be more useful for planning health services since they offer the ability to identify characteristics about different areas and allow targeting of certain areas for health promotion purposes. However, in this work the aim is to look at how deprivation indices can be used to look at inequalities in the presentation of amblyopia and should provide a starting point for further localised in-depth investigation. One of the aims of this thesis is to show how measures can be used in a practical way to monitor future presentation patterns in amblyopia and other areas and potentially incorporated into audit. Hence an easily available, well-recognised, easy to understand measure is needed and for this reason the Townsend score is used. This is not to dispute the importance of multi-dimensional measures and their use would be a logical extension of this work. These multi-dimensional classifications have now become extremely easy to access unlike predecessors based on the 1981 census data and hence in the future will be far easier to use. At the time of undertaking the work in this thesis, this was much more difficult.

3.9 SUMMARY AND CONCLUSIONS

In this chapter I have shown many problems in using deprivation measures of the types discussed. However many research projects employ them as they are currently the only means of measuring socio-economic and environmental factors and they have been shown to be a valuable tool in examining health differentials (Carstairs, 1995).

In order to minimise the effect of the problems discussed it is important that methods should be suitable for the context in which they are to be used. The most obvious example of inappropriate use is that of using the Jarman score, designed for resource allocation purposes, to investigate deprivation and health. It is important to choose a score specifically intended to look at deprivation for this purpose. Since this thesis is looking to see if there is a relationship between presentation of amblyopia and deprivation, the Townsend score has been chosen since it has been designed for health research to look at material deprivation. Although composite indicators lose the multi-dimensionality of the subject, they are useful for providing summaries of the data, presenting data in a comprehensible way and can also be used to investigate trends within the data. Since the Townsend score is frequently used it also enables comparison and 'academic peer review'. For these reasons, in this thesis I use the Townsend score to investigate inequalities in the presentation of amblyopia.

Information on deprivation from the census may be an improvement on using social class data but there are still limitations with this data. Underenumeration can be a major problem particularly among young males from minority ethnic groups. This is perceived to be less of a problem in this research since it is based on families where enumeration is thought to be more complete. The problem of 'ecological fallacy' has also been raised in this chapter, illustrating the problems of assuming area level characteristics to exist at an individual level. In this thesis I also investigate this under-researched area and look at whether area measures are underestimating the effect of deprivation found at an individual level.

In chapter 4 I investigate deprivation in Leicestershire, the main setting for the work in this thesis, using the Townsend score and other deprivation scores. I then use the Townsend score in chapters 6, 7 and 8 to quantify the levels of inequalities in the presentation of amblyopia before and after the introduction of an intervention to improve screening for vision problems.

The issue of ecological fallacy raised in this chapter is then addressed by comparing area level and individual level data in chapters 9 to 12.

CHAPTER 4

PATTERNS OF DEPRIVATION IN LEICESTERSHIRE

4.1 AIMS OF THE CHAPTER

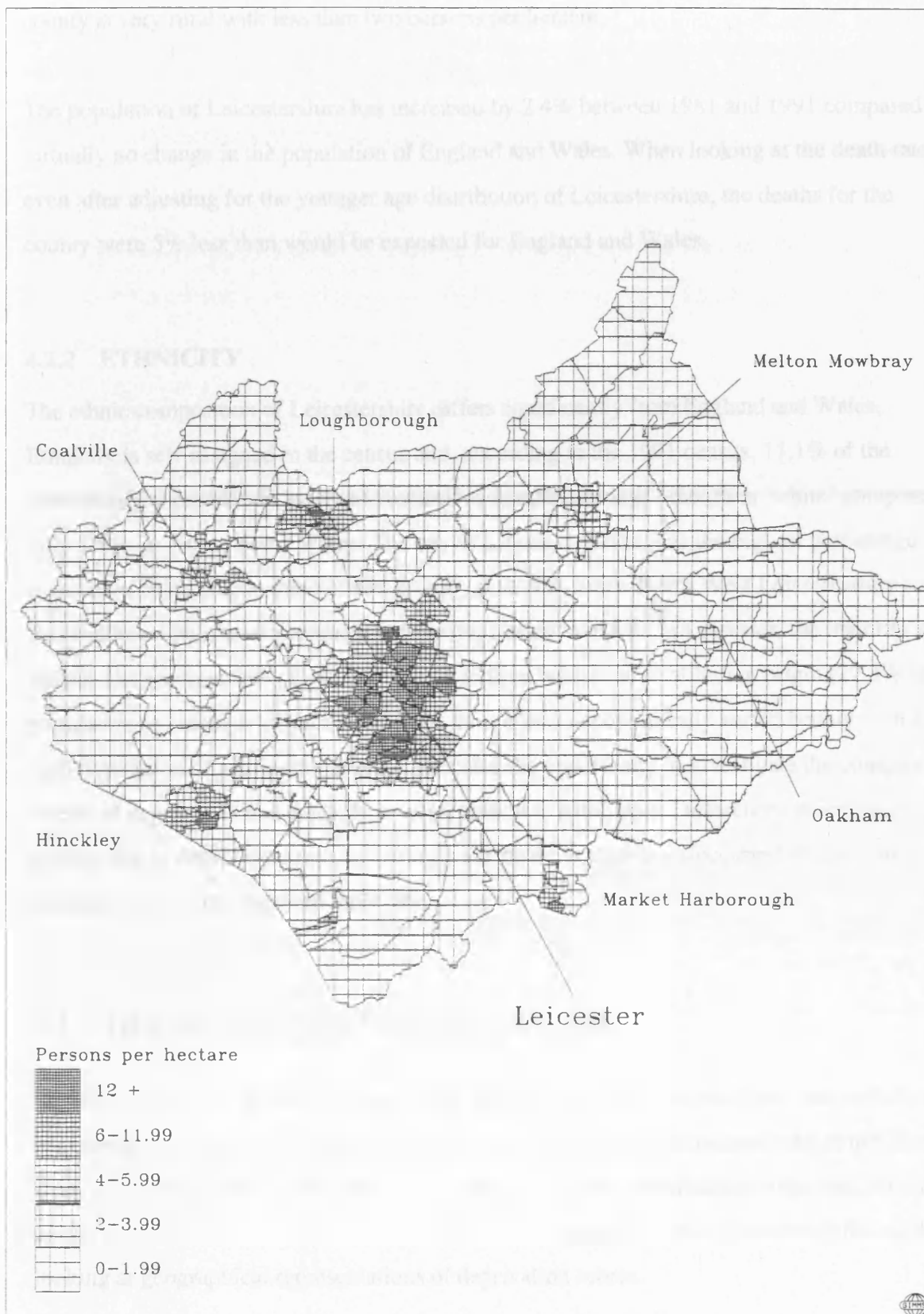
The previous chapter discussed various measures of deprivation that have been developed. In this chapter I look at the distribution of deprivation using these measures in the context of the main study setting. The majority of work documented in this thesis takes place in Leicestershire. The distribution of the population of Leicestershire is discussed and is compared with England and Wales. This is then followed by a comparison of the patterns of deprivation when measured by different methods, which were discussed in chapter 3, focusing on the Townsend score. It concentrates on data from the 1991 census but changes over time are discussed, by comparing this with data from the 1981 census. The aim is not to justify the choice of the Townsend measure for use in this thesis but to show how using different measures could affect the outcome of the research.

4.2 POPULATION OF LEICESTERSHIRE

4.2.1 POPULATION DENSITY

Leicestershire is a typical shire county in the East Midlands. With a population of approximately 865,000 (census 1991), about a third of the population live in Leicester city and about half of the population live within five miles of the city centre. The rest of the population is concentrated in the surrounding towns of Loughborough, Market Harborough, Oakham, Melton Mowbray, Hinckley and Coalville. This is highlighted by the population density in Leicester district being 36.9 persons per hectare (176.9 in Spinney Hill ward) compared to 3.4 persons per hectare for the county as a whole and 3.2 persons per hectare in England and Wales. These patterns can be seen in figure 4.1 where the population per hectare is mapped for the wards of Leicestershire. GIMMS (Carruthers and Waugh, 1992), a geographical mapping package was used to produce this figure.

Figure 4.1: Population density of Leicestershire wards (persons per hectare)



This map illustrates how the concentration of the population is in the city of Leicester with extremely high population density. The rest of the population can be seen to be concentrated in the six county towns, although this is at a much lower population density. The rest of the county is very rural with less than two persons per hectare.

The population of Leicestershire has increased by 2.4% between 1981 and 1991 compared to virtually no change in the population of England and Wales. When looking at the death rates, even after adjusting for the younger age distribution of Leicestershire, the deaths for the county were 5% less than would be expected for England and Wales.

4.2.2 ETHNICITY

The ethnic composition of Leicestershire differs significantly from England and Wales. Ethnicity is self-assigned in the census and, according to the 1991 census, 11.1% of the Leicestershire population assigned themselves to ethnic groups other than 'white' compared to only 5.5 % in England and Wales. The city of Leicester district has the highest percentage of population belonging to these ethnic groups, at 28.5%, while North West Leicestershire and Rutland have the lowest percentage of the population, at 0.8%. In Leicester, the majority of the population classifying themselves as other than 'white' were of Indian origin (22.3% of the population of Leicester). Since Leicestershire is a socio-economically varied county with a high population of ethnic minorities, it provides the opportunity to investigate the complex effects of deprivation and ethnicity on presentation in amblyopia. Differences in amblyopia referral due to deprivation may be wrongly attributed to ethnicity since most of the ethnic minorities live in the deprived inner city.

4.3 DEPRIVATION IN LEICESTERSHIRE

In order to assess the levels and distribution of deprivation in Leicestershire, various indicators of material deprivation are mapped and discussed. This work does to some extent replicate the work of Carstairs and Morris (1991) who compared various deprivation indices but it focuses on the setting of much of the work in this thesis, Leicestershire, and also extends the work to looking at geographical representations of deprivation scores.

4.3.1 MATERIAL DEPRIVATION INDICATORS IN LEICESTERSHIRE

Leicestershire is made up of approximately 350,000 dwellings. The housing in Leicester city comprises about 34,000 dwellings built before 1918 (mostly terraced housing in the inner city area built after 1870), 32,000 built between 1918 and 1945 (mainly semi-detached) and 40,000 built since 1945. The housing policies of the city council in the 1980's were to improve both city council estates and older housing by renovation and then to build new housing to meet the needs of the community.

The census describes the population in terms of households rather than houses. A household is defined in the census as 'either a person living alone or a group of persons (who may or may not be related) living at the same address with common housekeeping' (Denham and Rhind 1983). Therefore a house may comprise of more than one household. Of Leicestershire households, 17% were rented from a local authority in 1991, 72.5 % were privately owned and the percentage of owner-occupiers varied greatly between district with 57.5% of households being privately owned in Leicester compared to 82.5 % in Oadby and Wigston.

Overcrowding is used as a measure of material deprivation in the Townsend score and is assessed in terms of households with more than one person per room. This indicator is shown to be decreasing with a reduction from 3.4% of households with more than one person per room in 1981 to 2.0% in 1991, with the most overcrowded areas being in the city.

Car ownership is thought by Townsend to be a surrogate measure of current income but it is also linked to the rurality of residence. Therefore, rural areas generally have a high level of car ownership, although other characteristics of deprivation may be much higher. In 1991, 29% of Leicestershire households had no access to a car. This was highest in Leicester city district at 45% but much lower in all of the other districts with Blaby and Harborough being the lowest at less than 18%.

Unemployment is used by Townsend to represent a general lack of material resources. Unemployment in 1991 in Leicestershire was 7.9% for men and 4.2% for women. It was highest in Leicester (13.9% men and 7.0 % women) and lowest in the rural districts of Melton and Rutland (4.5 % and 2.6%).

There is a consistent pattern for all these indicators with higher levels of deprivation indicators in the city of Leicester than the more rural surrounding districts.

4.3.2 COMPARING DEPRIVATION SCORES

Figures 4.2-4.5 show the geographical distribution of deprivation in Leicestershire wards according to the four most frequently used deprivation scores, the Townsend deprivation score, the Carstairs deprivation score, the Jarman score of underprivileged areas and the Department of Environment score. The work in this thesis is based primarily on deprivation at ward level since enumeration district level data was hard to obtain from the 1981 census. Therefore ward level data have been used in this chapter to illustrate patterns of deprivation. For this chapter the four different deprivation scores were obtained from work by Jane Eimermann held in a file at Manchester Computing Centre. Several scores for the Department of the Environment index and for the Townsend index were missing from this file due to the absence of either overcrowded households and/or households lacking basic amenities in certain wards. Therefore there are a small number of missing scores for wards in the maps presented here.

For each of these measures, the deprivation scores were ranked and then divided into population weighted quintiles of wards ranging from the most affluent 20% of the Leicestershire population to the most deprived 20% of the population. These maps have been produced purely as an attempt to illustrate how deprivation is distributed in Leicestershire and how all of these deprivation measures show very similar patterns. However there are limitations with this type of map that must be kept in mind. Rhind (1983) has pointed out that mapping can be informative but can also mislead. People are very unevenly distributed across Leicestershire and different areas may appear grossly under or over represented if equal area maps are used rather than those adjusting for differences in populations (Dorling, 1995). This must be appreciated since rural areas may appear large on the map but only have a small population. However these maps do give a general impression of the distribution of deprivation in Leicestershire with the most deprived areas concentrated in the city. It also shows the similarity of the four deprivation measures. This is particularly evident in the deprived areas of the inner city. There are some differences between the scores particularly in the urban areas outside of the city where areas are generally less deprived.

Fig 4.2 Quintiles of Carstairs deprivation score for Leicestershire wards

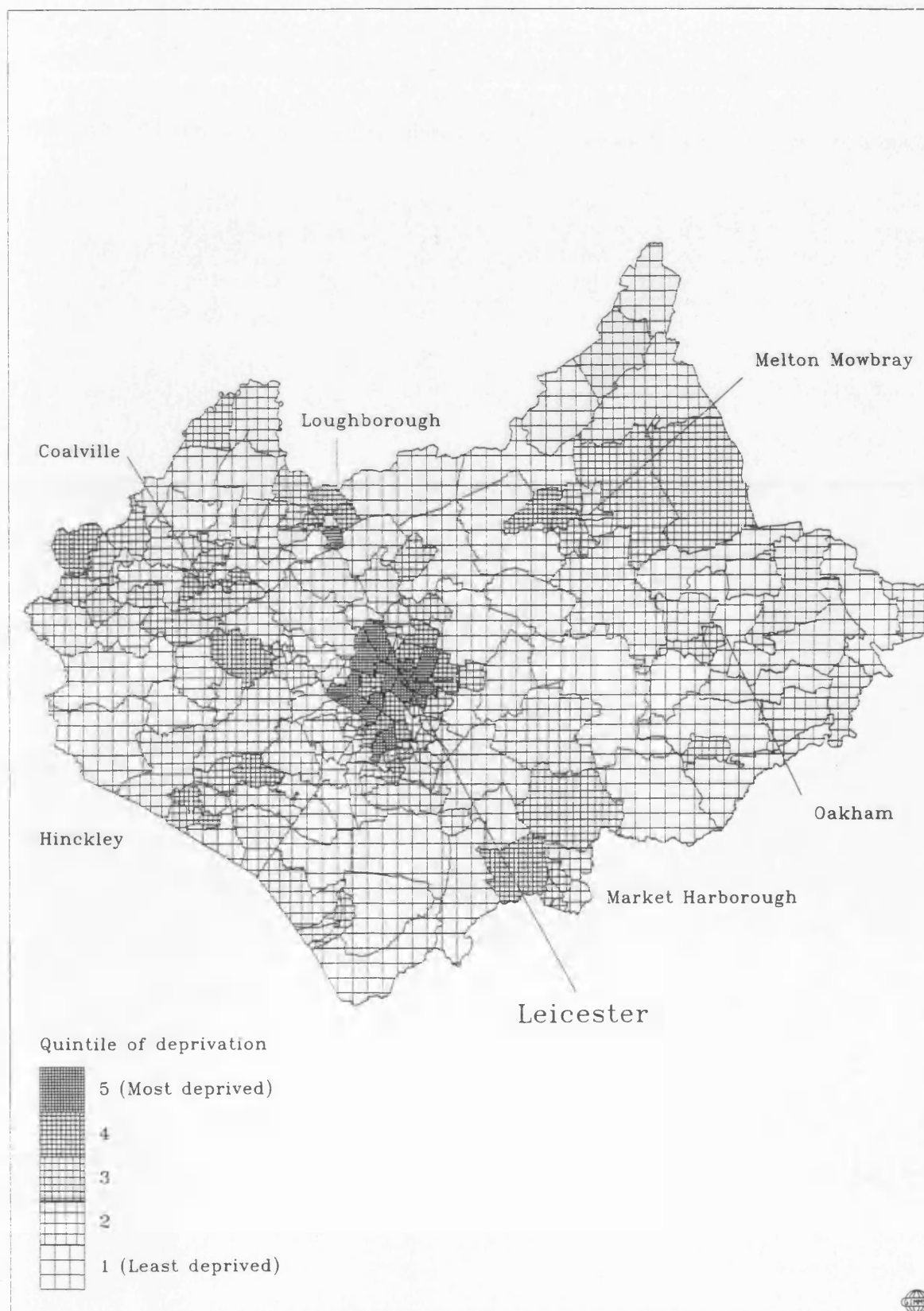


Fig 4.3 Quintiles of DOE deprivation score for Leicestershire wards

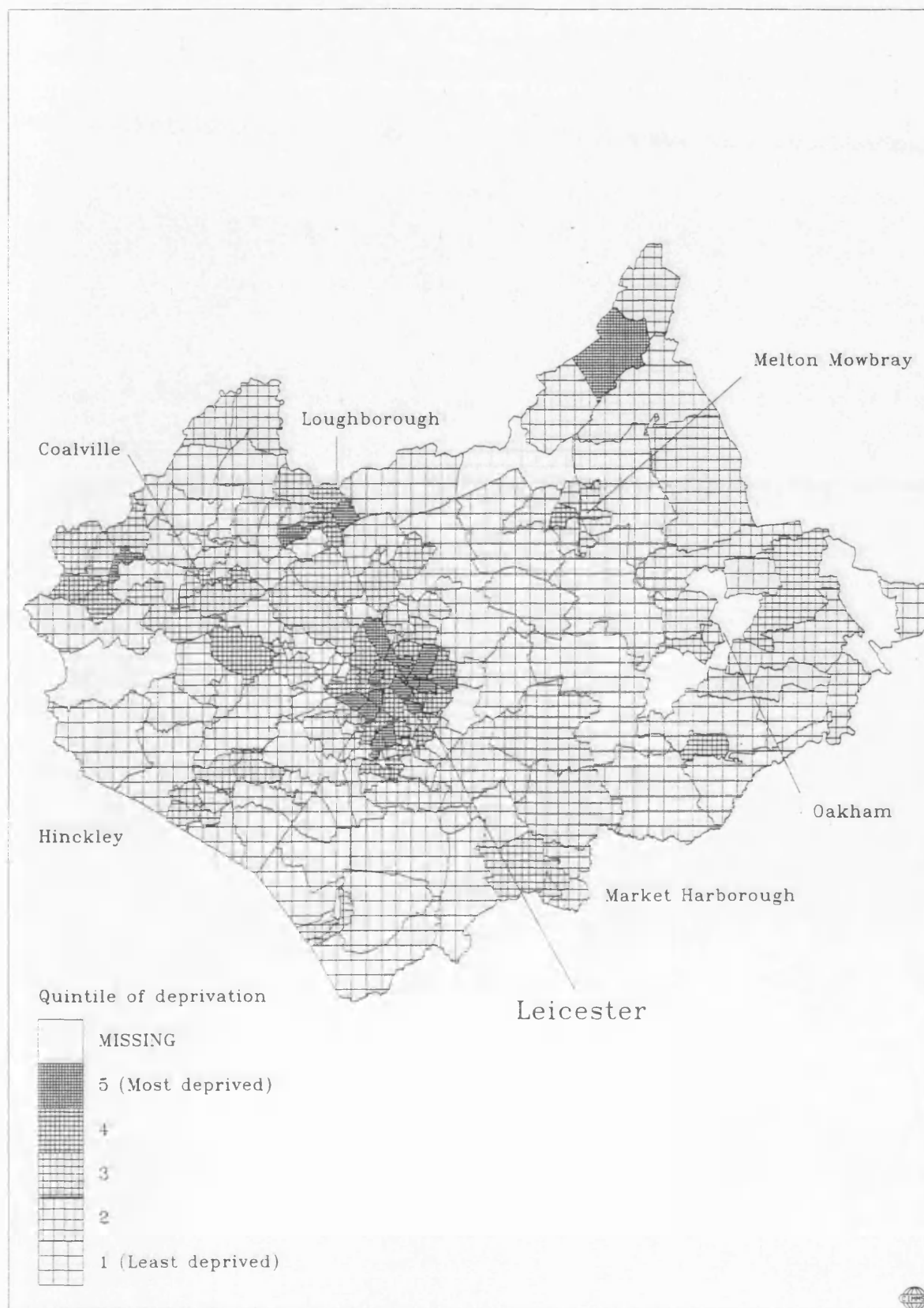


Fig 4.4 Quintiles of Jarman deprivation score for Leicestershire wards

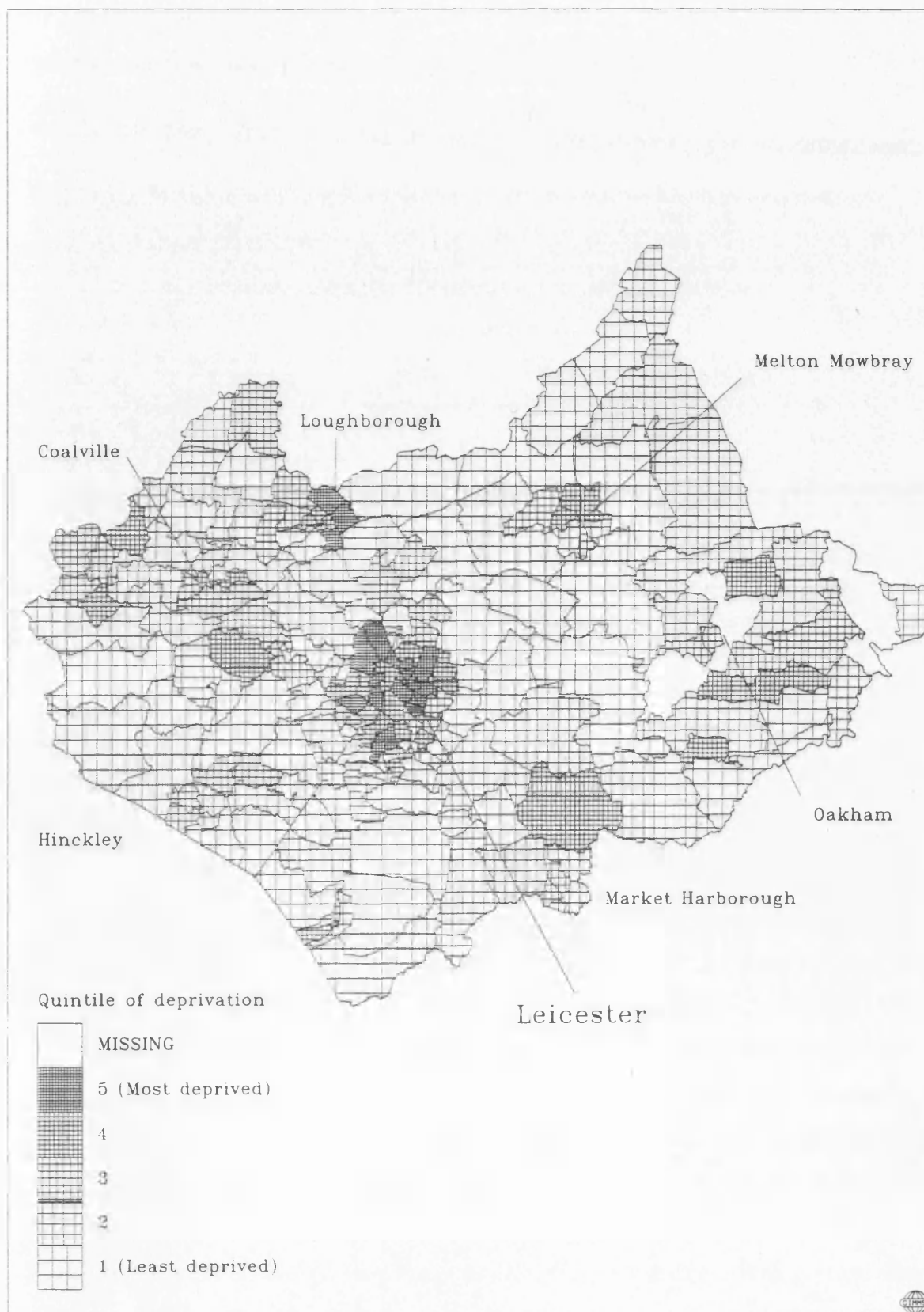
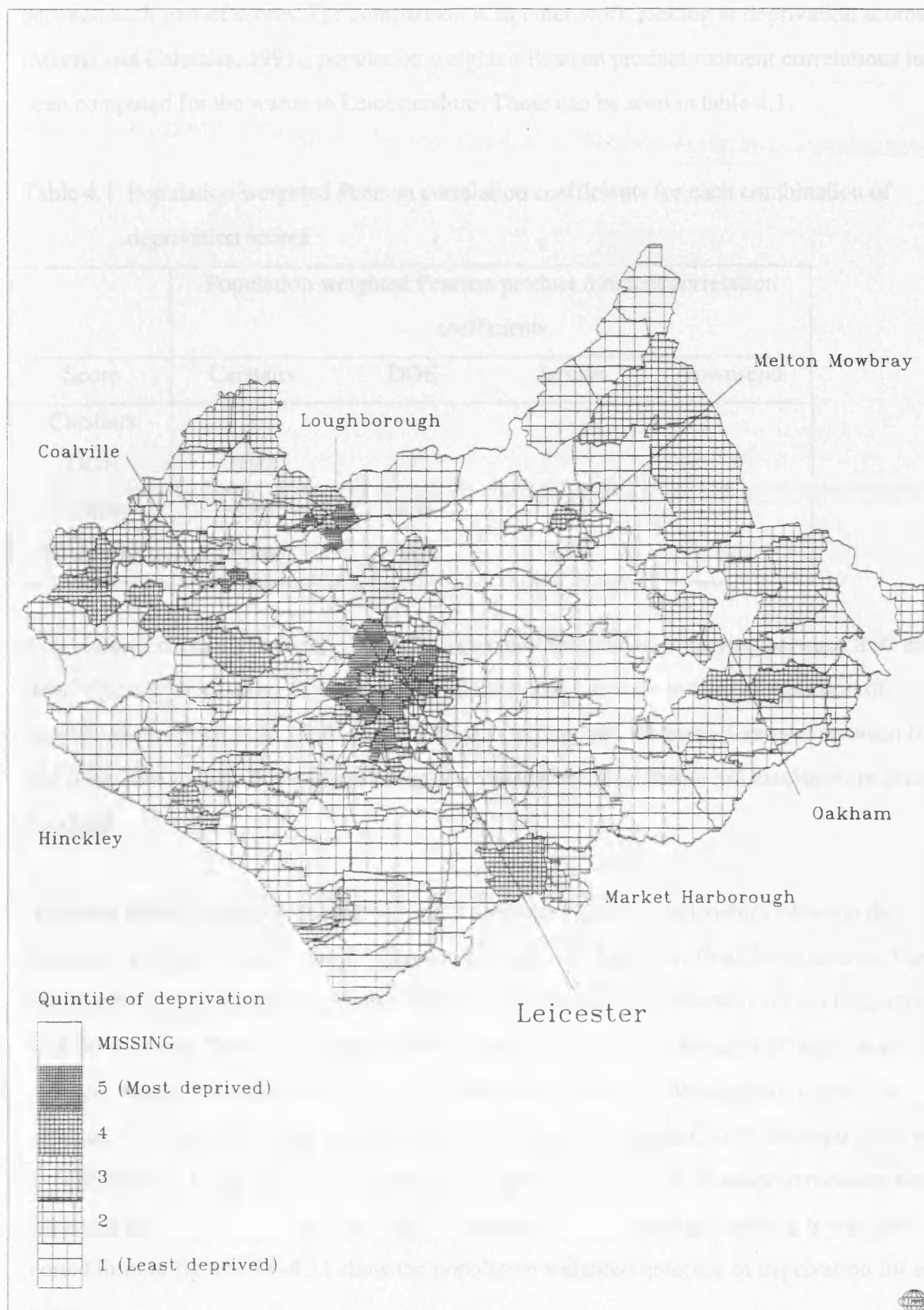


Fig 4.5 Quintiles of Townsend deprivation score for Leicestershire wards



In order to compare the differences in the distribution of deprivation, as defined by the four different measures, scatterplots of the ward deprivation scores were made for each pair of scores. These can be seen in figures 4.6-4.13. These show a general strong relationship between each pair of scores. For comparison with other work looking at deprivation scores (Morris and Carstairs, 1991), population weighted Pearson product moment correlations have been computed for the wards in Leicestershire. These can be seen in table 4.1.

Table 4.1: Population weighted Pearson correlation coefficients for each combination of deprivation scores.

	Population weighted Pearson product moment correlation coefficients			
Score	Carstairs	DOE	Jarman	Townsend
Carstairs	1			
DOE	0.92	1		
Jarman	0.95	0.96	1	
Townsend	0.96	0.95	0.96	1

This shows extremely high correlation between the four indicators varying between 0.92 and 0.96 compared with those demonstrated by Morris and Carstairs in their comparison of deprivation scores for the 1010 postcode sectors in Scotland where they ranged between 0.80 and 0.96. This may be due to there being less variability within wards in Leicestershire than in Scotland.

Although these correlation coefficients show extremely strong relationships between the scores, it is of interest how people's deprivation ranking changes with different scores. The deprivation scores are continuous but frequently deprivation scores are divided into groups such as quintiles. Therefore, when looking at the how these four measures of deprivation differed, one of the main points of interest was the how the quintile assigned to a person changed with different deprivation scores. For example, the proportion of the population who are classified as being in the same quintile whichever score is used. In order to measure this degree of agreement between deprivation measures, firstly a graphical approach was used. The dotted lines in figures 4.6-4.11 show the population weighted quintiles of deprivation for each score.

Figures 4.6-4.11: Scatterplots of ward deprivation for Leicestershire for each pair of deprivation scores (Dotted lines indicate population weighted quintiles of deprivation for each score. Positive scores relating to more deprived areas)

Figure 4.6: Carstairs Index and DOE Index

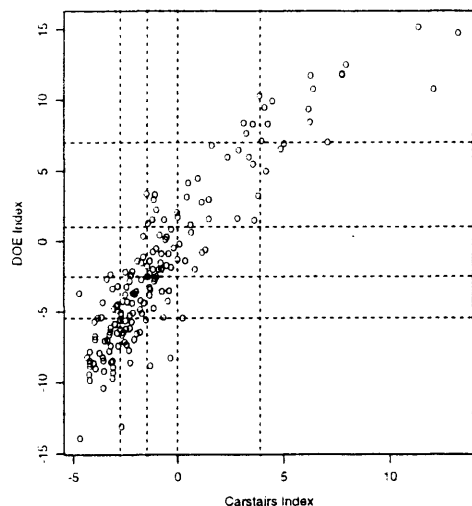


Figure 4.8: Carstairs and Townsend Index

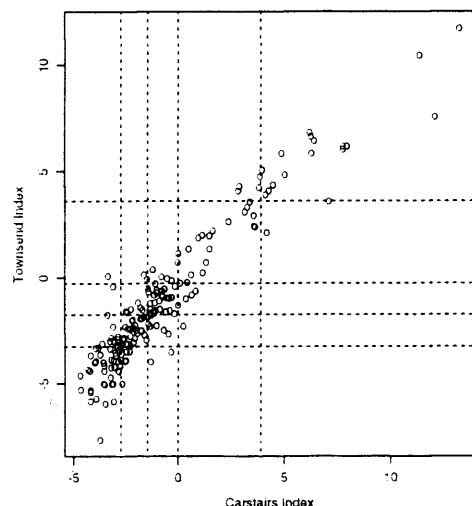


Figure 4.10: DOE Index and Townsend Index

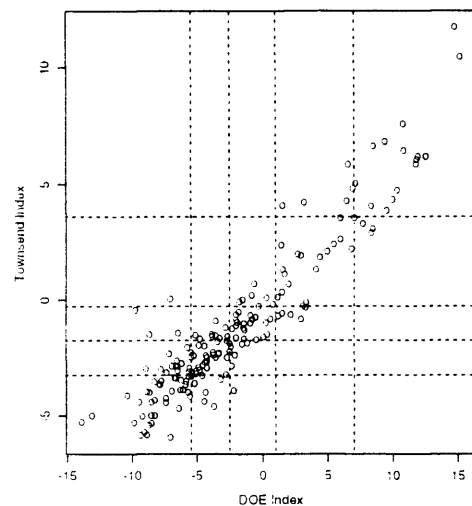


Figure 4.7: Carstairs Index and Jarman Index

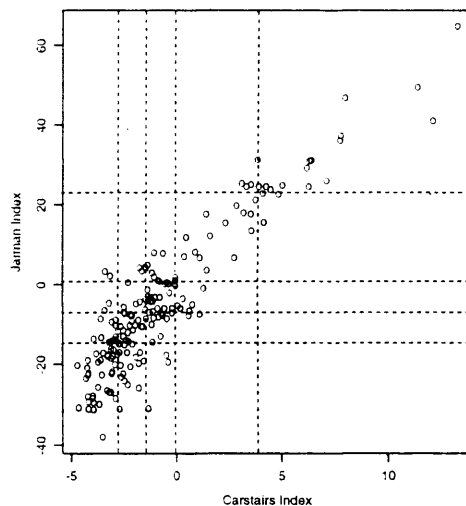


Figure 4.9: DOE Index and Jarman Index

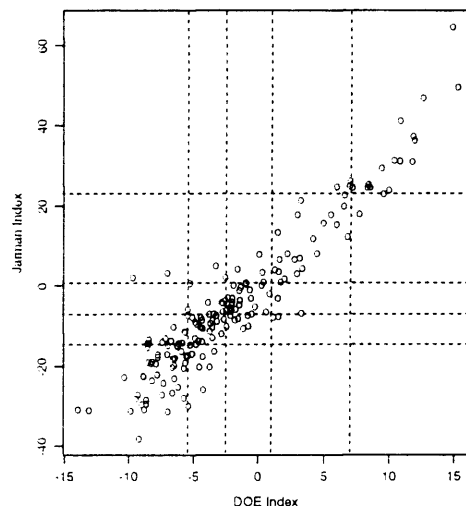
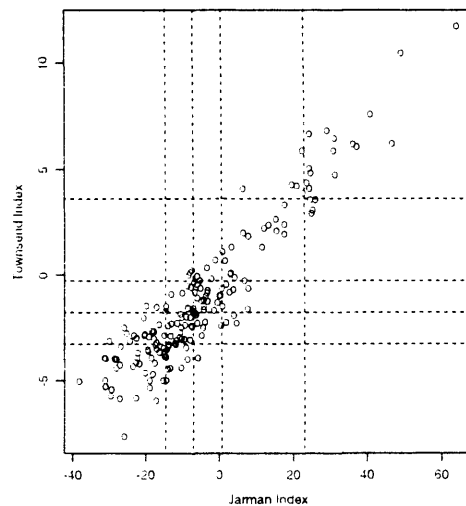


Figure 4.11: Jarman Index and Townsend Index



There are fewer wards in the more deprived quintiles as these wards are much bigger than the more rural affluent wards and the quintiles are based on groups of equal population rather than groups with an equal number of wards. If there was perfect agreement between the scores, with everybody being classified into the same quintile whichever deprivation score was used, all of the data points would lie in the central diagonal boxes. Wards lying outside of this main diagonal imply disagreement in classification between scores. This shows for most pairs that there is rarely disagreement by more than one quintile. There appears to be less disagreement between the Townsend score and the Carstairs score and between the DOE score and the Jarman score. As deprivation increases, the variability decreases with less disagreement. Thus, these scores appear to be consistent at classifying people as very deprived but less consistent at classifying the less deprived people. In order to quantify the degree of agreement between the scores, firstly the simple approach of looking at the percentage of exact agreements that were observed was used. Table 4.2 shows for each combination of the four deprivation scores, the percentage of the Leicestershire population classified into the same quintile of deprivation for both scores.

This confirms the pattern of highest agreement between the Carstairs and Townsend scores, and the DOE and Jarman scores. This would be expected since the components of these two pairs of scores are very similar i.e. Carstairs and Townsend are based on direct indicators alone while DOE and Jarman are based on a combination of direct and indirect measures. The exact agreement between deprivation scores is lowest for comparing measures based on direct indicators alone with those based on a combination of direct and indirect measures. Over a third of the population are assigned a different quintile when comparing these measures. The Townsend score seems to have consistently higher agreement.

Table 4.2: The percentage of exact agreement for 15-65 year olds classified into quintiles of deprivation based on four different deprivation measures

	Percentage of exact agreement			
	Carstairs	DOE	Jarman	Townsend
Carstairs	100%			
DOE	59%	100%		
Jarman	60%	71%	100%	
Townsend	75%	66%	65%	100%

Although this method gives some indication of agreement it takes no account of whether it could be due to chance since some agreement is expected. The kappa statistic is used to assess agreement and adjusts the level of agreement by the amount that would be expected to occur by chance (Cohen, 1960). A value of 1 indicates perfect agreement while a value of zero indicates no agreement better than chance and a negative value indicates worse than chance agreement which would be very unlikely in this situation. This measure is still only looking at exact agreement, e.g. assigned quintile 1 by both Townsend and Jarman scores. Table 4.3 shows the kappa values for each combination of deprivation scores. The confidence intervals are not displayed since the large population size led to standard errors of less than 0.005 in all cases.

Table 4.3: Simple kappa statistics for each combination pair of deprivation measures

	Simple Kappa statistic			
Score	Carstairs	DOE	Jarman	Townsend
Carstairs	1			
DOE	0.48	1		
Jarman	0.51	0.64	1	
Townsend	0.69	0.58	0.57	1

Guidelines given by Altman (1991) for the interpretation of these kappa statistics are used here. This again shows good agreement between Carstairs and Townsend, and between DOE and Jarman, while other combinations show only moderate agreement. A weakness of this simple kappa statistic is that it does not use any information other than the exact agreement. All disagreements are treated equally and yet it may be more sensible in this case to weight the disagreement according to the level of discrepancy. The weighted kappa statistic puts more weight on observations near to the main diagonal (i.e. exact agreement) and less weight to those further away (Altman 1991). Table 4.4 shows the weighted kappa values for each combination of deprivation scores. This particular weighted kappa statistic is based on the linear distance from the main diagonal as opposed to other methods which use the squared distance. Here the weights are 1 for the main diagonal and then 0.75, 0.5, 0.25 and 0 for the respective diagonals. The formulae for this and the simple kappa statistic can be seen in Altman (1991).

Table 4.4: Weighted kappa statistics for each combination pair of deprivation measures.

Score	Weighted Kappa statistic			
	Carstairs	DOE	Jarman	Townsend
Carstairs	1			
DOE	0.70	1		
Jarman	0.72	0.78	1	
Townsend	0.83	0.76	0.76	1

This shows good agreement between the scores when allowing for different weights depending on the degree of difference. Higher values are expected when comparing the weighted kappa statistic with the simple kappa statistic since disagreements are more likely to be small. It must be remembered that the Kappa statistic is strongly influenced by the proportion of subjects or prevalence of each category. Hence if quartiles or deciles were used instead of quintiles the results would differ. An extension of this work could be to look at loglinear models to assess agreement as done by Agresti (1988), but a sophisticated analysis of agreement was unnecessary in this case and kappa is suitable to give an estimate of the agreement between the four measures.

In conclusion when looking at deprivation scores, all four scores investigated generally appeared to be very similar. Although the correlations show extremely strong association between the scores, when comparing agreement in terms of quintiles the association is less apparent. For each pair, a third of individuals were assigned a different quintile, but most of these were assigned into the adjacent quintile. The Townsend score has the highest level of correlation and agreement with all of the other scores. It was expected that the Townsend score would be highly related to the Carstairs score since it only differs by one variable. However it has a different construction to the DOE and Jarman scores with no indirect measures of deprivation and yet the rankings do not vary greatly.

There was more variability in the more affluent areas than in the more deprived areas. These scores appear to be more sensitive at picking out areas of extreme deprivation but less good at defining more affluent areas. This may be due to the particular distribution of Leicestershire, where the most deprived areas are geographically confined to Leicester city. Therefore, these

scores appear to be consistent at differentiating the city from the rest of the county, but less consistent at differentiating the deprived non-city areas from the more affluent areas.

4.4 SUMMARY AND CONCLUSIONS

This chapter has shown the value of mapping techniques to illustrate patterns of area deprivation. Leicestershire has been shown to be a county of great socio-economic variability. Areas of high population density are generally more deprived. With a high population of Asian compared to England and Wales as a whole, there is the opportunity to investigate the complex relationships between ethnicity and deprivation.

There are problems in using maps based on area rather than on equal populations but they are useful in showing the distribution of deprivation to some extent and illustrate the city/rural differences.

This chapter has shown there to be little difference in the four measures discussed despite their different construction. As discussed in chapter 3, the Townsend score is used in the analyses that make up the rest of this thesis because of the theory behind its construction.

CHAPTER 5

AMBLYOPIA AND VISION SCREENING

5.1 AIMS OF THE CHAPTER

The previous chapters have discussed inequalities in health and how to measure deprivation in order to quantify these inequalities. In this thesis I investigate inequalities in the age at presentation of amblyopia, a common visual disability. Here I present an overview of this condition and discuss issues of presentation and diagnosis related to inequalities in health.

5.2 WHAT IS AMBLYOPIA ?

‘The observer saw nothing and the patient very little’

(A von Graefe (Von Noorden, 1990))

Amblyopia, often known as ‘lazy eye’, is the most common visual disability in children. A loss of vision occurs in an otherwise healthy eye which is caused by an abnormality preventing normal use of the eyes during visual development. Untreated amblyopia can mean irreversible, serious visual defects and the loss of depth of perception and binocular vision. Although amblyopia is potentially reversible if identified at an early stage, there is controversy over the effectiveness of screening, diagnosis and treatment.

5.2.1 AMBLYOPIA AND VISUAL DEVELOPMENT

Vision develops from birth and through early childhood and is an unstable changing state. Although babies can see from birth, vision improves with increased use of the eyes. If the eyes are not used to their capacity in early life then vision can deteriorate. The visual system is thought to be most sensitive to conditions such as amblyopia in the first two to three years of life with sensitivity decreasing with age (Von Noorden, 1990). The development of the visual system is thought to be complete by the age of about nine or ten years and after this time it is thought that poor visual ability cannot be improved. This indicates a need to detect amblyopia at an early age.

When both eyes are functioning normally each eye sends an image to the brain and these two images are blended into a single three dimensional image. Amblyopia can be caused by any condition which affects normal use of the eyes and is the result of prolonged suppression in one eye. The most common type of amblyopia is caused by strabismus, where the eyes are misaligned. In a child with strabismus one eye turns out (diverging) or turns in (converging). Two different images are seen by the eyes and the child suppresses vision in the strabismic eye to avoid double vision. Vision in this eye may deteriorate through disuse and it becomes amblyopic. There appears to be no association between the severity of amblyopia and size of angle of strabismus but amblyopia most commonly occurs with an esotropia, (a large angle convergent strabismus) where patients strongly favour one eye for fixation over the other. Amblyopia more rarely occurs with exotropia, (a large angle divergent strabismus), and microtropia (a small angle of strabismus).

Amblyopia can also be caused by anisometropia where there is a difference in refractive error between the two eyes. This leads to one eye being out of focus because it is more near-sighted, far-sighted or astigmatic than the other. The eye with the less clear image may then be suppressed and vision deteriorates leading to amblyopia. The eyes outwardly appear normal. There is no evidence of a direct relationship between the amount of anisometropia and the severity of amblyopia (Abrahamsson, Fabian, Andersson and Sjostrand, 1990). A one dioptre difference in the refractive error between the eyes is often used as an arbitrary cut-off for defining significant anisometropia. Strabismus and anisometropia may both be present in a child with amblyopia and although visual loss often follows the onset of strabismus, strabismus can be a consequence of poor vision in one eye due to anisometropia.

There are other less common types of amblyopia associated with congenital cataract or vitreous haemorrhage which are generally irreversible. Occlusion amblyopia is caused by prolonged occlusion of the unaffected eye in order to treat amblyopia. This leads to loss of vision in the occluded eye, but is almost always reversible. These types of amblyopia have a known cause and are not strictly true cases of functional amblyopia and will not be discussed further in this thesis.

5.3 PRESENTATION AND DIAGNOSIS OF AMBLYOPIA

Parents often notice large angle strabismus in their child (Wang et al, 1990) since there is a visible deviation and any associated amblyopia is then detected when they seek medical attention. Poor vision is much harder to detect (Wang et al, 1990). Since amblyopia associated with a very small angle of strabismus (microtropia) or with anisometropia offers no obvious outward signs, it is usually only detected by an orthoptist or ophthalmologist. Amblyopia associated with a large angle of strabismus is usually detected at a much earlier age compared to amblyopia associated with anisometropia (Shaw et al, 1988; Hiscox et al, 1992), with children with pure strabismic amblyopia presenting on average up to three years earlier than those with pure anisometropic amblyopia.

There has been very little research into other factors affecting age at presentation. Shaw et al (1988) found that for all types of amblyopia, males presented later than females and Asians presented later than Caucasians. However this study failed to take into account the relationship between ethnicity and deprivation. Campbell and Charney (1991) found that the age at diagnosis depended on family history of strabismus, degree of squint, level of maternal education and degree of parental suspicion of a problem. However this study did not look at these factors in a multivariate way.

A Glasgow study (Williamson et al 1995) has shown higher default rates for vision screening in more deprived areas. Bowman et al (1996) showed that children from more deprived areas were less likely to attend their paediatric ophthalmology outpatient appointments than those from less deprived areas. Hence this indicates that the age at presentation of children from these more deprived areas may be greater. Furthermore, this study indicated that a shorter waiting time for a paediatric appointment has led to improved attendance rates. Since patients from non-fundholding general practitioners are thought to generally have longer waiting times and these practices are more often in inner city areas this may also be contributing to inequalities in access to services.

5.4 SCREENING FOR AMBLYOPIA

The work in this thesis concentrates on the presentation of amblyopia. Therefore detection of the problem and the pathway to care is of importance. As discussed earlier, although large angle strabismus in a child is usually noticed by their parents and any associated amblyopia detected when they seek medical help, cases of amblyopia associated with microtropia or no strabismus are generally only detected by screening (Wang et al, 1990). The use of vision screening for the detection of amblyopia is under constant debate. Stewart-Brown et al (1988) found that 94% of health authorities in England and Wales had vision screening programmes in operation. The Hall report (Hall, 1989) reviewed all screening services for pre-school children and highlighted the fact that many new vision screening programmes were introduced before their benefit had been established. The report questioned their continuation. The Royal College of Ophthalmology and the British Paediatric Society (1994) called for the auditing of amblyopia screening programmes to assess their effectiveness.

Wilson and Junger (1968) set out the main requisites for any screening programme to be effective. These are that

- 1) the condition should be of public health importance
- 2) there should be effective treatment available for patients with recognised disease.
- 3) facilities for diagnosis and treatment should be available and shown to be effective
- 4) there should be a latent or early symptomatic stage of the condition
- 5) there should be a suitable test or examination that is simple, valid, reasonably priced, repeatable, sensitive and specific and acceptable to the majority of the population
- 6) the natural history of the condition and of conditions which may mimic it should be understood
- 7) there should be an agreed definition of what is meant by a case of the condition.
- 8) treatment at the early, latent or presymptomatic phase should favourably influence prognosis.
- 9) the cost of screening should be economically balanced in relation to expenditure on the care and treatment of persons with the condition and to medical care as a whole
- 10) case finding may need to be a continuous process and not a once and for all project with explicit justification for each stage of screening.

These requirements are now discussed in more detail with respect to vision screening.

5.4.1 PUBLIC HEALTH IMPORTANCE

The first of the requisites given by Wilson and Junger is that the condition should be of public health importance. The prevalence of amblyopia shows this to be the case. Estimates of prevalence vary widely depending on the study subjects, from 1% to 5% (Von Noorden, 1990). Some of the wide variation in the estimated prevalence is due to the fact that there is no agreed definition of amblyopia. A further problem is that studies investigating amblyopia have used different criteria for defining disease (Thompson et al, 1991; Hillis et al, 1983; Hillis, 1986). A recent local study of amblyopia in Leicestershire, estimated that 3 per cent of Leicestershire children develop the condition (Thompson et al, 1991). Two thirds of these patients underwent unsuccessful treatment, and carried amblyopia into later life.

Although the prevalence of the condition is high, there is controversy over whether it is a significant disability. Most sufferers may only be conscious of a minimal amount of disability, relying on the vision in their 'good' eye. However the condition affects binocular vision and depth of perception restricting sufferers' choice of lifestyles such as occupation. Injury or disease in the non-amblyopic eye can lead to extremely poor vision and sometimes blindness. Vereecken (1984) showed that on loss of the 'good' eye, vision in the amblyopic eye improved in only a quarter of cases with or without treatment. One of the greatest causes of added morbidity from amblyopia may be the reluctance of surgeons to operate on cataract patients who have only one good eye (Thompson et al, 1991). Hence although some sufferers perceive minimal disability, it is apparent that there are many ways it can affect health and lifestyle making the condition of public health importance.

5.4.2 AVAILABILITY OF TREATMENT

The requirements of a screening programme also demand that there should be accepted treatments to deal with the condition at an early stage. Correcting the cause of amblyopia i.e. the anisometropia or strabismus, does not cure the visual loss and the amblyopia must be treated separately. Where strabismus is present it must be surgically treated. If there is a significant refractive error then glasses are prescribed to correct it. If this does not improve the child's visual acuity then amblyopia treatment is started.

Amblyopia treatment is based on the fact that vision has deteriorated in one eye through disuse. The weaker eye is forced into use by reducing the vision in the stronger eye over a period of weeks or months. Reducing the vision in the stronger eye is known as occlusion and

is generally done by placing a sticky patch over it, although other methods such as atropine drops which blur the image in the non-amblyopic eye (penalisation), or a frosted lens in a pair of glasses are also used. By impairing the vision in the stronger eye the amblyopic eye is forced into use and occlusion is used to try and bring the visual acuity in the weak eye up to that of the stronger eye. A possible alternative treatment for amblyopia (the CAM machine) was introduced (Watson et al, 1978, Campbell et al, 1978) but found to be unsuccessful. Patching is still the preferred method of treatment.

5.4.3 DIAGNOSIS AND TREATMENT EFFICACY

Diagnosis can be made by an ophthalmologist when a child with a suspected vision problem is referred to them. All regions undertaking vision screening have an ophthalmology service to refer children with suspected problems for diagnosis.

As discussed previously there is a recognised treatment for amblyopia which is readily available since it is relatively cheap and easily prescribed. However, although treatment appears simple, its success is less assured since compliance with treatment is a major problem. Children do not usually like patching because they are stopped from using their favoured eye making their vision worse and they have to endure wearing a sticky patch on their face. There is debate over the amount of occlusion which produces the best results. Full-time patching is thought to be most effective but is associated with low compliance rates. Part-time patching is less effective but is more frequently prescribed. Monitoring patients after treatment may also be necessary since Scott and Dickey (1988) found 25% of patients had a drop in visual acuity after treatment while Levartovsky (1992) found that 55% of patients deteriorated six years after cessation of treatment.

There is extensive variation in the range of success rates of amblyopia treatment reported by different studies, ranging from between 30% (Flynn and Cassady, 1978) and 90% (Scott and Dickey, 1988). A Leicestershire study showed only one third of patients underwent successful treatment (Thompson et al, 1991). However a recent Cambridge study showed that 87.2% of straight-eyed amblyopes and 64.3% of strabismic amblyopes detected by screening at age 3 1/2 had successful treatment (a vision of 6/9 or better) (Newman et al, 1996).

Some of this variation in success can be explained by the use of different exclusion criteria with some studies excluding upto 59% of patients due to non-attendance or non-compliance to treatment. Compliance is related to treatment outcome (Lithander, 1991) and many studies exclude non-compliers in the analysis of results leading to superficially high success rates. The variation in treatment outcome may also be due to different criteria for success which are constantly being proposed (Meyer et al, 1991). The need for a universally used outcome measure for amblyopia treatment is emphasised by Romano (1991). Defining treatment outcome is also complicated by the variation in visual acuity when measured by different vision tests with wide variation both between and within children (The tests used to test vision are discussed further in section 5.4.5).

5.4.4 LATENT OR EARLY SYMPTOMATIC STAGE OF THE CONDITION

As discussed earlier, Wang et al (1990) have shown that parents often notice large angle strabismus in their child. Since there is a visible deviation, parents will generally seek medical attention and amblyopia can then be detected. Poor vision is much harder to detect and therefore amblyopia which is not associated with a visible problem can generally only be detected by a professional. Hence, there is a asymptomatic period during which the condition may be identified.

5.4.5 TESTS FOR DETECTING THE CONDITION

The fifth requisite of Wilson and Junger indicate that there should be a simple, valid acceptable and efficacious procedure for detecting the condition at a sufficiently early stage to permit intervention. Vision screening programmes in Britain are based on parental observation, family history, visual acuity tests and tests for strabismus. Strabismus is screened for by looking for asymmetry of the corneal reflexes using a hand held light and looking at the reflections. Strabismus has to be distinguished from pseudostrabismus in which the eyes are aligned but epicanthal folds or a broad bridge of the nose give a false impression of squint. The cover test is a definitive test for strabismus and looks for deviation in the eye but experience is needed to achieve reliable results. Refractions are rarely carried out and as refractive error may be a risk factor for amblyopia (Ingram, Walker et al, 1986), there is a call for orthoptists to be taught how to refract children.

Amblyopia is detected by noting a difference in the visual acuity between the two eyes by using a visual acuity test. Testing visual acuity in the very young is extremely difficult and if used solely as a screening test would lead to a great many false positives and false negatives. If a child is literate then diagnosis can be made based on their visual acuity assessed by a standard chart. The Snellen chart is the most commonly used and has a series of graded letters which get progressively smaller down the chart. The test is performed 6 metres away from the chart and the visual acuity is recorded in terms of the smallest line of letters that the child can read. The 6/6 line refers to the size of letters that someone with normal vision could read at six metres while the 6/60 line indicates that someone of normal vision would be able to read at a distance of 60 metres. For example a child who could only read the 6/60 line at a distance of 6 metres would have to be ten times closer to the letters to read them than a child with normal vision.

In children below reading age, the Sheridan Gardiner test is often used. The child holds a card with several letters and from a distance of six metres the examiner holds a single letter (Singles test) or a line of letters pointing to one of them (linear test) and asks the child to point at the letter they are being shown on their own card. Children as young as two and a half are often able to complete this test. Picture tests may be used for very young children who have verbal abilities.

In pre-verbal children diagnosis is extremely difficult. Visual acuity is estimated by covering one eye and seeing how well the child follows an object with the uncovered eye. This is good at diagnosing severe cases where the child is almost blind on covering the stronger eye but less severe cases are more difficult to diagnose. There has recently been research into electrophysiological tests to assess visual acuity, but as yet these methods are not in full use.

In children with amblyopia, objects viewed through the amblyopic eye appear to continuously fade in and out of focus. Amblyopic patients therefore seem to have a much better visual acuity if they read slowly. They also find it easier to see single letters than a letter surrounded by others as they tend to blend into one another. This effect is known as the crowding phenomenon. This feature means that visual acuity tested by a Sheridan Gardiner singles test where just one letter is held up for the child to identify will overestimate the true visual acuity. Picture charts may also give an inaccurate assessment of visual acuity since the different

symbols are so dissimilar that it is easier to distinguish between them than between letters. The recorded visual acuity may also be dependent on the child's reading age and can appear to deteriorate as they change from one vision test to another.

Perhaps most importantly detection of amblyopia and strabismus depends on who is performing the testing. In most areas, health visitors, community paediatricians or general practitioners perform vision screening. Many present tests are not successful at detecting disease since they are performed by inadequately trained people leading to high numbers of false positive and false negative cases. Community orthoptists have been introduced in some regions to perform screening. Jarvis et al (1990) found community orthoptists were more effective at screening than health visitors or primary care screeners at 35 months with increased sensitivity and detected incidence of target conditions, particularly for anisometropic amblyopia. No differences were found in younger children with all screeners giving poor results. MacLellan and Harker (1979) found orthoptists performing primary screening reduced inappropriate referrals by 25%. Although costs are higher for orthoptic primary screeners, orthoptists have been shown to be better at diagnosing amblyopia and strabismus than health visitors or clinical medical officers (Edwards et al, 1989). Hall (1989) does not recommend formal tests for squint and visual acuity prior to school entry unless undertaken by orthoptists. Secondary orthoptic screening services are recommended rather than primary orthoptic screening to examine children referred by parents or professional staff, selecting out those who require detailed assessment by an ophthalmologist.

5.4.6 NATURAL HISTORY OF THE CONDITION

The natural history of the condition is fairly well documented. Visual loss is known to occur in childhood when the eye is suppressed. As discussed earlier it is recognised as occurring in conjunction with anisometropia or strabismus. The eyes are thought to stop developing by the age of about nine or ten years of age and the condition will advance until that age. It is thought to be untreatable past this time. Untreated, the condition will not improve although Vereecken (1984) showed that on loss of the 'good' eye, vision in the amblyopic eye improved in a quarter of cases with or without treatment.

5.4.7 DEFINITION OF A CASE

There is an agreed definition of amblyopia as defective vision in an eye with no organic cause. However, there is no consensus as to the degree of visual loss which this definition refers to. Studies have used various definitions, some based on a difference in vision between the two eyes e.g. two lines difference in visual acuity between the eyes, while others are based on a certain level of vision in the poorer eye e.g. 6/9 vision or worse in the poorer eye. Therefore there may be variation in children classed as amblyopic in these studies.

5.4.8 EARLY TREATMENT SHOULD IMPROVE OUTCOME

There is conflicting evidence as to whether earlier treatment leads to better outcome. The visual system is thought to be more responsive in younger children. Eibschitz et al (1978), and Epelbaum et al (1993) show better outcome in younger children while Levartovsky (1992), Kutschke et al (1991), Hiscox et al (1992), and Hardman-Lea et al (1989), found no such relationship. In an analysis of 23 studies of the results of treatment Wick et al (1992) found that success seemed possible at any age.

5.4.9 COST OF SCREENING

Vision screening is generally done as part of child health surveillance programmes and as resources are already allocated to child health surveillance in most areas the implementation of vision screening as part of these programmes generally does not affect the allocation of resources. Treatment of the condition is relatively cheap and does not require extensive treatment programmes. However, in terms of inequalities in health, there is debate over the implementation of child health surveillance programmes which cover a wide range of conditions of which amblyopia is just one. The targeting of resources at an individual level is questioned since they may be spent more effectively at a collective level to relieve greater causes of ill-health such as poverty (Butler, 1989).

5.4.10 SCREENING SHOULD BE A CONTINUOUS PROCESS

Many studies suggest different optimal screening ages (Sjostrand et al, 1990; Neumann et al, 1987). In Britain most children are screened for vision defects at birth and at around 6 weeks, 8 months, 18 months and 3½ years. Anisometric amblyopia is checked for at 3½ years while strabismus is checked from 8 months onwards. Taylor (1987) recommends the introduction of a test for strabismus and amblyopia at 6 weeks, but current screening tests are inadequate for this. There is controversy over screening at 3½ years. Beardsell (1989)

recommends it for a quick and accurate assessment for strabismus and amblyopia but Ingram, Holland et al (1986) do not support it, and Taylor (1985,1987) found it detected an inadequate proportion of existing visual defects and that it is too late for optimal treatment of any amblyopia. The idea that anisometropic amblyopia can be treated successfully at any age (Wick et al, 1992) offers little support to screening.

5.4.11 SUMMARY OF SCREENING REQUIREMENTS.

In terms of whether amblyopia is of public health importance, my discussion in section 5.4.1 has shown this not to be in dispute. It is a relatively common condition and untreated can lead to poor vision, limited lifestyle opportunities and even blindness. The availability of effective treatment is less clear cut (see 5.4.2, 5.4.3). Although effective treatment is available, compliance reduces the success rates with Thompson et al (1991) showing only one third of patients to undergo successful treatment. The use of vision tests for amblyopia screening are evidently safe yet their efficacy in detecting the condition in young children is less certain if not undertaken by a trained orthoptist (see 5.4.5). Further there is little evidence of whether earlier treatment is related to improved outcome (5.4.8).

All of the criteria are therefore not met for a screening programme to be effective and efficient. However since nearly all regions are performing vision screening in one form or another, the service that is provided should be equitable and therefore need evaluation. Furthermore section screening for amblyopia is recommended by the Royal College of Ophthalmologists (see 5.4.12) and the use of orthoptists as secondary screeners to improve the screening service has been recommended by Hall (1989) in his review of child health surveillance practices. If such services are introduced then equity of access to services must be monitored and maintained. If there were inequalities in access to care for amblyopia this may result in amblyopia being detected later among children from more deprived areas. It is this aspect of amblyopia that is focused on in this thesis.

5.4.12 CURRENT RECOMMENDATION FOR AMBLYOPIA SCREENING

A joint working party of the Royal College of Ophthalmology and the British Paediatric Society (1994) was set up to 'review ophthalmic services for children in the United Kingdom'. With the introduction of the purchaser/provider split it was important to reassess the value of ongoing vision screening programmes, particularly as many had previously been developed in

an ad hoc fashion. They recommended screening at 6 weeks for strabismus by a general practitioner or clinical medical officer. This should be followed by screening at 7 to 8 months by a health visitor to check parents concerns about their child's vision and to advise parents on what to look for in terms of strabismus and what to do if they are concerned. The report left districts to choose whether to screen at 3¹/₂ years but advised those who did to run primary orthoptic screening. School entry was recommended as a time point to test children's visual acuity in both eyes together and each separately by a trained school nurse, orthoptist or other trained professional to identify amblyopia. The need for suitable training for those performing vision screening was highlighted and an emphasis was placed on monitoring and auditing ongoing programmes.

5.4.13 VISION SCREENING IN LEICESTERSHIRE

Present practices were set up in 1991 when the Leicestershire working party produced their outline programme, although some changes came in gradually before this time. Vision screening is performed as part of the child health surveillance programme at 6 to 8 weeks, 7¹/₂ to 10 months, 18 to 24 months and 3 to 3¹/₂ years with computer records of the results of each screen. Parents are asked about any concerns regarding their child's vision. This is followed by an examination of the child's vision and their eyes (see appendix A). Strabismus is screened for using the cover test and corneal reflexions, while amblyopia is screened for using a visual acuity test. Informal vision screening is performed at a variety of times before starting school.

Referral processes in children under two months old have not radically changed with referral straight to an ophthalmologist. In older children, practices have changed. In 1989 a secondary orthoptic screening service was introduced, as recommended by Hall (1989), for use by health visitors when they suspected a child of having strabismus or amblyopia. Health visitors had previously been unable to refer children straight to a consultant ophthalmologist, having to refer to the general practitioner who would then refer on to the consultant ophthalmologist where necessary. A secondary orthoptic screening service was thought to be more accessible to health visitors and would reduce the number of false positive referrals to the consultant ophthalmologists and decrease the number of false negatives. Children found to have any defects by the orthoptists would then be referred on to the consultant ophthalmologist. The effectiveness of these changes to the vision screening programmes need to be assessed and their effect on inequalities in health evaluated.

5.5 SUMMARY AND CONCLUSIONS

This chapter has described the condition of amblyopia, the focus of this thesis and discussed various contentious issues in its diagnosis and treatment particularly concentrating on aspects related to screening. Amblyopia is the most common visual disability in children and as it is potentially reversible this makes the condition of significant importance to public health. The difficulty of testing visual acuity in very young children leads to problems in diagnosing and screening for amblyopia, particularly among children with amblyopia associated with microtropia or anisometropia who are consequently referred later. It also leads to problems in monitoring progress and in assessing the results of treatment.

The effectiveness of vision screening is controversial in pre-school children and is an important issue in investigating inequalities in health. Although the criteria for screening are not adequately fulfilled, many regions have introduced screening programmes and it is important to review current methods. This is particularly important in the areas where there is no clear evidence of benefits, notably at 3 $\frac{1}{2}$ years. The introduction of secondary orthoptic screening in Leicestershire needs to be assessed to see whether it is a relatively inexpensive way of increasing the efficacy of screening tests and increasing the rate of detection of amblyopia.

Evidence discussed in chapter 2 has shown that people from more deprived areas are less likely to attend children's immunisation programmes (Reading, Jarvis and Openshaw, 1993; Lynch, 1995) and are less likely to attend vision screening programmes (Williamson et al, 1995) and ophthalmology outpatient appointments (Bowman et al, 1996). As shown in this chapter, there has been little research into factors affecting the presentation of amblyopia and further research is needed to assess whether there are inequalities in access to amblyopia treatment by studying the relationship between deprivation and age at presentation of amblyopia. Therefore, in this thesis, I use data from a multi-centre study of children treated for amblyopia (Chapter 6) to assess the relationship between deprivation and the age at presentation in seven centres. I then describe the design and analysis of a study I undertook to assess the effect of changes made to screening in Leicestershire on this relationship (Chapters 7 and 8) and indicate whether these methods could be used to monitor inequalities in health in the future.

CHAPTER 6

AGE AT PRESENTATION IN AMBLYOPIA AND DEPRIVATION: ANALYSIS OF A MULTICENTRE STUDY

6.1 AIMS OF THE CHAPTER

The first chapters of this thesis have reviewed the background to research in inequalities in health and shown evidence of access to health care being related to deprivation. Research has shown attendance at children's immunisation clinics, vision screening and paediatric ophthalmology clinics to be related to deprivation (see chapter 2). Chapter 5 showed that amblyopia is a condition of public health importance that relies on screening for its detection. However of the little research into presentation of amblyopia there has been no research to look at inequalities in presentation. In this chapter I seek to use the Townsend score, reviewed in chapter 3, to look at the relationship between deprivation and the age at presentation of children with amblyopia in seven UK orthoptic clinics. Shaw et al (1988) found that Asians presented later than Caucasians and so I assess whether these data confirm this finding. It was also highlighted in chapter 5 that some types of amblyopia are easier to detect than others. Therefore the relationship between age and deprivation is investigated for each type of amblyopia.

6.2 PRESENTATION OF AMBLYOPIA

As discussed in chapter 5, large angle strabismus (i.e. esotropia or exotropia) in a child is usually noticed by their parents, and when they seek medical help any associated amblyopia is likely to be detected. In contrast, amblyopia associated with a very small angle of strabismus (microtropia) or with no strabismus offers no obvious outward signs and is usually only detected by a vision screening test. Little is known about the factors affecting the age of presentation of amblyopia. Shaw et al (1988) found that for all types of amblyopia, males

presented later than females and Asians presented later than Caucasians. Campbell and Charney (1991) found that the age at diagnosis depended on family history of strabismus, degree of squint, level of maternal education and degree of parental suspicion of a problem, but this study did not distinguish between different types of amblyopia.

Reading, Jarvis and Openshaw (1993), Marsh and Channing (1986), Lynch (1995) and Williamson et al (1995) have suggested that many screening services fail to properly serve deprived areas. Hence this could lead to inequalities in the age at presentation for conditions which rely on screening for their detection. Wang et al (1990) found that parents were more likely to notice strabismus while poor vision was much harder to identify. Here I hypothesise that deprivation may affect the age at presentation of children with microtropia or no strabismus but not those with a large angle of strabismus. Deprivation should not affect the parents ability to detect a large angle of strabismus but could affect a child's access to the screening necessary to detect amblyopia associated with anisometropia or microtropia. In this chapter I analyse data from a multi-centre study of children treated for amblyopia to test this hypothesis. My analysis of these data in a slightly different format can be seen in Smith et al (1994) (See paper attached in appendix B).

6.3 BACKGROUND TO THE MULTI-CENTRE STUDY

6.3.1 REASONS FOR THE STUDY

Although there have been studies of amblyopia in recruited soldiers and in patients with ophthalmic disorders, there have been few large studies of amblyopia in children, particularly in Britain. Studies involving soldiers and ophthalmic patients often lead to biased results, and frequently patients are excluded from the results of treatment for non-compliance. Hence the selection of patients and their clinical presentation are major factors responsible for the variable results of amblyopia treatment reported in the literature with success rates ranging from 30% to over 90%. There is very little information about the presentation of children with amblyopia in the UK so that the relevance to British practice of the different results of treatment claimed in the literature is not clear. Until recently most studies on children have been small, with previous epidemiological studies based on clinic data from a single centre.

6.3.2 STUDY DESIGN

The aim of the multicentre study was to obtain a more complete picture of children with amblyopia in the UK by looking at various aspects of amblyopia, principally presentation, treatment, and compliance. The study was based on all children first presenting with amblyopia in 1983. Centres around England were invited to participate in the study providing they could identify children referred in 1983 and trace their orthoptic records up to the end of their treatment. Seven centres participated in the final study: Leicester, Birmingham, Bristol, Leeds, Nottingham, Sunderland and Worthing. Although the study was designed to look at many aspects of amblyopia, these data also provide an ideal opportunity to look at the effect of deprivation on age at presentation of amblyopia and to see whether any effects of deprivation are similar in different parts of the country.

At each of the seven volunteering centres, a local co-ordinator completed a questionnaire about screening policies, type of hospital, catchment area, and the annual number of new referrals. To ensure consistency one research orthoptist and a research assistant visited each centre and collected data on every patient who had their first appointment in the orthoptic department in 1983. Children were included if they had not had any previous treatment, were under 15 years at the first appointment and had been prescribed occlusion treatment for amblyopia at any time either at the first appointment or subsequently. As definitions of amblyopia vary, and visual acuity is notoriously hard to measure, it was decided that all children treated for amblyopia should be included in the study rather than including all children who at some point had a two line difference in visual acuity between the eyes, or those with vision of 6/12 or worse in one eye. Therefore the study is not strictly about children with amblyopia but children treated for amblyopia.

A total of 981 children were included in the study. General information was collected where possible on sex, age, address, source of referral, type of squint, type of amblyopia, age at onset and any developmental delay. Then information on every outpatient appointment made for each child up until 1992 was taken from the notes, irrespective of whether they attended or whether they were being treated by occlusion at that time. This included the date of appointment, whether the child attended, any visual acuity recorded at that visit and the test used (with and without glasses), any occlusion prescribed and the amount and type prescribed, where appropriate whether co-operation with treatment was good or the reason for stopping

occlusion, and any information on refractions performed. Finally, information was obtained on every eye operation mentioned in the orthoptic notes, with the date, the eye operated on and the type of operation (e.g. squint). All of this information was then entered into a database. The same research orthoptist determined the diagnosis from the information in the orthoptic records in the seven different centres. As the study was retrospective the only way of determining ethnicity was to assess whether the patient had an Asian surname or forename which has been shown to be an extremely accurate measure for identifying Asians and non-Asians (Nicholl et al, 1986) but cannot detect any other minority ethnic group. In Leicestershire, this was not a major problem since the census data in chapter 4 has shown that the majority of people classifying their ethnicity as other than 'white' were of Indian origin.

6.4 PRELIMINARY ANALYSIS

General information on the presentation of the children in this study and my analyses of other aspects of presentation of amblyopia can be seen in Woodruff et al (1994a) (see appendix C). Here the effect of deprivation on age at presentation will be reported. For this analysis, children who were not treated within two years of first attending the orthoptic clinic were excluded as they were unlikely to have had amblyopia at the time of referral. Also children with stimulus deprivation amblyopia were excluded from the analysis since they are not 'true' cases of amblyopia as they have a known cause. There were also two children for whom a Townsend deprivation score could not be calculated due to missing address data. This resulted in data on 916 children being available for analysis.

In order to investigate the relationship between deprivation and age at presentation, initially a univariate analysis was performed on the data. Possible explanatory factors such as ethnicity and presence of anisometropia were also investigated for each type of strabismus: large angle strabismus, microtropia, and no strabismus.

6.4.1 CLASSIFICATION OF AMBLYOPIA

The patients were classified by their refractive error and the degree of strabismus. Strabismus was defined as manifest strabismus on cover testing with large angle strabismus defined as 5 degrees or more (esotropia or exotropia), while microtropia was manifest strabismus of less than 5 degrees. Anisometropia was defined as a difference of one dioptre or more of either

sphere or cylinder between the two eyes. Children with amblyopia associated with a large angle of strabismus presented significantly younger (mean 3.4 years n=636) than children with amblyopia associated with microtropia (mean 5.7 years n=117) or no strabismus (mean 5.6 years n=163) (Analysis of variance $P=0.0001$). This confirms the pattern seen in other studies.

6.4.2 CLINICS

Table 6.1 compares age at presentation for the seven clinics and shows a range of 18 months in the mean age at presentation for children with a microtropia and 22 months in the mean age of those with no strabismus compared to 12 months for those with a large angle of strabismus. There appears to be more variability in the age at presentation of types of amblyopia which are dependent on screening and it is of interest to note that the youngest patients with microtropia or no strabismus come from Sunderland which was the only area to run an orthoptic secondary screening service at the time of the study.

Table 6.1: Mean age in years, 95% confidence interval and number of children attending the seven orthoptic clinics with different levels of strabismus

Clinic	Large angle strabismus			Microtropia			No strabismus		
	Mean	95% C.I.	N	Mean	95% C.I.	N	Mean	95% C.I.	N
Birmingham	3.0	2.61,3.43	83	5.6	4.55,6.74	20	5.2	4.28,6.14	21
Bristol	4.0	3.48,4.50	70	5.5	4.59,6.49	24	5.1	4.53,5.57	29
Leeds	3.1	2.48,3.61	45	4.9	3.13,6.68	4	5.5	4.02,6.96	4
Leicester	3.4	3.08,3.64	157	5.9	5.09,6.61	32	6.1	5.62,6.67	51
Nottingham	3.6	3.27,3.85	144	6.3	5.52,7.14	24	5.5	4.61,6.37	24
Sunderland	3.3	2.96,3.62	109	4.8	3.78,5.77	10	4.8	4.17,5.47	23
Worthing	3.3	2.63,4.04	29	5.3	1.00,9.69	3	6.9	5.73,8.15	11

6.4.3 TOWNSEND DEPRIVATION SCORE

The degree of deprivation of each child was estimated by the Townsend deprivation score. A 1981 postcode directory linked the patient's postcode to the wards in which they lived and the Townsend score was calculated by linking these wards to 1981 census data using SASPAC (London Research Centre, 1992). As the clinics were spread around England it was decided to

base the Townsend scores on the mean and standard deviation of each variable for England (Table 6.2).

Table 6.2: Summary statistics of the components of the Townsend score for England from the 1981 census (Log transformed values in parenthesis)

Townsend score components	Mean	Standard Deviation
% households with no car	39.3	15.93
% households with > 1 person per room	3.4 (1.36)	2.52 (0.49)
% households not owner occupied	42.9	21.3
% economically active people unemployed	9.6 (2.28)	5.21 (0.45)

Table 6.3 shows the mean and standard deviation of the Townsend score for children in each study centre, showing a wide variety in the deprivation in those attending the seven centres with some very deprived and some more affluent centres.

Table 6.3: Summary statistics of the Townsend score for children attending each orthoptic clinic

Centre	Mean	Standard deviation	Least deprived area	Most deprived area	Number of children
Birmingham	1.82	3.88	-7.4	9.1	123
Bristol	-1.90	3.10	-7.7	5.5	123
Leeds	0.04	3.35	-4.6	6.9	53
Leicester	0.49	4.00	-6.4	8.1	240
Nottingham	0.57	3.44	-5.8	7.5	192
Sunderland	4.11	2.75	-5.0	8.9	142
Worthing	-2.48	1.82	-6.2	3.1	43

The main analysis of this data is based on the Townsend score as a continuous variable but for tabulation purposes in this preliminary analysis children at each clinic were classified by the Townsend deprivation score of the ward in which they lived into quintiles of deprivation *within* their clinic. These quintiles ranged from the 20% of children living in the most affluent areas to the 20% living in the most deprived areas within each clinic. The postcode directory contained co-ordinates for each postcode enabling the distance from the orthoptic centre to be

calculated. Quintiles of Townsend scores within clinics were used rather than quintiles of deprivation across all clinics for tabulation purposes since there are evidently large differences between clinics and tabulating the deprivations scores into quintiles across all clinics would lead to any effect of deprivation being masked by the differences between clinics. Table 6.4 shows quintile of deprivation within clinic by age at presentation for different levels of strabismus and anisometropia.

Table 6.4: Mean age, 95% confidence interval and number of children by level of strabismus and quintile of Townsend deprivation score where 1 is the least deprived quintile and 5 is the most deprived quintile.

Townsend deprivation quintile (within clinic)	Large angle strabismus			Microtropia			No strabismus		
	Mean	95% C.I.	N	Mean	95% C.I.	N	Mean	95% C.I.	N
1	3.4	3.06,3.79	124	5.6	4.96,6.25	26	5.2	4.63,5.71	36
2	3.2	2.96,3.52	130	5.1	3.98,6.14	20	5.3	4.66,5.90	28
3	3.6	3.21,3.90	120	5.7	4.54,6.77	19	5.5	4.78,6.25	38
4	3.4	3.06,3.71	138	5.9	5.29,6.53	26	5.9	5.35,6.44	29
5	3.4	3.08,3.68	124	6.2	5.15,7.18	26	6.1	5.36,6.87	32
ALL	3.4	3.25,3.53	636	5.7	5.33,6.09	117	5.6	5.30,5.87	163
Test for trend	P=0.917*			P=0.151*			P=0.015*		

*Test for trend statistic is based on a regression of quintile of deprivation on age at presentation

From this table it can be seen that among those with a microtropia or no strabismus there is a trend of increasing age at presentation with increasing deprivation within each clinic. There is no similar pattern among those with a large angle of strabismus. The test for a trend of increasing age with increasing deprivation showed no effect of deprivation among those with a large angle of strabismus, a nonsignificant trend among those with a microtropia and a significant trend among those with no strabismus. Figures 6.1-6.7 illustrate the relationship between age and Townsend score for each centre for children with a microtropia or no strabismus. This general trend of increasing age with increasing Townsend score can be seen at some of the centres although two of the centres have very few patients. These graphs are not adjusted for any of the other explanatory factors.

Figures 6.1-6.7: Townsend deprivation score by age at presentation for children with microtropia or no strabismus for each of the seven orthoptic clinics.

Figure 6.1: Birmingham

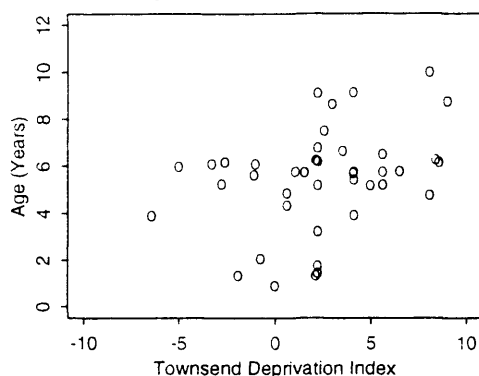


Figure 6.2: Bristol

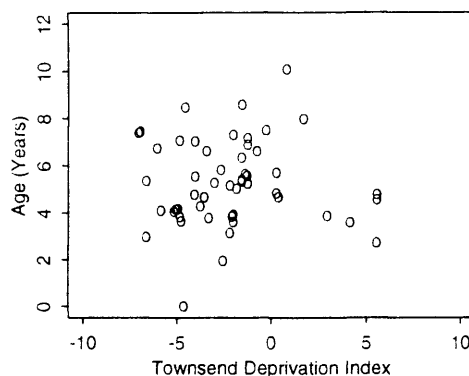


Figure 6.3: Leeds

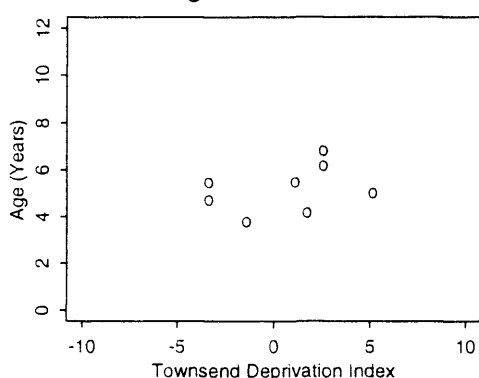


Figure 6.4: Leicester

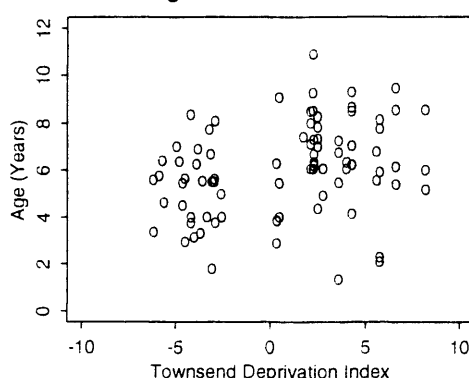


Figure 6.5: Nottingham

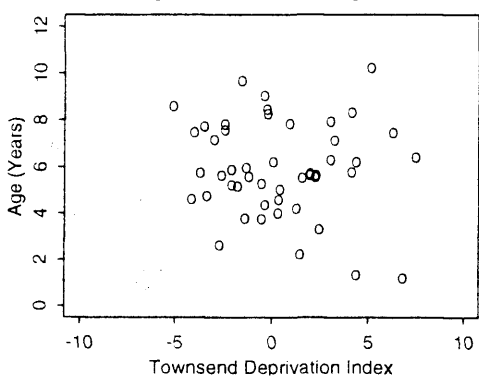


Figure 6.6: Sunderland

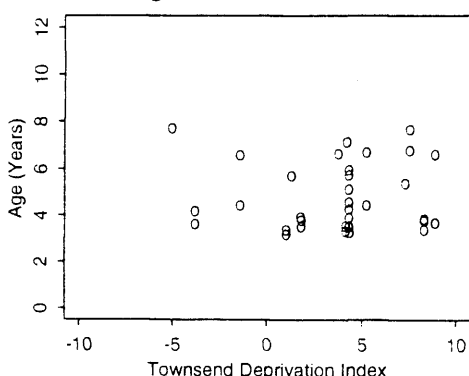
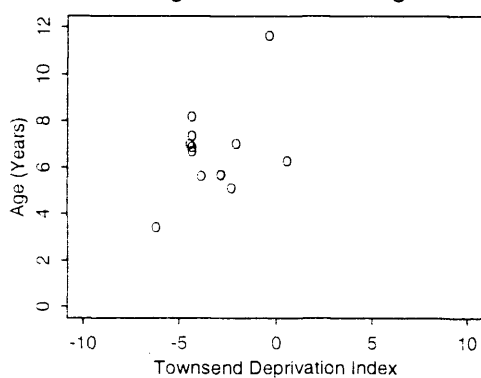


Figure 6.7: Worthing



6.4.4 OTHER EXPLANATORY VARIABLES

Children with an Asian name who have microtropia or no strabismus on average present over a year later than those with a non-Asian name (table 6.5). This is reversed among children with a large angle of strabismus where there is a slightly younger mean age for Asian named children. However the number of Asians in the study is very small (4%, 41 children), and the proportion varies greatly between centres with most Asian children coming from Leicester and Birmingham.

Table 6.5: Mean age, 95% Confidence interval and number of children by level of strabismus and Asian/Non-Asian names

Asian name	Large angle strabismus			Microtropia			No strabismus		
	Mean	95% C.I.	N	Mean	95% C.I.	N	Mean	95% C.I.	N
Asian	3.0	2.18,3.92	23	7.3	5.65,8.98	5	6.6	5.67,7.57	13
Non-Asian	3.4	3.26,3.55	613	5.6	5.25,6.03	112	5.5	5.19,5.79	150

Table 6.6 shows that children with a significant amount of anisometropia appear to present later than those with less than one dioptre and this exists within each level of strabismus. This may be due to this type of amblyopia having a later onset. It appears from this table that a few patients with no strabismus were also classified as not having a significant amount of anisometropia. These patients did have anisometropia but it was below the cut-off.

Table 6.6: Mean age, 95% confidence interval and number of children by level of strabismus and anisometropia

Degree of anisometropia	Large angle strabismus			Microtropia			No strabismus		
	Mean	95% C.I.	N	Mean	95% C.I.	N	Mean	95% C.I.	N
≥1 dioptre	3.9	3.66,4.22	188	6.1	5.65,6.62	68	5.7	5.39,5.97	154
<1 dioptre	3.2	3.00,3.33	448	5.1	4.54,5.71	49	3.9	2.69,5.11	9

Table 6.7 shows that the age at presentation for males and females. There are no apparent differences in presentation between the two sexes when assessed univariately.

Table 6.7: Mean age, 95% confidence interval and number of children by level of strabismus and sex

Sex	Large angle strabismus			Microtropia			No strabismus		
	Mean	95% C.I.	N	Mean	95% C.I.	N	Mean	95% C.I.	N
Female	3.4	3.21,3.63	303	5.8	5.20,6.48	50	5.6	5.15,5.96	85
Male	3.4	3.17,3.57	333	5.6	5.15,6.09	67	5.6	5.20,6.02	70

Another factor that may be important in the presentation of children with amblyopia is their visual acuity. However many children did not have a visual acuity recorded at presentation and for those children who did have a visual acuity recorded many different tests were used and so the visual acuity recorded by these different methods is not comparable. Also children are better at performing the tests as they get older. Therefore this data was not analysed since conclusions could not be drawn from the results.

This preliminary analysis shows the similarity in presentation of those children with a microtropia and those with no strabismus. These groups have a similar age of presentation and trend of increasing age of presentation with increasing deprivation. There is also more variability in the age of presentation at different centres in these two groups compared to those with a large angle of strabismus. The detection process is similar for amblyopia associated with a microtropia or no strabismus since the children's eyes appear straight and detection relies on screening. In order to analyse the data using regression analysis children with microtropia or no strabismus were grouped together while those with a large angle of strabismus were analysed separately.

6.5 REGRESSION ANALYSIS

A Normal errors regression analysis was performed on the two sets of data to investigate the relationship between age at presentation and the other explanatory factors of interest and the interactions between them. Formal model checking was carried out and is discussed for the final fitted model. The analysis of variance in tables 6.8 and 6.9 indicate the models that were fitted to the data.

Table 6.8 shows the analysis of variance of models fitted for children with microtropia or no strabismus. It appears that when looking at main effects separately there is no significant effect of deprivation, sex or distance from the clinic. Children with an Asian name present differently to those with a non-Asian name and there is a difference between the clinics and between those with and without anisometropia.

Fitting further models demonstrates that after adjusting the effect of clinic and anisometropia there is a significant deprivation effect. The effect of Asian/non-Asian name is no longer significant after adjusting for this deprivation effect. Looking at interactions between the main effects shows the effect of deprivation to be consistent within each clinic and within each level of anisometropia. Also the effect of anisometropia is consistent within each clinic. The effect of deprivation appears to be linear with no significant quadratic term. Therefore the model including the effects of deprivation, clinic and anisometropia is interpreted and discussed.

Table 6.9 shows the analysis of deviance for children with a large angle of strabismus. The models fitting each main effect separately show evidence of a difference between clinics and between those with and without anisometropia. There is no evidence of any other significant effects. Even after adjusting for the differences between clinics and those with and without anisometropia there is no evidence of an effect of deprivation, distance to the clinic or between Asians and non-Asians. The effect of anisometropia appears to be constant between clinics since there is no evidence of a significant interaction. Therefore the most sensible model to adopt appears to be just including the effect of anisometropia and clinic. However in order to compare the effect of deprivation with those with a microtropia or no strabismus the model which is discussed in detail is that including the effects of deprivation, anisometropia and clinic.

Table 6.8 Analysis of variance table of regression model for microtropia or no strabismus

Model terms fitted	Previous model	F statistic for change in fit	P value for change in fit
ANISO	NULL	9.86 _{1,278}	0.002
CLINIC	NULL	2.87 _{6,273}	0.010
SEX	NULL	0.04 _{1,278}	0.840
TOWNSEND	NULL	2.57 _{1,278}	0.110
ASIAN	NULL	7.25 _{1,278}	0.008
DISTANCE	NULL	0.05 _{1,278}	0.828
CLINIC+TOWNSEND	CLINIC	7.52 _{1,272}	0.006
CLINIC+TOWNSEND+ASIAN	CLINIC+TOWNSEND	2.30 _{1,271}	0.131
CLINIC+TOWNSEND+ANISO	CLINIC+TOWNSEND	9.05_{1,271}	0.003
CLINIC+TOWNSEND+ANISO+DISTANCE	CLINIC+TOWNSEND+ANISO	1.14 _{1,270}	0.287
CLINIC+TOWNSEND+ANISO+CLINIC.TOWNSEND	CLINIC+TOWNSEND+ANISO	1.59 _{1,265}	0.150
CLINIC+TOWNSEND+ANISO+CLINIC.ANISO	CLINIC+TOWNSEND+ANISO	1.17 _{1,265}	0.323
CLINIC+TOWNSEND+ANISO+TOWNSEND.ANISO	CLINIC+TOWNSEND+ANISO	0.73 _{1,270}	0.395
CLINIC+TOWNSEND+ANISO+(TOWNSEND) ²	CLINIC+TOWNSEND+ANISO	0.03 _{1,270}	0.863

Key to parameters in regression models TOWNSEND: Townsend deprivation score CLINIC: Orthoptic clinic attended DISTANCE: Distance from orthoptic clinic

Italics indicate final chosen model ASIAN: Asian surname or forename ANISO: Presence of ≥ 1 dioptre of anisometropia SEX: Male/Female

Table 6.9 Analysis of variance table of regression model for large angle strabismus

Model terms fitted	Previous model	F statistic for change in fit	P value for change in fit
ANISO	NULL	24.6 _{1,634}	0.000
CLINIC	NULL	2.37 _{6,629}	0.029
SEX	NULL	0.11 _{1,634}	0.736
TOWNSEND	NULL	0.45 _{1,634}	0.501
ASIAN	NULL	0.84 _{1,634}	0.360
DISTANCE**(2 distances missing)	NULL	0.47 _{1,632}	0.491
<i>ANISO+CLINIC</i>	<i>ANISO</i>	<i>1.79_{1,628}</i>	<i>0.099</i>
ANISO+CLINIC+TOWNSEND	ANISO+CLINIC	0.10 _{1,627}	0.667
ANISO+CLINIC+ASIAN	ANISO+CLINIC	0.99 _{1,627}	0.320
ANISO+CLINIC+DISTANCE**	ANISO+CLINIC	0.28 _{1,625}	0.599
ANISO+CLINIC+ANISO.CLINIC	ANISO+CLINIC	1.71 _{1,622}	0.116

Key to parameters in regression models

TOWNSEND: Townsend deprivation score CLINIC: Orthoptic clinic attended DISTANCE: Distance from orthoptic clinic

Italics indicate final chosen model

ASIAN: Asian surname or forename

ANISO: Presence of ≥ 1 dioptre of anisometropia

SEX: Male/Female

Formal model checking procedures were performed as described by McCullagh and Nelder (1988), on the model including deprivation, clinic and anisometropia for each amblyopic group. This involved analysis of the Studentised residuals to check for Normality (by producing a histogram and a Normal scores plot). Their distribution was also investigated with respect to the fitted values and deprivation. Influential variables were assessed using the leverage of each observation, i.e. locating observations that are influential in determining the position of the fitted model and influential on the parameter estimates. These are the diagonal elements of the projection ('Hat') matrix $H = X(X^T X)^{-1} X^T$, which is the ratio of the covariance matrix of the fitted values to the covariance matrix of the observed data. The average value of h_i is p/N where p is the number of parameters and N is the number of observations. $2(p+1)/N$ was used as an arbitrary cut-off point to indicate high influence. Cook's D statistic was also used to look at the effect of observations on the parameter estimates by comparing the parameter estimates with and without each individual observation (McCullagh and Nelder 1988).

Figures 6.8-6.13 illustrate the checking procedures for children with microtropia or no strabismus. There was no distinct deviations from Normality among the residuals with the histogram indicating a Normal distribution (figure 6.8) and the Normal scores plot indicating a straight line (figure 6.9). There was no apparent increase in the variation of residuals with increasing age (figure 6.10) or increasing deprivation (figure 6.11) and there was no evidence of non-linearity of the Townsend score. A plot of the leverage values (figure 6.12) showed 22 of the 280 observations to have influential values (i.e. $\text{leverage} > 2 \cdot (p+1)/n = 0.064$ where $p=8$ parameters and $n=280$ observations). The apparent systematic pattern in this plot is due to the id numbers being sorted by clinic and the influential H values can be seen to be clustered within clinics. Weighting these observations out of the regression resulted in a significant change in the clinic estimates for the two smallest clinics, Leeds and Worthing. As the clinic estimates were not of primary interest and the deprivation and anisometropia estimates changed by less than 10% these observations were kept in the analysis. A plot of the Cook's D statistic (figure 6.13) shows some influential observations. Only their effect on the estimate of deprivation was of interest. Removing the influential values showed these observations to affect the clinic estimates or to increase the estimate for deprivation. Hence this gave further evidence of a significant effect of deprivation.

Figures 6.8-6.13: Model checking for children with microtropia or no strabismus

Figure 6.8: Histogram of Studentized Residuals

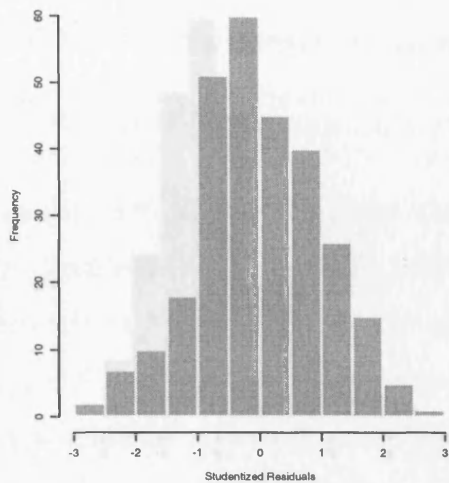


Figure 6.9: Normal Plot of Studentized Residuals

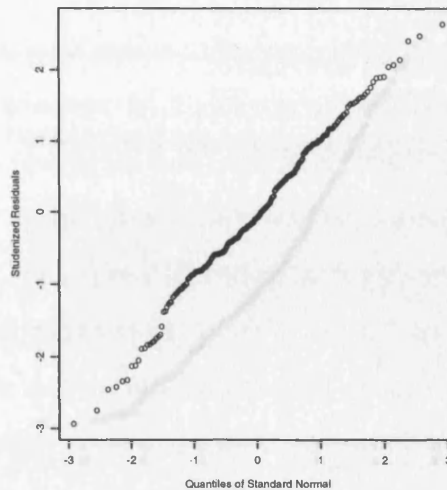


Figure 6.10: Predicted Values vs Studentized Residuals

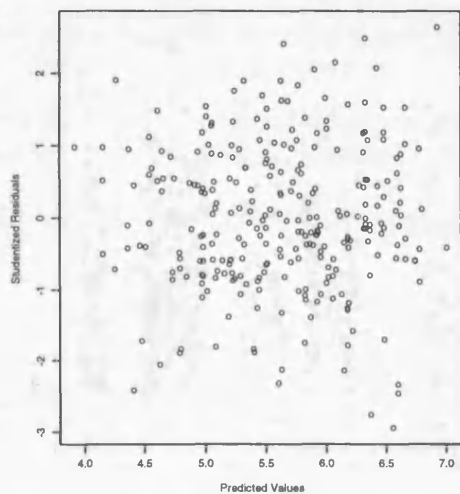


Figure 6.11: Plot of Studentized Residuals vs Townsend Index

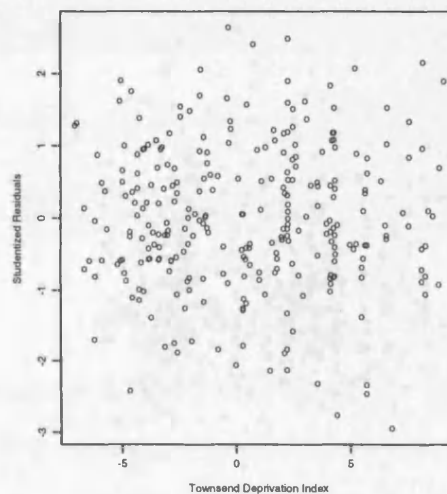


Figure 6.12: Plot of Leverage Values

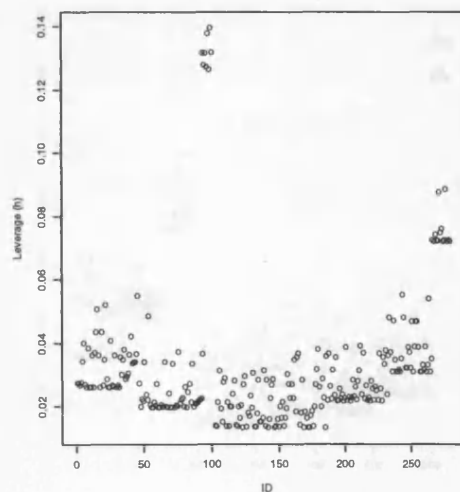
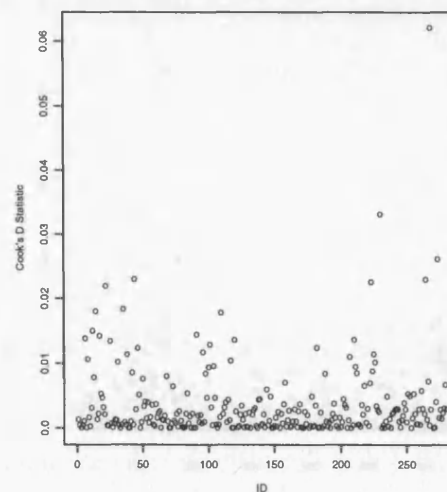


Figure 6.13: Plot of Cook's D Statistic



Figures 6.14-6.19: Model checking for children with large angle strabismus

Figure 6.14: Histogram of Studentized Residuals

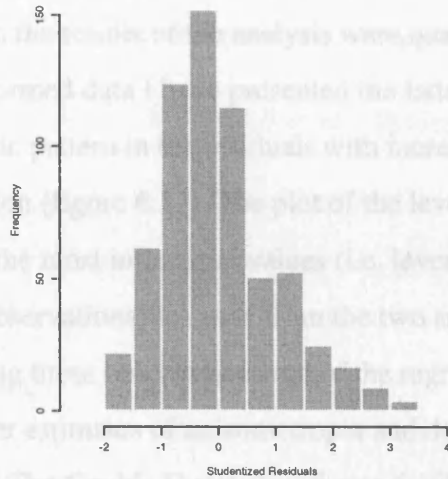


Figure 6.15: Normal Plot of Studentized Residuals

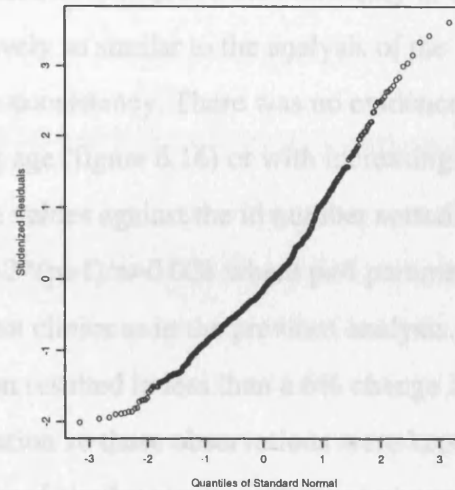


Figure 6.16: Predicted Values vs Studentized Residuals

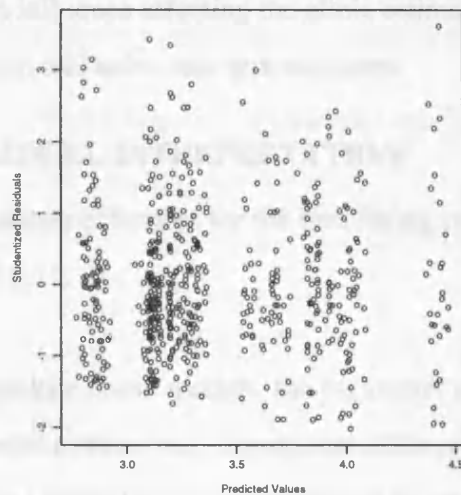


Figure 6.17: Plot of Studentized Residuals vs Townsend Index

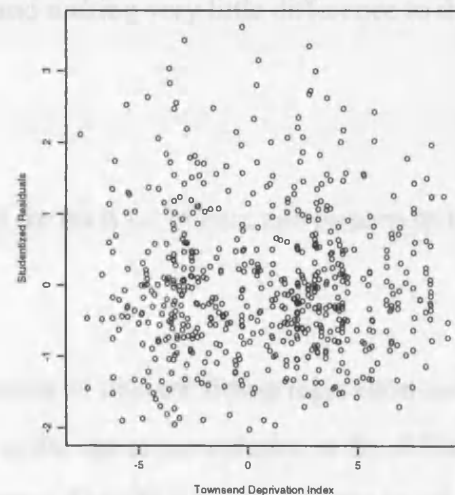


Figure 6.18: Plot of Leverage Values

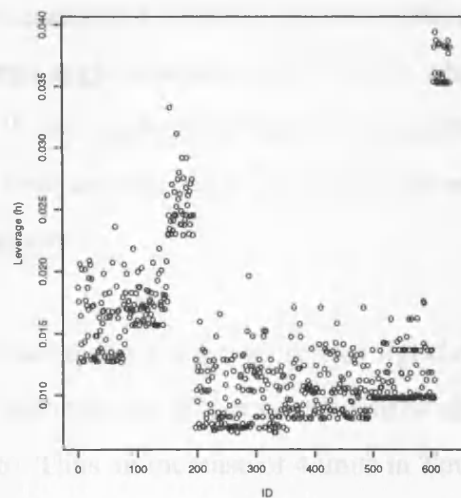
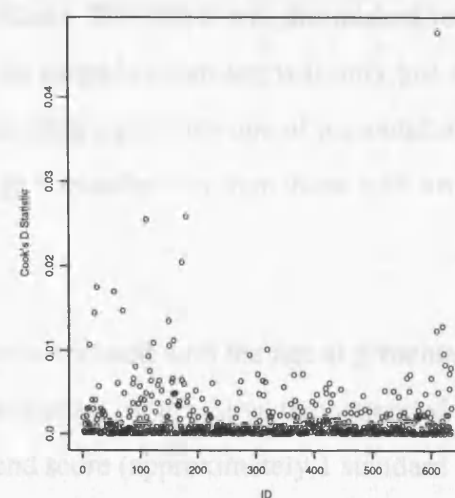


Figure 6.19: Plot of Cook's D Statistic



Figures 6.14-6.19 illustrate the checking procedures for the model fitted for children with a large angle of strabismus. The histogram of the residuals showed slightly skewed data (figure 6.14) and the Normal scores plot showed a very slight curvature (figure 6.15). When a log-transformation of age was used as the outcome variable to improve the Normality of the residuals, the results of the analysis were qualitatively so similar to the analysis of the untransformed data I have presented the latter for consistency. There was no evidence of a systematic pattern in the residuals with increasing age (figure 6.16) or with increasing deprivation (figure 6.17). The plot of the leverage values against the id number sorted by clinic showed the most influential values (i.e. $\text{leverage} > 2 \cdot (p+1)/n = 0.028$ where $p=8$ parameters and $n=636$ observations) to come from the two smallest clinics as in the previous analysis. Weighting these observations out of the regression resulted in less than a 6% change in the parameter estimates of anisometropia and deprivation so these observations were kept in the analysis. The Cook's D statistic (figure 6.18) showed similar results with those observations with high influence affecting the clinic estimates and making very little difference to the deprivation and anisometropia estimates.

6.5.1 MODEL INTERPRETATION

The parameter estimates for the best fitting model for each set of data can be seen in tables 6.14 and 6.15.

Having checked both models, the parameter estimates of the best fitting regression models confirmed that there were significant differences in the age at presentation at the different clinics for children with microtropia or no strabismus ($P=0.0001$), with differences of over two years in the adjusted means of the most extreme clinics. This effect was diminished for those with a large angle of strabismus ($P=0.09$) where the range in mean age was only just over 6 months. Within each group there was a significant difference in the age of presentation of children with anisometropia, presenting on average 9 months later than those with no anisometropia.

The Townsend deprivation score was significantly associated with the age at presentation for children with microtropia or no strabismus after adjusting for the clinic they attended ($P=0.016$). Thus an increase of 4 units in Townsend score (approximately 1 standard deviation) led to an average delay in presentation of 4 months, with a difference of 15 months

between those from the least deprived and most deprived areas of the study after adjusting for the clinic they attended. No similar relationship was observed in children with a large angle of strabismus ($P=0.66$) with a nonsignificant difference of 2 months between those from the least deprived and most deprived areas. In order to assess whether there was a significant difference in the effect of deprivation between the different types of amblyopia a regression model was fitted to the combined data and an interaction between deprivation and type of amblyopia was included in the model. This gave a P value of 0.06 indicating further that there is a significant difference in the referral of those with an obvious squint and those with apparently straight eyes. Among those with microtropia or strabismus there was a significant difference between those with an Asian name and those with a non-Asian name. However this effect was no longer significant after adjusting for deprivation and clinic. This effect seems to have been due to the fact that most children with Asian names were from two centres and lived in the more deprived areas within those centres and that it is deprivation that explains the apparent relationship between ethnicity and age at presentation.

Table 6.14: Parameter estimates for regression model of age of presentation of children with a microtropia or strabismus

Parameter	Parameter estimate	Standard Error	T value	P value
INTERCEPT	6.970	0.507	13.75	0.0001
TOWNSEND	0.077	0.032	2.41	0.0164
ANISO				
<1 dioptre	-0.832	0.276	-3.01	0.0028
>1 dioptre	-	-	-	-
CLINIC				
Birmingham	-1.507	0.604	-2.50	0.0131
Bristol	-1.366	0.558	-2.45	0.0150
Leeds	-1.610	0.832	-1.94	0.0539
Leicester	-0.806	0.552	-1.46	0.1450
Nottingham	-0.934	0.575	-1.62	0.1058
Sunderland	-2.312	0.631	-3.67	0.0003
Worthing	-	-	-	-

Table 6.15: Parameter estimates for regression model of age of presentation of children with a large angle of strabismus

Parameter	Parameter estimate	Standard Error	T value	P value
INTERCEPT	3.855	0.355	10.86	0.0001
TOWNSEND	0.009	0.021	0.45	0.6554
ANISO				
<1 dioptre	-0.7250	0.1581	-4.59	0.0001
>1 dioptre	-	-	-	-
CLINIC				
Birmingham	0.2933	0.3960	-0.74	0.4593
Bristol	0.5746	0.3977	1.45	0.1489
Leeds	0.2920	0.4304	-0.68	0.4978
Leicester	0.0451	0.3672	0.12	0.9022
Nottingham	0.1630	0.3707	0.44	0.6603
Sunderland	-0.0621	0.3987	-0.16	0.9941
Worthing	-	-	-	-

6.5.2 ALTERNATIVE MODELS

A reanalysis of these data by type of amblyopia i.e. pure strabismic amblyopia, mixed strabismic and anisometropic amblyopia and pure anisometropic amblyopia, can be seen in Smith et al (1994) (see paper in appendix B). This analysis showed similar results with a significant effect of deprivation in those with pure anisometropic amblyopia, but not among those with strabismic or mixed amblyopia. As the data were divided into three groups there were not enough data to test for an interaction between degree of strabismus and the relationship with deprivation. Furthermore it was not possible to look at whether children with anisometropia presented later than those without, for children with different levels of strabismus.

One limitation of the analysis presented in this chapter is that it does not allow for the expected similarities between individuals within the same ward i.e. that children from the same ward might be expected to have similar ages to each other than to children in other wards. In order to assess whether this would have affected the results of the analysis I fitted a model

allowing for correlation between individuals in the same ward using a multi-level modelling approach as suggested by Goldstein (1995). This involves fitting a model to allow for within and between ward variation:

$$y_{ij} = \beta_0 + \beta_1 * (Townsend) + \beta_2 * (Aniso) + \gamma_k + u_i + e_{ij}$$

where y_{ij} is the age of child i from area j

and γ_k is the estimate for clinic k

and $u_i \sim N(0, \sigma^2_B)$ the between area variation

and $e_{ij} \sim N(0, \sigma^2)$ the within area variation

The parameter estimates for this model were extremely similar to those from the original analysis and the standard errors for these parameter estimates were very slightly raised (parameter estimate for the Townsend score: original model 0.077 (s.e. 0.032) $P=0.016$, multilevel model 0.075 (s.e. 0.034) $P=0.029$). The intraclass correlation coefficient $((\sigma^2_B/(\sigma^2_B + \sigma^2)) = (0.444/0.444 + 3.003) = 0.129)$ was low indicating little similarity between children of the same ward. Therefore the results and interpretation of this model were qualitatively so similar to the original results that the latter is discussed here.

6.6 DISCUSSION OF RESULTS

Wang et al (1990) showed that parents are often the first to notice visual problems in their children. Squints are often readily apparent but poor vision is harder to identify. I hypothesised here that patients from more deprived areas would have amblyopia detected later than children from less deprived areas. These results show that this is true of amblyopia associated with microtropia or no strabismus, conditions which are generally asymptomatic, but not true of amblyopia associated with a large angle of strabismus. These results also showed that the age at presentation of children with a large angle strabismus was much lower than for those with microtropia or no strabismus. Those with microtropia or no strabismus were on average above school age when they presented with amblyopia. This finding is consistent with those of Shaw et al (1988) and Hiscox et al (1992). In children with all levels of strabismus those with anisometropia presented significantly later than those without. This perhaps implies that in these cases the strabismus follows the anisometropia, since a large angle strabismus should be detected irrespective of any associated anisometropia.

The analysis showed a much wider difference in age at presentation between centres for children with a microtropia or no strabismus than for those with a large angle of strabismus. This may be because factors associated with local screening and health care provision play a more important role in the referral of children with amblyopia associated with a microtropia or no strabismus because it is usually asymptomatic. Unfortunately information is not available on the source of referral of the children in this study. For example it is not known whether individuals were referred by a screening service. The sources of referral for the seven clinics are known to have varied greatly in 1983 with some clinics running extensive screening programmes while others had none. Sunderland, which had the lowest age of presentation of children with a microtropia or no strabismus, was known to be the only clinic to employ orthoptists as primary screeners in screening for the detection of amblyopia. Despite this, I found no evidence of an interaction between clinic and deprivation, with the effect of social deprivation on age at presentation being of the same magnitude in each centre despite their different sources of referral.

As discussed in chapter 5 there has been little research into factors affecting presentation but Shaw et al (1988) found Asians presented later than Caucasians. In our study there was a difference in age at presentation between those with and without an Asian name but this was no longer significant after adjusting for deprivation. This indicates that the effect was due to the Asians in the study generally coming from more deprived areas and thus it was the effect of living in a deprived area rather than having an Asian name that was related to the delay in attendance.

Shaw et al (1988) also found that boys presented later than girls but could offer no explanation for this finding. No similar relationship between sex and age at presentation was observed in this study. Campbell and Charney (1991) found age at diagnosis depended on degree of parental suspicion and level of maternal education for all types of amblyopia. This study could not measure this specifically but the Townsend score is thought to reflect these factors.

This cohort represents one of the largest series of children with amblyopia ever studied. The children were referred in 1983 and their treatment followed through to 1991. The referral patterns thus relate to a period of up to 14 years ago and it is possible that the importance of

deprivation has either increased or decreased since that time. Missing data are usually a problem with retrospective studies although in this study it was minimal.

Deprivation was measured using the Townsend deprivation score based on census data from electoral ward areas. However there may be a stronger underlying relationship between individual deprivation and age at presentation than I have shown by using this relatively insensitive measure. The issue of ecological fallacy discussed in chapter 3 must be born in mind since a relationship found at the area level may not necessarily exist at the individual level. Although there was no individual level measures of deprivation available in this study, this problem is further investigated later in the thesis by analysing data at both area and individual levels.

6.7 IMPLICATIONS OF THE FINDINGS

This analysis has shown inequalities in the presentation of asymptomatic amblyopia but not in symptomatic amblyopia indicating differential use of or access to screening services for the detection of asymptomatic amblyopia. This shows a similar relationship to the patterns of lower uptake of immunisation and screening services by people of lower social class shown by Essen and Wedge (1982) and Blaxter (1981) using data from the National Child Development Study. Reading, Jarvis and Openshaw (1993), Marsh and Channing (1986) and Williamson et al (1995) have also shown similar patterns using measures of deprivation akin those used in my analysis.

Macintyre (1989) has argued that if differentially used, screening or treatment programmes could influence cure of conditions between different social groups. Since the review of literature in chapter 5 has shown some evidence of improved outcome for earlier treatment, later detection of asymptomatic amblyopia among children from more deprived areas could mean poorer outcome of treatment in these individuals.

To reduce these inequalities an understanding of the underlying explanations is needed. In order to aid understanding of the implications of this finding I use the model constructed by Feinstein (1993) to illustrate factors explaining health inequalities as shown in table 6.16.

Table 6.16: Feinstein's conceptual decomposition of factors explaining health inequalities

		Source of inequality	
		Life span	Access to and utilization of health care system
Type of explanation	Materialistic (access to resources)	Housing, overcrowding, sanitation, transit mode, occupational hazards, environmental hazards	Ability to purchase health care, ability to purchase pharmaceuticals, regular physician
	Behavioural (psychological, genetic, cultural)	Diet, smoking, exercise regime, leisure activities, risk taking, alcohol and substance abuse.	Comprehensive medical information, 'playing the system', following instructions, self-diagnosis, and awareness of recurrence

It is apparent from this conceptual decomposition that in this case there are inequalities in access to and utilisation of the health care system and that it is materialist or structural explanations that affect the ability to access screening services.

Most districts of Britain implement extensive visual screening of children with the early detection of asymptomatic amblyopia as one of the main objectives. Based on Feinstein's conceptual decomposition, if interventions are to be made to tackle these inequalities, an approach needs to be made from a structural and not a behavioural perspective i.e. not through changing people's behaviour but through changing the structure of existing services to make them more accessible to people with less access to resources. Reading, Colver, Openshaw and Jarvis (1994) have shown that improving the uptake of screening services does not necessarily reduce inequalities but could even make them relatively greater and so this must also be remembered when deciding on an intervention strategy.

6.8 OTHER ASPECTS OF AMBLYOPIA

In this chapter I have presented the investigation of a very specific hypothesis about the presentation of amblyopia and deprivation. However, I have performed further analyses on the data to look at other aspects of interest and these papers are enclosed at the end of the thesis. Woodruff et al (1994a) (Appendix C) and (1994b) (Appendix D) include my analysis of data on factors relating to the presentation of amblyopia and the outcome of treatment for amblyopia for this group of children. Analysis of the final visual acuity after treatment for

amblyopia (1994b) showed no evidence of a relationship with Townsend score even though I have shown in this chapter that age at presentation is related to deprivation. Furthermore, there was no evidence of a relationship between age at presentation and final visual acuity. If this is the case then the inequalities seen in the age at presentation may not be affecting final outcome of treatment. However these relationships may be confounded by the problems of visual acuity measurement discussed in chapter 5.

Non-compliance has been shown by analysis of this study to reduce the improvement a child could achieve from treatment (Woodruff et al, 1994b). I also used the data from this study to look at factors affecting compliance with amblyopia treatment (Smith, Thompson, Woodruff and Hiscox, 1995) (Appendix E). Clinic attendance was used as a proxy for compliance which has been used by other studies in amblyopia (Nucci et al, 1992; Oliver et al 1986). This showed a significant decrease in compliance with increased deprivation ($P < 0.0001$) with compliance rates over 50% better in the least deprived areas compared with the most deprived areas. This suggests that compliance may be related to factors associated with deprivation such as the ability to attend the clinic, access to transportation, family support and motivation. Further study is necessary to investigate this relationship between compliance and deprivation and to understand the degree of association between compliance to treatment and attendance at clinic.

6.9 SUMMARY AND CONCLUSIONS

In this chapter I have sought to assess whether inequalities in access to care discussed in chapter 2 exist in the case of amblyopia. This analysis has indicated that for types of amblyopia where screening is the main method of detection there is a relationship between age at presentation and deprivation. This is not the case for more easily detectable types of amblyopia.

These findings have pointed to structural differences in the accessibility of services. In order to tackle inequalities in the presentation of asymptomatic amblyopia, interventions at a structural level are necessary. As discussed in chapter 5, structural changes have been made to screening practices in Leicestershire since this study was undertaken although they were not explicitly

made to reduce inequalities in access. In chapters 7 and 8, I describe the design and analysis of a study to evaluate the effect of these structural changes on the inequalities in presentation of amblyopia in Leicestershire.

The importance of adjusting for deprivation when looking at ethnicity has also been shown by this analysis since apparent differences in age at presentation between Asians and non-Asians seem to be due to differences in deprivation. A limitation of the analysis is the issue of ecological fallacy discussed in chapter 3 since area data are being used to estimate an individual's level of deprivation. The extent of this problem is assessed by using data collected in the second study of amblyopia on both individual and area level deprivation. The relationship between area and individual deprivation is then discussed in chapters 9 to 12.

CHAPTER 7

DESIGN OF A STUDY TO INVESTIGATE THE EFFECT OF CHANGES IN VISION SCREENING IN LEICESTERSHIRE

7.1 AIMS OF THIS CHAPTER

In the previous chapter, I showed inequalities in the age at presentation of asymptomatic amblyopia, with those from more deprived areas presenting later. There was no similar pattern among those with symptomatic amblyopia. This finding pointed to a structural difference in the provision of services and indicated that interventions needed to be made from a structural rather than a behavioural perspective. As discussed in chapter 5, since the multi-centre study was undertaken, major structural changes were made to the screening programmes in Leicestershire particularly to vision screening. In this chapter I describe my design of a study to investigate whether these changes were associated with changes in the presentation of children with amblyopia. A limitation of the analysis of chapter 6 is the issue of ecological fallacy where area-level deprivation measures are used to assess deprivation at an individual level. I discuss the design of the study to fulfil the secondary aim to collect individual level deprivation data to assess how good area measures are as a proxy for individual level deprivation.

7.2 STUDY OBJECTIVES

7.2.1 AGE AT PRESENTATION AND DEPRIVATION

Chapter 6 discussed the analysis of a multicentre study of amblyopia which showed a relationship between age of presentation and the Townsend deprivation score for the ward in which the child lived at each of the seven clinics. Among children with amblyopia associated with microtropia or no strabismus, those from more deprived areas were seen to present at an older age. There was no similar relationship found among those children with amblyopia associated with a large angle of strabismus. An explanation for this pattern is that amblyopia

associated with a large angle of strabismus is easier to detect but that amblyopia associated with microtropia or no strabismus relies on screening for its detection. Although deprivation does not affect the ability of a parent to detect a large angle of strabismus, it does affect their access to screening, pointing to structural differences in the provision of services.

Screening plays a large role in the detection of asymptomatic amblyopia. In the 1983 multicentre study the only clinic where an orthoptist was known to perform primary screening also had the youngest age of referral for amblyopia associated with microtropia or no strabismus. The variability in screening practices in the seven clinics is perhaps also reflected in the high variability in the mean age at presentation of this type of amblyopia at the clinics. As discussed in chapter 5, between 1987 and 1991 radical changes were made to the child health surveillance programme in Leicestershire. Over the last twenty years a system has been in place in Leicestershire for children to be screened for amblyopia and strabismus, at 6 weeks, 7 1/2 months, 18 months (since 1991) and 3 1/2 years. Initially health visitors were responsible for most of the screening with children thought to have strabismus or amblyopia being referred to a consultant ophthalmologist via their GP. Between 1988 and 1991 radical changes were made to the county's vision screening. A major development was the introduction of a secondary orthoptic screening service which allows a much more prompt and readily available referral service for children suspected of having amblyopia or strabismus at the initial screening. Although most of the initial screening is still carried out by health visitors, responsibility of child health surveillance has been transferred to GPs who are required to make a return for every child screened.

The Hall report (1989) reviewed all screening services for pre-school children and highlighted the fact that many new screening programmes were introduced before their benefit had been established. Although the report noted the widespread practice of pre-school vision screening it found no evidence of health gains to support this practice and questioned the continuation of screening tests for amblyopia and strabismus. Orthoptic based pre-school screening has been shown to be more effective but there has been little research to investigate the effectiveness of secondary orthoptic screening. Research is needed to investigate whether these changes in screening have an effect on the referral of amblyopia.

One of the principal aims of the study was to investigate the effect of these changes to screening in Leicestershire on the presentation patterns of amblyopia. Of particular interest is the effect of these changes on amblyopia associated with a microtropia or no strabismus. I aimed to investigate whether the changes would lead to a reduction in the effect of deprivation on age at presentation and whether the overall age at presentation would be affected. By comparing children referred in 1983 to children referred after the changes to screening had been made it would be possible to assess these questions. Therefore, the aim of the study was to use the Leicestershire data collected for the 1983 multicentre study and, using the same methodology, to collect similar data for children referred after the changes were introduced. This was possible using patient orthoptic notes only.

7.2.2 AREA AND INDIVIDUAL DEPRIVATION

The main conclusion drawn from the multicentre study was that age of presentation was either related to area deprivation representing the availability of local health care, or individual deprivation representing lower educational standards and less ability to make use of available facilities. It is most likely to be a combination of both area and individual deprivation. In chapter 3 I highlighted the problem of ecological fallacy where an effect at the area level is wrongly implied to be present at the individual level. The Townsend score estimates deprivation in a geographical area rather than at an individual level and the degree of deprivation can vary widely across a ward, particularly since wards range in size from 500 households to 15000 households. Hence the Townsend score may not accurately reflect the deprivation of an individual. Mertens (1993) has shown that random misclassification of a study exposure reduces the observed strength of an association. Therefore if individual deprivation is important then there is likely to be a stronger underlying relationship between deprivation and age at presentation than I was able to demonstrate using the data from the 1983 multicentre cohort and the Townsend score at ward level.

There are possible ways of estimating the variability within wards. However, these only give a guide as to the accuracy of the ward level estimate and the methods are approximate. Some methods involve weighting patients according to the estimated variability. These methods do not give us an estimate of individual deprivation and it is preferable to weight all patients equally. Several studies have investigated the effectiveness of deprivation scores with respect

to health measures but the relationship between ward level deprivation scores and individual deprivation has not been fully investigated.

When analysing the data from the multicentre cohort it was only possible to look at ward level deprivation. Since 1991 it has become possible to relate postcodes directly to enumeration districts (ED) which contain only about 200 households. This study provides information in order to be able to compare ward level and ED level estimates of deprivation with individual level information to see how good these area measures are for use in such studies. The study aimed to estimate the deprivation of the individual's household by looking at measures of deprivation at a household level and form an individual deprivation score for the household. The collection of this type of information required personal interviews with the parents of those children participating in the study.

7.2.3 FURTHER ASPECTS

Several other aspects of amblyopia were also investigated in this study although they will not be reported in this thesis. The pattern of referral of amblyopia patients to different health professionals varies greatly, with some being referred straight to hospital by a GP, while others are sent from one expert to another before being seen by an ophthalmologist or orthoptist therefore varying the wait before seeing an eye specialist. One of the aims of this study was to make a comprehensive record of the precise referral pattern from when the child first had a suspected problem until they received appropriate care.

Treatment non-compliance is an important issue in the success of treatment for amblyopia. As commented on in chapter 6, clinic attendance is frequently used as a measure of compliance. This study also aimed to compare clinic attendance with parental response about compliance and to use this to understand more factors related to poor compliance.

7.3 STUDY DESIGN

7.3.1 OVERVIEW OF STUDY DESIGN

The principal aim of the study was to compare age at presentation before and after the introduction of changes to screening. Therefore since information on children with amblyopia

who first presented to Leicester orthoptic clinic in 1983 was available, information on a similar cohort was necessary after the changes had been made. In order to make valid comparisons exactly the same methodology was employed. Hence to answer this primary question the study was designed to be a cohort of all children treated for amblyopia who first attended in a given period with a review of the orthoptic notes for all children in the study necessary.

The second aim of the study was to assess individual and area deprivation. This required more in-depth information which was only available from the parents of the children in the study. Hence for this part of the study a detailed questionnaire was administered to the parents.

The study was based on a cohort of children treated for amblyopia. Their parents were interviewed about topics covering referral, compliance and deprivation. The best location to interview the parents was thought to be the orthoptic clinic, when their child attended for a routine orthoptic appointment regarding their amblyopia. Children who were discharged before the start of the study or those who failed to attend the clinic were approached by post and interviewed over the telephone. Children from families with no telephone or those who would not permit the release of their telephone number were asked to complete a postal questionnaire. Routine information was retrieved from orthoptic notes using the same methodology as the multi-centre study.

I was the main investigator for this study. Although the interviewing and data entry were undertaken by a research clerk, I was in charge of designing the study, compiling and testing the questionnaire, day-to-day management of the study, liaison with the orthoptic and ophthalmology departments, supervision of data entry and validation, and analysis of the study data.

7.3.2 SAMPLE SIZE AND POWER CALCULATIONS

Sample size calculations were based on investigating the effect of changes in screening on the relationship between deprivation and age at presentation among children with microtropia or no strabismus. In order to assess the number of patients needed for the study two probabilities need to be minimised. Firstly the chance of rejecting the null hypothesis of no change in the effect of deprivation between the studies given the null hypothesis is true (α). Secondly, the probability of obtaining a non-significant result when there has been a true reduction in the

effect of deprivation (β where $100(1-\beta)\%$ is the power of the study). Power calculations were performed using simulation studies. Since the final data would comprise of two groups, before and after changes to screening, the data were to be analysed using analysis of covariance. The sample size needed to be large enough to detect an interaction between cohort and deprivation if it exists at a given level of significance.

The data on the period before changes to screening were made had already been collected as part of the 1983 study. The Townsend scores for the children in 1983 attending the Leicester clinic were re-calculated based on the mean and standard deviation of all of the wards in a 10 kilometre radius of the clinic. These scores were to be calculated using the 1991 census for the second cohort, using the mean and standard deviation for 1991. It was therefore expected that the Townsend scores for each year would have a similar distribution. Investigating the distribution of the Townsend scores for the children presenting in 1983 showed them to come from an approximately Normal distribution with mean 0 and standard deviation 3.8.

A regression model was then fitted to the Leicestershire 1983 data with age at presentation as the outcome variable and the Townsend score as an explanatory variable. Only those children who lived in a 10 kilometre radius of the orthoptic clinic and those who had a delay of less than two years before treatment were included in the model (70 children). This model showed the parameter estimate of the Townsend deprivation score to be 0.20 (standard error 0.05) which is larger than that in the audit of all seven clinics. The standard deviation of the residuals was found to be 1.5. The required sample size was estimated by simulating a varying number (n) of children with a Townsend score from an $N(0,3.8)$ distribution. Their age at presentation was then calculated based on the equation

$$AGE = 6.337 + b \cdot TOWNSEND + \epsilon$$

where 6.337 is the estimated intercept based on the 1983 data, b is an estimate of the effect of deprivation in the new cohort (compared to an estimate of 0.2 in 1983) and ϵ is an error term from an $N(0, \sigma^2)$ distribution. A model was then fitted to the combined data from the 1983 study and the simulation, to test the effect of an interaction between deprivation and cohort. 1000 simulations were performed in SAS for each combination of estimates of b (0,0.05,0.1), σ (1.5,1.75) and n (70,90,110). Therefore an estimate of 0.05 for b represented a reduction in the effect of deprivation of 0.15 from 0.2 in 1983 to 0.05 in the new cohort. Although the validity of the analysis of covariance relies on similar variances in both groups two values for

the standard deviation were used to allow for inaccuracy in the estimate. The power of each combination of b , n , and σ was then calculated for two significance levels ($\alpha=0.05, 0.01$) by looking at the number of simulated studies showing a significant P-value at the α level for the interaction in each model. Table 7.1 shows the results of these simulations.

Table 7.1: A table of corresponding power for studies of varying sample size (n), standard deviation of errors (σ) and estimates of deprivation in the second study (b) and significance level (α)

n	σ	b	$\alpha=0.05$	$\alpha=0.01$
70	1.5	0.00	91.6	67.8
90	1.5	0.00	95.2	76.6
110	1.5	0.00	98.1	84.3
70	1.5	0.05	63.3	29.9
90	1.5	0.05	72.5	34.9
110	1.5	0.05	76.2	38.8
70	1.5	0.10	23.0	5.4
90	1.5	0.10	27.6	5.4
110	1.5	0.10	27.5	4.5
70	1.75	0.00	81.6	52.4
90	1.75	0.00	88.7	62.7
110	1.75	0.00	93.0	68.0
70	1.75	0.05	52.4	20.5
90	1.75	0.05	59.3	23.8
110	1.75	0.05	62.3	27.0
70	1.75	0.10	21.7	4.8
90	1.75	0.10	24.7	5.6
110	1.75	0.10	21.4	3.5

From the estimates in table 7.1, a similar sized cohort to 1983 would give a study with a power of 70% at the significance level of 0.05 if the variance was similar to the 1983 data and the estimate of the effect of deprivation was 0 in the new study (compared to 0.2 in 1983). An increase in the sample size would mean a power of greater than 70% at a significance level of

0.05 and a deprivation estimate of 0.05. However, if there was less of a difference in the effect of deprivation between the two years, then the power would be greatly reduced unless a great many children were included in the study. It has been shown by other studies that using orthoptists in the screening process increases the detection of 'straight-eyed amblyopia' so it is likely that the number of cases would be greater than 70 in one year. An estimate of deprivation of 0.1 in the new cohort was thought to be of less clinical significance and so it was decided to use one year of patients and then if there were fewer cases to increase the sample.

This analysis is based on data collected retrospectively from orthoptic notes and so non-response would have no effect on the power of the study in terms of detecting a significant change in the effect of deprivation on age at presentation. The issue of non-response in answering other questions posed by the study was not thought to be a problem. It was thought that the questionnaire and the situation in which it was asked would lead to a very high response rate. The study population was therefore set to be all children who first attended the orthoptic clinic at Leicester Royal Infirmary in one year and who were subsequently treated with occlusion. It was necessary to choose a year close to the last census for up-to-date census data. It was also necessary to have a recent year so that many of the children were still attending clinic regularly. As the study was organised to start in September 1993 this meant children in the study would have to have their first appointment at the clinic between 1991 and 1993. Therefore all children first attending the orthoptic clinic in 1992 were to be included in the study. These would have been referred after the changes were made to screening between 1987 and 1991.

7.3.3 ETHICAL CONSIDERATIONS

Since the study involved looking at patient's orthoptic notes, ethical approval for the study was sought from the Leicestershire ethical committee which was subsequently granted. Patient confidentiality was ensured by anonymisation of all questionnaires and all data entered into the study database from the orthoptic notes. Parents were asked for their consent to be involved in the study and always had the right to refuse to take part.

7.3.4 QUESTIONNAIRE DESIGN

The main aim of the questionnaire was to collect data on individual levels of deprivation. It was also designed to collect information on compliance and referral processes. The design of the questionnaire was important since it had to be suitable for three different methods of administration, interviewer administered questionnaire in the orthoptic clinic, interviewer administered questionnaire over the telephone, and self-completed postal questionnaire. The questionnaire was only to be answered by the child's parents and all information remained confidential. The questionnaire compiled for the study can be seen in appendix F.

The questionnaire recorded information relating directly to the child and their condition of amblyopia. The primary aim was to record the pattern of referral from when the child's condition was first noticed until they were treated. The questions were designed to assess the degree of contact the child had with health professionals, the exact path of referral the child took to get to the clinic and whether the problem had gone unnoticed in vision screening. Any family history of the condition was also noted to see if these children present earlier.

A further aim was to get a basic understanding of the child's compliance with treatment. This was assessed by asking questions about the ease of patching the child and also compliance in attendance at clinic with questions to assess how easy it was for the parents to attend the clinic with their child.

In order to measure individual deprivation various questions were incorporated into the questionnaire. Each of the four variables that comprise the Townsend score were assessed at an individual level, i.e. whether the parents owned their house, the ratio of persons in the household to rooms, whether the household had access to a car and whether the parents were unemployed. The study recorded occupations of the parents to assess social class for comparison. It also aimed to look at the effect of one-parent families and ethnicity on deprivation. Since ED and ward level Townsend deprivation scores were obtained using postcode data, the phrasing of the questions was copied from the census since these questions have been thoroughly checked for validity and would be consistent with the data collected in the census. Information was also recorded on the birth order of the child, where the mother was living when the child was born, whether the child comes from a single parent family and what language was spoken in the home.

There were also questions relating to the parents of the child. It recorded ethnic background of the parents, country of birth, level of education and occupation using the exact questions used in the census. Occupation was recorded for subsequent social class classification using the Registrar General's classification scheme.

A pilot study ran in June and July 1993 to look at the response to the questionnaire and any problems that occurred and to estimate the length of interviews. Twenty pilot interviews were carried out by the interviewer to familiarise herself with the questionnaire. The interviews lasted between 5 and 10 minutes, with no refusers. Several questions regarding compliance were altered slightly for clarity but no other questions were problematic.

7.3.5 METHODS AND TIMETABLE

The same methods used for the 1983 multi-centre audit were used to collect data from the orthoptic notes for comparability. An orthoptist identified all patients who had their first appointment in 1992, underwent patching treatment and were under fifteen years old. For each patient a record was made of the presence of manifest strabismus on cover testing (Large angle strabismus: more than 5 degrees, Microtropia: 5 degrees or less) and the presence of anisometropia (a difference between the two eyes of at least one dioptre in either sphere or cylinder).

A research clerk then collected information from the orthoptic notes of these children on attendance at orthoptic screening, age at presentation, age at starting treatment, address and postcode, Asian/non-Asian name, and visual acuity at presentation, starting treatment and discharge. When the child next attended the clinic the research clerk approached those who accompanied them. If the parents accompanied the child then they were asked to participate in the study by answering the questions in the questionnaire administered by the clerk. If they did not have the time but agreed to being in the study, they were interviewed at their next appointment. If the child was not accompanied by his or her parents then the questionnaire was not administered and the parents were approached at the next visit or by telephone.

Parents of children who had been discharged before the start of the study or those who failed to turn up for any appointments during the study period were informed about the study by letter and then they were telephoned by the clerk who administered the same questionnaire

over the telephone. For parents who were not contactable by telephone a postal questionnaire was sent out with an explanatory letter and prepaid envelope (appendix G) with a maximum of three postings if no positive or negative reply was received. Addresses of children not interviewed were checked using the Family Health Services Association register and by contacting the child's general practitioner. Those patients who refused to enter the study were to be followed up, looking at orthoptic notes and deprivation scores at enumeration district level via their postcode.

All of this information was then entered into a data base for statistical analysis. All patients first attending for amblyopia treatment in 1992 were classified in May 1993. The research clerk began interviewing in October 1993 after finalising the questionnaire and producing the appointment diary. The orthoptist rechecked the notes in April 1994 to locate children who at the time of the first checking had not undergone occlusion therapy but went on to receive occlusion at a later date and these were included in the study and interviewed. The majority of clinic interviews were completed in the first six months of the study but the final telephone interviews and postal questionnaires were not completed until a year later due to problems with patient's home addresses and response to postal questionnaires. The data was all collected by October 1994.

Data were collected on 218 children first presenting to the orthoptic clinic in Leicester in 1992. Since these data were to be used to investigate age at presentation of amblyopia, only 202 were eligible for inclusion in the analyses described. Children were excluded because they attended for over two years before treatment commenced as it was doubtful that they had amblyopia on presentation. Children were also excluded if they lived over 10 kilometres from the clinic to avoid the problems of other orthoptic clinics affecting the data. The overall response rate for the questionnaire for these children was 91% with only 19 non-responders. Two of these nonresponders were not under the supervision of their parents and so they were not eligible to complete the questionnaire, and one further child had moved out of the area with no forwarding address. Of those interviewed 59% (108) were clinic interviews, 35% (63) were telephone interviews and 7% (12) were postal questionnaires.

7.4 SUMMARY AND CONCLUSIONS

This chapter has described the design of a study to investigate the effect of changes to screening on inequalities in the presentation of amblyopia. It has illustrated my role in the design and undertaking of this study and identified the main study hypotheses.

The data from this study will be used in this thesis to investigate two main hypotheses. Initially I use the data to assess the effect of changes made to vision screening on the relationship between age at presentation of amblyopia and deprivation. The data on age at presentation of amblyopia from this study will be analysed in conjunction with the data on Leicestershire from the multi-centre study in chapter 8. The use of identical methodology permits direct comparison of data in 1983 and 1992, before and after the changes were made to screening.

The questionnaire was designed to look at aspects of individual deprivation as well as collecting information on referral processes and compliance with treatment. In chapter 10, I analyse these data to compare area and individual measures of deprivation.

Although data on other aspects of presentation and compliance were collected in this study, these will not be discussed in this thesis but offer other future possible avenues of investigation.

CHAPTER 8

ASSESSING THE EFFECT OF SCREENING CHANGES ON AGE AT PRESENTATION OF AMBLYOPIA

8.1 AIMS OF THE CHAPTER

This chapter looks at the effect of changes made to vision screening on the relationship between deprivation and age at presentation of amblyopia. Here I present my analyses of data on the presentation of amblyopia before and after the changes were made to screening in Leicestershire.

As previously discussed in chapter 5, detection of amblyopia associated with anisometropia and microtropia relies on screening while amblyopia associated with a large angle of strabismus is identified when a parent notices the squint. In chapter 6, my analyses showed that in the 1983 multi-centre study the age at presentation of children with microtropia or no strabismus was related to deprivation, but that there was no similar relationship among children with large angle strabismus. I hypothesised that this may be because deprivation does not affect the detection of large angle strabismus but does affect a child's access to the screening necessary to detect amblyopia associated with microtropia or no strabismus. These findings have pointed to structural differences in the accessibility of services. In order to tackle inequalities in the presentation of asymptomatic amblyopia, interventions at a structural level are necessary. As discussed in chapter 5, structural changes have been made to screening practices in Leicestershire between 1988 and 1991, although they were not explicitly made to reduce inequalities in access. A secondary orthoptic screening service was introduced and responsibility of child health surveillance was transferred to general practitioners.

The Hall report (1989) reviewed screening services for pre-school children and highlighted the fact that many new screening programmes were introduced before their benefit had been established. Although the report noted the widespread practice of pre-school vision screening, it found no evidence of health gain to support this practice and questioned the continuation of

screening tests for amblyopia and strabismus. Orthoptic based pre-school screening has been shown to be more effective (Jarvis et al, 1990; Edwards et al, 1989; MacLellan and Harker, 1979). Hall has recommended that if pre-school vision screening is to be undertaken, secondary vision screening by an orthoptist is a more cost effective alternative. However there has been little research to investigate the effectiveness of secondary orthoptic screening. The study described in chapter 7 was designed to provide data to investigate the effect of changes made to the screening programme in Leicestershire to the pattern of amblyopia presentation. In this chapter I analyse the data from children presenting before and after changes were made to screening to assess the effect on the relationship between deprivation and age at presentation. This work is also reported in Smith, Thompson and Woodruff (1995) (Appendix H).

8.2 HYPOTHESIS

This chapter investigates the changes in the age of detection of amblyopia over a nine year period in Leicester during which time there have been major changes in the screening services. I hypothesised that among children with microtropia or no strabismus, the introduction of improved referral processes would reduce the effect of deprivation on age at presentation. No change was expected among children with large angle strabismus. I analyse data from the 1983 and 1992 studies for Leicester to investigate my hypothesis.

8.3 STUDY DATA

The study was based on the two cohorts of children presenting at Leicester orthoptic department in 1983 and 1992. Comprehensive data were collected as described in Chapters 6 and 7. Children who were first treated for amblyopia more than two years after first attending the orthoptic clinic were excluded as being unlikely to have had amblyopia at the time of presentation. To avoid the confounding effect of children attending one of the six outlying clinics at Oakham, Melton Mowbray, Coalville, Hinckley, Loughborough and Market Harborough, only children from within a 10 kilometre radius of the hospital are included in this analysis.

8.3.1 DEPRIVATION

Each child's deprivation score was estimated using the Townsend deprivation score for the ward in which they lived. A 1981 and 1991 computerised postcode directory linked the patient's postcode to the ward in which they lived. Data from the 1981 census were used to calculate the Townsend scores for the 1983 cohort and data from the 1991 census were used for the 1992 cohort. All wards within a 10 kilometre radius of the orthoptic clinic were used to calculate the mean and standard deviation of each variable separately for each year. Changes in patterns of deprivation between the 1981 and 1991 census led to a decrease in the percentage prevalence of all four indicators but very little difference in the variance as shown in table 8.1. This is similar to the trends in these indicators for the whole of England and Wales as shown by Dolan et al (1995).

Table 8.1: Summary statistics of the components of the Townsend score for wards in a 10km radius of Leicester orthoptic department from the 1981 and 1991 censuses (Log transformed values in parenthesis).

	1981 census		1991 census	
Townsend score components	Mean	Standard Deviation	Mean	Standard Deviation
% households with no car	40.5	16.61	34.3	16.22
% households with > 1 person per room	4.5 (1.53)	3.63 (0.59)	2.7 (1.16)	2.32 (0.53)
% households not owner occupied	37.9	21.44	30.5	20.32
% economically active people unemployed	10.3 (2.29)	6.02 (0.51)	9.7 (2.22)	6.27 (0.53)

In order to avoid any problems due to changes in the distribution of deprivation over time, the analyses in this chapter are based on Townsend scores specifically calculated for each year separately based on the summary statistics in table 8.1. Hence absolute deprivation scores for 1981 cannot be compared with deprivation scores for 1991. To construct quintiles of deprivation for each census year I took all the wards in the 10 kilometre radius and ranked them by their Townsend score. I then divided all of the wards into quintiles of deprivation based on the total population of children under 15 obtained from census data. This was done

separately for each year. Thus if the study sample were a representative sample of children under 15, 20% of the study sample would be expected to live in each quintile.

Children were classified by the degree of strabismus (large angle strabismus, microtropia, no strabismus) and the amount of anisometropia (less than one dioptre, at least one dioptre) as defined in the 1983 multi-centre audit reported in chapter 6.

Of the 240 children from the 1983 study, 31 were excluded since they had a delay of over two years before treatment or because they lived outside the 10 kilometre radius of the orthoptic clinic. Of these children, 70 had amblyopia associated with microtropia or no strabismus while 139 had a large angle of strabismus. In comparison there were 218 children in 1992 of which 16 were excluded. Of these children, 92 had amblyopia associated with a microtropia or no strabismus while 110 had a large angle of strabismus.

8.4 PRELIMINARY RESULTS

Table 8.2 shows the mean age at presentation by quintile of Townsend deprivation for each cohort.

Table 8.2: Mean age at presentation, 95% confidence interval and number of children by quintile of Townsend deprivation score, year and type of amblyopia.

Quintile of Townsend deprivation score	TYPE OF AMBLYOPIA											
	Microtropia or no strabismus						Large angle of strabismus					
	1983			1992			1983			1992		
	(Mean, 95% C.I., n)			(Mean, 95% C.I., n)			(Mean, 95% C.I., n)			(Mean, 95% C.I., n)		
1	5.4	4.22,6.47	10	4.5	3.81,5.22	17	2.8	2.03,3.61	19	3.8	2.60,4.91	14
2	5.4	4.51,6.32	8	4.8	4.23,5.31	20	3.2	2.67,3.80	28	3.3	2.37,4.24	18
3	6.5	5.65,7.37	16	5.4	4.17,6.63	21	3.3	2.84,3.77	28	3.3	2.35,4.29	25
4	7.3	6.47,8.07	17	5.3	4.62,6.05	20	3.4	2.80,4.01	42	3.1	2.43,3.74	25
5	7.1	6.37,7.88	19	4.7	3.89,5.48	14	3.5	2.82,4.13	22	3.2	2.52,3.93	28
ALL	6.6	6.17,6.97	70	5.0	4.61,5.34	92	3.3	3.01,3.55	139	3.3	2.93,3.65	110

The overall mean age of children with a large angle of strabismus is 3.3 years in both 1983 and 1992. There has been a significant reduction in the mean age at presentation of those with a microtropia or no strabismus of 19 months, from 6.6 years in 1983 to 5.0 years in 1992 ($P=0.0001$, 95% confidence interval (13-26 months)). There has been an increase in detected cases of amblyopia associated with a microtropia or no strabismus with a corresponding decrease in cases associated with a large angle of strabismus. The proportion of children treated for amblyopia with microtropia or no strabismus has increased from 33.5% to 45.5% (a change of 12 %, 95% confidence interval (2.5%,21.2%)).

Table 8.2 also shows that for children with microtropia or no strabismus there is an increase in mean age with increasing deprivation in 1983 but not in 1992. In order to further investigate the effect of deprivation on the age at presentation of children with microtropia or no strabismus, plots of age at presentation by Townsend deprivation score for each year can be seen in figures 8.1 and 8.2. The regression lines are discussed later. These graphs illustrate evidence of an effect of deprivation among those with microtropia or no strabismus in 1983 but a more variable pattern in 1992.

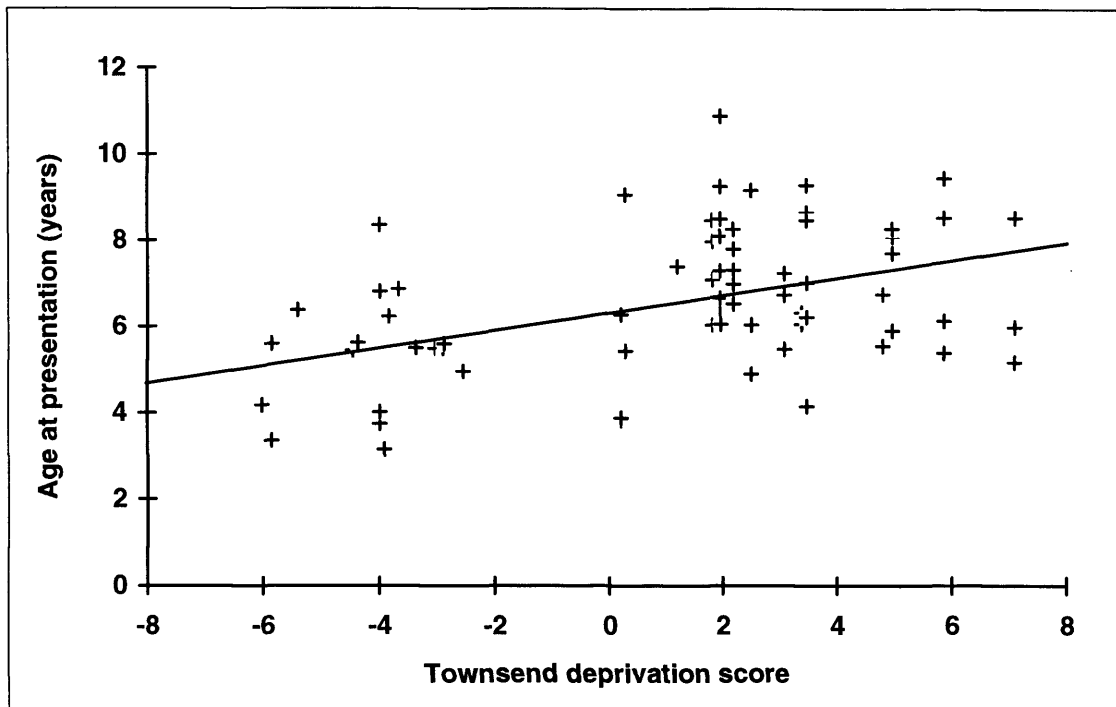
For children with a large angle of strabismus there is a similar but far less pronounced relationship with deprivation. In 1992, a higher proportion of children with a large angle of strabismus came from the more deprived areas than in 1983 but this was not significant ($\chi^2_4 = 4.51$, $P=0.34$).

8.5 REGRESSION ANALYSES

8.5.1 MICROTROPIA OR NO STRABISMUS

A Normal errors regression was performed on the data for the 162 children with a microtropia or no strabismus to look at the factors affecting age at presentation. An analysis of variance for the models fitted can be seen in table 8.3. This shows the only factors to affect age at presentation are the year of the cohort and the child's Townsend deprivation score. There is an interaction between deprivation and year which approaches formal significance which indicates a difference in the effect of deprivation between the two cohorts.

Figures 8.1: Plot of Townsend deprivation score by age at presentation for children with microtropia or no strabismus presenting in 1983



Figures 8.2: Plot of Townsend deprivation score by age at presentation for children with microtropia or no strabismus presenting in 1992.

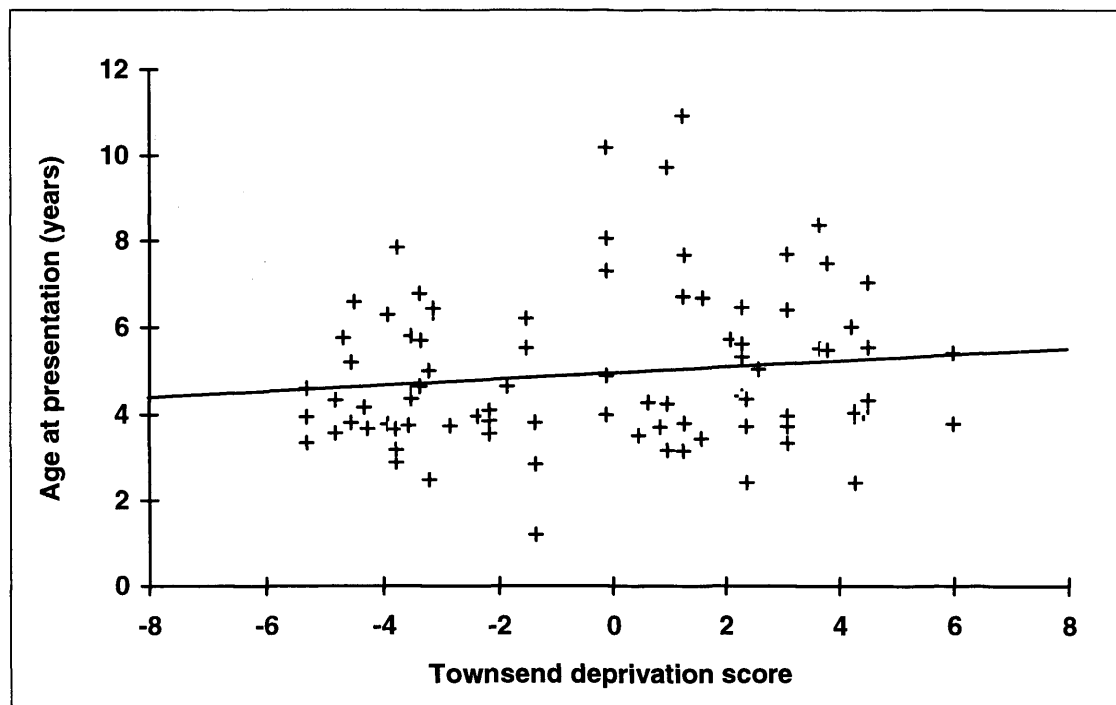


Table 8.3: Analysis of variance table of regression model for microtropia or no strabismus

Model terms fitted	Previous model	F statistic for change in fit	P value
YEAR	NULL	34.2 _{1,160}	0.000
TOWNSEND	NULL	15.3 _{1,160}	0.000
ANISO	NULL	0.00 _{1,160}	0.961
ASIAN	NULL	0.91 _{1,160}	0.341
YEAR + TOWNSEND	YEAR	12.1 _{1,159}	0.001
YEAR + TOWNSEND + ANISO	YEAR + TOWNSEND	0.66 _{1,158}	0.418
YEAR + TOWNSEND + ASIAN	YEAR + TOWNSEND	0.31 _{1,158}	0.578
YEAR + TOWNSEND + YEAR.TOWNSEND	YEAR + TOWNSEND	3.35₁₁₅₈	0.069
YEAR + TOWNSEND + YEAR.TOWNSEND + ANISO	YEAR + TOWNSEND + YEAR.TOWNSEND	0.39 _{1,157}	0.535
YEAR + TOWNSEND + YEAR.TOWNSEND + ASIAN	YEAR + TOWNSEND + YEAR.TOWNSEND	0.34 _{1,157}	0.561

Key to parameters in regression models

TOWNSEND: Townsend deprivation score

ANISO: >1 dioptre of anisometropia

ASIAN: Asian surname or forename

YEAR: Year of cohort

Table 8.4: Analysis of variance table of regression model for large angle strabismus

Model terms fitted	Previous model	F statistic for change in fit	P value
YEAR	NULL	0.00 _{1,247}	0.956
TOWNSEND	NULL	0.03 _{1,247}	0.859
<i>ANISO</i>	<i>NULL</i>	6.92 _{1,247}	0.009
ASIAN	NULL	0.16 _{1,247}	0.688
ANISO + YEAR	ANISO	0.00 _{1,246}	0.994
ANISO + TOWNSEND	ANISO	0.18 _{1,246}	0.674
ANISO + ASIAN	ANISO	0.15 _{1,246}	0.703
ANISO + YEAR + TOWNSEND + YEAR.TOWNSEND	ANISO + YEAR	1.16 _{1,244}	0.316

Key to parameters in regression models

TOWNSEND: Townsend deprivation score

ANISO: >1 dioptre of anisometropia

ASIAN: Asian surname or forename

YEAR: Year of cohort

Formal model checking indicated that there was greater variation in the residuals in 1992 than in 1983. Homogeneity of variance is one of the standard requirements of normal regression analysis. The failure of this basic assumption means that heterogeneity of variance needs to be allowed for explicitly in the analysis. A model which allowed different variances for each year was fitted in GLIM using a method described by Aitkin (1987). Since performing this analysis the statistical package SAS (1992) has introduced PROC MIXED which avoids all of the extra programming that is needed for GLIM. The analysis using GLIM will be reported here.

8.5.1.1 Modelling Variance Heterogeneity in GLIM

For the normal regression model

$$y_i = \beta' x_i + e_i$$

with $e_i \sim N(0, \sigma^2)$, under homogeneity, Aitkin proposes the model for heterogeneity

$$\text{var}(e_i) = \sigma_i^2 = \exp(\lambda' z_i)$$

where z_i may contain some or all of the explanatory variables in x_i and other variables not included in x_i ; z_i is assumed to contain a constant 1. The log-linear form ensures that σ^2 remains positive. This model can be fitted by maximum likelihood methods and the mean regression parameter β can be estimated simultaneously. Here the parameter estimated in the log-linear model is year as it appears that there is a difference in variability between the two years. This method was applied to the data, firstly fitting a model with Normal errors assuming homogeneous variances.

Table 8.5 shows the parameter estimates and standard errors for this Normal errors regression model assuming homogeneous variance.

Table 8.5: Parameter estimates of Normal regression model fitted for children with a microtropia or no strabismus (assuming homogeneous variance)

Parameter	Parameter Estimate	Standard Error	t statistic	P value
Intercept	6.336	0.2076	30.5	0.0001
TOWNSEND	0.204	0.0555	3.67	0.0003
YEAR(92)	-1.373	0.2699	5.09	0.0001
YEAR(92)*TOWNSEND	-0.134	0.0733	1.83	0.0691

The squared residuals from this model were used to investigate the difference in variance between the two years. A log-linear model with gamma errors was fitted to the squared residuals. Fitting the year parameter looked for evidence of different variation between the two years. These fitted values were then used to adjust the original analysis to account for differences in the variation of age between the two cohorts. After several iterations the model converged and can be seen in table 8.6. This showed a difference between the two years, which although not formally significant was adjusted for in the analysis. The normal regression model was refitted with weights $1/(\text{fitted values of log-linear model})$ giving the model in table 8.7.

Table 8.6: Parameter estimates of log-linear model with Gamma errors

Parameter	Parameter Estimate	Standard Error	t statistic	P value
Intercept	0.7952	0.1690	4.71	0.0001
YEAR(92)	0.3047	0.2243	1.36	0.1758

Table 8.7: Parameter estimates of Normal regression model fitted for children with a microtropia or no strabismus (assuming heterogeneous variance)

Parameter	Parameter Estimate	Standard Error	t statistic	P value
Intercept	6.336	0.1870	33.9	0.0001
TOWNSEND	0.204	0.0499	4.09	0.0001
YEAR(92)	-1.373	0.2602	5.28	0.0001
YEAR(92)*TOWNSEND	-0.134	0.0708	1.89	0.0606

Comparing the model before and after allowing for variance heterogeneity, shows no change in the parameter estimates but a reduction in the standard errors resulting in smaller P-values. A reanalysis allowing for this difference in the variance did not lead to any change in the conclusions of the choice of model seen in table 8.3. This model was then checked using the same formal model checking procedures as discussed in chapter 6 and can be seen in figures 8.3-8.8.

Figures 8.3-8.8: Formal model checking procedures for children with microtropia or no strabismus

Figure 8.3: Histogram of Studentized Residuals

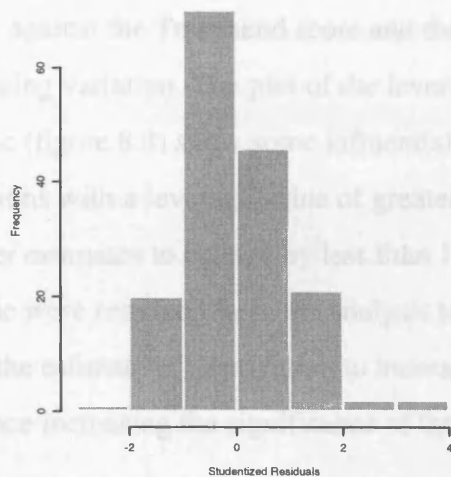


Figure 8.4: Normal Plot of Studentized Residuals

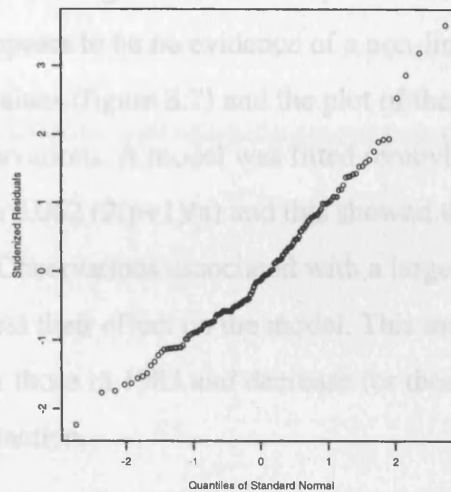


Figure 8.5: Predicted Values vs Studentized Residuals

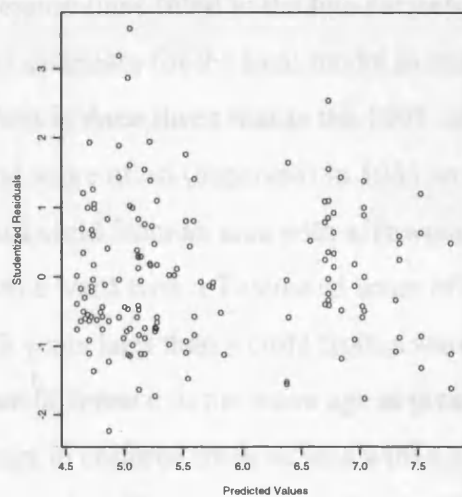


Figure 8.6: Plot of Studentized Residuals vs Townsend Index

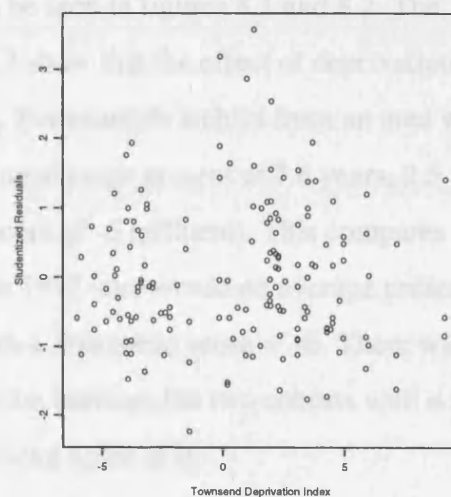


Figure 8.7: Plot of Leverage Values

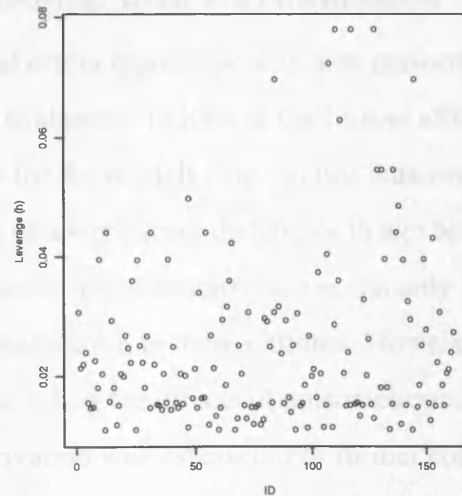


Figure 8.8: Plot of Cook's D Statistic

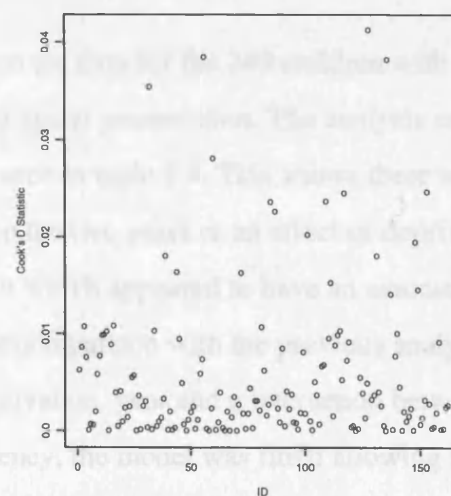


Figure 8.3 and 8.4 illustrate the distribution of the Studentised residuals for the final fitted model. There is no evidence of any distinct deviations from Normality. In order to look at the variation of the residuals figure 8.5 shows a plot of the residuals by the predicted values and there is no longer any evidence of systematic variation. Figure 8.6 shows a plot of the residuals against the Townsend score and there appears to be no evidence of a non-linear trend or increasing variation. The plot of the leverage values (figure 8.7) and the plot of the Cook's D statistic (figure 8.8) show some influential observations. A model was fitted removing observations with a leverage value of greater than $0.062 (2(p+1)/n)$ and this showed the parameter estimates to change by less than 10%. Observations associated with a large Cook's D statistic were removed from the analysis to assess their effect on the model. This model showed the estimate of deprivation to increase for those in 1983 and decrease for those in 1992 hence increasing the significance of the interaction.

The regression lines fitted to the two datasets can be seen in figures 8.1 and 8.2. The parameter estimates for the final model in table 8.7 show that the effect of deprivation in the 1983 cohort is three times that in the 1992 cohort. For example a child from an area with a Townsend score of +6 (deprived) in 1983 would on average present at 7.6 years, 2.5 years later than a child from an area with a Townsend score of -6 (affluent). This compares with a child from a ward with a Townsend score of +6 in 1992 who would on average present at 5.4 years, 0.9 years later than a child from a ward with a Townsend score of -6. There was also a significant difference in the mean age at presentation between the two cohorts with a reduction of 1.4 years in children from an area with a Townsend score of 0.

8.5.2 LARGE ANGLE STRABISMUS

A Normal errors regression was then performed on the data for the 249 children with a large angle of strabismus to look at the factors affecting age at presentation. The analysis of deviance for the models fitted to this data can be seen in table 8.4. This shows there was no evidence of a significant difference in age between the two years or an effect of deprivation. The presence of anisometropia was the only factor which appeared to have an association with age at presentation in these children. However for comparison with the previous analysis the model including the effects of anisometropia, deprivation, year and a interaction between year and deprivation was assessed. For further consistency, the model was fitted allowing for heterogeneous variances between the two cohorts. Using the GLIM method described

previously the model was initially fitted without weights and the parameter estimates and their associated standard errors can be seen in table 8.8

Table 8.8: Parameter estimates of Normal regression model fitted for children with a large angle strabismus (assuming homogeneous variance)

Parameter	Parameter Estimate	Standard Error	t statistic	P value
Intercept	1.758	2.217	0.793	0.4290
ANISO	0.6409	0.2381	2.692	0.0079
TOWNSEND	0.8425	0.5687	1.481	0.1406
YEAR(92)	0.0080	0.0254	0.315	0.7532
YEAR(92)*TOWNSEND	-0.0095	0.0065	1.462	0.1457

The squared residuals from this model were then used to investigate the difference in variance between the two years. The parameter estimates for the log-linear model with gamma errors for the squared residuals can be seen in table 8.9 showing some differences between the two cohorts. These fitted values were then used to adjust the original analysis to account for heterogeneous variation between the two cohorts and the final parameter estimates for the model can be seen in table 8.10.

Table 8.9: Log-linear model with Gamma errors to study difference in variation between cohorts

Parameter	Parameter Estimate	Standard Error	t statistic	P value
Intercept	0.9067	0.1200	7.556	0.0001
YEAR(92)	0.3294	0.1805	1.825	0.0700

Table 8.10: Parameter estimates of Normal regression model fitted for children with a large angle strabismus (assuming heterogeneous variance)

Parameter	Parameter Estimate	Standard Error	t statistic	P value
Intercept	1.814	2.226	0.815	0.4163
ANISO	0.5889	0.2330	2.527	0.0125
TOWNSEND	0.8386	0.5647	1.485	0.1395
YEAR(92)	0.0081	0.0257	0.315	0.7532
YEAR(92)*TOWNSEND	-0.0095	0.0065	1.462	0.1457

This shows a slight change in the parameter estimates of the model. Further model fitting allowing for other factors did not change the conclusions from the previous model fitting based on homogeneous variances.

This model was then checked using the same procedures as before (figures 8.9-8.14). Figures 8.9 and 8.10 show the distribution of the residuals for the final fitted model. There is no evidence of any distinct deviations from Normality. In order to look at the variation of the residuals figure 8.11 shows a plot of the residuals by the predicted values and there is no evidence of any systematic variation. There is no apparent pattern among the residuals with deprivation (figure 8.12). The plot of the leverage values (Figure 8.13) and the plot of the Cook's D statistic (Figure 8.14) show some influential observations. When influential leverage values ($2*(p+1)/n=0.048$) were weighted out of the regression there was no significant change in the parameter estimates. Similarly when the most influential D statistics were weighted out the conclusions of the analysis did not alter.

The parameter estimates for the final model in table 8.10 show that children with a significant amount of anisometropia present 6 months later than those without. This is consistent with the findings of the previous multicentre study and may imply a different onset pattern. There was no significant effect of deprivation on age at presentation and the mean age at presentation did not vary between the two cohorts.

8.6 DISCUSSION OF REGRESSION ANALYSIS

Between 1988 and 1991 the organisation of vision screening in Leicestershire for pre-school children has been radically changed with the introduction of secondary orthoptic screening, an emphasis on GPs managing screening and formal monitoring of children as part of the child health surveillance programme. Based on the analyses performed in this chapter it can be seen that over this period of change in screening, there has been a dramatic alteration in the referral pattern of children with amblyopia including an increase in cases of asymptomatic amblyopia, a decrease in cases of symptomatic amblyopia, a reduction (although not statistically significant) in the effect of deprivation on the age at presentation of asymptomatic amblyopia and a reduction in the mean age at presentation of asymptomatic amblyopia.

Figures 8.9-8.14: Formal model checking of model fitted for children with a large angle of strabismus

Figure 8.9: Histogram of Studentized Residuals

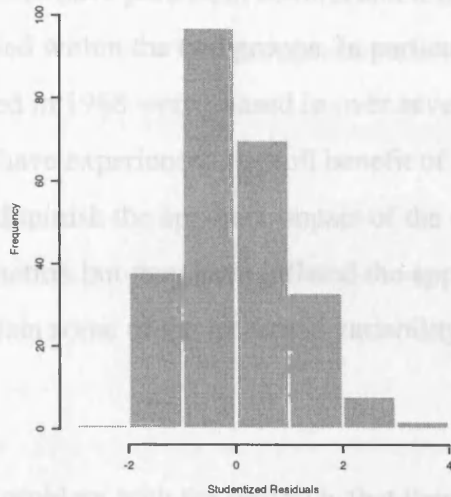


Figure 8.10: Normal Plot of Studentized Residuals

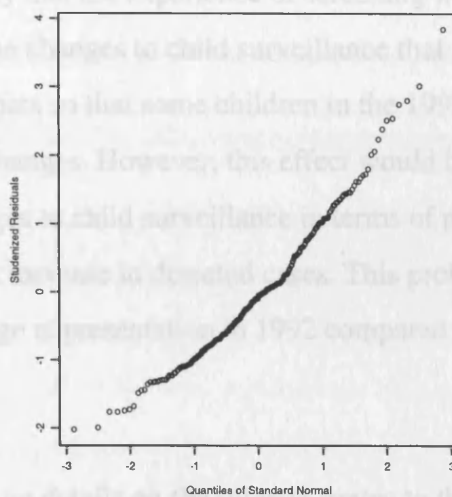


Figure 8.11: Predicted Values vs Studentized Residuals

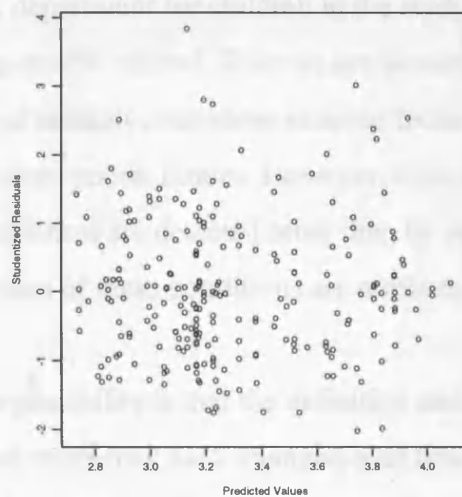


Figure 8.12: Plot of Studentized Residuals vs Townsend Index

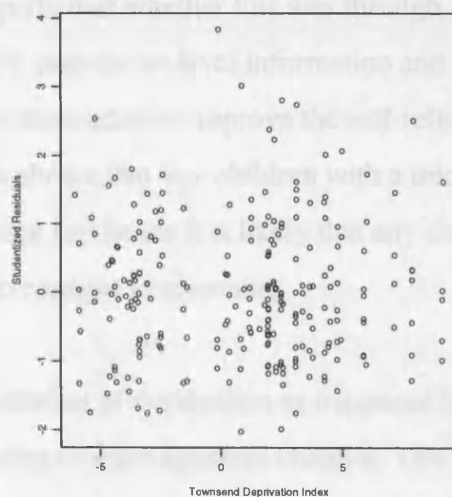


Figure 8.13: Plot of Leverage Values

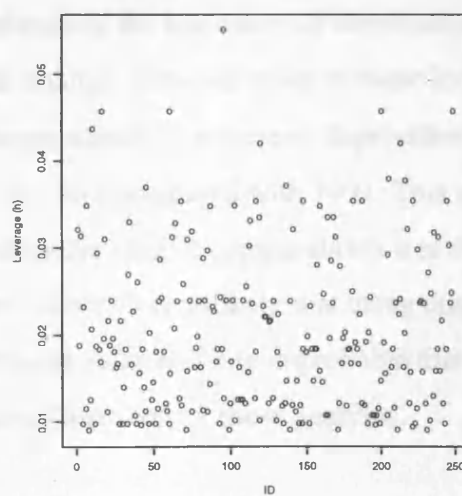
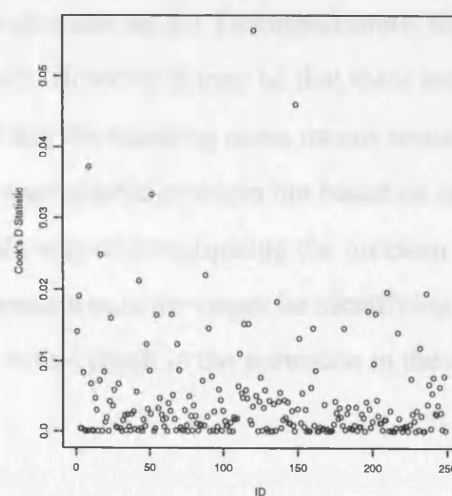


Figure 8.14: Plot of Cook's D Statistic



There are several possible explanations for these changes in the pattern of amblyopia presentation as well as the changes being attributable to the screening reforms. Firstly, they could be due to the design of the study. By taking children who started treatment in a given year I do not have pure birth cohorts and it is likely that the experience of screening would have varied within the two groups. In particular the changes to child surveillance that were introduced in 1988 were phased in over several years so that some children in the 1992 group may not have experienced the full benefit of the changes. However, this effect would have acted to diminish the apparent impact of the changes to child surveillance in terms of mean age at presentation but may have inflated the apparent increase in detected cases. This problem may explain some of the increased variability in age at presentation in 1992 compared with 1983.

Another problem with the design is that there are no details on the referral routes to the orthoptic department for children in the study, in particular whether this was through screening or self-referral. Thus we are working with population level information and it is possible, if unlikely, that some external factor may have acted to improve the self-referral of children from poorer homes. However, it has been shown that few children with a microtropia or no strabismus are detected other than by screening and hence it is likely that any changes to the detection of these conditions are attributable to changes in screening.

A further possibility is that the definition and distribution of deprivation as measured by the Townsend score may have changed over time leading to these apparent changes. This has been discussed in section 8.3.1 where it was illustrated that although there has been a decrease in the prevalence of the indicators of deprivation which make up the Townsend score, there has been little change in the variation of these indicators. However it may be that these indicators are no longer a useful measure of deprivation and that the resulting score means something different in 1981 compared with 1991. This is an unavoidable problem but based on current work and for the sake of comparability it is the only way of investigating the problem using a continuous score. It is unlikely that using this measure would no longer be identifying people from deprived areas and it is improbable that this would result in the reduction in the effect of deprivation illustrated by these analyses.

From this discussion of possible alternative explanations, I feel it is unlikely that they could explain the dramatic changes that have occurred in the presentation of amblyopia between 1983 and 1992. It appears that improved organisation of child health surveillance is the most likely reason for the changes in the pattern of presentation of amblyopia. Prior to 1988 health visitors were required to refer children suspected of a vision problem to their GP who could refer them on to an ophthalmologist. This process offered the opportunity for delay, drop-out and error. The current system whereby children are referred directly from primary screening to the secondary orthoptic screen reduces delay, cuts down the possibility of a child dropping out of the system and offers a trained assessment of the child's problem.

There were a third more cases of amblyopia associated with microtropia or no strabismus treated in 1992 compared to 1983, with no change in the size of the population. It is possible that in 1983 a smaller proportion of those detected with amblyopia were treated, or that there has been a rise in the incidence of amblyopia associated with microtropia or anisometropia but this is unlikely. Some of the increase may be due to the problems of study design as discussed above but it is probable that there has been an increase in detection with children being treated today who would previously have gone undetected.

There are no apparent changes in the age at presentation of children with large angle strabismus but there has been a reduction in the number of cases treated. It is possible that this is because some children with large angle strabismus are now detected before amblyopia has developed. If this is so screening may actually be lowering the incidence of amblyopia. By looking at the figures in each quintile of Townsend deprivation it appeared that those referred in 1992 were more likely to come from a deprived area but this was not formally significant.

Based on these data it can be estimated that about 40 children born in Leicestershire each year will now have amblyopia associated with microtropia or no strabismus detected that would previously have been missed. This is based on an increase of 23 cases in a 10 kilometre radius of Leicester in 1992 compared to 1983. Assuming the incidence of amblyopia to occur similarly in both groups, this figure of 23 new cases is based on a birth cohort of 7000. Relating these figures to the 12,000 births a year in Leicestershire gives 40 cases detected that would previously have been missed. In addition those children who are picked up at screening now present on average 19 months earlier than before. Finally there is the possibility that

screening may have reduced the number of children with strabismus going on to develop strabismic amblyopia. The effect of deprivation on age at presentation is still present although to a significantly lesser degree and further efforts need to be made to reduce this.

Evidently general NHS reforms have also had a part to play. In particular, financial incentives have been offered to GPs achieving high coverage rates for Child Health Surveillance. The effect of this is reflected in the changes in the percentage of the population of children in Leicestershire screened. This has increased at 3½ years from 80% in 1983 to 88% in 1992 and similarly those screened at 7½ months has increased from 91% in 1983 to 95% in 1992. This compares with a rate of just 57% at 3½ year screening in Glasgow (Williamson et al, 1995). However Reading, Colver et al (1994) have shown that improvements in the coverage of screening may not lead to reductions in inequalities. It is evident that in Leicestershire there are still differences in the uptake of screening between children from more deprived and less deprived areas despite these overall increases as supported by the findings of the National Child Development Study highlighted in chapter 2. The residual relationship between deprivation and age at presentation in 1992 may be linked to inequalities in child health surveillance coverage. Also Bowman et al (1996) showed that in a study of children referred by their general practitioner with suspected amblyopia or strabismus, that patients from deprived areas were less likely to attend their first ophthalmology hospital outpatient appointment than those from less deprived areas. Therefore there are still likely to be inequalities in presentation which were not detectable by this study because we have no information on children who were referred to the orthoptic department but never attend.

Although there are possible alternative theories to explain the changes in presentation of amblyopia, it is likely that changes to vision screening have had an impact and that one of the major effects is a reduction in the effect of deprivation albeit not a statistically significant change. This appears to be a rare example of a reduction in inequalities in health. I believe that it is the restructuring of vision screening with the introduction of a formal method of referral for health visitors that may have brought about this reduction in the effect of deprivation.

Bowman et al (1996) showed that longer waiting times for appointments were associated with poorer attendance, particularly among patients from more deprived areas. The waiting time for a orthoptic screening appointment is anecdotally much shorter than that for an appointment with the hospital consultant and this may lead to higher rates of attendance after initial referral.

The system no longer relies on parents taking the advice of the health visitor to see their GP but gives them a set appointment at the orthoptic screening clinic which may also lead to improved referral. Since one of the problems of this study is the fact that the changes were brought in over a long period of time, there is a need to look at children presenting in the orthoptic department who have all benefited from the new system to see if the patterns have improved further.

This study has pointed to the positive effects of using orthoptists in the vision screening process. Other studies have found primary orthoptic screening to show similar patterns of increased rates of detection of asymptomatic amblyopia (Jarvis et al, 1990; Edwards et al, 1989; MacLellan and Harker, 1979). However the use of orthoptists as secondary screeners is thought to be more cost effective and has been recommended by Hall (1989). This analysis supports his findings that secondary orthoptic screening improves the vision screening service but more work is necessary on the actual changes in the costs for introducing this service to assess the actual costs of identifying patients whose condition may previously have gone undetected.

8.7 SUMMARY AND CONCLUSIONS

In chapter 6, I showed through the analysis of data from the multicentre study, a relationship between deprivation and the age at presentation of asymptomatic amblyopia. This pointed to structural differences in the accessibility of services. In this chapter I have analysed data on children presenting with amblyopia before and after structural changes were made to vision screening in Leicestershire. This analysis has shown that there have been several major changes in the presentation of amblyopia during this period:

- 1) An increase in cases of asymptomatic amblyopia
- 2) A decrease in cases of symptomatic amblyopia
- 3) A reduction in the effect of deprivation on the age at presentation of asymptomatic amblyopia
- 4) A reduction in the mean age at presentation of asymptomatic amblyopia

Although these changes in the pattern of presentation of amblyopia may be attributable to factors such as study design or variations in exposure to screening, it is apparent that the

structural changes in the organisation of vision screening have had an impact. The new system seems to be more effective in terms of higher and earlier detection and may have had some effect on the relationship between age at presentation and deprivation.

One problem with this analysis highlighted earlier is that of ecological fallacy where an effect at the area level is assumed to exist at the individual level. Further it may be possible that the area level deprivation measure is relatively insensitive and that there still exists a relationship between age at presentation and deprivation when measured at the individual level. The rest of this thesis concentrates on looking at this problem of ecological fallacy and attempts to overcome it.

CHAPTER 9

ASSESSING THE PROBLEM OF ECOLOGICAL FALLACY

9.1 AIMS OF THE CHAPTER

One of the issues highlighted in the analysis of the amblyopia studies is that the measure of deprivation is based on area level data. These proxies may be leading to biased estimates of the effect of deprivation on age at presentation when compared to that which would be observed using data at an individual level. This problem is known as the ecological fallacy and was highlighted in section 3.4.5.

Individuals are not homogeneous within areas and aggregating data assumes homogeneity. Sloggett and Joshi (1994) have discussed the effects of individual indicators of deprivation and an area based deprivation score on morbidity recorded in the OPCS longitudinal study. They found that there was no relationship between area deprivation and mortality in males after adjusting for individual indicators of deprivation. Ecob (1996) recently found that when using a multi-level modelling approach to analysing individual and area deprivation data that for many health outcomes there were no area level differences after adjusting for differences between individuals. Research using a multi-dimensional area measure (Shouls et al, 1996) has shown again using a multi-level modelling approach that most of the variation in limiting long term illness could be explained by individual level deprivation factors but that there were some remaining contextual effects.

In the 1992 study of amblyopia, individual level data on the indicators of deprivation used in the Townsend score were collected to provide the opportunity to investigate the relationship between area and individual level deprivation. However since this dataset is relatively small, in this chapter I use two larger datasets to illustrate the possible effects of using area level data and then apply these methods to the amblyopia study data in chapter 10 to assess the issue of ecological fallacy in my study. The aim of this work is not to replicate the work of those such as Sloggett and Joshi (1994) as discussed above but to assess whether area level measures are

underestimating the effects of deprivation at the individual level (this chapter and chapter 10) and to then investigate whether information about the relationship between area and individual level deprivation can be used to adjust the effect of deprivation in studies where only area level information is available (chapters 11 and 12).

In this chapter I firstly look at a large national dataset from the 1991 census comparing the relationship between morbidity as assessed by limiting long-term illness and deprivation at both the individual and area level. I then use a local dataset from a study of perinatal mortality in Leicestershire to investigate the relationship between various aspects of child health and deprivation at the individual and area level.

9.2 SAMPLES OF ANONYMISED CENSUS RECORDS

To compare area and individual deprivation a large dataset was required with information available at the individual level. Morbidity data were also needed to see how the apparent extent of inequalities in health differs when measured at the area and individual level. The 1991 census data at area level does not provide the opportunity to crosstabulate the four variables of the Townsend score. Therefore it is not possible to assess the relationship between them and understand how many households possess none, one, two, three or all four of the characteristics. However as well as producing data at area levels, the census provides samples of anonymised records (SAR) which provide individual level data on a 1 % sample of individuals and a 2 % sample of households. To retain confidentiality of the data, these samples are anonymised. The district of residence is recorded for the individual sample while the county of residence is recorded for the household sample. This dataset provides the possibility of investigating area and individual level data although the geographical size of the area recorded on the SAR records is larger than ideal. With knowledge about the relationship of deprivation scores between districts, EDs and wards it is possible to draw up some informed hypotheses from these data.

To assess whether the estimate of the association between deprivation and morbidity was biased by area level data, a measure of health was needed. Although mortality statistics were available at an area level using alternative datasets, the introduction of a question on limiting

long-term illness to the 1991 census meant that the SAR data included individual health data that could be related directly to individual level deprivation measures. The census asks

'Does the person have any long-term illness, health problem or handicap which limits his/her daily activities or the work he/she can do? Include problems which are due to old age.'

The response to this is subjective and may lead to great variation in illness reported. Using data from the whole country should eradicate this problem unless there is systematic bias in reporting, i.e. that deprived people were more likely to report the same condition as limiting than less deprived people.

Since individual SAR data were only available at district level it was decided to collect SAR information for the whole of England rather than just Leicestershire and to compare individual and district level deprivation scores with limiting long-term illness. Multivariate analysis could then be used to look at the relationship for different age groups and gender.

This analysis attempts to understand the relationship between area level and individual level data and morbidity. Area level data may be biasing the effect of deprivation. The effect of deprivation on limiting long-term illness is assessed by looking at area and individual data. Secondly it is probable that there is a more complex relationship between area and individual deprivation than has previously been assumed. For example, there may be differences in the health experiences of people living in the same area but with different individual levels of deprivation. A very deprived person living in a less deprived area and similarly a less deprived person living in a more deprived area may have very different health experiences to other people in their area. These effects tend to change with age since the distribution of deprivation indicators varies with age with retired people less likely to own a car and families being more likely to own a car, or own a house, etc. These factors also are not necessarily similar for both sexes with many women having employment breaks. Hence different age and gender strata need to be investigated and the effect of area data will be assessed after adjusting for individual data.

9.2.1 DESCRIPTION OF THE DATA

Information was extracted from the 1991 census samples of anonymised records on males and females in five age bands, 16-24, 25-34, 35-44, 45-54, 55-64 years. It was decided to only look at males and females between the ages of 16 and 64 years since there are many problems associated with data on children, and retired people. It was decided to only include economically active people as percentage unemployment in the Townsend score is based only on economically active individuals. Economically inactive individuals maybe inactive due to illness and counting them as **not** 'unemployed' could bias the results.

To relate individual level data to the previous analyses using the Townsend deprivation score the individual data on the four components of the Townsend score: car ownership, house ownership, overcrowding and unemployment, were used. These data were then combined with deprivation indicators at the district level extracted using SASPAC. Unemployment was based on the employment of the nominated head of the household. This is slightly different from the definition used in the Townsend score but was due to limitations of the data.

In order to construct an individual deprivation score for each person, it was decided to use the same construction as for the previous work with area level data, using an adaptation of the Townsend score. Therefore the mean and standard deviation for each variable were calculated for the national district level data. Then Normal scores were calculated for each variable after log transforming unemployment and overcrowding, and then summing them to an overall score. This was done at both district and individual level. Presence of one of the aspects of deprivation at an individual level was interpreted as 100% prevalence for that factor, while absence was recorded as 0%.

The individual deprivation scores were then ranked and divided into groups for tabulation and analysis purposes. When grouping area level data, as was done in chapters 6 and 8, it is possible to use even sized groups such as quintiles or deciles. However this was not possible for the individual level deprivation scores since such a high percentage of people had no characteristics of deprivation. Therefore the individual deprivation scores were grouped according to the number of deprivation characteristics present (i.e. none, 1, 2, 3 or 4). This did not affect the rankings of the scores since scores relating to the same number of deprivation characteristics had very similar rankings. This is extremely similar to the individual deprivation

score used by Shouls et al (1996). The district level deprivation scores were also ranked and grouped into five ordered groups corresponding to the same proportions as those people with none, one, two, three or four of the deprivation characteristics at an individual level. These groupings were used rather than quintiles since it provided the opportunity to compare the same percentage of people classified as most deprived at the individual level with those classified as most deprived at the area level. Evidently this means that a lot of information is being lost since so many people are being grouped together as least deprived. However although the primary analyses are based on these groupings, the data are also analysed using district level deprivation as a continuous variable and grouped into quintiles.

9.2.2 COMPARISON OF INDIVIDUAL AND AREA DATA

Table 9.1 shows the distribution of individual deprivation scores for males and females aged 16-64 years.

Table 9.1: Percentage of people with different numbers of individual deprivation characteristics by sex and age

NUMBER OF DEPRIVATION FACTORS	MALES Age at 1991 census					FEMALES Age at 1991 census				
	16-24	25-34	35-44	45-54	55-64	16-24	25-34	35-44	45-54	55-64
0	54%	61%	70%	72%	65%	56%	65%	73%	72%	64%
1	28%	25%	20%	20%	24%	27%	23%	19%	20%	25%
2	14%	11%	8%	7%	9%	13%	9%	7%	7%	11%
3	3%	2%	2%	1%	2%	3%	2%	1%	1%	1%
4	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
ALL	45632	68853	62830	50443	32906	38116	49036	48702	38814	18485

The zeros in this table are due to rounding errors. It can be seen that the majority of people are in the least deprived category. The distribution of deprivation characteristics appears to change with age with the prevalence of these characteristics decreasing with age as would be expected. There is a rise in the prevalence of deprivation in the highest age group which could relate to many of the more affluent individuals in this age group being classed as retired and economically inactive and hence not included in the analysis while those who have no pension other than state allowances may be classed as unemployed. Males and females show very

similar patterns. This would be expected since the deprivation score is based on the household of the individuals and unemployment is based on the head of the household. Therefore for a couple in a household there would be no difference in the individual deprivation score.

Table 9.2 shows a cross tabulation of deprivation based on individual and district data. The apparent lack of people with all four characteristics is due to rounding errors since only 0.2% of people are in this group. There are only 54% of individuals lying on the main diagonal of the table, which represents perfect agreement. The table shows that of individuals with all four characteristics of deprivation over 30% are classified as least deprived by their area score (285/911). This percentage rises for those with two or three deprivation characteristics.

Table 9.2: Number of people with different levels of area deprivation for each level of individual deprivation (0=least deprived group, 4=most deprived group).

		Area deprivation					Total
		0	1	2	3	4	
Individual deprivation	0	212180	63765	19834	2111	276	298166
	1	63166	24579	13012	2486	407	103650
	2	18892	12938	8416	2217	398	42861
	3	2999	2779	1845	480	126	8229
	4	285	254	241	78	53	911
Total		297522	104315	43348	7372	1260	453817

In order to assess the proportional agreement adjusted for chance, the kappa statistic was used (see Chapter 4). The simple kappa statistic for table 9.2 was 0.096 while the weighted kappa statistic was 0.159. This shows extremely poor agreement between the area and individual level deprivation measures which is hardly better than chance. This was similarly poor for all age and sex strata.

This high degree of misclassification when looking at individual and district level data could be partially due to the fact that districts are very large areas and that deprivation based on ward level data would show smaller differences. A table of district versus ward deprivation was drawn up based on all economically active individuals using the 1991 census small area statistics (table 9.3). The district and ward deprivation scores were ordered and divided into

the same proportions as the individual deprivation scores, 65.7% (least deprived), 22.8%, 9.4%, 1.8% and 0.2% (most deprived).

Table 9.3: Number of economically active individuals with each level of district deprivation and each level of ward deprivation

District deprivation	Ward deprivation					ALL
	0	1	2	3	4	
0	7079136	1278707	163026	3120	0	8523989
1	1246025	1215350	465031	46418	0	2972824
2	196194	449987	402876	111911	28161	1189129
3	2537	32275	146983	86285	5833	273913
4	0	0	7927	26352	6566	40845
ALL	8523892	2976319	1185843	274086	40560	13000700

Table 9.3 compares district and ward deprivation scores. In this case 68% of individuals lie on the diagonal and are predicted the same level of deprivation by both methods compared to 54% when comparing district and individual data. Ninety-six per cent are predicted the same group or the next ranked group. The simple kappa statistic for this table is 0.364, with a weighted kappa of 0.491 showing moderate agreement between district and ward level deprivation. Using district deprivation is likely to reduce the degree of agreement seen between ward level and individual data. However, there is less misclassification at this level than when looking at district and individual data. The relationship between ward level and individual level data is unlikely to be much improved in comparison to the relationship between district and individual level data. This problem is further investigated in section 9.3 when looking at the data from the perinatal mortality study.

9.2.3 DEPRIVATION AND LIMITING LONG TERM ILLNESS

The comparison of district level and individual deprivation has shown a high degree of misclassification. The next step was to assess the degree to which this would affect the observed relationship between deprivation and limiting long-term illness. Table 9.4 shows the prevalence of limiting long-term illness as recorded in the census for females and males from districts with different levels of deprivation.

Table 9.4: Percentage of reported limiting long-term illness by district deprivation (0=least deprived, 4=most deprived), age and sex

District deprivation		Age group				
MALES	16-24	25-34	35-44	45-54	55-64	Total (N)
0	2.4%	2.5%	3.4%	5.0%	9.1%	171395
1	2.7%	2.9%	4.2%	6.3%	10.7%	59920
2	2.8%	3.0%	4.3%	7.0%	11.3%	24572
3	2.0%	4.1%	4.7%	7.9%	10.5%	4016
4	2.2%	3.1%	3.8%	10.5%	9.9%	761
Total	45632	68853	62830	50443	32906	260664
FEMALES						
0	2.1%	2.1%	2.9%	4.7%	7.0%	126127
1	2.1%	2.3%	3.5%	5.4%	9.2%	44395
2	2.2%	3.0%	3.4%	6.2%	8.9%	18776
3	3.0%	3.5%	5.2%	8.7%	9.7%	3356
4	2.1%	2.0%	2.4%	8.2%	7.8%	499
Total	38116	49036	48702	38814	18485	193153

As would be expected the prevalence of long-term illness increases with age. This table also shows that in general the percentage of reported long-term illness increases with increasing area deprivation, but this trend is much greater as age increases. The deprivation groups 3 and 4 contain very few individuals and the estimates for these groups fluctuate because of random error. There are much greater differences among the 55-65 year age group compared to the 16-25 group. Females generally appear to have lower rates of reported illness. This may be because women who have a long-term illness may be more likely to define themselves as 'looking after the house' in the census and then classed as economically inactive and not included in this dataset. Men in this situation may be more likely to define themselves as unemployed and therefore economically active and included in the dataset. However when figures for England on limiting long-term illness were obtained it showed that for all females (economically active and inactive) the rate of disease prevalence is 1% lower at 8.1% than males 9.0%. Therefore it seems that females generally report less illness.

Prevalence of limiting long-term illness was then tabulated by individual deprivation in table 9.5 to see if there was any difference in the rates compared to area deprivation.

Table 9.5: Percentage of reported long-term illness by individual deprivation, age and sex

Individual deprivation		Age group					
MALES		16-24	25-34	35-44	45-54	55-64	Total
	0	1.9%	1.9%	2.6%	4.2%	7.9%	169130
	1	2.8%	3.4%	5.2%	7.3%	11.8%	60597
	2	3.7%	5.3%	7.7%	11.9%	15.2%	25087
	3	4.4%	5.2%	9.0%	14.6%	18.2%	5226
	4	5.2%	7.9%	8.1%	11.8%	14.7%	624
	Total	45632	68853	62830	50443	32906	260664
FEMALES		16-24	25-34	35-44	45-54	55-64	Total
	0	1.6%	1.7%	2.4%	4.0%	6.3%	129036
	1	2.3%	2.7%	4.0%	6.7%	8.8%	43053
	2	3.3%	4.6%	6.5%	10.1%	12.6%	17774
	3	4.2%	5.1%	8.2%	13.1%	19.1%	3003
	4	3.2%	4.7%	3.8%	0.0%	10.0%	287
	Total	38116	49036	48702	38814	18485	193153

It can be seen that again there is evidence of increased reporting of long-term limiting illness with age and also a trend of increased reporting of long-term illness with increased deprivation. However in this case there are greater differences between those with the most deprived characteristics and those with the least deprived. Those with no individual deprivation characteristics have a lower rate of illness than those classified as being in the least deprived districts, but for individuals with more deprivation characteristics their rate of illness is much higher than those in the deprived districts. This effect is best shown graphically. For this purpose the individuals were divided into not deprived and deprived based on both their individual and district deprivation. The rate of illness in group 0 was compared to that in groups 1, 2, 3 and 4 combined for both individual and district deprivation to give similar sized groups and avoid the problem of the very small deprivation groups (Figures 9.1 and 9.2).

Figure 9.1: Percentage of economically active males reporting limiting long-term illness by level of area deprivation

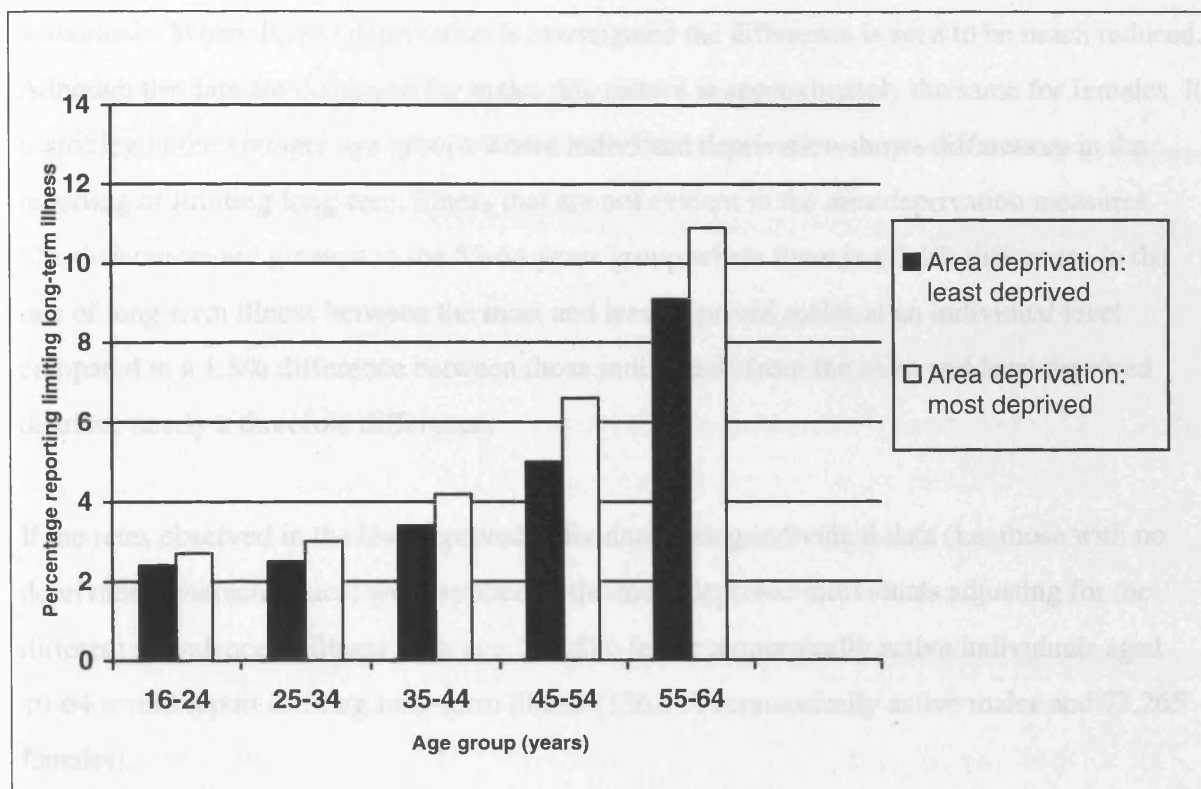
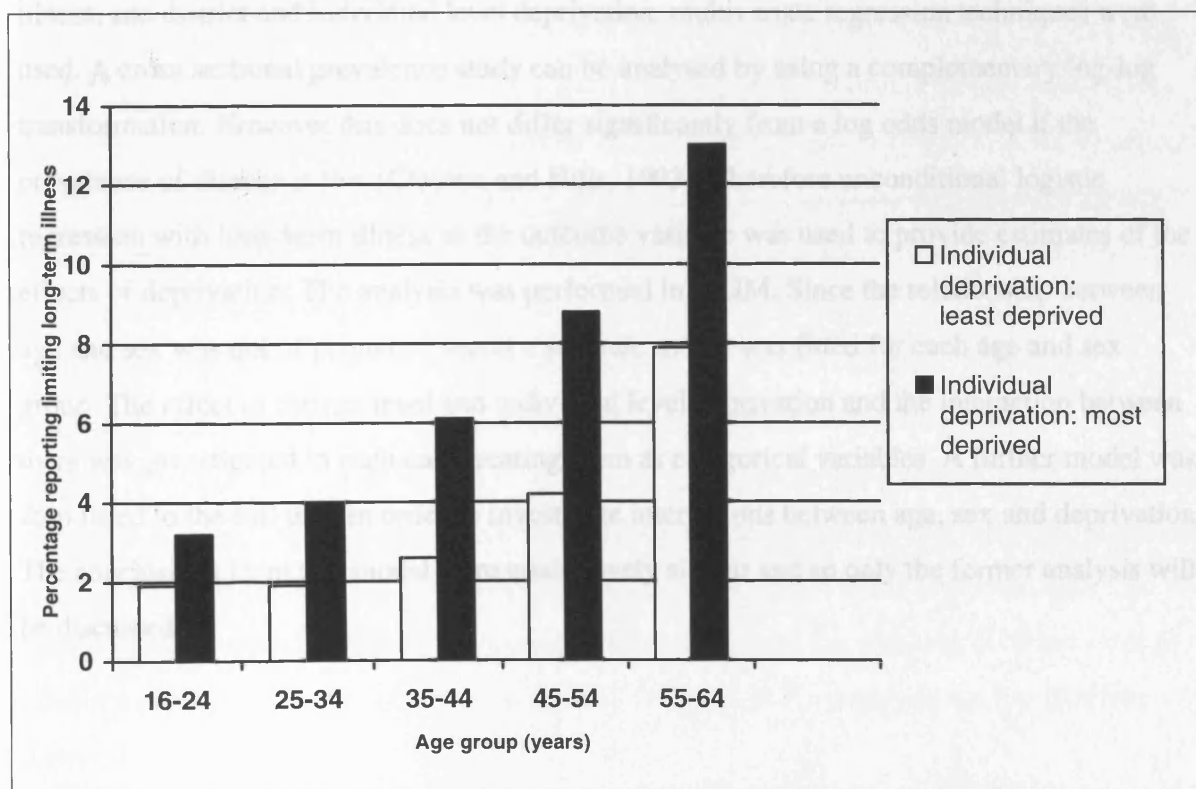


Figure 9.2: Percentage of economically active males reporting limiting long-term illness by level of individual deprivation



Figures 9.1 and 9.2 show for males that when deprivation is measured at an individual level there are large differences in the rates of reported illness between the least and most deprived individuals. When district deprivation is investigated the difference is seen to be much reduced. Although the data are displayed for males this pattern is approximately the same for females. It is striking in the younger age groups where individual deprivation shows differences in the reporting of limiting long-term illness that are not evident in the area deprivation measures. The differences are greatest in the 55-64 years group where there is a 5.1% difference in the rate of long-term illness between the most and least deprived males at an individual level compared to a 1.8% difference between those individuals from the most and least deprived districts, nearly a threefold difference.

If the rates observed in the less deprived individuals using individual data (i.e. those with no deprivation characteristics) were applied to the more deprived individuals adjusting for the different prevalence of illness with age, 209,586 fewer economically active individuals aged 16-64 would report limiting long-term illness (136,321 economically active males and 73,265 females).

9.2.4 REGRESSION ANALYSES

In order to investigate these observed relationships between reported limiting long-term illness, and district and individual level deprivation, multivariate regression techniques were used. A cross sectional prevalence study can be analysed by using a complementary log-log transformation. However this does not differ significantly from a log odds model if the prevalence of disease is low (Clayton and Hills, 1993). Therefore unconditional logistic regression with long-term illness as the outcome variable was used to provide estimates of the effects of deprivation. The analysis was performed in GLIM. Since the relationship between age and sex was not of primary interest a separate model was fitted for each age and sex group. The effect of district level and individual level deprivation and the interaction between them was investigated in each case treating them as categorical variables. A further model was then fitted to the full data in order to investigate interactions between age, sex and deprivation. The conclusions from this model were qualitatively similar and so only the former analysis will be discussed.

A major problem with the analysis was deciding how to assess the significance of explanatory variables in the model, particularly complex interactions. In logistic regression, changes in deviance of the model when introducing a new term are related to the χ^2 distribution. Since the full dataset was so large with over 450,000 observations, extremely small clinically non-significant effects would be found to be statistically significant if using an hypothesis testing approach with an arbitrary cut off point of $P=0.05$. Furthermore, there are problems with repeated significance testing since the analysis is being split into 10 strata which increases the chance of a type I error, i.e. the probability of obtaining a significant result and rejecting the null hypothesis when the null hypothesis is true. One method of adjusting for repeated significance testing would be to apply a Bonferroni correction which would mean testing at the $0.05/10 = 0.005$ level (Altman, 1991) although this is thought to be conservative in large sample studies. It was therefore decided to base inferences on changes in the parameter estimates as well as observing changes in deviance in the model. The change in the parameter estimates with and without the term of interest was used as a method for deciding the importance of effects. For each age and sex stratum, a model was fitted with the area and individual level deprivation effects. Tables 9.6 and 9.7 show the parameter estimates for the models including both area and individual deprivation. The changes in deviance to the model after including these effects are also tabulated.

In males and females aged 25-64 the effect of area deprivation before adjusting for individual deprivation is significant at the 5% level even after applying the repeated testing adjustment. There is no evidence of a significant effect in 16-24 year old males or females. After adjusting for individual deprivation in the models, area deprivation becomes non-significant at the 5% level in all cases apart from 55-64 year old females. In this case the P value was 0.02 and the only significant odds ratio is for females from area deprivation group 1. This implies that females from these areas are 25% more likely to be long-term ill compared to those in deprivation areas 0 no matter what their individual characteristics are. However this effect must be investigated since there are the inherent problems of multiple testing. If the Bonferroni correction was applied here, this effect would not be formally significant. Further, on removal of the area effect, none of the estimates of individual deprivation change by more than 5%. Individual deprivation was shown to explain a large amount of the variation in illness rates in all of the models. This effect appears to increase with age, but is relatively similar between males and females.

Table 9.6: MALES: Parameter estimates with standard errors for model including individual and area deprivation for each age stratum

Age		16-24		25-34		35-44		45-54		55-64	
		Parameter estimate	Standard error	Parameter estimate	Standard error	Parameter estimate	Standard error	Parameter estimate	Standard error	Parameter estimate	Standard error
	Intercept	-3.94	0.050	-3.95	0.039	-3.64	0.033	-3.15	0.029	-2.47	0.028
Individual deprivation	0	-	-	-	-	-	-	-	-	-	-
	1	0.378	0.071	0.594	0.055	0.726	0.050	0.570	0.047	0.437	0.043
	2	0.686	0.083	1.066	0.065	1.134	0.063	1.096	0.060	0.732	0.058
	3	0.880	0.141	1.044	0.118	1.312	0.115	1.332	0.115	0.942	0.113
	4	1.065	0.367	1.499	0.264	1.207	0.304	1.042	0.347	0.760	0.489
Area deprivation	0	-	-	-	-	-	-	-	-	-	-
	1	0.021	0.071	0.028	0.056	0.094	0.051	0.115	0.047	0.085	0.046
	2	-0.015	0.099	-0.063	0.076	-0.027	0.074	0.078	0.068	0.060	0.063
	3	-0.460	0.273	0.125	0.140	-0.122	0.167	0.032	0.158	-0.164	0.151
	4	-0.439	0.578	-0.267	0.387	-0.458	0.418	0.163	0.317	-0.320	0.308
Change in deviance for area deprivation not adjusted for individual deprivation		$\chi^2_4=5.08$ P=0.280		$\chi^2_4=19.2$ P<0.001		$\chi^2_4=27.7$ P<0.001		$\chi^2_4=57.3$ P<0.001		$\chi^2_4=25.74$ P<0.001	
Change in deviance for area deprivation adjusted for individual deprivation		$\chi^2_4=4.10$ P=0.393		$\chi^2_4=2.57$ P=0.632		$\chi^2_4=6.30$ P=0.178		$\chi^2_4=6.32$ P=0.176		$\chi^2_4=6.94$ P=0.139	
Change in deviance for individual deprivation adjusted for area deprivation		$\chi^2_4=92.1$ P<0.001		$\chi^2_4=320.8$ P<0.001		$\chi^2_4=444.3$ P<0.001		$\chi^2_4=421.2$ P<0.001		$\chi^2_4=230.1$ P<0.001	

Table 9.7: FEMALES: Parameter estimates with standard errors for model including individual and area deprivation for each age stratum

Age		16-24		25-34		35-44		45-54		55-64	
		Parameter estimate	Standard error	Parameter estimate	Standard error	Parameter estimate	Standard error	Parameter estimate	Standard error	Parameter estimate	Standard error
	Intercept	-4.08	0.059	-4.09	0.049	-3.71	0.038	-3.21	0.034	-2.75	0.042
Individual deprivation	0	-	-	-	-	-	-	-	-	-	-
	1	0.382	0.085	0.484	0.072	0.511	0.064	0.548	0.056	0.349	0.065
	2	0.753	0.098	1.001	0.086	1.019	0.080	0.971	0.072	0.730	0.080
	3	1.033	0.162	1.118	0.161	1.257	0.161	1.255	0.169	1.220	0.223
	4	0.752	0.503	0.998	0.511	0.436	0.714	-2.476	2.668	0.492	1.057
Area deprivation	0	-	-	-	-	-	-	-	-	-	-
	1	-0.104	0.087	0.024	0.074	0.119	0.062	0.059	0.057	0.226	0.067
	2	-0.148	0.117	0.133	0.091	-0.066	0.095	0.078	0.081	0.093	0.093
	3	0.050	0.234	0.169	0.165	0.219	0.180	0.268	0.166	0.074	0.198
	4	-0.307	0.579	-0.434	0.581	-0.561	0.714	0.147	0.430	-0.194	0.523
Change in deviance for area deprivation not adjusted for individual deprivation		$\chi^2_4=2.78$ P=0.595		$\chi^2_4=24.1$ P<0.001		$\chi^2_4=20.44$ P<0.001		$\chi^2_4=32.08$ P<0.001		$\chi^2_4=25.64$ P<0.001	
Change in deviance for area deprivation adjusted for individual deprivation		$\chi^2_4=2.87$ P=0.580		$\chi^2_4=3.50$ P=0.478		$\chi^2_4=6.78$ P=0.148		$\chi^2_4=3.71$ P=0.447		$\chi^2_4=11.55$ P=0.021	
Change in deviance for individual deprivation adjusted for area deprivation		$\chi^2_4=78.1$ P<0.001		$\chi^2_4=154.5$ P<0.001		$\chi^2_4=198.9$ P<0.001		$\chi^2_4=235.3$ P<0.001		$\chi^2_4=102.1$ P<0.001	

9.2.5 SENSITIVITY ANALYSES

In each model, the interaction between area and individual deprivation was investigated. In three of the analyses the interactions are formally significant testing at the 5% level with no adjustment for repeated testing. The biggest change in deviance was seen in females aged 16-24 ($\chi^2_{16}=32.65$ $P=0.008$). To assess the effect of the interaction on the parameter estimates, the most significant parameter estimate of the interaction was studied. This was judged by calculating the ratio between the estimates and their standard errors and choosing the largest. This was shown to be an interaction between group 3 of individual deprivation and group 3 of area deprivation. The odds ratio for having these characteristics was then calculated for the model with and without the interaction term. This gave an odds ratio of 1.94 for the model with the interaction term and 1.83 for the model without. This represents a change in the estimate of the odds ratio of just 6% which was deemed not to be clinically significant. Since this was the most significant parameter estimate it was decided that the interaction terms could be justifiably removed from all of the 10 models.

The analysis had not taken account of the ordered nature of the deprivation data. Therefore for each model the linearity of the individual deprivation term was investigated by assuming the groups to be equally spaced on a linear scale. The change in deviance between the models with the categorical variable fitted and with the linear variable can be assessed with respect to the change in degrees of freedom. For example, in males aged 16-24 the change in deviance for deprivation as a five level factor was 93 on 4 degrees of freedom, while the change in deviance when fitting the linear effect was 91 on 1 degree of freedom. The difference in deviance of 2 can be related to the chi-squared distribution on $4-1=3$ degrees of freedom. Since this is non-significant then it can be assumed that most of the observed variation between the groups can be attributed to a linear trend. However in over half of the models there was significant variation not associated with a linear trend. In most cases the trend was linear for deprivation groups 0, 1, 2, and 3 but was very variable for group 4. For each stratum the model using categorical data will be interpreted. A re-analysis using the midpoints of the corresponding deprivation score for each group, i.e. -15, -5, 5, 15, 30 did not change any of the conclusions of the previous model.

One of the drawbacks to the analysis is that in studies using area deprivation scores, the groupings used in this analysis would not be used and information may be lost here because deprivation can be measured on a continuous scale. The analyses were repeated using area deprivation as a continuous score based on the original scores before grouping, and also grouped into quintiles. This made no difference to the effect of area deprivation after adjusting for individual deprivation.

The full model fitted to all of the data confirmed the above conclusions, with no evidence of an area and individual deprivation interaction, and no effect of area deprivation after adjusting for individual deprivation. It also confirmed that there was a difference in the rate of illness between males and females with females having lower rates. Rates of illness also significantly rose with age and the relationship between individual deprivation and illness increased with age.

An extension to these analyses would be to take account of the clustering of individuals within areas since it would be expected that individuals from the same area would be more similar than individuals from different areas. This could be done using a multi-level modelling approach (Goldstein, 1995). However because this dataset was so large I did not use standard errors to assess the significance of terms in the logistic regression model and based interpretations on parameter estimates. It is unlikely that this alternative technique to allow for within area correlation would lead to clinically significant changes in the parameter estimates and there would have to be an extremely high degree of similarity between individuals within areas to increase the standard errors by a large enough amount to affect the conclusions of my analyses.

9.2.6 INTERPRETATION OF REGRESSION ANALYSIS

Tables 9.8 and 9.9 show for each age and sex stratum, the odds ratios for the model with just individual deprivation and for the model with just district deprivation so that the analyses can be compared in terms of the size of the effect observed.

Table 9.8 MALES Odds ratios with 95% confidence intervals for individual deprivation not adjusted for area deprivation, and for area deprivation not adjusted for individual deprivation

Age		16-24		25-34		35-44		45-54		55-64	
		Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval
Individual deprivation	0	1	-	1	-	1	-	1	-	1	-
	1	1.45	1.26-1.67	1.81	1.62-2.01	2.06	1.87-2.27	1.79	1.63-1.96	1.55	1.43-1.69
	2	1.95	1.67-2.29	2.90	2.56-3.28	3.10	2.75-3.50	3.08	2.75-3.45	2.09	1.87-2.34
	3	2.36	1.80-3.10	2.84	2.26-3.57	3.71	2.97-4.64	3.90	3.12-4.88	2.59	2.08-3.22
	4	2.80	1.41-5.95	4.47	2.67-7.50	3.27	1.81-5.91	3.04	1.57-5.89	2.01	0.78-5.19
Area deprivation	0	1	-	1	-	1	-	1	-	1	-
	1	1.12	0.97-1.28	1.15	1.03-1.28	1.25	1.13-1.38	1.28	1.17-1.40	1.20	1.10-1.31
	2	1.18	0.98-1.43	1.19	1.03-1.38	1.28	1.11-1.47	1.43	1.26-1.63	1.27	1.12-1.43
	3	0.85	0.50-1.45	1.65	1.26-2.15	1.40	1.02-1.94	1.63	1.21-2.22	1.17	0.88-1.57
	4	0.90	0.29-2.78	1.21	0.57-2.56	1.13	0.50-2.54	2.23	1.23-4.07	1.10	0.61-1.99

Table 9.9 FEMALES Odds ratios with 95% confidence intervals for individual deprivation not adjusted for area deprivation, and for area deprivation not adjusted for individual deprivation

Age		16-24		25-34		35-44		45-54		55-64	
		Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval
146	Individual deprivation	0	1 -	1 -	1 -	1 -	1 -	1 -	1 -	1 -	1 -
		1	1.45 1.23-1.71	1.64 1.43-1.89	1.67 1.47-1.89	1.75 1.57-1.95	1.44 1.27-1.63				
		2	2.06 1.71-2.49	2.80 2.38-3.30	2.80 2.38-3.26	2.72 2.37-3.12	2.14 1.84-2.49				
		3	2.71 1.99-3.69	3.18 2.33-4.34	3.55 2.60-4.85	3.67 2.65-5.08	3.52 2.28-5.43				
		4	2.03 0.76-5.43	2.86 1.05-7.78	1.60 0.40-6.47	0.09 0.00-17.1	1.65 0.21-13.1				
	Area deprivation	0	1 -	1 -	1 -	1 -	1 -	1 -	1 -	1 -	1 -
		1	1.00 0.85-1.18	1.12 0.97-1.30	1.22 1.08-1.38	1.17 1.05-1.30	1.35 1.18-1.53				
		2	1.07 0.85-1.33	1.45 1.22-1.73	1.18 0.98-1.41	1.36 1.17-1.59	1.30 1.09-1.55				
		3	1.47 0.94-2.32	1.71 1.24-2.34	1.85 1.31-2.61	1.94 1.41-2.67	1.43 0.98-2.10				
		4	1.04 0.34-3.20	0.95 0.30-2.97	0.85 0.21-3.41	1.83 0.79-4.23	1.13 0.41-3.15				

From the analyses it is known that the effect of area deprivation is no longer important after adjusting for individual deprivation, but these tables will enable the assessment of how area deprivation analyses underestimate the underlying effects shown at an individual level. These odds ratios show that there is a strong relationship between individual deprivation and rate of long-term illness in all age and sex strata. Deprivation seems to play a more prominent role in older individuals. For example, compared to those with no deprivation characteristics, males aged 16-24 with 2 characteristics had an excess risk of disease of 95% compared to males aged 45-54 years who had an excess risk of 208%. This effect is not present in the 55-64 age group and this may be due to the previously discussed problems of there being less individuals with characteristics of deprivation.

When looking at the area deprivation odds ratios, it can be seen that the effect of deprivation is much reduced. In all cases the excess risk carried by living in areas of deprivation level 1 when compared to level 0, varies between a minimum of 0% and a maximum of 35%. This compares with effects between at minimum 44% and at maximum 106% excess risk when looking at individual deprivation. This difference is most prominent in 16-25 year olds where the estimated effect for area deprivation is nonsignificant for both males and females at all levels, compared with odds ratios of 2 and above for individuals with two deprivation characteristics compared to no characteristics.

9.2.7 ETHNICITY AND DEPRIVATION

The aim of the analyses discussed here was to compare the effect of deprivation at the individual and area level based on the Townsend score on limiting long-term illness. No other factors affecting limiting long-term illness were investigated other than age and sex. However in many analyses of morbidity data, ethnicity is investigated and in many cases effects may be wrongly attributed to ethnicity instead of deprivation since the majority of people who classify their ethnicity as other than 'white' live in deprived areas. I decided to look at the effect of ethnicity before and after the effect of deprivation. Ethnicity was based on reported ethnicity in the census and this was crudely grouped into those classifying themselves as 'White' and those classifying themselves as other than 'White'. Individual level ethnicity data was available from the SAR data.

Table 9.10 shows the odds ratios for ethnicity for each age and sex group before and after adjusting for deprivation at the individual and area levels.

Table 9.10: Odds ratios with 95% confidence intervals for ethnicity (baseline = 'white') for each age and sex stratum.

	Age group				
	16-24	25-34	35-44	45-54	55-64
MALES					
Unadjusted	1.07 (0.84,1.36)	1.06 (0.89,1.27)	1.10 (0.92-1.31)	1.45 (1.24,1.70)	1.32 (1.12,1.56)
Adjusted for area deprivation	1.03 (0.80,1.32)	0.98 (0.82,1.18)	1.02 (0.85,1.22)	1.29 (1.10,1.52)	1.24 (1.05,1.47)
Adjusted for individual deprivation	0.93 (0.73,1.19)	0.87 (0.73,1.05)	0.86 (0.72,1.03)	1.15 (0.97,1.35)	1.12 (0.95,1.33)
FEMALES					
Unadjusted	1.24 (0.94,1.64)	1.43 (1.17,1.75)	1.27 (1.03,1.57)	1.26 (1.02,1.56)	1.85 (1.44,2.39)
Adjusted for area deprivation	1.20 (0.90,1.60)	1.28 (1.04,1.58)	1.19 (0.95,1.47)	1.11 (0.89,1.34)	1.69 (1.30,2.20)
Adjusted for individual deprivation	1.06 (0.80,1.39)	1.16 (0.95,1.43)	1.05 (0.85,1.31)	1.08 (0.87,1.34)	1.63 (1.26,2.11)

This shows that before deprivation was adjusted for at either the individual or area level the effect of ethnicity was formally significant in 6 of the 10 strata. The odds ratio of not being 'white' ranged from 1.06 in males aged 25-34 to 1.85 in females aged 55-64. After adjusting for area deprivation the odds ratio for ethnicity was reduced in all cases but remained significant in 4 strata. After adjusting for individual deprivation the odds ratios were significantly reduced and only one remained significant in females aged 55-64 years. This single result was surprising and showed a 63% excess risk of increased reporting of limiting long-term illness for women who classified their ethnicity as other than 'white'. Evidently there is the problem of multiple testing with ten different strata. However the size of the effect is large with a relatively small confidence interval. One possibility is that these women are more likely to report an illness as long-term than their 'white' counterparts. This is extremely

unlikely and would be expected to be evident in younger women as well. Alternatively they may be more likely to class themselves as economically active while the 'white' females may be more likely to class themselves as retired or 'looking after the home'. However these women may just be more ill than their 'white' counterparts. The effects of deprivation and ethnicity appear to be working in combination among these women leading to significantly worse health experiences. The reasons for this are not immediately obvious and need further investigation.

9.2.8 CONCLUSIONS OF SAR ANALYSES

These analyses have shown that area deprivation is failing to pick up major differences in morbidity between the deprived and less deprived, particularly in younger males and females. When looking at deprivation measured at an individual level, the rates of reported limiting long-term illness are consistently higher among the more deprived individuals, with up to four fold differences between the most deprived and least deprived. The differences observed when using area deprivation are sharply reduced and are non-significant in the 16-25 year old males and females. Analyses using area measures may be seriously underestimating the true size of inequalities in health.

These analyses are only based on economically active people aged 16-64 years. Economically inactive individuals were excluded as they may be inactive due to illness and would not be counted as unemployed and may have affected the results. If the economically inactive were included as well, the differences seen may have been even greater.

After adjusting for individual deprivation there was no evidence of an effect of area deprivation. This appears to show that it is individual characteristics that are most important and surroundings play less of a role in long-term illness. There was no evidence of a clinically significant interaction between area and individual deprivation. Furthermore, although the area measure used was grouped and hence some information was lost, the reanalysis using the original score did not affect the results. It is most likely that both area and individual factors play a role in the prevalence of long-term illness. However, the Townsend deprivation score does not appear to be a proxy for these area effects.

Sloggett and Joshi (1994) showed that ward level deprivation did not explain a significant amount of variation in mortality after adjusting for individual deprivation characteristics. This analysis has confirmed these findings but has gone further to show area based scores are underestimating deprivation effects.

Sloggett and Joshi used a slightly different measure of deprivation, using combined indicators from the Carstairs and Townsend score which was based on unemployment, lack of car ownership, lack of house ownership and proportion of people in social class IV or V. This replaced overcrowding, which is becoming less and less common for social class. The analyses in this paper were performed again using this alternative measure and this showed extremely similar results with an increasing effect of area deprivation with increasing age and a constant effect of individual deprivation throughout.

This work also supports the findings of Ecob (1996) who showed through multilevel modelling of information on limiting long-term illness that after controlling for individual effects of deprivation, there was no evidence of differences between areas of differing deprivation levels. His analyses were based on postcode sectors in Scotland which represent smaller areas than the districts used here. This therefore perhaps also indicates that had ED or ward level data been available, similar results would have been seen here. Shouls et al (1996) found a similar result with individual deprivation explaining most of the variation in limiting long-term illness. However they also found some residual contextual effects. These analyses were based on a multi-dimensional area measure of deprivation which may be measuring something different to the area level Townsend scores. The aim of this analyses is to look at the use of continuous unidimensional measures such as the Townsend score in health research and to assess whether adjustments for measurement error can be made. Their findings of residual contextual effects are important and of interest in social policy planning but they represent a different area of interest to that which is being pursued here.

The use of self-reported limiting long-term illness from the 1991 census as a measure of morbidity needs further investigation since the inequalities in health seen here may just be differences in reporting of illness between people from different socio-economic groups. The response to this is very subjective leading to great variation in what is actually reported. This type of work needs validation using a more robust measure of morbidity.

Further investigation is needed to understand why people with these individual deprivation characteristics have a much worse health experience than the less deprived. The measure used is thought to convey overall material deprivation and the psychological factors that go with loss of earnings and the instability of renting accommodation. More work is needed using alternative datasets such as the longitudinal study to relate smaller area level data such as wards to individual measures of deprivation.

A criticism of this work may be that it just reflects what is already known, that it is intuitive that when measured at an individual level the effect of deprivation on mortality will be seen to be greater. However most policy decisions are made on the evidence of reports based on area data. This work shows poor agreement between area and individual deprivation and perhaps indicates the need to target resources at deprived individuals rather than deprived areas.

Since area deprivation has been categorised into five groups based on the population with individual characteristics, it is possible that some information is being lost. This method of looking at individual characteristics in the Townsend score has also led to a lack of differentiation between the majority of individuals who have none of the four characteristics. However if future studies aimed to collect deprivation characteristics at an individual level, then it would be possible to ask for further information that would allow differentiation between these people.

It is probable that the differences between using area and individual scores are slightly over emphasised in my analyses since it was only possible to look at district level data. By using ward data the effect may be reduced but it is likely that the relationship between deprivation and health is still being underestimated. Further work using datasets combining data at the individual level and ward and enumeration district level are now investigated in section 9.3 and chapter 10.

9.3 PERINATAL MORTALITY DATA

9.3.1 DESCRIPTION OF THE DATA

These data are from a study of perinatal mortality in Leicestershire carried out by the University of Leicester's department of Epidemiology and Public Health. It is a large ongoing case control study of perinatal mortality. Still births of at least 24 weeks gestation and births of any gestation who do not survive the first 7 days of life were included as cases. Data for this analysis were available for 5 years from 1990 to 1994 with characteristics of deprivation collected at an individual level and the postcode available to be linked to area census data. Data were also available on indicators of child health. In order to look at the relationship between deprivation and child health only the controls are used since they provide a random sample of Leicestershire births between 1990 and 1994. Indicators of child health chosen were birth weight, use of ante-natal facilities before 18 weeks gestation and mother's age at the time of the birth.

With these data it was possible to link individual measures of deprivation with ED-level, ward level and district level area deprivation thus overcoming some of the limitations of the analyses of the SAR data. This provides an opportunity to investigate relationships between all levels of area data.

Data were available on 911 births with a recorded postcode, from which the enumeration district, ward and district codes were determined. For each study subject an individual household deprivation status was calculated as in the previous chapter, basing the individual score on the Townsend score. Since overcrowding was not recorded on this dataset deprivation was just based on unemployment, house ownership and car ownership. In the analyses of SAR data the individual deprivation scores were grouped according to the number of deprivation characteristics with five different groups. For this work however the number of individuals in the study was much smaller and so it was decided to divide them into two groups, those with no deprivation characteristics being classified as 'not deprived' and those with one or more characteristics being classified as 'deprived'. Deprivation scores were then calculated at enumeration district, ward and district level based on the same three deprivation characteristics. A cut off point was then calculated at each level so that the proportion

deprived and not deprived at each level of deprivation were similar to the proportion of 'deprived' and 'not deprived' at the individual level. This was more difficult at the district level since there were only 9 districts and hence it was not possible to have exactly the same number in each category. These classifications showed 42% to be deprived at the individual level and 58% to be classified as not deprived.

This is evidently an extremely crude way of classifying the data losing an enormous amount of information and disregarding the fact that the scores were designed to be ranked along a continuum. However the analyses of SAR data showed that using the area level deprivation score as a continuous measure or grouped into quintiles did not affect the conclusions of the analyses. Therefore this method was thought to provide a useful way of looking at the data given the size limitation of the dataset.

9.3.2 ASSESSING AGREEMENT

Crossclassification tables were drawn up to compare the misclassification at ED, ward and district levels. In order to assess the degree of agreement, the percentage of exact agreement and the kappa statistic were used. Table 9.11 shows the percentage agreement between the individual, ED, ward and district deprivation. There is higher agreement between the different levels of area based scores than between individual deprivation and area scores. There is little difference in the percentage of exact agreement between individual deprivation and deprivation based on different sizes of area with ED level deprivation having slightly higher exact agreement than district level deprivation.

Table 9.11: Percentage of exact agreement between different levels of deprivation.

Deprivation	Percentage of exact agreement			
	Individual	ED	Ward	District
Individual	100%			
ED	72%	100%		
Ward	71%	84%	100%	
District	67%	77%	86%	100%

Table 9.12 shows the kappa statistic for each combination of individual and area level measures. This study of the agreement appears to show similarly moderate agreement between

individual deprivation and each level of area deprivation, with very little improvement when the size of the area is reduced.

Table 9.12: Simple kappa statistic for different area level measures of deprivation

	Simple kappa statistic			
	Individual	ED	Ward	District
Individual	1			
ED	0.43 (0.37-0.49)	1		
Ward	0.40 (0.34-0.46)	0.67 (0.62-0.71)	1	
District	0.32 (0.26-0.38)	0.54 (0.48-0.59)	0.72 (0.67-0.76)	1

An alternative way of assessment was then used to see how well area level deprivation predicted individual deprivation. To do this the sensitivity and specificity of the area level measures to predict individual deprivation were calculated, where the sensitivity is the probability of being classified as deprived at area level given they are classified as deprived at individual level, and specificity is the probability of being classified as not deprived at area level given they are not deprived at the individual level.

Table 9.13 shows the sensitivity and specificity for ED, ward and district data at predicting individual deprivation. This shows the sensitivity to be similar whichever area level measure is used. However the specificity at district level is lower than for ED and ward level data. This means in Leicestershire, district level data is less good at predicting those who are not individually deprived. This may be because of the way the majority of the Leicestershire population is concentrated into the more deprived Leicester city district.

Table 9.13: Sensitivity and specificity of ED, ward and district deprivation for predicting individual deprivation for the perinatal mortality data

Deprived		ED deprivation			Ward deprivation			District deprivation		
		Yes	No	Total	Yes	No	Total	Yes	No	Total
Individual deprivation	Yes	257	127	384	251	133	384	249	135	384
	No	127	400	527	134	393	527	170	357	527
	Total	384	527	911	385	526	911	419	492	911
Sensitivity		67%			65%			65%		
95% Confidence Interval		(62%-72%)			(61%-70%)			(60%-70%)		
Specificity		76%			75%			68%		
95% Confidence Interval		(72%-80%)			(71%-78%)			(64%-72%)		

9.3.3 DEPRIVATION AND MEASURES OF CHILD HEALTH

After comparing the misclassification of deprivation at area level the next step was to assess its effect on the relationship between deprivation and several indicators of child health. The child health outcome measures used were the percentage of children with birthweight < 2.5kg, the percentage of children whose mothers were under 18 years at the time of birth, and the percentage of children whose mothers had not contacted their GP or midwife for antenatal care before 18 weeks gestation. Tables 9.14-9.16 show these three indicators by individual and area deprivation. As before the effect of ethnicity was investigated (Asian/non-Asian).

Table 9.14: Percentage of children (and numbers) whose mothers had not contacted their GP or midwife by 18 weeks gestation by level of deprivation.

	Individual deprivation	ED level deprivation	Ward level deprivation	District level deprivation
Deprived	14.3% (55/384)	10.9% (42/384)	10.6% (41/385)	10.3% (43/419)
Not deprived	5.1% (27/527)	7.6% (40/527)	7.8% (41/526)	7.9% (39/492)
χ^2 (df=1)	22.95	3.039	2.212	1.507
	P<0.001	P=0.081	P=0.137	P=0.220

Table 9.14 shows significantly more children are born to mothers who have not contacted their GP or midwife for ante-natal care before 18 weeks from deprived households than from less

deprived households. This effect is reduced when looking at area level data and is not formally significant.

These data were then investigated using logistic regression to adjust for area and individual deprivation. After adjusting for individual deprivation there was no significant effect of ED, ward or district level deprivation. This model estimated the odds ratio of being from a deprived household as 3.10 (95% confidence interval (1.91-5.01)), which means that children from deprived households had a 210% increase in their odds of being born to mothers who had not registered for antenatal care before 18 weeks gestation. There was no difference between Asian and non-Asian mothers before or after adjusting for deprivation (P=0.801 and P=0.931 respectively).

Table 9.15 shows the percentage of children born to mothers under 18 years to be significantly greater in deprived households. This effect is reduced and not significant when looking at area level data no matter which level of data is used.

Table 9.15: Percentage of children (and numbers) born to mothers who were under 18 years at the time of birth

	Individual deprivation	ED level deprivation	Ward level deprivation	District level deprivation
Deprived	5.5% (21/384)	3.1% (12/384)	3.4% (13/385)	3.1% (13/419)
Not deprived	0.2% (1/527)	1.9% (10/527)	1.7% (9/526)	1.8% (9/492)
χ^2 (df=1)	26.27	1.420	2.617	1.557
	P<0.001	P=0.233	P=0.106	P=0.212

Again these data were investigated using a logistic regression and showed no significant effect of deprivation at the area level whichever area measure was used, before or after adjusting for individual deprivation. The regression fitted gave an odds ratio of 35.3 (95% confidence interval (4.79-260.1) for individual deprivation which shows a greatly increased odds of children being born to a young mother if they are from a deprived household. The confidence interval for this effect is extremely wide because of the small number of mothers under 18 (22 mothers), and so the odds ratio cannot be accurately estimated but there is a significant difference between children from deprived households and those from households classified as

not deprived. It was not possible to assess the effect of ethnicity in this model since all of the 22 mothers who were under 18 were non-Asian.

Table 9.16 shows an apparently different pattern between birthweight and deprivation to the previous indicators of child health. Here there appears to be a significant effect of deprivation at individual, ward, and district level with more low birthweight children from deprived areas and households.

Table 9.16: Percentage of children (and numbers) with birthweight < 2.5kg by level of deprivation.

	Individual deprivation	ED level deprivation	Ward level deprivation	District level deprivation
Deprived	9.6% (37/384)	8.1% (31/384)	9.9% (38/385)	9.5% (40/419)
Not deprived	4.9% (26/527)	6.1% (32/527)	4.8% (25/526)	4.7% (23/492)
χ^2 (df=1)	7.629 P=0.006	1.381 P=0.240	9.043 P=0.003	8.343 P=0.004

When a logistic regression was performed, there was still a significant effect of district level deprivation, after adjusting for individual deprivation, and the effect of individual deprivation was reduced. There was no evidence of a deprivation effect at the ED or ward level. However after adjusting for ethnicity the effect of district deprivation was no longer significant (P=0.268) but there was a significant effect of ethnicity (P=0.0009) and individual deprivation (P=0.009). This model gave an odds ratio of 2.70 (95% confidence interval (1.50,4.86)) for being an Asian mother which relates to a 170% increase in the odds of having a low birthweight baby if the mother is Asian compared with those who are not Asian. In terms of deprivation the model estimated the odds ratio as 2.01 (95% confidence interval (1.19,3.39)) for mothers from deprived households which relates to an increase of 101% in the odds of having a low birthweight baby if the mother comes from a deprived household compared with a non-deprived household. At district level, Leicester city is the only deprived district. Since nearly all of the Asian mothers came from the city district this explains why there was an apparent district deprivation effect before adjusting for ethnicity.

9.3.4 SENSITIVITY ANALYSIS

One of the main criticisms that could be made about these analyses is that the data is too crudely grouped. This loses much information on area deprivation and ignores the fact that the Townsend score was designed to be used as a continuous scale. Therefore sensitivity analyses were performed to assess the extent of this problem. For each of the three outcome variables the analyses were repeated, assessing the effect of deprivation at the ED, ward and district level and using each area measure as a continuous variable and categorised into quintiles. Since so 40% of patients were from one deprived district the fourth and fifth quintile of deprivation scores based on district level deprivation had to be grouped together. Table 9.17-9.19 detail the results of these sensitivity analyses. Since there was a large difference in the percentage of low birthweight babies being born to Asian mothers the analyses of low birthweight were adjusted for whether or not the mother was Asian.

Table 9.17: Sensitivity analysis of the relationship between area deprivation and the probability of the mother not registering for antenatal care before 18 weeks gestation

Area level		Analysis based on quintiles of deprivation			Analysis based on continuous deprivation score		
		ED	Ward	District	ED	Ward	District
Unadjusted for individual deprivation	Change in deviance	$\chi^2_4=8.699$	$\chi^2_4=6.491$	$\chi^2_3=3.182$	$\chi^2_1=2.846$	$\chi^2_1=3.799$	$\chi^2_1=1.531$
	P value	0.069	0.165	0.364	0.092	0.051	0.216
Adjusted for individual deprivation	Change in deviance	$\chi^2_4=6.321$	$\chi^2_4=2.785$	$\chi^2_3=0.337$	$\chi^2_1=0.543$	$\chi^2_1=0.018$	$\chi^2_1=0.346$
	P value	0.176	0.594	0.953	0.461	0.893	0.556

Table 9.18: Sensitivity analysis of the relationship between area deprivation and the probability of the mother being under 18 years

Area level		Analysis based on quintiles of deprivation			Analysis based on continuous deprivation score		
		ED	Ward	District	ED	Ward	District
Unadjusted for individual deprivation	Change in deviance	$\chi^2_4=3.523$	$\chi^2_4=3.208$	$\chi^2_3=6.244$	$\chi^2_1=2.050$	$\chi^2_1=2.195$	$\chi^2_1=1.388$
	P value	0.474	0.524	0.100	0.152	0.139	0.239
Adjusted for individual deprivation	Change in deviance	$\chi^2_4=2.860$	$\chi^2_4=1.410$	$\chi^2_3=2.825$	$\chi^2_1=0.885$	$\chi^2_1=0.302$	$\chi^2_1=0.292$
	P value	0.582	0.843	0.419	0.347	0.583	0.589

Table 9.19: Sensitivity analysis of the relationship between area deprivation and the probability of the child having a birthweight of <2.5kg (Model adjusted for whether the mother was Asian)

Area level		Analysis based on quintiles of deprivation			Analysis based on continuous deprivation score		
		ED	Ward	District	ED	Ward	District
Unadjusted for individual deprivation	Change in deviance	$\chi^2_4=11.26$	$\chi^2_4=4.889$	$\chi^2_3=6.154$	$\chi^2_1=3.994$	$\chi^2_1=2.284$	$\chi^2_1=3.940$
	P value	0.024	0.299	0.104	0.046	0.131	0.047
Adjusted for individual deprivation	Change in deviance	$\chi^2_4=9.189$	$\chi^2_4=2.254$	$\chi^2_3=3.235$	$\chi^2_1=0.652$	$\chi^2_1=0.199$	$\chi^2_1=1.264$
	P value	0.057	0.689	0.357	0.419	0.655	0.261

These analyses do not affect the interpretations of the analyses based on categorising the deprivation scores into deprived/not deprived. In no cases was there a significant effect of area level deprivation after adjusting for individual level deprivation. Further there was no evidence of a significant relationship between any of the outcome measures and area level deprivation

before adjusting for individual deprivation in any of the analyses except for low birthweight using ED level data indicating that no important relationships were being overlooked by basing the analyses on a crude binary categorisation.

9.3.5 CONCLUSIONS OF ANALYSES OF PERINATAL MORTALITY DATA

This analysis of Leicestershire data from the perinatal mortality study has confirmed the poor agreement between area and individual deprivation measures seen when looking at the SAR data. Using data on smaller areas to calculate deprivation scores, i.e. ED instead of district or ward level information, led to very little improvement in the level of agreement. ED and ward level deprivation showed very similar agreement with individual data while district level data showed slightly worse agreement. Agreement between area measures of deprivation e.g. ward and ED, was much higher. This pattern may be unusual to Leicestershire since the majority of the population is concentrated in the city and are classified as deprived. In areas with a different population distribution the pattern may differ.

When looking at indicators of child health, there was no evidence of a significant relationship between area deprivation and the percentage of children born to mothers under 18 years or the percentage of children born to mothers who had not seen a GP or midwife for antenatal care. But when deprivation was measured at an individual level a significant difference was seen with increased risk of being born to a young mother or a mother who did not register early for antenatal care if the child came from a deprived household. When looking at low birthweight there appeared to be a geographical effect of increased low birthweight babies in the city. However this was due to the relationship between ethnicity and birthweight and the high concentration of Asian mothers in the deprived city district. After adjusting for ethnicity there was evidence of increased odds of having a low birthweight baby if the mother came from a deprived household as when looking at antenatal care and mother's age. These results confirm the findings of the previous chapter that area measures are underestimating the relationship between deprivation and health that are seen when using individual measures of deprivation.

It could be argued that the analyses performed in section 9.2 based on the SAR data were of extremely limited use because there are no ward or ED identifiers and that deprivation at district level is of little use. However the analyses of data from the perinatal mortality study have indicated very little differences in inequalities in health when based on ED, ward or

district level deprivation indicating that particularly in Leicestershire basing analyses on reduced area size makes little difference. The percentage of exact agreement between individual level deprivation and area level deprivation was extremely similar for enumeration districts, wards and districts and it is between area level and individual level that major differences can be observed.

These analyses, like those based on the SAR data, indicate that knowing whether or not a person has one of the four characteristics of deprivation that make up the Townsend score can explain much more variation in health than knowing the area level deprivation score. Furthermore the inequalities in health identified at the individual level in the analyses of child health data were not evident at the area level indicating that inequalities in health may be overlooked if analyses are based on area measures alone.

9.4 SUMMARY AND CONCLUSIONS

In this chapter two datasets have been used to look at the problem of ecological fallacy in estimating inequalities in health. The ecological fallacy argument indicates that it is wrong to assume that a relationship between health and deprivation as measured at area level would necessarily exist between health and deprivation measured at the individual level since not all people living in a deprived area are deprived. However the analyses in this chapter have shown that far from being non-existent, the relationship between health and deprivation is underestimated when using area level measures. This confirms the opinion of MacRae (1994) that sceptics of ecological correlations between deprivation and health cannot seek the support of the ecological fallacy argument. Further, area deprivation measures may even be failing to detect inequalities in health present at the individual level.

These findings indicate two important problems. Firstly, information on individual level deprivation may alter the results of the analyses of the 1992 amblyopia study. In chapter 10 I reanalyse the data from the 1992 amblyopia study to assess whether the additional information on individual levels of deprivation indicates a relationship between age at presentation and deprivation at the individual level.

Secondly this chapter has indicated that area deprivation measures underestimate the effect of deprivation on several measures of health. Information on the relationship between area and individual level deprivation may help us to estimate the effect of deprivation at the individual level in studies where only area level deprivation is available. The area deprivation score assigned to an individual evidently measures their deprivation at the individual level with error. This type of exposure measurement error has been shown to lead to bias in the estimation of the exposure effect as observed here. Methods for adjusting for exposure measurement error have been developed for use in other research areas. In chapter 11 I adapt several of these methods in order to apply them to the SAR data. I look at whether it is possible to adjust relative risk estimates for deprivation obtained using area data to estimate the relative risk of deprivation at the individual level in other studies using these census data. I then discuss this work on adjustment for exposure measurement error in terms of the analysis of the amblyopia studies in chapter 12.

CHAPTER 10

COMPARING THE EFFECT OF DEPRIVATION AT INDIVIDUAL AND AREA LEVEL ON THE AMBLYOPIA STUDY RESULTS

10.1 AIMS OF THE CHAPTER

In chapters 6 and 8 I analysed data on the presentation of amblyopia and investigated the relationship between age at presentation of amblyopia and area level deprivation as measured by the Townsend score. However the problem of ecological fallacy was identified in the interpretation of these analyses and others based on area-level measures. This asserts that it is wrong to assume that an effect observed using area level data necessarily exists at the individual level.

In chapter 9 I used census data and data on a random sample of births to investigate this problem. I compared the effect of area level and individual level deprivation on a range of health outcome measures. These analyses showed that far from the relationship at the individual level being fallacious, area level measures were serving to underestimate inequalities in health seen when individual level measures were used. These findings supported those of Ecob (1996) and Sloggett and Joshi (1994) who showed that for similar health measures there was little or no residual effect of area level deprivation after adjusting for individual level measures.

Therefore based on these analyses it would appear possible that the effect of deprivation on the age at presentation of amblyopia would be greater than that observed in chapters 6 and 8 if individual level measures of deprivation had been available for analysis. As discussed in chapter 7, in the 1992 study of amblyopia information was collected on individual and area level deprivation. In this chapter these data are analysed to assess the effect of deprivation on age at presentation at both an area and individual level.

10.2 DATA COLLECTED

As discussed in chapter 7 there were 202 children who presented to Leicester Royal Infirmary orthoptic department in 1992 and were eligible for study. This excluded all children who lived more than 10 kilometres from the clinic and those who were not treated within two years of presentation. As reported earlier these comprised 92 with a microtropia or no strabismus and 110 with a large angle of strabismus. Although the data from orthoptic notes were available for all study patients, there was a non-response rate of 9% (19 children) for the more detailed questionnaire among the 202 eligible children. The majority of parents were interviewed in the clinic or by telephone. Data were therefore available on 183 children with information on individual deprivation.

10.3 INFORMATION ON INDIVIDUAL DEPRIVATION

Data were collected from the parents of children with amblyopia about various aspects of material deprivation, educational level and occupation which are also used as measures of deprivation. In order to relate individual level data to the previous analyses using the Townsend deprivation score, an adaptation of the Townsend score at the individual level was used as in chapter 9. This differed slightly as information on unemployment was based on whether the parents were unemployed rather than the household head as in the previous analysis of SAR census data.

10.4 COMPARING INDIVIDUAL AND AREA LEVEL DATA

Since numbers were fairly small it was decided to classify patients as 'deprived' if they had at least one of the deprivation characteristics and 'not deprived' otherwise. This evidently loses some information but the numbers with higher levels of deprivation are small. Further the analyses in chapter 9 of the data from the perinatal mortality study showed very little difference in the results whether the analyses were based on this binary classification, quintiles of deprivation or a continuous score. This method classified 54% of patients as deprived and 46% as not deprived. The postcodes of the children were used to obtain area level census data. Since only children within a 10km radius of Leicester Royal Infirmary were studied,

district level information was not investigated since the majority were in the city district. The ED, and ward level deprivation scores for the patients were divided into the same proportions so that those in the most deprived 54% of areas were classified as deprived and the remaining 46% were classified as not deprived. Table 10.1 shows the percentage agreement between the individual, ED, and ward deprivation. There is higher agreement between the different levels of area based scores than between individual deprivation and area scores. There is little difference in the percentage of exact agreement between individual deprivation and deprivation based on ED and ward data.

Table 10.1: The percentage of exact agreement for children treated for amblyopia based on different levels of deprivation measure

	Percentage of exact agreement		
	Individual	ED	Ward
Individual	100%		
ED	74%	100%	
Ward	72%	90%	100%

Table 10.2 shows the kappa statistics for this data. In general this shows fair to moderate agreement between individual deprivation and the two measures of area deprivation and good agreement between the different levels of area deprivation.

Table 10.2: The kappa statistics for children treated for amblyopia based on different levels of deprivation measure (95% Confidence Intervals in parentheses)

	Simple kappa statistic		
	Individual	ED	Ward
Individual	1		
ED	0.47 (0.34-0.60)	1	
Ward	0.44 (0.31-0.57)	0.79 (0.70-0.88)	1

There appears to be similar agreement between individual deprivation and each level of area deprivation, with very little improvement when the size of the area is reduced. This seems to indicate that when using this crude measure of deprived/ not deprived, there is no gain from using ED level data or ward level. An alternative way to assess the degree to which area level deprivation agrees with individual deprivation is measure the sensitivity and specificity of the area level measures for predicting individual deprivation, where the sensitivity is the probability of being classified as deprived at area level given they are classified as deprived at individual level, and specificity is the probability of being classified as not deprived at area level given they are not deprived at the individual level. Table 10.3 shows the sensitivity and specificity for ED and ward level data at predicting individual deprivation. The sensitivity is very similar for ward and ED data with the specificity being slightly lower than the sensitivity. For this 10km radius of Leicester the sensitivity is higher than the data for the whole of Leicestershire from the perinatal mortality study. The specificity is lower for the amblyopia data. It appears for this smaller region of Leicestershire, area data is better at identifying deprived areas but less good at identifying the less deprived areas.

Table 10.3: Sensitivity and specificity of ED, ward and district deprivation for predicting individual deprivation

Deprived		ED deprivation			Ward deprivation		
		Yes	No	Total	Yes	No	Total
Individual deprivation	Yes	76	24	100	74	26	100
	No	24	59	83	25	58	83
	Total	100	83	183	99	84	183
Sensitivity		76%			74%		
95% Confidence Interval		(68%-84%)			(65%-83%)		
Specificity		71%			70%		
95% Confidence Interval		(61%-81%)			(60%-80%)		

10.5 ANALYSIS OF AGE AT PRESENTATION AND DEPRIVATION

The analysis of the amblyopia data in chapter 8 has shown that the relationship between deprivation and age of presentation of children with microtropia or no strabismus reduced between 1983 and 1992. This conclusion was based on deprivation measured at ward level. The information available at individual household level from the study in 1992 is used here to see if the ward level deprivation was underestimating the effect of deprivation at the individual level. The age of presentation of those classified as deprived and not deprived was compared using individual, ED and ward level deprivation. Table 10.4 shows the mean age of presentation for each type of amblyopia for those classified as deprived and not deprived at both area and individual levels.

Table 10.4: Summary statistics for age at presentation for different types of amblyopia and different levels of deprivation measure.

		Age at presentation					
		Microtropia / No strabismus			Large angle strabismus		
Deprived		Mean	Standard deviation	N	Mean	Standard deviation	N
Individual deprivation	Yes	4.84	1.85	38	3.05	1.87	62
	No	4.80	1.54	46	3.70	1.97	37
		P=0.904*			P=0.103*		
ED level deprivation	Yes	5.14	1.97	42	3.26	1.86	58
	No	4.50	1.26	42	3.34	2.04	41
		P=0.078*			P=0.824*		
Ward level deprivation	Yes	5.07	1.81	40	3.29	1.86	59
	No	4.59	1.53	44	3.30	2.04	40
		P = 0.191*			P=0.985*		
* Test based on independent two sample t-test							

For children with microtropia or no strabismus there is no significant difference in the age at presentation of those from deprived households compared with those from households classified as 'not deprived'. At ED level and ward level however there appears to be a difference with those from deprived areas presenting later. A Normal errors regression to look

at the relationships between age at presentation and deprivation did not alter these conclusions with no significant effect of individual level or ward level deprivation and an effect of ED level deprivation which approached significance. Based on previous hypotheses it would be expected that the relationship between deprivation and age would be increased at the individual level. In this case the pattern is reversed. There was no effect of ethnicity before or after adjusting for deprivation ($P=0.646$ and $P=0.643$ respectively).

For the children with a large angle of strabismus there is no evidence of a relationship between age at presentation and deprivation at the area level. At the individual level, those from more deprived households are younger, although this is not formally significant. This pattern is again different to what would be expected based on previous hypotheses. Again there was no difference between Asians and non-Asians before or after adjusting for deprivation ($P=0.786$ and $P=0.823$ respectively)

It is also apparent that those with a microtropia or no strabismus tend to come from more deprived households than those with a large angle of strabismus ($\chi^2_1=5.544$ $P=0.019$). This is less evident when looking at deprivation based on area measures, with no significant difference at ED level ($\chi^2_1=1.352$ $P=0.245$) or ward level ($\chi^2_1=2.625$ $P=0.105$). This appears to indicate some individual deprivation effect on the incidence of large angle strabismus.

10.6 DISCUSSION OF THE FINDINGS

The analyses in this chapter have indicated two points of interest. Firstly they have confirmed the patterns shown by analysis of the perinatal mortality data that the relationship between area level and individual level data is similarly poor whether ED or ward level data are used. Recent research has confirmed the patterns seen here of little improvement in agreement when using ED level rather than ward level deprivation (Carr-Hill et al 1995). This indicates that the degree of underestimation of inequalities in health shown in the analysis of the SAR data in chapter 9 may not have been dramatically reduced if deprivation had been based on ED rather than district level deprivation. However this may be a particular phenomenon to Leicestershire because of the concentration of areas of high deprivation in the city district.

Secondly this chapter has shown no evidence of an increased relationship between deprivation and age at presentation of amblyopia when analyses are based on individual rather than area level deprivation. There were no significant effects of deprivation on age at presentation when deprivation was measured at any level. However there appeared to be a slight non-significant increase in age with increasing area deprivation which was not apparent at the individual level. This analysis seems to indicate area deprivation differences rather than individual differences. Since the changes made to screening were introduced between 1987 and 1991 some of the children in this study will not have passed through all of the vision screening programme since the changes were introduced. It is possible that some of the more deprived areas of Leicester were slightly later at implementing the changes than some of the less deprived areas. This effect is not seen at the individual level since there is such great disparity between area and individual deprivation. Individual factors play an important role in the referral of children with amblyopia but the indicators used here do not appear to be proxies for the individual factors affecting presentation. In order to more fully understand the processes that are occurring here a cohort of children need to be followed who were involved in the vision screening programme only after all of the changes had been made.

This finding is in contrast to the analyses of census data and birth data in chapter 9. This may be due to the differences in the outcomes used for analysis which in this chapter concentrated on inequalities in access to health care whereas in the previous chapter analyses were predominantly based on inequalities in health. In chapter 6 I concluded that the relationship between age at presentation of asymptomatic amblyopia and deprivation indicated structural differences in screening delivery and not behavioural differences in the uptake of screening. This conclusion would be consistent with the findings of the analysis in this chapter since it is apparent that there may be some small residual area effects of deprivation on age at presentation which are not evident at the individual level.

10.7 SUMMARY AND CONCLUSIONS

In this chapter I have further investigated the relationship between age at presentation of amblyopia and deprivation using measures of deprivation at both individual and area levels. The main findings of interest from these analyses were:

- 1) no evidence of an increased relationship between age at presentation and deprivation when analyses were based on individual level rather than area level measures.
- 2) the confirmation of a poor relationship between area level and individual level deprivation measures,
- 3) the lack of improvement in this relationship when the size of area was reduced

This analysis therefore appears to give further evidence that there is no longer a relationship between deprivation and age at presentation of amblyopia in Leicestershire. This may be partly attributable to the changes in screening although wider NHS reforms and study design issues discussed in chapter 8 may also be responsible. These data have also further confirmed the poor correlation between area and individual deprivation and this appears to improve very little on reducing the area size that the scores are based on in Leicestershire. This agrees with the findings of Carr-Hill et al (1995) who showed that ward level data was as useful as enumeration district level data in measuring deprivation. However Leicestershire may be unusual in the distribution of deprivation and this pattern may not be similar for other areas of the country.

Two issues were highlighted by the analyses in chapter 9. The first, whether individual deprivation was related to age at presentation of amblyopia in 1992 has been investigated in this chapter. The second is the issue of whether knowledge of the relationship between area and individual deprivation can be used to estimate the relationship between health and individual deprivation in studies where only area deprivation was available such as the 1983 amblyopia study. Methods for doing this are discussed in chapter 11.

CHAPTER 11

ADJUSTING THE EFFECT OF DEPRIVATION FOR MEASUREMENT ERROR

11.1 AIMS OF THE CHAPTER

In chapter 9 I have shown area level measures of deprivation to underestimate the relationship between health and deprivation seen when measured at the individual level. The relationship between area and individual measures was not significantly improved when the area measures were based on smaller areas, with ED level deprivation showing similarly high levels of measurement error as district level deprivation. These analyses have shown that error in the measurement of deprivation is leading to bias in the estimated effect of deprivation on health consistently underestimating the exposure effect.

Faced with the problem of error in the measurement of exposures, researchers have taken a variety of approaches (Bross, 1954). At one extreme some assume that in the long run the problems of measurement error will cancel themselves out and therefore they use the usual analysis methods. At the other extreme some believe that the data are biased and that no conclusions can be drawn. In this chapter I take an intermediate position and look at methods that have been developed to adjust for measurement error to estimate the underlying true exposure effect. They offer the opportunity to use information about the relationship between area and individual level deprivation to adjust the exposure effect in studies where information is only available at the area level. In this chapter I explore several possible methods for adjusting for measurement error using the SAR data analysed in chapter 9 to assess whether they are successful in adjusting the effect of deprivation on limiting long-term illness seen at the area level and hence estimate the effect that would be observed if individual level deprivation were used. I look at the possibility of using the information about the relationship between area and individual deprivation to adjust the effect of deprivation in other studies which only have information at area level.

11.2 MEASUREMENT ERROR

Measurement error of an exposure leading to bias is a frequent problem in epidemiological studies and can arise from researchers, their subjects or the methods used to collect data (Sitthi-amorn and Poshyachinda, 1993). Many cohort or case-control studies use self-reported exposure status rather than a more reliable clinical measure since this often requires less resources, and may be less intrusive and painful to the subjects. In the case of deprivation it is the measure used to assess deprivation that is being measured in error. Ecological studies of deprivation are appealing to undertake since it is far easier to obtain an individual's postcode and get an area deprivation score from census data than it is to interview patients and obtain data on deprivation at an individual level. One major effect of measurement error in a study is that it leads to bias in the measure of association between the exposure and the outcome of interest.

Measurement error can occur where the exposure is a continuous or a categorical variable and in the latter case this is frequently termed misclassification since it results in the exposure being misclassified from one exposure category to another. In this chapter I concentrate on adjusting for measurement error based on a categorical exposure. Based on the analyses in chapter 9 it is apparent that the majority of individuals all had the same deprivation score at the individual level because they had none of the four characteristics of deprivation. Therefore the deprivation data were categorised for the analyses. Furthermore very few individuals fell into the very deprived categories at the individual level. The analyses in this chapter are hence based on a binary categorisation of deprivation as used in the analysis of perinatal mortality data and amblyopia data in chapters 9 and 10, with people categorised as deprived at the individual level if they have any of the four characteristics of deprivation and the rest are classified as not deprived. Area level deprivation was categorised into two groups of similar proportions to the individual level deprivation groups. Evidently this loses much information and may ignore the fact that the area scores were set up to rank areas. However because of the distribution of deprivation at the individual level and the fact that using this binary categorisation led to very little difference in the results of the analyses in chapter 9 this is a useful method for looking at the measurement error problem.

11.3 THE EFFECTS OF MISCLASSIFICATION

In this thesis the concern is with misclassification of the exposure and not of the disease. There are two possible types of exposure misclassification and these can be explained in terms of deprivation measurement. Firstly there are people who are deprived at the individual level who are classified as not deprived at the individual level and these are called false negatives.

Secondly there are individuals who are not deprived at the individual level who are classified as deprived at the area level, known as false positives. The degree of misclassification is generally measured by the sensitivity and specificity of the exposure measure (Mertens, 1993) as in chapter 10 where:

sensitivity: the probability of being recorded as exposed by the less reliable measure given the subject is exposed as measured by the gold standard
(the probability of being classified as deprived at the area level given the subject is deprived at the individual level)

specificity: the probability of being recorded as unexposed by the less reliable measure given the individual is unexposed,
(the probability of being classified as not deprived at the area level given the subject is not deprived at the individual level).

Misclassification can lead to biased estimation of the exposure effect and the type of bias depends on these two measures.

11.3.1 NON-DIFFERENTIAL MISCLASSIFICATION

Non-differential misclassification refers to errors in the categorisation of the exposure unrelated to the outcome of interest. This means the sensitivity and specificity will be equal for the diseased group and the undiseased group. For example, in the analysis of the SAR data it would mean that people were as likely to be misclassified if they had a limiting long-term illness as if they did not. In epidemiological studies where this assumption holds, it has been argued that misclassification will underestimate the association between the exposure and the outcome of interest and will not 'cross-over' the null value of 1 (Mertens, 1993). Bross (1954) indicates that this type of misclassification should not lead to an increased risk of making a type I error (i.e. falsely rejecting the null hypothesis of no exposure effect) but that the power of the study will be reduced. Even in cases where the sensitivity and specificity of the exposure

measure are extremely high this can have a considerable attenuation effect on the estimates of the association (Armstrong et al, 1994).

However this assumption that non-differential misclassification always leads to an underestimation of the exposure effect does not hold if the less reliable exposure measure is so poorly associated with true exposure status that an individual is as likely to be misclassified as to be correctly classified, (i.e. sensitivity+specificity ≤ 1). In these cases the observed association may be equal to the null or actually cross over the null value and give an apparent inverse effect (Flegal et al, 1986). The effect of the sensitivity and specificity on the observed relative risk is summarised in table 11.1 (Flegal et al, 1986).

Table 11.1: General effects of true relative risk (R), sensitivity (U), and specificity (V) on apparent relative risk (R*) (from Flegal et al (1986), table 2).

True risk	Sum U+V	Apparent risk	Comment
R>1.0	U+V>1.0	$1.0 < R^* < R$	Underestimation of association
	U+V=1.0	$R^*=1$	Underestimation of association
	U+V<1.0	$R^* < 1$	Reversal of direction
R=1.0		$R^*=1.0$	No bias when no association
R<1.0	U+V>1.0	$R < R^* < 1.0$	Underestimation of association
	U+V=1.0	$R^*=1.0$	Underestimation of association
	U+V<1.0	$1.0 < R^*$	Reversal of direction

However in most epidemiological studies for an instrument to be considered a measure of the true exposure it should correctly classify subjects with a higher probability than chance (Armstrong et al, 1994). Effects will be similar on other measures of association such as the odds ratio. There are further scenarios where non-differential misclassification does not lead to underestimation of the exposure effect related to polychotomous exposures rather than dichotomous exposures (Dosemeci et al, 1990) but these are not of direct relevance here.

11.3.2 DIFFERENTIAL MISCLASSIFICATION

Differential misclassification occurs when exposure misclassification differs according to disease status i.e. the assumption of equal sensitivity and specificity across disease groups is violated. The assumption of non-differential misclassification has been shown not always to be

justifiable (Fleiss, 1981) and the consequences of this are more serious since the effect on the estimate of the exposure effect is far less well defined. Risk estimates may be distorted in either direction and may often lead to overestimation of the effect (Mertens, 1993).

11.3.3 ADJUSTING FOR MISCLASSIFICATION

Having discussed the possible effects of misclassification on the estimated exposure effect, several researchers have developed methods to correct for misclassification. Although interpretation of the adjusted estimates for the effect of exposure should be made with caution, they can be informative on the potential size of the exposure effect. In this work the emphasis is on looking at supplemental data to adjust the relative risk estimate for deprivation at the area level, although the methods will be validated using data from the same source. Although methods have been developed to look at multiple category exposures and outcomes (Reade-Christopher and Kupper, 1991), I investigate three methods to look at adjusting for the misclassification of a dichotomous exposure (deprivation) with a dichotomous outcome (limiting long-term illness). Flegal et al (1986) developed an equation that adjusts the estimate of the exposure effect given information on the sensitivity and specificity of the misclassification and the prevalence of the true exposure. Espeland, Hui and Odoroff (Espeland and Hui, 1987; Espeland and Odoroff, 1985) and Ekholm, Palmgren and Green (Ekholm, 1991; Ekholm and Palmgren, 1987; Palmgren and Ekholm, 1987; Ekholm, Green and Palmgren, 1986; Ekholm and Green, 1995) both use model fitting approaches to solve this problem. These three methods of adjustment will be used to look at the SAR data and assess how useful they are in studying deprivation and morbidity.

These methods all assume non-differential misclassification i.e. that the probability of being misclassified does not vary by disease status. Since the aim of this work is to see whether information from the SAR on the relationship between individual and area data could be used to adjust estimates in other studies where only area level data was available, this assumption of non-differential misclassification would have to be made. However the consequences of making this assumption are discussed in detail in section 11.8.

11.3.4 USE OF RELATIVE RISK AND ODDS RATIO

The analysis of SAR data in chapter 9 focused on estimating odds ratios. However the majority of work on misclassification is based on its effect on the estimates of relative risk.

Since the SAR data is a cross-sectional survey, and the prevalence of disease is low at around 4% in economically active individuals, the odds ratio will be similar to the estimate of the relative risk provided the disease does not alter the exposure. In order for comparison with other work on misclassification, I analyse the data in terms of relative risks rather than odds ratios. Deprivation at area and individual level is collapsed into two groups, deprived and not-deprived. Relative risk estimates are based on the increased risk of being deprived as opposed to not deprived.

Using the SAR data, the relative risk for individual level and area level deprivation is known. Therefore the estimate of the relative risk calculated using these methods of adjusting for misclassification can be compared with the known estimate of deprivation from the individual data.

11.3.5 TERMINOLOGY

The same terminology is used in this chapter to demonstrate the three methods of adjustment for misclassification. The analyses use 'doubly sampled' data from the SAR data. The disease status (limiting long-term illness) is denoted by D . Exposure status is denoted by E for the gold standard measure (individual deprivation) and E^* for the less reliable measure (area deprivation). Presence of a characteristic is indicated by + and absence by -. The SAR data provides information on three two-way tables, illness status by area deprivation (D by E^*), illness status by individual deprivation (D by E) and area deprivation by individual deprivation (E^* by E). The methods used in this chapter concentrate on incorporating the information from the D by E^* table and the E by E^* table in order to estimate the D by E table. This estimate will then be compared with the observed D by E table from the SAR data to validate the methods. The sensitivity U and the specificity V of the misclassification table are used to adjust the relative risk observed at the area level R^* in order to estimate the relative risk R at the individual level. The estimate of the relative risk will be referred to as R' . These parameters can be written in the following form:

$$\text{Sensitivity} = U = P(E^*+|E+) \quad (11.1)$$

$$\text{Specificity} = V = P(E^*-|E-) \quad (11.2)$$

$$\text{Relative risk (area)} = R^* = P(D+|E^*+) / P(D+|E^*-) \quad (11.3)$$

$$\text{(individual)} = R = P(D+|E+) / P(D+|E-) \quad (11.4)$$

where $P(E+)$ is the probability of E being positive and $P(E-)$ is the probability of E being negative.

11.4 FLEGAL ET AL'S METHOD

Flegal et al (1986) have developed an equation (11.5) to show how the apparent relative risk based on the less reliable measure (R^*) is related to the relative risk based on the gold standard (R), the sensitivity (U) and specificity (V) of the misclassification table and the prevalence of the exposure ($P(E+)$). R'_F is the estimate of the relative risk based on the Flegal method.

$$R^* = \frac{[UR'_F P(E+) + (1-V)P(E-)][(1-U)P(E+) + VP(E-)]}{[UP(E+) + (1-V)P(E-)][(1-U)R'_F P(E+) + VP(E-)]} \quad (11.5)$$

This is derived from the fact that

$$R^* = \frac{P(D+/E^*+)}{P(D+/E^*-)} \quad (11.6)$$

Expanding the numerator gives:

$$\begin{aligned} P(D+/E^*+) &= \frac{P(D+E^*+)}{P(E^*+)} = \frac{P(D+E^+/E+)P(E+) + P(D+E^+/E-)P(E-)}{P(E^+/E+)P(E+) + P(E^+/E-)P(E-)} \\ &= \frac{P(E^+/D+E+)P(D+E+) + P(E^+/D+E-)P(D+E-)}{P(E^+/E+)P(E+) + P(E^+/E-)P(E-)} \end{aligned} \quad (11.7)$$

Flegal et al then make the assumption of non-differential misclassification i.e. that the probability of being misclassified does not vary by disease status, such that:

$$P(E^+/D+E+) = P(E^+/D-E+) = P(E^+/E+) \quad (11.8)$$

Which substituting into

$$\begin{aligned} P(D+/E^*+) &= \frac{P(E^+/E+)P(D+E+) + P(E^+/E-)P(D+E-)}{P(E^+/E+)P(E+) + P(E^+/E-)P(E-)} \\ &= \frac{UP(D+E+) + (1-V)P(D+E-)}{UP(E+) + (1-V)P(E-)} \end{aligned} \quad (11.9)$$

Similarly the denominator of equation 11.6 can be rewritten:

$$P(D+/E^*-) = \frac{(1-U)P(D+E+) + VP(D+E-)}{(1-U)P(E+) + VP(E-)} \quad (11.10)$$

Substituting 11.9 and 11.10 back into 11.6 gives

$$R^* = \frac{(UP(D+E+) + (1-V)P(D+E-))((1-U)P(E+) + VP(E-))}{(UP(E+) + (1-V)P(E-))((1-U)P(D+E+) + VP(D+E-))} \quad (11.11)$$

If $P(D+E+)$ is then rewritten in terms of conditional probabilities

$$P(D+E+) = \frac{P(D+/E+)}{P(D+/E-)} P(D+E-) \frac{P(E+)}{P(E-)} = RP(D+E-) \frac{P(E+)}{P(E-)} \quad (11.12)$$

This can then be substituted into 11.11 to give equation 11.5.

I have rewritten equation 11.5 in order to estimate R'_F in terms of R^* , U , V and $P(E+)$. This equation can be seen in (11.13)

$$R'_F = \frac{\left[\frac{(1-U)P(E+) + VP(E-)}{UP(E+) + (1-V)P(E-)} \right] (1-V)P(E-) - VP(E-)}{R^* (1-U)P(E+) - \left[\frac{(1-U)P(E+) + VP(E-)}{UP(E+) + (1-V)P(E-)} \right] UP(E+)} \quad (11.13)$$

From this equation the adjusted relative risks R'_F can then be estimated. Table 11.2 shows the relative risks before and after adjustment, along with the relative risk for individual deprivation estimated from the known data. Individual deprivation scores were divided into 'deprived' if the individual had one or more deprivation characteristics, and 'not deprived' if they had no deprivation characteristics. The area deprivation scores were divided into two groups, based on the same proportions as seen in the individual data and called 'deprived' and 'not deprived'.

The sensitivity is evidently much lower than the specificity. This means that area deprivation is not good at predicting individuals who are deprived, with over 50% wrongly classified. It is slightly better at predicting individuals who are not deprived but 30% are still wrongly classified. The pattern of the sensitivities and specificities is very similar for males and females which would be expected as individual deprivation is based on the household rather than the individual. There is a slight decrease in sensitivity and an increase in specificity with increasing age. The prevalence of deprivation generally decreases with age but is higher in the 55-64 year old age group. It is very similar for males and females.

The table also demonstrates an increasing relative risk with age at the area level for both males and females. The relative risks are higher for individual deprivation as seen in the analyses of odds ratios in chapter 9. The pattern with age is similar, although some differences may be due to changes in the prevalence of deprivation with age. The estimated relative risks using equation 11.13 (R'_F) are much higher than those using the known data about individual

deprivation and illness (R). The method appears to be overadjusting for the effect of deprivation at the individual level in all but one case (females 16-24 years). The difference is over threefold in females aged 55-64 years. The over-adjustment may be due to strong assumptions made by the method. However before the assumption of non-differential misclassification is investigated further, the two other methods will be discussed.

Table 11.2: Table of sensitivity (U), specificity (V), prevalence of exposure ($P(E+)$), relative risk estimated using area deprivation R^* , relative risk estimated using individual deprivation R , and estimated relative risk based on equation (11.13) R'_F for each age and sex strata.

	Sensitivity U	Specificity V	Exposure prevalence $P(E+)$	R^* (area)	R (individual)	R'_F
Males						
16-24	45%	70%	46%	1.12	1.65	2.00
25-34	46%	69%	39%	1.19	2.14	3.04
35-44	44%	72%	30%	1.25	2.34	3.81
45-54	44%	74%	28%	1.32	2.09	4.33
55-65	44%	74%	35%	1.19	1.64	2.53
Females						
16-24	47%	70%	44%	1.04	1.70	1.25
25-34	49%	67%	35%	1.25	2.00	4.06
35-44	43%	72%	27%	1.23	1.98	3.80
45-54	43%	73%	28%	1.24	1.97	3.61
55-65	42%	73%	36%	1.31	1.61	5.20

A drawback with this method, other than the apparent over adjustment, is that there is no easy method of obtaining standard errors for the relative risk estimate. The delta method could be used to estimate the standard errors but this would be complex and a modelling method which fitted a model and automatically calculated an adjusted standard error would be more useful. Also it is not possible using this method to look at more than one stratum at a time and use models to collapse over strata. For example, a model fitting procedure would allow the possibility to test if making the assumption that the sensitivity was the same for males and

females would reduce how well the model fitted. The following two methods are based on model fitting. These methods can be used to develop the analysis further and produce standard errors for the relative risk estimates. Also Flegal et al (1986) found that when there is misclassification of exposure, variation in the prevalence of the true exposure can lead to spurious trends in a third variable such as age. Although the prevalence of deprivation did vary with age, the trend of increasing relative risk of deprivation with increasing age was also apparent when looking at individual deprivation.

11.5 ESPELAND AND HUI'S METHOD

Espeland and Hui (1987) present a method for analysing epidemiological data in the presence of misclassification. They propose the use of log-linear models and maximum likelihood estimation to adjust estimates of exposure effect, incorporating information from the table of misclassification. The notation defined in 11.3.5 is used to explain the method. They fit a model to the data given by the two by two tables D by E^* (i.e. long-term illness by area deprivation) and E by E^* (i.e. individual deprivation by area deprivation), to produce a table of estimates for D by E (i.e. long-term illness by individual deprivation) and estimate the relative risk at the individual level (R'_{ES}). They use data gathered by Diamond and Lillienfeld (1962) to demonstrate their method. Although they use data from a case control study, the principle of adjusting for misclassification is the same but here a relative risk will be calculated instead of an odds ratio.

The method is based on writing the data in the form shown in table 11.3. The $x=[x_{ijkl}]$ refer to the complete cross-classification table which is unknown where i indicates the level of D , j indicates the level of E , k indicates the level of E^* and l indicates the level of L which indicates whether the data originates from the original study or the resampled misclassification data. Here the study data and supplemental data are from the same study.

The process uses a loglinear model to partition the 8 observed cells y into the 16 unobserved completely cross classified cells x where $y_1=x_{1111}+x_{1211}$, $y_2=x_{2111}+x_{2211}$ etc. They refer to this model as the miscategorization model. Fisher's scoring algorithm is used to estimate x which is assumed to be from a Poisson or multinomial distribution. This can be performed in SAS using

the program seen in appendix I. It provides an estimate of the fully crossclassified table which can be collapsed over E^* to give the table D by E and an estimate of the relative risk at the individual level R'_{Es} .

Table 11.3: Espeland and Hui's partitioning of the study data.

Area deprivation E^*	Individual deprivation E	Study data $l=1$ D		Supplemental data $l=2$ D	
		III (+) $i=1$	Not ill (-) $i=2$	III (+) $i=1$	Not ill (-) $i=2$
Yes (+) $k=1$	Yes (+) $j=1$	x_{1111}	x_{2111}	x_{1112}	+ x_{2112}
		+	+		
	No (-) $j=2$	x_{1211}	x_{2211}	x_{1212}	+ x_{2212}
No (-) $k=2$	Yes (+) $j=1$	x_{1121}	x_{2121}	x_{1122}	+ x_{2122}
		+	+		
	No (-) $j=2$	x_{1221}	x_{2221}	x_{1222}	+ x_{2222}

However in a reader's discussion of this method, Ekholm (1991) points out that this method is fundamentally flawed because it relies on an incorrect assumption. They verify a general rule by which the fitted values for x are produced. Equation 11.7 shows this rule written in terms of probabilities.

$$P(DEE^*) = P(DE^*)P(EE^*)/P(E^*) \quad (11.14)$$

The only way of satisfying this rule is to assume that D and E are conditionally independent given E^* . This is shown in the following steps:

$$\begin{aligned}
 P(DEE^*) &= P(DE|E^*)P(E^*) \\
 &= P(D|E^*)P(E|E^*)P(E^*) \text{ (assuming } D \text{ and } E \text{ conditionally independent)} \\
 &= P(DE^*)P(EE^*)/P(E^*)
 \end{aligned}$$

This assumption means that for a given level of area deprivation, individual deprivation provides no further information about illness status. However this is incorrect since chapter 9 has shown that after adjusting for individual deprivation, area deprivation adds no further information about illness status. Therefore the assumption that should be made is that D and

E^* are conditionally independent given E . The effect of using this incorrect assumption will be demonstrated by applying the model to the SAR data.

11.5.1 THE ESPELAND AND HUI MODEL APPLIED TO SAR DATA

A full description of the partitioning applied to males aged 16-24 is shown here and then final results for all age and gender strata are displayed. Table 11.4 gives the observed data for males aged 16-24 years as written by Espeland and Hui.

Table 11.4: SAR data for males aged 16-24 tabulated in Espeland and Hui format

Area	Individual	Study data		Supplemental data	
deprivation	deprivation	$L=1$		$L=2$	
E^*	E	D		D	
		Ill	Not ill	Ill	Not ill
Yes	Yes	447	16307	9384	
	No			7370	
	Yes	689	28189	11404	
No	No			17474	

Using the program seen in Appendix I the matrix x is obtained. The values of x can be used to draw up table 11.5 of maximum likelihood estimates as follows

Table 11.5: Maximum likelihood estimates of x for males aged 16-24

				D		D	
				+	-	+	-
E^*	+	E	+	250.37	9133.63	250.37	9133.63
			-	196.63	7173.37	196.63	7173.37
	-	E	+	272.09	11131.91	272.09	11131.91
			-	416.91	17057.09	416.91	17057.09
				$L=1$		$L=2$	

The estimates are the same for $L=1$ and $L=2$ since the supplemental data is from the same study. These would be different if using misclassification data from a supplemental study. Collapsing the data across factor E^* leads to the estimated frequency table D by E (table 11.6) which can be used to calculate the estimate of the relative risk at the individual deprivation level, R'_{Es} .

Table 11.6: Estimated table D by E for males aged 16-24 years

Individual deprivation (E)	Illness status (D)	
	Ill	Not ill
Yes	522	20266
No	614	24230
	Relative risk $R'_{Es}=1.02$	

Table 11.7 shows the results of all ten strata with the area level relative risk R^* , the known individual level relative risk R , and the estimate of R'_{Es} given by Espeland and Hui's method.

Table 11.7: Table of relative risk estimates based on area deprivation (R^*), individual deprivation (R) and estimate based on the Espeland and Hui model (R'_{Es})

Age group	Males			Females		
	R^*	R	R'_{Es}	R^*	R	R'_{Es}
	(area)	(individual)		(area)	(individual)	
16-24	1.12	1.65	1.02	1.04	1.70	1.00
25-34	1.19	2.14	1.03	1.25	2.00	1.04
35-44	1.25	2.34	1.04	1.23	1.98	1.02
45-54	1.32	2.09	1.05	1.24	1.97	1.03
55-65	1.19	1.64	1.03	1.31	1.61	1.04

All of the estimates based on the loglinear model by Espeland and Hui demonstrate a reduction in the estimate of the relative risk, tending in each case towards the null which is opposite to the observed trend. It also contradicts the conclusions of both Flegal et al (1986) who found that using nondifferential misclassified data led to an underestimate not an overestimate of the true relative risk and Mertens (1993) who showed a similar pattern. The method is adjusting the estimates in the opposite way to Flegal et al (1986). This is due to the incorrect

assumption of the data shown by Ekholm (1991), that factors D and E are conditionally independent given the value of E^* . Ekholm discusses this incorrect assumption and suggests an alternative method which is discussed next.

11.6 EKHOLM'S METHOD

In Ekholm's (1991) critique of the method adopted by Espeland and Hui to analyse misclassification problems, he introduces the use of exponential family non-linear models. This model fitting approach provides the opportunity to estimate standard errors for the relative risk estimates and also to test various hypotheses about different strata. A full description of the use of non-linear models is given in (Ekholm and Palmgren, 1987; Palmgren and Ekholm, 1987; Ekholm, Green and Palmgren, 1986; Ekholm and Green, 1995), but here an introduction to the type of models used is given and then the method is adapted to analyse the SAR data.

11.6.1 PARAMETERISATION OF MODEL

The terminology described in 11.3.5 is adopted here. As before D represents disease status i.e. long-term illness, E represents the true exposure status, i.e. individual deprivation and E^* represents the exposure status measured with error, i.e. area deprivation, where + denotes presence and - denotes absence of the characteristic. The method is based on parameterising the frequency tables and then estimating these parameters using the data available. Firstly five structural and conditional probabilities are defined as follows:

$$\pi = P(E+) \quad \text{Probability of being individually deprived} \quad (11.15)$$

$$\delta^+ = P(E^+|E+) \quad \text{Sensitivity} \quad (11.16)$$

$$\delta^- = P(E^+|E-) \quad \text{1-Specificity} \quad (11.17)$$

$$\gamma^+ = P(D+|E+) \quad \text{Probability of being diseased given individually deprived} \quad (11.18)$$

$$\gamma^- = P(D+|E-) \quad \text{Probability of being diseased given not individually deprived} \quad (11.19)$$

where (γ^+/γ^-) = adjusted relative risk for individual deprivation

The counts of the tables D by E^* and E^* by E can then be written as seen in tables 11.8 and 11.9 in terms of probabilities 11.15-11.19, where N_1 and N_2 are the total number of observations in the two tables.

Table 11.8: Parameterisation of table D by E^* using Ekholm's method

E^*	D	
	+	-
+	$N_1(\pi\delta^+\gamma^+(1-\pi)\delta^-\gamma^-)$	$N_1(\pi\delta^+(1-\gamma^+)+(1-\pi)\delta^-(1-\gamma^-))$
-	$N_1(\pi(1-\delta^+)\gamma^+(1-\pi)(1-\delta^-)\gamma^-)$	$N_1(\pi(1-\delta^+)(1-\gamma^+)+(1-\pi)(1-\delta^-)(1-\gamma^-))$

Table 11.9: Parameterisation of table E by E^* using Ekholm's method

E^*	E	
	+	-
+	$N_2\pi\delta^+$	$N_2(1-\pi)\delta^-$
-	$N_2\pi(1-\delta^+)$	$N_2(1-\pi)(1-\delta^-)$

These parameterisations are based on general probability assumptions based on Bayes theorem, where X , Y and Z are binary events :

$$P(XY)=P(X)*P(Y|X)=P(Y)*P(X|Y)$$

$$P(Y|X)=P(Y)*P(X|Y)/P(X)$$

$$P(Y|X)=\sum(P(Y|XZ)P(Z)) \quad Z=1,2$$

The parameterisation relies on the assumption that E^* is conditionally independent of D given E i.e. given individual deprivation, area deprivation adds no information about illness status.

For example:

$$\begin{aligned}
 P(E^*+D+) &= P(E^*+D+|E+)P(E+)+P(E^*+D+|E-)P(E-) \\
 &= P(E^*+|E+)P(D+|E+)P(E+) + P(E^*+|E-)P(D+|E-)P(E-) \\
 &= \pi\delta^+\gamma^++(1-\pi)\delta^-\gamma^-
 \end{aligned}$$

For the purpose of this model fitting however, the model will be reparameterised in terms of probabilities 11.15-11.18 and the logarithm of the relative risk:

$$\Gamma = \log(\gamma^+/\gamma^-) = \log(P(D+|E+)/P(D+|E-)) \quad (11.20)$$

Therefore γ^- in tables 11.8 and 11.9 is replaced with $\gamma^+/\exp(\Gamma)$. The model will then give immediately a standard error for the logarithm of the relative risk. These parameterisations shown in tables 11.8 and 11.9 show combinations of additive and multiplicative relationships. Therefore it is not possible to fit a generalised linear model (GLM) to this data. Therefore

Ekholm introduces the concept of non-linear models to fit the data. The principles behind these models are described in the following section.

11.6.2 DESCRIPTION OF EXPONENTIAL FAMILY NON-LINEAR MODELS

In a generalised linear model (GLM) as described by McCullagh and Nelder (1988) the aim is to describe the observed data Y_i , $i=1, \dots, n$ (response variable) in terms of a function of a vector of known and fixed explanatory variables $\mathbf{x}_i=(x_{i1}, \dots, x_{ip})^T$ (covariates) and a vector of unknown parameters $\boldsymbol{\beta}=(\beta_1, \dots, \beta_p)^T$. Such a model is defined by three components:

- 1) a probability function $f(\mathbf{y})$ for the response variable \mathbf{y} , where \mathbf{y} is dependent on the mean μ and other possible parameters.
- 2) a linear predictor in the p explanatory variables

$$\eta_i = \mathbf{x}_i^T \boldsymbol{\beta} = x_{i0}\beta_0 + x_{i1}\beta_1 + \dots + x_{ip}\beta_p \quad (x_{i0}=1)$$

- 3) a link function $g(\mu)=\eta$ which relates the linear predictor to the mean μ .

The mean μ can be written in terms of a function $h(\cdot)$ such that:

$$\mu_i = h(\mathbf{x}_i^T \boldsymbol{\beta}) \quad i=1, \dots, n$$

where each μ_i is connected to a linear expression of the parameters.

In an exponential family non-linear model the predictor is a not-necessarily-linear function of \mathbf{x} and $\boldsymbol{\beta}$ such that

$$\eta_i = c(\mathbf{x}_i, \boldsymbol{\beta}) \text{ with } \mu_i = h(\mathbf{x}_i, \boldsymbol{\beta}) \quad i=1, \dots, n$$

As seen in the tables 11.8 and 11.9 it is not possible to write the Y_i from the misclassification data in the form of a linear function since they are a combination of additive and multiplicative terms. These η_i are referred to as 'observationwise defined' since the function $c(\cdot)$ can vary with each observation.

It has been shown by Ekholm and Palmgren (1987) that the maximum likelihood estimates of $\boldsymbol{\beta}$ can be obtained by using iterative reweighted least squares (IRLS) which solves the linear equations of the form

$$(\mathbf{D}_t^T \mathbf{V}_t \mathbf{D}_t) \boldsymbol{\beta}_{t+1} = \mathbf{D}_t^T \mathbf{V}_t \mathbf{z}_t \quad t=0, 1, 2, \dots$$

where \mathbf{D}_t is an n by p local model matrix $d_{iv} = \delta \mu_i / \delta \beta_v$, $i=1, \dots, n$; $v=1, \dots, p$, \mathbf{V}_t is a local weight matrix and \mathbf{z}_t is an adjusted dependent variable. All of these are functions of the current

parameter values of β_i . If y and μ are substituted for the observed and fitted values of the observed data $Y=(Y_1, \dots, Y_n)^T$ then z_i can be written

$$z_i = D' \beta_i + (y - \mu_i)$$

The weight matrix V is a diagonal matrix $n \times n$ with elements $v_i = 1/\text{Var}(Y_i)$. So each iteration of the scoring algorithm can be expressed as a weighted least squares regression of the adjusted dependent variable z_i on the explanatory variables x with weight vector V . The elements of D , d_{iv} can be factorised as follows

$$\delta \mu_i / \delta \beta_v = \delta \mu_i / \delta \eta_i * \delta \eta_i / \delta \beta_v$$

In a generalised linear model this equates to

$$\delta \mu_i / \delta \beta_v = h'(x_i^T \beta) x_{iv}$$

where $\delta \eta_i / \delta \beta_v$ are just the explanatory variables x . However in a non-linear model the $\delta \eta_i / \delta \beta_v$'s are dependent on β and they must be updated at each iteration. The method of Ekholm and Palmgren (1987) uses approximations of the numerical derivatives to get over this problem, calculated from the equation

$$\delta \mu_i / \delta \beta_v \approx (h_i(x_i, \beta_i + \delta_v) - h_i(x_i, \beta_i)) / 2\delta_v \quad i=1, \dots, n \text{ and } v=1, \dots, p$$

where $\delta_v = (0, \dots, 0, \delta, 0, \dots, 0)^T$ is a $p \times 1$ vector with v th element δ_v . In this particular example as in most non-linear models, the link $g(\mu)$ is the identity link and so $\delta \mu_i / \delta \beta_v = \delta \eta_i / \delta \beta_v$ and

$$\delta \eta_i / \delta \beta_v \approx (c_i(x_i, \beta_i + \delta_v) - c_i(x_i, \beta_i)) / 2\delta_v \quad i=1, \dots, n \text{ and } v=1, \dots, p$$

11.6.3 FITTING THE MODEL IN GLIM

Ekholm et al (Ekholm, Green and Palmgren, 1986; Ekholm and Green, 1995) have devised macros for the statistical package GLIM to fit this type of non-linear model. I have adapted these macros for use in this example and these can be seen in appendix J. Since the data are from two contingency tables, the Poisson error function is used. I wrote the macro ETA for the misclassification problem which specifies the function $\eta_i = c(x_i, \beta)$. Using an identity link, where $\eta_i = \mu_i$, the macro is a GLIM coding of the model formula based on the parameterisation in equations 11.15-11.18 and 11.20 with 5 parameters for each age and sex strata. These parameters are referred to as P_i rather than X_i . For example, the observed number of individuals who are ill and area deprived can be written

$$\begin{aligned} N_i * P(D+E*+) &= N_i (\pi \delta^+ \gamma^+ + (1-\pi) \delta^- \gamma^-) \\ &= N_i * (P(1) * P(2) * P(4) + (1-P(1)) * P(3) * \exp(P(4)/P(5))) \end{aligned}$$

The macro NLIN finds the number of elements in the vector P of initial estimates and sets up the first version of the local model matrix P , (referred to previously as D) by setting up P vectors of length n using the macro DER which is run p times. This sets up the local model matrix of size n by p , using the initial estimates of P that I supplied. Then macro NDER calculates the numerical derivatives approximated using the approximation equation for each vector $P(1)$ upto $P(p)$. The value of δ_v is initially set to 0.001 and the upper and lower bounds are calculated from macro ETA. The model is then fitted using FIRST to set up the initial estimates. Then M1 is used to estimate β and the numerical approximations are calculated until convergence. The macros had to be adapted from those presented by Ekholm and Green (1995) since I needed to model a large number of parameters which GLIM does not permit. Therefore I had to copy the vectors $P(i)$ into a matrix m . In terms of fitting a model GLIM regarded this matrix as a single parameter although it outputs estimates for each $P(i)$ and this adaptation permitted the inclusion of a virtually unlimited number of parameters in the model. Table 11.10 shows the results of the model fitting for males aged 16-24.

Table 11.10: Parameter estimates for Males aged 16-24 years based on Ekholm model

		Parameter estimate	Standard error
π	$P(E+)$	0.456	0.0023
δ^+	$P(E^+ E+)$	0.451	0.0030
δ^-	$P(E^+ E-)$	0.297	0.0025
γ^+	$P(D+ E+)$	0.034	0.0052
$\log(\gamma^+/\gamma^-)$	$\log(P(D+ E+)/P(D+ E-))$	0.692	0.3949

This process was then repeated for each stratum. I then extended the program seen in appendix J for males and females of all age groups, fitting 5 parameters for each stratum. Table 11.11 shows the estimates and 95% confidence intervals for the relative risk for each stratum. Like the method used by Flegal et al (1986), this method is again overestimating the relative risk in each case. When looking at the confidence intervals it can be seen that they are extremely wide and offer little certainty. The next step would be to refine the model fitted and look at the deterioration in fit when parameters were fixed across age or sex or linear functions of age were introduced. However since the adjusted relative risks are all overestimating the exposure effect further investigation of the principal assumption of non-differential misclassification is necessary before any further modelling is performed.

Table 11.11: Relative risk for individual data and estimated relative risk adjusted for misclassification for all strata

Age group	Males			Females		
	R^* (area)	R (individual)	R'_{EK} 95% CI	R^* (area)	R (individual)	R'_{EK} 95% CI
16-24	1.12	1.65	2.00 (0.92-4.33)	1.04	1.70	1.25 (0.57-2.75)
25-34	1.19	2.14	3.04 (1.59-5.84)	1.25	2.00	4.06 (1.77-9.34)
35-44	1.25	2.34	3.81 (2.26-6.43)	1.23	1.98	3.80 (1.94-7.44)
45-54	1.32	2.09	4.33 (2.83-6.63)	1.24	1.97	3.61 (2.09-6.23)
55-64	1.19	1.64	2.52 (1.74-3.65)	1.31	1.61	5.20 (2.28-11.9)

11.7 SUMMARY OF RESULTS FROM THE THREE METHODS

The table 11.12 shows the relative risk estimates obtained from the three methods demonstrated in these analyses, alongside the estimate from the original data. This table shows that the method of Espeland and Hui, because of its incorrect assumption that given area deprivation, individual deprivation is independent of illness, is adjusting the relative risk in the wrong direction. The methods of Ekholm et al and Flegal et al, using the correct assumption that given individual deprivation, area deprivation is independent of illness, are adjusting the relative risk in the right direction. They also give identical estimates since they are built on the same assumptions. The method of Ekholm extends the work of Flegal enabling the estimation of confidence intervals for the relative risk. However both of these methods are overestimating the effect of individual deprivation when compared with the known data, by up to four times. The principal assumption of these methods is that of non-differential misclassification, i.e. that the misclassification is not different for diseased and non-diseased. This assumption and its effect on the data will now be investigated.

Table 11.12: Relative risk estimates based on individual deprivation (R), area deprivation (R^*) and area deprivation adjusted for misclassification by the Flegal et al method (R'_F), Espeland and Hui method (R'_{ES}) and Ekholm method (R'_{EK}).

	Individual data	Area data	Flegal	Espeland	Ekholm
Males	R	R^*	R'_F	R'_{ES}	R'_{EK}
16-24	1.65	1.12	2.00	1.02	2.00
25-34	2.14	1.19	3.04	1.03	3.04
35-44	2.34	1.25	3.81	1.04	3.81
45-54	2.09	1.32	4.33	1.05	4.33
55-64	1.64	1.19	2.53	1.03	2.52
Females					
16-24	1.70	1.04	1.25	1.00	1.25
25-34	2.00	1.25	4.06	1.04	4.06
35-44	1.98	1.23	3.80	1.02	3.80
45-54	1.97	1.24	3.61	1.03	3.61
55-64	1.61	1.31	5.20	1.04	5.20

11.8 NONDIFFERENTIAL MISCLASSIFICATION ASSUMPTION

One of the reasons for the over adjustment by Flegal's and Ekholm's methods could be the strong assumption of non-differential misclassification. This means that the degree of misclassification i.e. the sensitivity and specificity of the misclassification table should be the same for diseased and non-diseased. For example, it is possible that people with long-term illness become unemployed due to their illness, increasing their individual deprivation, but do not change their area of residence. This and other similar scenarios could lead to differential misclassification.

Table 11.13 shows the specificity and sensitivity for the misclassification table by each age and sex stratum and also by illness status. This shows that when looking at the overall sensitivity collapsed over age and sex, it is slightly higher for the long-term ill than those who have no illness and the difference is formally significant ($\chi^2_1 = 14.6$ $P < 0.001$), but the effect appears to be relatively small with an overall difference of 2% in sensitivity. In terms of specificity, there

is a much more variable pattern between the groups, but overall there is no evidence of a difference with only 1.4% difference between those who are ill and those who are not ($\chi^2_1 = 2.7$ $P=0.102$). The sensitivities and specificities for those who are ill are based on much lower numbers than for those who are not ill and in all strata the percentage difference between the two groups (e.g. sensitivity in the diseased / sensitivity in the not diseased * 100), is less than 10%.

Table 11.13: Sensitivity and specificity by disease status for each age and sex strata

	Sensitivity		Specificity	
	Diseased	Not disease	Diseased	Not diseased
Males				
16-24	45.5%	45.1%	69.2%	70.4%
25-34	49.7%	46.2%	70.2%	67.3%
35-44	45.9%	43.9%	70.0%	71.9%
45-54	47.6%	43.5%	71.2%	71.0%
55-64	46.1%	43.5%	72.2%	73.9%
Females				
16-24	49.0%	47.1%	75.2%	69.7%
25-34	51.6%	48.7%	64.7%	67.3%
35-44	46.0%	42.7%	70.1%	72.0%
45-54	47.1%	42.9%	71.4%	72.7%
55-64	44.3%	42.2%	67.0%	73.8%

In order to see whether it is the slight changes in sensitivity and specificity that lead to these large differences between the estimate of the relative risk at individual level R and the relative risk adjusted for misclassification R' , I have adapted the equation developed by Flegal et al (11.5) to allow for differential sensitivity and specificity. This relative risk adjusted for differential misclassification is denoted by R'_d . The equation used by Flegal et al is the same as that employed by the model based method of Ekholm et al and so modifying the equation will give the same results as adapting the model fitting method to allow for differential misclassification. Equation 11.5 for R^* can be seen below.

$$R^* = \frac{[UR'_F P(E+) + (1-V)P(E-)][(1-U)P(E+) + VP(E-)]}{[UP(E+) + (1-V)P(E-)][(1-U)R'_F P(E+) + VP(E-)]} \quad (11.5)$$

In order to adjust for differential misclassification I have rewritten equation 11.5 to allow for differences in sensitivity and specificity between those who are ill and those who are not.

Equations 11.7 and 11.8 defined the sensitivity U as $P(E^+|E+)$ and the specificity V as $P(E^-|E-)$.

The sensitivity and specificity in the diseased group denoted by U_1 and V_1 respectively can be defined as:

$$U_1 = P(E^+|E+D+) \quad (11.21)$$

$$V_1 = P(E^-|E-D+) \quad (11.22)$$

The relative risk based on area data can be written as follows

$$R^* = \frac{P(D+E^+)}{P(D+E^-)} = \frac{P(D+E^+)/P(E^+)}{P(D+E^-)/P(E^-)} \quad (11.23)$$

The elements of equation 11.23 can then be rewritten as follows

$$P(D+E^+) = P(E^+|D+E+)P(D+E+) + P(E^+|D+E-)P(D+E-) \quad (11.24)$$

$$P(D+E^-) = P(E^-|D+E+)P(D+E+) + P(E^-|D+E-)P(D+E-) \quad (11.25)$$

$$P(E^+) = P(D+E^+) + P(D-E^+) \quad (11.26)$$

Since

$$P(E^+|E+) = P(E^+|E+D+)P(D+E+) + P(E^+|E+D-)P(D-E+) \quad (11.27)$$

$$P(E^-|E+) = P(E^-|E+D+)P(D+E+) + P(E^-|E+D-)P(D-E+) \quad (11.28)$$

therefore

$$P(E^+) = P(E^+|E+)P(E+) + P(E^-|E+)P(E-) = UP(E+) + (1-V)P(E-) \quad (11.29)$$

If $P(D+E^+)$ (11.24) and $P(D+E^-)$ (11.25) are then divided by $P(D+E-)$ and substituted into equation 11.23 then

$$R^* = \frac{P(E^+|D+E+) \frac{P(D+E+)}{P(D+E-)} P(E+) + P(E^+|D+E-) P(E-)}{P(E^-|D+E+) \frac{P(D+E+)}{P(D+E-)} P(E+) + P(E^-|D+E-) P(E-)} \quad (11.30)$$

If U , V , U_1 and V_1 are substituted into this equation along with R'_d then equation 11.31 is obtained.

$$R^* = \frac{[U_1 R'_d P(E+) + (1-V_1)P(E-)][(1-U)P(E+) + VP(E-)]}{[UP(E+) + (1-V)P(E-)][(1-U_1)R'_d P(E+) + V_1 P(E-)]} \quad (11.31)$$

This can be then be written for calculation of the adjusted relative risk R'_d in (11.32)

$$R' = \frac{\left[\frac{UP(E+) + (1-V)P(E-)}{(1-U)P(E+) + VP(E-)} \right] V_1 P(E-) R^* - (1-V_1) P(E-)}{U_1 P(E+) - \left[\frac{UP(E+) + (1-V)P(E-)}{UP(E+) + (1-V)P(E-)} \right] (1-U_1) P(E+) R^*} \quad (11.32)$$

If the values for each stratum are entered into this equation the results shown in table 11.14 are obtained.

Table 11.14: Estimates of relative risk based on individual deprivation (R), based on area deprivation adjusted for nondifferential misclassification (R'_F) and based on area deprivation adjusted for differential misclassification (R'_d)

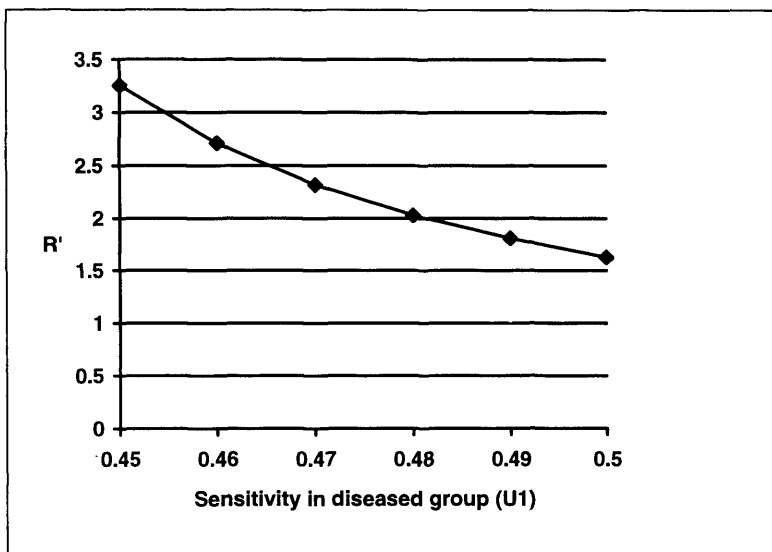
	Males			Females		
Age group	R	R' _F	R' _d	R	R' _F	R' _d
16-24	1.65	2.00	1.64	1.70	1.25	1.70
25-34	2.14	3.04	2.14	2.00	4.06	2.00
35-44	2.34	3.81	2.34	1.98	3.80	1.98
45-54	2.09	4.33	2.09	1.97	3.61	1.97
55-64	1.64	2.53	1.62	1.61	5.20	1.61

This shows that if the differential misclassification is allowed for in the equation then the estimate of the relative risk is extremely accurate. Hence, it appears that the assumption of nondifferential misclassification was the reason for the poor estimate of the relative risk and that the differences are due to differences in sensitivity and specificity. However the effect of very small amounts of differential misclassification has a very large effect on the adjusted relative risk. For example in males aged 16-24 years where the sensitivity differs between the diseased and the not diseased groups by only 0.4% and the specificity differs by 1.2% this leads to a relative risk estimate of 2.00 instead of 1.65 which is overestimating the relative risk by a factor of 1.2. In other strata, slightly bigger differences in sensitivity and specificity lead to 2 and 3 fold difference between the relative risk adjusted for misclassification and the relative risk based on individual data.

In most types of epidemiological studies, where the aim is to adjust for misclassification, it would be extremely difficult to estimate the sensitivity and specificity for both the diseased and not diseased groups to such a precise level. Furthermore, it is rare to be able to assess whether

there is differential misclassification since the full table of D by E by E^* (illness by individual deprivation by area deprivation) needs to be known to assess it. If only a small sample of individuals are assessed to give this three way table to adjust the D by E^* (illness by area deprivation) it would be hard to estimate the misclassification to such a precise degree. The small differences in sensitivity and specificity seen in males aged 16-24 years suggests that this method is of little use. Figure 11.1 demonstrates in a more general case the effect of differential misclassification on the relative risk estimate.

Figure 11.1: Estimate of adjusted relative risk R' for different values of sensitivity in the diseased group U_1 , where $U=0.45$, $V=0.7$, $R^*=1.2$, $R=3.25$, $P(E)=0.38$



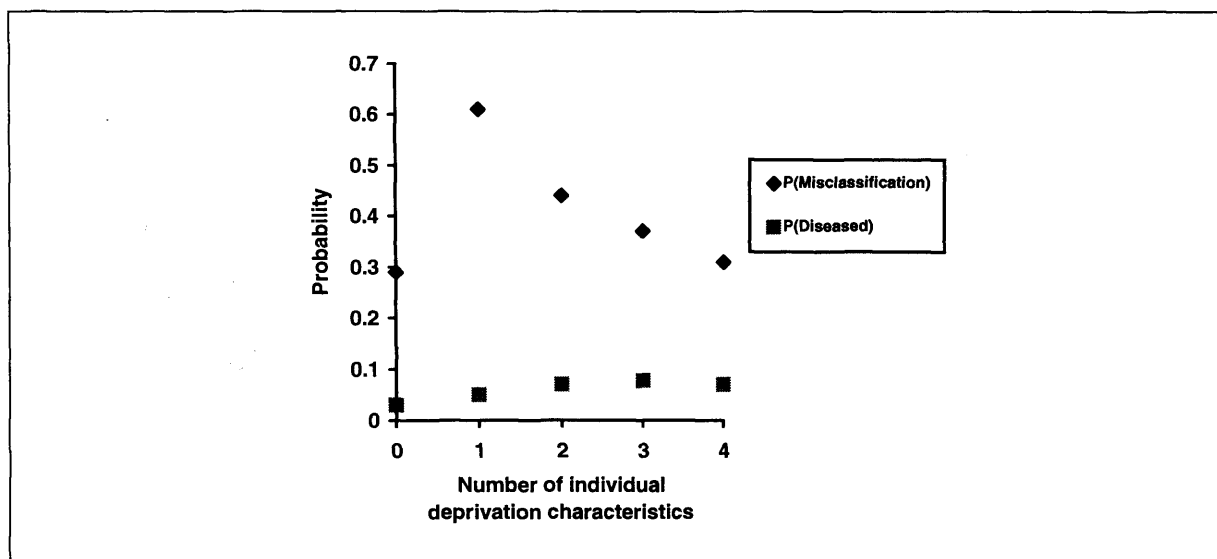
This shows how small differences in the sensitivity lead to large changes in the estimated relative risk R' . A relative sensitivity (sensitivity in the diseased/sensitivity in the undiseased) of 1.1 leads to halving the estimate of the relative risk. Specificity had an even greater effect but working in the reverse direction. For comparison, a relative specificity of 0.9 lead to an estimate a third of the size of that assuming nondifferential misclassification. So both a relative sensitivity of greater than one and relative specificity of less than one can easily lead to a relative risk a quarter of the size estimated by assuming nondifferential misclassification.

A possible reason for these large differences in the estimated relative risks from small changes in sensitivity and specificity could be due to the low level of sensitivity and specificity of deprivation classification. Therefore similar calculations were made for higher levels of sensitivity and specificity for the same level of relative sensitivity and specificity based on

$R^*=1.2$. For an overall sensitivity of 80%, a relative sensitivity of 1.1 leads to a reduction of 30% in the relative risk compared to 50% when the sensitivity was 45%, and an overall sensitivity of 90% leads to a 27% reduction. Therefore the effect is diminished with higher sensitivity but is still considerable.

It is possible that the differential misclassification is due to true differences between the diseased and not diseased. For example as discussed before, those who become ill may be more likely to become individually deprived through unemployment but they may stay in the same area of residence. However it is possible that the differential misclassification seen here has arisen from nondifferential errors in exposure measurement. Flegal et al (1991) have shown that the categorisation of a continuous exposure variable can lead to differential misclassification. The probability of being individually deprived is plotted against the probability of misclassification in figure 11.2.

Figure 11.2: Probability of misclassification and illness for different levels of individual deprivation.



As individual deprivation was reduced to a dichotomous variable, individuals near the cut-off are more likely to be misclassified. For example, people with one characteristic are more likely to be misclassified than people with four characteristics. When this is combined with the increasing probability of illness with deprivation, differential misclassification can occur. This is because among those who are classified as individually deprived, individuals near the cut-off point will have a lower probability of disease but a high probability of misclassification, while

those with four characteristics and away from the cut-off will have a higher probability of disease but a lower probability of misclassification. Therefore individuals with the disease will be less likely to be misclassified than individuals without the disease. Hence this leads to differential misclassification. Among those who are not deprived, differential misclassification is less of a problem since only people with no characteristics were included.

11.9 DISCUSSION

This chapter has looked at the possibility of adjusting observed relative risk estimates for misclassification using three different methods. The previous chapters had shown that the degree of misclassification is relatively high and that this is leading to an underestimation of the relative risk of deprivation.

The method of loglinear modelling proposed by Espeland and Hui (1987) has been shown to have a fundamental flaw since it assumes that individual deprivation is independent of disease, i.e. offers no information on disease status, if area deprivation is known, which is the opposite case to the situation that needs to be modelled. This incorrect assumption has led to the adjusted relative risks tending towards the null, the opposite direction to the known relative risk. This problem has been indicated by Ekholm (1991) and my analysis has confirmed their findings. The use of this method seems to be doubtful in this particular situation.

The methods proposed by Ekholm (1986) and Flegal et al (1986) are both based on the correct assumption that area deprivation is independent of illness if individual deprivation is known. Their methods link the sensitivity and specificity of the misclassification table, the prevalence of exposure and the observed relative risk at the area level to produce an estimate of the relative risk at the individual level. Flegal et al (1986) use an equation to estimate the relative risk while Ekholm et al (1986) use exponential family non-linear models to do this. The latter method offers the opportunity of calculating standard errors for the estimated relative risks very simply and also provides the facility to test the fit of different models such as a model with fixed sensitivities for both genders or a linear relationship between prevalence of deprivation and age.

These two methods adjusted the relative risk in the correct direction but they appeared to overestimate the effect of deprivation at the individual level. With further investigation, it seems that the reason for this overestimation is due to differential misclassification. Very small differences in the rates of sensitivity and specificity between the diseased and not diseased led to large differences in the estimate of the relative risk. This made the estimate of the relative risk of no practical use since it would be very difficult to be able to calculate the misclassification in each group so precisely. In studies with higher levels of sensitivity and specificity it appears that the effect of differential misclassification would be lessened but would still be a considerable effect.

Although it is possible that there is true differential misclassification with those who are long-term ill becoming unemployed but not moving from their area of residence, it is likely that the differential misclassification has arisen from nondifferential errors in exposure measurement. Higher levels of misclassification combined with lower levels of illness lead to differential misclassification. Flegal et al (1991) does not recommend the use of methods to correct for misclassification assuming nondifferential misclassification where categories have been formed from a continuous covariate. This is because the assumption that random nondifferential error leads to nondifferential misclassification may result in estimates of the relative risk which are far higher than the true relative risk. This has been confirmed in this analysis with two fold differences due to the assumption of nondifferential misclassification..

These results confirm the opinions of Flegal et al (1991) indicating that the methods used to adjust relative risks for misclassification rely on the vital assumption of nondifferential misclassification. It is extremely difficult to verify that this assumption is true and incorrect usage of methods assuming nondifferential misclassification can lead to large errors in relative risk estimation. In general, estimation of relative risks is far more reliable if an attempt is made to try and improve the data that is collected for the study. Therefore it is better to collect data at an individual level and be more assured in the calculation of the relative risks of deprivation than to collect misclassified data and try to adjust for the misclassification. Reducing the size of area used to calculate deprivation scores has been shown by chapter 9 to be of limited value and so individual data seems to be the only answer for reliable relative risk estimates. In studies which are only based on area data it is important to discuss the level of

misclassification and how this in general leads to underestimation of the relative risk, but not make an attempt to calculate a corrected relative risk.

The results of this work are not only of importance in the measurement of deprivation. The work of Flegal et al (1991) shows how common differential misclassification is. Differential misclassification does not just arise from clear differences in the misclassification of the population between the diseased and not diseased. It is very likely to occur from the categorisation of a nondifferential continuous exposure variable. The use of methods for adjusting the relative risk for misclassification in many situations in epidemiological research needs extremely careful investigation. These methods seem appealing at the outset with the promise of adjusting for misclassification without having to go and do all the hard work of getting the true exposure levels. However they rely heavily on this one assumption and their poor performance when there is differential misclassification point the way to collecting the true exposure right from the start.

11.10 SUMMARY AND CONCLUSIONS

Previous chapters have shown area level measures of deprivation to underestimate the effect of deprivation on health measures present when using individual level deprivation measures. In this chapter I have investigated the possibility of adjusting estimates of the relative risk for measurement error in the exposure by using information on the relationship between area and individual level deprivation. This has shown these methods to be of little use since extremely small amounts of differential misclassification can lead to poor estimation of the underlying relative risk at the individual level. In most studies it would be very hard to estimate the levels of misclassification in both the diseased and non-diseased groups accurately enough to make these methods of any practical use. In chapter 12 I apply these methods to the estimates of the effect of deprivation in the study of amblyopia to demonstrate their effect on these data.

CHAPTER 12

APPLICATION OF MISCLASSIFICATION ADJUSTMENT METHODS TO THE AMBLYOPIA DATA

12.1 AIMS OF THE CHAPTER

In chapter 11, I described three possible methods of adjusting the relative risk for exposure misclassification. I adapted these methods and applied them to census data to see whether they were suitable for adjusting the effect of area deprivation on long-term illness. This showed these methods were unsuitable for this particular case because there was a small degree of differential misclassification. The relationship between area and individual deprivation differs between those who were diseased and those who were not. Limiting long-term illness appeared to affect the misclassification relationship. This may be due to the illness being of a long-term nature leading to changes over time which would affect deprivation at an individual level but not at an area level. Alternatively it may be due to categorisation of the deprivation data or a combination of the two.

If this differential misclassification is related to the nature of long-term illness, then other conditions may be less likely to affect the relationship between individual and area deprivation. The methods for adjusting for misclassification may be more useful in these cases. However if the differential misclassification was due to grouping the deprivation data then analyses based on these data will result in poor estimation of the underlying relative risk.

Earlier in this thesis, I investigated the relationship between deprivation and age at presentation of amblyopia. I now apply the method of Ekholm et al (1987) discussed in chapter 11, to the amblyopia data, in order to assess the effect of misclassification on the relationship between area deprivation and age at presentation and to estimate the effect of deprivation at the individual level. The work in chapter 11 was based on an extremely large study. Here, I use the data from the 1983 and 1992 Leicestershire amblyopia studies where the sample size is much smaller in order to see how this affects the results. In the case of

amblyopia, differential misclassification is unlikely. Most epidemiological studies do not have samples of such vast size as seen in the SAR data and this work may be a more realistic application of the method. Here, 'disease' status is age at presentation with young referrals being successes (i.e. not 'diseased') and older referrals being failures (i.e. 'diseased'). The distribution of individual deprivation characteristics among older children with amblyopia is unlikely to differ from younger children with amblyopia in the same area. However, there is still the problem of differential misclassification due to categorising deprivation. The results of this analysis and its implications are discussed here.

12.2 ANALYSIS OF AMBLYOPIA DATA

Data were available on the age at presentation of amblyopia for children presenting to Leicester Orthoptic Department in 1983 and 1992, as analysed in chapter 8. The analyses presented in chapter 11, comparing area and individual deprivation and adjusting the deprivation effect for misclassification were based on categorical disease and exposure status. Therefore the amblyopia data needed to be grouped in a similar way.

In this analysis, the exposure is deprivation as before but 'disease' status is derived from age at presentation with older children being failures and younger children being successes. Evidently a different cut-off for age at presentation was needed for the different types of amblyopia because of the different age distributions.

For children with microtropia or no strabismus, an appropriate cut-off would be five years of age since this would identify children detected by pre-school screening. Another alternative cut-off was seven years since children presenting later than this are unlikely to undergo successful treatment. Theoretically, the cut-off should be chosen before viewing the data since it is possible to choose a cut-off to suit the hypothesis in question. However there was a vast difference in the age distributions of the two cohorts and choosing either five or seven years as a cut-off led to very small numbers in one of the cohorts, with only 11 out of 70 children presenting under five in 1983 and only 12 out of 92 presenting over seven in 1992. Methods adjusting for area level data are likely to estimate a more extreme relationship between age and deprivation and result in a relative risk tending to infinity which would be uninformative.

Therefore a cut-off of six years was chosen for both cohorts. For children with a large angle of strabismus a cut-off of three years was chosen since this discriminated between those detected before and after the 3-3½ years screen. The age distribution of the two cohorts was very similar so there were no problems in choosing this cut-off.

To adjust the area deprivation effect for misclassification, a table of the relationship between area and individual deprivation was needed. In chapter 11, SAR census data were used comparing district level and individual level deprivation of economically active adults aged 16-64. The amblyopia data to be analysed here is based on children from Leicestershire. These children are likely to show a different distribution of deprivation to the previous analysis. Data were available from the 1992 study relating to individual and area deprivation so these data were used for the misclassification table needed to adjust for exposure misclassification. There were no data available for the 1983 study on individual deprivation and so it had to be assumed that the misclassification table would be similar although the prevalence of deprivation characteristics may have changed. A further problem was that the misclassification table was based on very small numbers of children in comparison to the previous analysis of SAR data. This would affect the size of the standard errors of the relative risk. It was not sensible to look at differential misclassification here since there are different illness criteria for the two types of amblyopia.

The relationship between area and individual deprivation for children involved in the 1992 amblyopia study was discussed in chapter 10. The data on individual and ward level deprivation is used in the analysis of 1983 and 1992 data. Table 12.1 shows the number of children with different levels of ward level and individual level deprivation.

Table 12.1: Ward level and individual level deprivation for children participating in the 1992 amblyopia study.

Deprived		Ward deprivation			Sensitivity 74% 95% C.I. (65%,83%)
Individual deprivation		Yes	No	Total	
	Yes	74	26	100	Specificity 70% 95% C.I. (60%,80%)
	No	25	58	83	
	Total	99	84	183	

This table is based on the 183 children in the amblyopia study whose parents participated in an interview. Children were classified as individually deprived if they had one of the four characteristics that make up the Townsend deprivation score. Ward level deprivation was based on the Townsend score for the ward in which the child lived using 1991 census data. Ward level deprivation scores were then divided into deprived and not deprived in the same proportions as the individual deprivation scores.

Area deprivation for each child was measured using the Townsend deprivation score based on data from the 1981 and 1991 censuses respectively. Deprivation was grouped into deprived and not deprived based on the proportion who were individually deprived and not deprived in the 1992 amblyopia study. Therefore the 55% of children from the most deprived wards were classified as 'deprived' and the remaining 45% classified as 'not deprived'. This was done separately for each year.

12.3 RESULTS OF ADJUSTING FOR MISCLASSIFICATION

Table 12.2 shows the children with each type of amblyopia classified by their age at presentation and their ward deprivation.

Table 12.2: Deprivation by age at presentation for each type of amblyopia and each cohort with relative risks of late presentation for deprived children

Year	Deprived	Microtropia or no strabismus			Large angle strabismus		
		<6 years	>6 years	Total	<3 years	>3 years	Total
1983	Yes	8	28	36	36	40	76
	No	17	17	34	31	32	63
	Total	25	45	70	67	72	139
	Relative risk (95% C.I.)	1.56 (1.07,2.28)			1.04 (0.88,1.22)		
1992	Deprived	<6 years	>6 years	Total	<3 years	>3 years	Total
	Yes	34	13	47	33	32	65
	No	35	10	45	18	27	45
	Total	69	23	92	51	59	110
	Relative risk (95% C.I.)	1.24 (0.60,2.53)			0.82 (0.69,0.98)		

For children with microtropia or no strabismus in 1983, there was a significant effect of deprivation with children from deprived areas having a 56% increase in their odds of presenting late compared to children from less deprived areas. This effect is smaller and non-significant in 1992. For children with a large angle of strabismus there is no significant effect of deprivation in 1983 and unusually, a reversed relationship in 1992. This pattern was not apparent when looking at age on a continuous scale in chapter 8. The regression analysis in chapter 8 is more reliable since it does not lose information on age by categorisation.

These data were then used in conjunction with the misclassification table in order to estimate the exposure effect adjusted for misclassification. The macros seen in appendix I, based on those of Ekholm and Green (1995), were adapted for these data. Table 12.3 shows the relative risk of deprivation before and after adjustment for misclassification for data from the 1983 and 1992 studies. It also shows the relative risk at the individual level for the 1992 study where both area and individual data were collected.

Table 12.3: Estimates of relative risk of late presentation for deprived children before and after adjusting for misclassification for both cohorts and type of amblyopia

	Estimate of Relative risk (95% confidence interval)					
	Microtropia or no strabismus			Large angle strabismus		
Year of cohort	R* Area	R' _{Ek}	R Individual	R* Area	R' _{Ek}	R Individual
1983	1.56 (1.07,2.28)	3.09 (0.86,11.1)	-	1.04 (0.88,1.22)	1.08 (0.52,2.28)	-
1992	1.24 (0.60,2.53)	1.68 (0.28,10.0)	0.77 (0.33,1.79)	0.82 (0.69,0.98)	0.64 (0.28,1.42)	0.67 (0.47,0.95)

For children with microtropia or no strabismus there is an increase in the relative risk of being a late referral if they were deprived after adjusting for misclassification. Neither relative risk is formally significant and the associated confidence intervals are wide owing to small numbers in the misclassification table. The relative risks show a larger effect of deprivation after adjustment in 1983 than in 1992. These estimates indicate that for children in 1983, those from a deprived area have a 209% excess risk of presenting after 6 years compared to those in less

deprived areas. This compares with an excess risk of 68% in 1992. However the relative risk after adjustment for 1992 is very different to that calculated using the observed data of individual deprivation and age at presentation. The adjustment is in the opposite direction.

For children with a large angle of strabismus in 1983 there is very little change in the estimate of the relative risk after adjustment. Deprived children would have an adjusted excess risk of 8% of being referred after 3 years compared to those from less deprived households. However in 1992 there is a decrease in the relative risk which since it is less than one relates to an increase in the effect of deprivation with less deprived children being referred later than those who are more deprived. This relates to a 56% excess risk of being referred over 3 years if a child comes from a less deprived household compared with a more deprived household. This estimate, unlike the children with microtropia or no strabismus, is very similar to the one calculated using individual data.

Several possible explanations for the discrepancy between the adjusted relative risk and the relative risk at the individual level for children with a microtropia in 1992 will now be discussed and illustrated.

12.4 DISCUSSION OF RESULTS

12.4.1 DIFFERENCES IN MISCLASSIFICATION BETWEEN TYPES OF AMBLYOPIA

It is possible that there are differences in misclassification between the types of amblyopia. It was seen in chapter 8 that there was a higher proportion of children with large angle strabismus from deprived areas. This may also be linked with a difference at the individual level. In order to assess whether there was differential misclassification between types of amblyopia, table 12.4 shows the sensitivity and specificity by type of amblyopia based on the interview data from the 1992 study.

Table 12.4: Sensitivity and specificity and associated 95% confidence intervals by type of amblyopia based on data from the 1992 study

	All types (n=183)	Microtropia or no strabismus (n=84)	Large angle strabismus (n=99)
Sensitivity	74% (65%,83%)	68% (53%,83%)	77% (67%,87%)
Specificity	70% (60%,80%)	70% (57%,83%)	70% (56%,84%)

This shows large differences in sensitivity between the groups but no difference in specificity. This differential misclassification is most likely to be due to the higher proportion of children with large angle strabismus coming from deprived areas. It has been shown in chapter 10 that in Leicestershire ward data is better at identifying deprived children than less deprived children since the deprived population are concentrated in the city. Therefore with more children from deprived areas the sensitivity would be expected to increase and the prevalence of deprivation would also be increased. The effect of these differences can be seen in table 12.5

Table 12.5: Relative risk of late presentation for deprivation (baseline = not deprived) by type of amblyopia. Adjusted relative risk based on separate misclassification table for each type of amblyopia.

	Estimate of Relative risk (95% confidence interval)					
	Microtropia or no strabismus			Large angle strabismus		
Year of cohort	R* Area	R' _{Ek}	R Individual	R* Area	R' _{Ek}	R Individual
1983	1.56	3.49	-	1.04	1.09	-
1992	1.24	1.77	0.77	0.82	0.66	0.67

Here the appropriate misclassification table has been applied to the data to recalculate the relative risks for late presentation. The differential misclassification appears to have changed the relative risks by a very small amount. Since there is a rise in sensitivity and a rise in the prevalence of deprivation among those with a large angle of strabismus, these factors seem to be working in opposite ways on the relative risk leading to surprisingly little change overall.

12.4.2 DIFFERENCES IN MISCLASSIFICATION AND AGE AT PRESENTATION

Another possible reason for biased estimates after adjustment is differential misclassification with respect to disease status, i.e. early referrals have different patterns of deprivation than later referrals. Dividing the amblyopia data by type of amblyopia and age at presentation led to very small numbers in each group which could easily lead to differences in sensitivity and specificity between the groups. Table 12.6 shows the relative risks for each type of amblyopia and cohort adjusted for misclassification based on different misclassification tables for each type of amblyopia and early and late referrals.

Table 12.6: Relative risks of late presentation for deprivation (baseline = not deprived) by type of amblyopia. Adjusted relative risks based on separate misclassification tables for each type of amblyopia and each age group.

	Estimate of Relative risk (95% confidence interval)					
	Microtropia or no strabismus			Large angle strabismus		
Year of cohort	R* Area	R' _{EK}	R Individual	R* Area	R' _{EK}	R Individual
1983	1.56	1.00	-	1.04	1.18	-
1992	1.24	0.61	0.77	0.82	0.67	0.67

Adjusting for misclassification based on different misclassification tables for age at referral and type of amblyopia leads to very different estimates of the relative risks. It would be expected that based on these misclassification tables, the estimates for 1992 after adjustment should be exactly the same as those based on individual data. For children with a large angle strabismus this is the case. However there is some discrepancy for children with microtropia or no strabismus. The reason for this is the small amount of missing data due to non-response. Since the numbers in the misclassification tables are small due to dividing them into many strata, the absence of a small number of children has led to this discrepancy in the relative risks. The incorrect estimation seen assuming non-differential misclassification was again due to the failure of this assumption. Some of this differential misclassification may have been due to the categorisation of the continuous individual deprivation variable. However it is more likely to be due to poor estimation of the sensitivity and specificity because of small numbers of children in the study.

The estimates of the relative risks for 1983 seen here are unlikely to be correct since the numbers in the four misclassification tables based on the 1992 data are very small. Further investigation of differences between cohorts will now be discussed.

12.4.3 DIFFERENCES IN MISCLASSIFICATION BETWEEN COHORTS

The general pattern of deprivation has changed greatly between the 1981 and 1991 census, with a general decrease in the prevalence of deprivation based on car ownership, house ownership, unemployment and overcrowding. Therefore differences in the prevalence between the cohorts could mean that using 1992 data to adjust the relative risks for 1983 will lead to misleading results. Table 12.7 shows the estimates of the relative risk for 1983 based on the overall misclassification table used initially in section 12.3 and different estimates of the prevalence of deprivation.

Table 12.7: Relative risks for 1983 cohort based for different levels of deprivation prevalence.

(Original adjustments were based on proportion of 0.55)

Prevalence of deprivation	Estimate of Relative risk of deprivation	
	Microtropia or no strabismus	Large angle strabismus
0.55	3.09	1.09
0.60	3.46	1.10
0.65	4.07	1.10
0.70	5.43	1.11

This shows the prevalence of deprivation to have a large effect on the relative risk for 1983 and a low estimate of the prevalence leads to underestimation of the relative risk. There may be further differences between cohorts in misclassification. The raised prevalence of deprivation in 1983 may also be associated with different sensitivities and specificities for the misclassification table. This could evidently lead to incorrect estimates of the relative risk. Furthermore, there has been an increase in detected cases of amblyopia associated with microtropia or no strabismus and a decrease in cases associated with a large angle strabismus in 1992. This seems to indicate that a different population is under study in 1992 compared to 1983. This may be associated with changes in the relationship between area and individual

deprivation between the two cohorts. All of these changes are likely to lead to poor estimation of the relative risk after adjustment for misclassification.

12.5 SUMMARY AND CONCLUSIONS

Although chapter 11 had shown misleading results obtained from adjusting for misclassification, I applied the method of Ekholm et al (1987) in this chapter to assess the effect of deprivation on age at presentation of amblyopia after adjusting for misclassification. This method relies on accurate assessment of the sensitivity and specificity and assumes non-differential misclassification. Differences in misclassification between types of amblyopia, age at presentation group, and cohort led to misleading estimates of the relative risk.

Flegal et al (1991) have shown how categorisation of an exposure can lead to differential misclassification. This will have also played a part in this analysis since it is based upon the same categorisation of individual deprivation. Since age at presentation was related to deprivation in 1983 this means that as the probability of being misclassified increased, the probability of late presentation increased. This could lead to differential misclassification and incorrect estimates of the relative risk.

This chapter has shown the need to be able to specify the misclassification table accurately and the need to take into consideration various levels of differential misclassification. For children in 1983 I would expect the effect of deprivation at the individual level to be higher than that shown by area data. However it appears the only way of estimating this effect of deprivation at the individual level in 1983 is by collecting individual level data. The method illustrated here has again been shown to be very sensitive to small changes in the data and of little practical use for this particular example.

CHAPTER 13

SUMMARY AND CONCLUSIONS

13.1 AIMS OF THE CHAPTER

In this final chapter, I summarise the findings of the thesis and discuss the implications of these findings in a wider context and indicate possible extensions to the work.

I had two main objectives in undertaking this thesis:

- i) to investigate inequalities in the age at presentation of amblyopia
- ii) to evaluate the effect of changes made to vision screening in Leicester on any inequalities

I also aimed to explore some methodological issues relating to these objectives:

- iii) to investigate the problem of ecological fallacy to see whether area measures of deprivation were leading to biased estimates of the exposure effect
- iv) to assess whether methods developed to adjust for measurement error could be used to estimate the exposure effect at the individual level

I now summarise my findings and discuss their implications with respect to these objectives.

13.2 USING A MEASURE OF DEPRIVATION TO INVESTIGATE INEQUALITIES IN THE PRESENTATION OF AMBLYOPIA

In attempting to assess the extent of inequalities in the age at presentation of amblyopia I first had to choose a measure of socio-economic status. In chapter 3 I highlighted the problems of traditional occupation-based social class measures and investigated census based deprivation scores. The Townsend score, a unidimensional area deprivation measure was chosen for the analyses in this thesis because it had been constructed in a theoretically informed manner, it was a frequently used measure recognised by academic peers and offered the opportunity to look at trends in the data.

I used this area based deprivation measure in the analysis of a large UK multi-centre study of amblyopia. Wang et al (1990) showed that parents are often the first to notice visual problems in their children. Squints are often readily apparent but poor vision is harder to identify. Previous research looking at presentation of amblyopia has failed to investigate its relationship with deprivation. I hypothesised that children from more deprived areas would have amblyopia detected later than children from less deprived areas. The results show that this is true for amblyopia associated with microtropia or no strabismus, conditions which are generally asymptomatic, but not true for amblyopia associated with a large angle of strabismus. There was also a wider difference in age at presentation between centres for children with a microtropia or no strabismus than for those with a large angle of strabismus.

There has been little other research on inequalities in the presentation of amblyopia. However Shaw et al (1988) looked at children with amblyopia in Leicester, one of the centres participating in the multi-centre study, and found Asians to present later than Caucasians. My analysis of the multi-centre study showed that this finding may have been due to differences in the deprivation of Asian children compared with non-Asians and that it is deprivation that is related to age at presentation and not ethnicity. This indicates the importance of adjusting for deprivation in studies of ethnicity and health since many people from minority ethnic groups have been shown to have higher rates of deprivation (Atri et al, 1996). This finding confirms the work of Spencer (1996) who has indicated that although in some cases genetic differences in ethnic origin influence child health, it is socio-economic and environmental factors which explain the majority of the variations in health both between and within ethnic groups.

My findings indicated that there was differential access to screening services to detect asymptomatic amblyopia. This is similar to the patterns of inequalities in access to immunisation services shown by the National Child Development Study (Essen and Wedge, 1982; Blaxter, 1981), and other regional studies (Reading, Jarvis and Openshaw, 1993; Marsh and Channing, 1986; Lynch, 1995) and vision screening services shown by Williamson et al (1995). Since differential use of screening programmes could lead to differences in the incidence and treatment of conditions between social groups (Macintyre, 1989) this could lead to differential outcome of treatment for asymptomatic amblyopia. Data available from the 1983 study on visual outcome was poor and did not offer conclusive evidence of a relationship between age at presentation and outcome. Thus this relationship needs further investigation.

Based on the work of Feinstein (1993), inequalities in access to and utilisation of the health care system point to materialist or structural explanations that affect the ability to access screening services rather than behavioural differences. This indicates a need for structural changes to vision screening and child health surveillance in order to reduce these inequalities in access. It has been shown that the majority of health authorities run vision screening services without evaluating their efficacy (Stewart-Brown et al, 1988). Current services need evaluation and ways of changing them structurally to improve services and reduce inequalities need to be explored. The 1983 multi-centre study showed wide variation in the age at presentation of asymptomatic amblyopia between centres. The centre with the youngest mean age was also the only centre known to be running an orthoptic primary screening service. This indicated that such services may be constructive in changing referral patterns.

In terms of evaluating services, there is a need for increased monitoring of screening services to assess equity. One step towards this would be to include social statistics in audits of these services. The orthoptic department in Leicester is introducing an audit system which will be able to investigate aspects of referral and treatment of children with amblyopia presenting to the department. Inclusion of data on deprivation is being encouraged in its implementation. Although work like this is viewed as academic research by the orthoptists, it should be considered as a fundamental part of audit to achieve an equitable service.

Another possible extension to this work would be to use a multi-dimensional measure of deprivation such as the GB profiler discussed in chapter 3. In this thesis I wanted to use a uni-dimensional measure of deprivation since it offered the opportunity to look at trends in the data and the GB profiler was less well developed when this work was undertaken. However it would now be interesting to assess whether the multi-dimensional measure indicated the types of areas that had higher ages at presentation to understand more about children who presented later. This information could supplement the work in this thesis to improve future screening practices.

13.3 EVALUATING CHANGES MADE TO VISION SCREENING

The analyses of data from the multi-centre study of amblyopia indicated the need for structural changes to screening services in order to reduce inequalities. The review of inequalities in health in chapter 2 pointed to the need for major multi-disciplinary national interventions at a governmental level to reduce inequalities. However it was also shown that these have not been forthcoming and local interventions have currently been the only steps taken. Having illustrated inequalities in the age at presentation of amblyopia, I then investigated a local intervention to assess its effect on these inequalities. I designed a study to look at the presentation of amblyopia in Leicester after improvements had been made to the vision screening service. Between 1988 and 1991 changes in the organisation of vision screening for pre-school children in Leicestershire were made. I designed and analysed a study to compare age at presentation before and after these changes, to see if the service became more equitable. This showed a dramatic alteration in amblyopia referral patterns. In children with asymptomatic amblyopia, the number of cases rose by 30%, the average age of presentation was reduced by 19 months and there was less evidence of a relationship with deprivation. In amblyopia associated with a large angle of strabismus there was no change in the average age of presentation but the number of cases dropped by a third.

Although in chapter 8 I discussed possible alternative explanations for the reduction in the relationship between age at presentation and deprivation, I concluded that the changes made to vision screening were likely to have been associated with this reduction. This points to the success of involving orthoptists in screening at the secondary level. Previous research had shown improvements in screening with the use of orthoptists as primary screeners (Jarvis et al, 1990; Edwards et al, 1989; MacLellan and Harker, 1979), but Hall (1989) had recommended their use as secondary screeners because of the reduced cost. Although the service in Leicester has not been evaluated from an economic perspective it can be seen the use of secondary orthoptic screening has had a positive outcome on referral patterns in Leicester. It is not possible to assess whether the gains would have been greater had primary orthoptic screening been introduced and it is unlikely that this comparison will be possible in Leicester in the future since there are no current plans to reform screening services.

This reduction in inequalities is contrary to the findings of Reading, Colver et al (1994) who showed that increasing uptake of child health screening services led to an increase in inequalities. In Leicester early results of a study of child health surveillance have shown that there is evidence that although child health surveillance uptake has increased there are still inequalities in uptake similar to those found by Reading, Colver et al (1994). However the changes made to vision screening in Leicestershire appear to have had the effect of improving the equity of the vision screening service once children are participating in the child health surveillance programme. The new system, including the introduction of secondary orthoptic screening, is thought to have dramatically reduced waiting times to see the consultant ophthalmologist since far fewer children are referred to the consultants with the orthoptists reducing the number of false positive cases. Bowman et al (1996) indicated that people from more deprived areas were less likely to attend their hospital ophthalmology appointment the longer they had to wait and so this may explain the increase in equity of the service. However as indicated there still appear to be inequalities in access to child health surveillance programmes and there is no information available on cases of amblyopia that are still not detected or those that are detected by secondary orthoptic screening but fail to attend the appointment with the consultant. In order to investigate some aspects of this issue, I have helped to design a study to look at a cohort of children participating in child health surveillance. This study will permit the investigation of the referral path from vision screening by the GP or health visitor, through referral to secondary orthoptic screening and final presentation to the consultant ophthalmologist. Following a cohort of children will provide the opportunity to look at children who present to the consultant ophthalmologist who have not been detected by screening services.

The number of cases of amblyopia associated with a large angle of strabismus was shown to decrease. It is possible that changes in screening have led to earlier referral of strabismus with a smaller proportion of cases now going on to develop amblyopia thus lowering the incidence of this type of amblyopia. To investigate this hypothesis, information is required on all children with strabismus as well as those treated for amblyopia which is not possible with my study data. The cohort study discussed above will be able to assess the number of children presenting with a large angle of strabismus. This will explore whether there are inequalities in the presentation of strabismus, as in 1992 there appeared to be more children with amblyopia coming from deprived areas.

The study of children presenting in 1992 to Leicester orthoptic department extended the work of the 1983 multi-centre study and provided information on referral pathways of these children and measures of compliance through interviews with the parents. These data are not analysed here but provide many opportunities to investigate the referral process further and investigate whether, although there no longer appear to be inequalities in age at presentation, children from more deprived areas have different referral pathways to care than those from less deprived areas.

With the introduction of an orthoptic audit system to the Leicester orthoptic department as discussed in section 13.2 further monitoring of the service will be made easier and these data will provide the opportunity to assess patterns of presentation in a period when all children screened will have benefited from the changes to the service.

13.4 THE ISSUE OF ECOLOGICAL FALLACY

A problem with this type of analyses based on area deprivation measures highlighted in chapter 3 is ecological fallacy where area level effects are wrongly assumed to exist at the individual level. Using samples of anonymised records from the national census I illustrated that area level measures were underestimating the relationship between deprivation and health seen at the individual level. Applying rates of illness seen among the least deprived individuals to those among the more deprived show the severe extent of these inequalities. It appeared that the variation in prevalence of limiting long-term illness was explained by individual deprivation characteristics rather than by area level factors.

Similar results were found in the analysis of limiting long-term illness data by Sloggett and Joshi (1994) and Ecob (1996) illustrating that area level deprivation does not explain a significant amount of variation in mortality after adjusting for individual deprivation characteristics. My analysis has confirmed these findings but has gone further to show area based scores are underestimating deprivation effects and that inequalities may be overlooked because of the relative insensitivity of area level measures.

As in the analyses of the amblyopia study, this analysis showed the importance of looking at the multi-variate relationships between ethnicity, deprivation and health. In many of the age and sex strata under investigation, there appeared to be higher prevalence of reported illness in those who classified themselves as other than 'white'. However this effect disappeared in all but one case after adjusting for deprivation.

Using data from a Leicestershire study of perinatal mortality, I compared district, ward and ED level data as the SAR data only provided information at district level. This showed that using smaller areas did not improve the relationship between area and individual data. This may only be true to Leicestershire where the majority of deprived areas are located in the city. Investigation on a national level is needed but the access to a dataset such as the longitudinal study is limited again by confidentiality. This work also showed similar underestimation of inequalities in child health when using area level deprivation as opposed to individual level data.

Investigating the effect of using area level and individual measures of deprivation in the amblyopia study showed very different results. There was no evidence of a greater effect of deprivation at the individual level in this study and this finding offered further evidence that there was no longer a relationship between deprivation and age at presentation of amblyopia since the improvements had been made to screening services.

The amblyopia study showed that for children with a microtropia or no strabismus, there was no difference in age at presentation with individual deprivation but a small nonsignificant effect of deprivation at the area level. This was opposite to the patterns shown in the analysis of census data and data from the perinatal mortality study. The most likely reason for this is the inherent differences in the condition under study. I hypothesised that the presentation of asymptomatic amblyopia to be related to screening and hence the Townsend area score appears to reflect differences in screening programmes and access to screening at the area level. Individual factors play an important role in the referral of children with amblyopia but the particular deprivation indicators under study do not appear to be proxies for the individual factors affecting presentation. Evidently in most cases a combination of area and individual characteristics will play a role in inequalities in health.

In asymptomatic amblyopia it appears that a residual effect of deprivation on age at presentation is related to area level factors. Screening practices may have been introduced at slightly different times throughout the period of change. The planned study of vision screening discussed in section 13.3 will provide the opportunity to investigate a cohort who were all eligible for screening at the four pre-school screening sessions and assess the effect of deprivation among these children.

Individual deprivation was based on the Townsend score and classified nearly half of the population as not deprived with no way of differentiating between them. This measure is also limited in some groups such as the elderly since there are differences in unemployment and car ownership. My work is a starting point for future research in this area to investigate more useful indicators. Ongoing work is looking at consumables, for example ownership of washing machines, cars etc. to measure levels of deprivation. These require validation in epidemiological research.

13.5 ADJUSTING FOR MISCLASSIFICATION OF DEPRIVATION

An investigation of methods to correct for misclassification of deprivation at the area level in order to estimate the exposure effect at the individual level showed the methods to be of little practical use. This was mainly due to small amounts of differential misclassification leading to extreme overestimation of the exposure effect.

Area deprivation misclassifies the deprivation seen at the individual level leading to underestimation of the exposure effect. In this thesis I reviewed several methods for adjusting for exposure misclassification using SAR data. The method of Espeland and Hui (1987) relied on an incorrect assumption and this led to adjustment of the relative risk in the wrong direction. The methods of Flegal et al (1986) and Ekholm et al (1987) were based on the same assumption, that given information on individual deprivation, area deprivation adds no extra information about illness status. The method proposed by Ekholm et al (1987) based on non-linear models provides the opportunity to test hypotheses by collapsing over strata. Both of these methods adjusted the relative risk in the right direction but the assumption of non-differential misclassification led to overestimation. These methods were extremely sensitive to

small amounts of differential misclassification making them of little practical use for adjusting area based estimates of deprivation for misclassification.

In this example, a small amount of differential misclassification led to violation of the assumptions and consequently incorrect results. These estimates of the degree of misclassification were based on an abnormally large dataset. The application of these methods to the amblyopia study showed a more realistic application of the results. Poor estimation of the sensitivity and specificity of the misclassification led to wide confidence intervals for the relative risk. Small changes in the estimates of the relative risk and any differential misclassification would have completely changed the estimates of the relative risk and so the results were unreliable. A possible extension of this work would be to include prior beliefs on accuracy of the data. Some investigation of this showed where prior beliefs were vague, it led to wide confidence intervals for the adjusted relative risk estimates and thus interpretational problems.

If levels of misclassification were lower then the effect of differential misclassification would be slightly reduced. It is likely that some of the differential misclassification has arisen from nondifferential errors in exposure measurement due to categories being formed from continuous exposure data. Flegal (1991) discourages the use of these adjustment methods in this scenario.

This work points to the need to collect individual data on deprivation rather than trying to adjust relative risk estimates based on area data for misclassification. Any research where only area deprivation is available should discuss the likely degree of underestimation of effect but not attempt to quantify it using these methods.

13.6 FURTHER APPLICATIONS

Although I have concentrated on the use of deprivation measures in amblyopia, the principles behind this work are applicable in many areas of epidemiology. The discussion of area based deprivation measures is made in general terms and the conclusions of my review would be similar for the investigation of inequalities in many other conditions.

The Townsend score was used in thesis to assess levels of inequality in health. Uni-dimensional deprivation measures have limitations in their interpretational value. The development of multi-dimensional indicators warrants more work to combine the positive aspects of both types of indicator.

In this thesis I looked at the relationship between age at presentation of amblyopia and deprivation. This work was based on a review of orthoptic notes and involved no patient contact, making it a cheap and relatively simple study to perform. This type of investigation would be straightforward for many hospital departments and could be easily incorporated into audit programmes for monitoring inequalities. The introduction of computerised records in the Leicester orthoptic department as part of the audit process will make monitoring of age at presentation and the results of treatment in the future very straightforward.

I have also shown the success of changes made to a vision screening programme at reducing inequalities in presentation. Previous work has shown the efficacy of primary orthoptic screening in increasing detection of amblyopia. Here I have shown how the introduction of secondary orthoptic screening, a much cheaper facility, has led to similar increases in children detected with asymptomatic amblyopia as well as a reduction in the overall age at presentation in Leicester. This type of intervention may be an easily affordable option to improve screening programmes in other regions. The importance of ongoing auditing of this service in Leicester must be stressed, in order to keep inequalities at a minimum.

Overall in this thesis I have shown the positive use of area deprivation measures but have also shown the importance of trying to measure deprivation at an individual level. More appropriate measures of individual deprivation are needed in order to assess inequalities in health and to be able to design interventions that can improve them. Although changes are needed at a national level to reduce inequalities, the importance of locally based interventions must not be overlooked. In health service provision, where the aim is to provide an equitable service, area deprivation measures should be used as a basic and cheap auditing tool. However, the problems of using these measures highlighted in this thesis should always be kept in mind.

APPENDICES

APPENDIX A

Vision screening programme (as part of child health surveillance) in Leicestershire

Age	Screeners	History	Examination
6-8 weeks	Medical practitioner	Ask parents about any concerns; enquire whether child looks at parents, follows people as they move, fixates on parents' face.	Inspect the eyes externally and check the red reflex
7 ¹ / ₂ -10 months	Health visitor	Ask the parent about concerns about vision, whether their child looks at the parents, follows moving objects with the eyes, fixates on small objects.	Observe visual behaviour and look for squint; External examination of the eye; Examine symmetry of corneal light reflexes for squint
18-24 months	Medical Practitioner (Some by a health visitor)	Ask the parent about any concerns or whether they have suspected a squint	Check the child appears to fixate on small objects; Examine symmetry of corneal light reflexes for squint
3-3 ¹ / ₂ years	Health visitor	Ask the parent about whether the child has difficulties with vision; whether they have suspected a squint	Observe for any evidence of squint; External examination of the eye; Assessment of visual acuity (Stycar Vision test); Examine corneal light reflexes for squint

Social deprivation and age at presentation in amblyopia

Lucy K. Smith, John R. Thompson, Geoffrey Woodruff
and Fiona Hiscox

Abstract

Background Amblyopia is the most common visual disability in children. Early treatment is thought to be more effective, and therefore factors affecting the age at presentation are important. A relationship between social deprivation and access to health care and screening services is well known. We hypothesized that social deprivation might be associated with later presentation of amblyopia, particularly of anisometropic amblyopia which depends on vision screening for referral.

Methods Data from a historical cohort of 897 children with amblyopia, from seven UK orthoptic clinics, were used to test this hypothesis. Social deprivation was measured by the Townsend score of the ward in which the child lived.

Results A relationship between social deprivation and age at presentation was found in children with anisometropic amblyopia even after adjusting for differences between clinics ($p = 0.01$) but no similar association was evident in children with amblyopia associated with strabismus. There was a difference of 22 months in the average age at presentation between children with anisometropic amblyopia in the most deprived and least deprived areas of the study.

Conclusions If screening for anisometropic amblyopia is to be undertaken, priority should be given to screening children from areas of social deprivation.

Introduction

Amblyopia is the most common visual disability in children, with a prevalence of between 2 and 5 per cent.^{1,2} In amblyopia, normal vision fails to develop because of a difference in vision between the two eyes in early life. Usually it is associated with either misalignment of the eyes (strabismus), or a refractive error that is greater in one eye than the other (anisometropia), or a combination of these factors. Susceptibility to amblyopia and responsiveness to treatment are greater the more immature the visual system; thus early treatment is more effective. Strabismus is often first detected by parents, whereas anisometropia is much more difficult to diagnose and is usually only identified when amblyopia is detected at a vision screening test.³

Little is known about the factors affecting the age of presentation of amblyopia. Shaw *et al.*⁴ found that for all types of amblyopia, males presented later than females and Asians presented later than Caucasians. Campbell and Charney⁵ found that the age at diagnosis depended on family history of strabismus, degree of squint, level of maternal education and degree of parental suspicion of a problem, but this study did not distinguish between different types of amblyopia. The Black report⁶ highlighted a relationship between social deprivation and health. Reading *et al.*⁷ have suggested that many screening services fail to serve socially deprived people properly. We hypothesized that social deprivation might cause children with amblyopia to present late and that this might be particularly true of those with anisometropic amblyopia, referral of which depends on screening. Data from a multicentre study of patients treated for amblyopia were used to investigate age at presentation, social deprivation and other possible explanatory variables.

Methods

Data were obtained from a follow-up study of amblyopia that included all children who started treatment in 1983, following them through to 1992 in each of seven volunteering English orthoptic clinics. Consistency of information between centres was achieved by using one research orthoptist to visit all seven clinics. Children who were not treated within 12 months of obtaining a visual acuity, those who were

Department of Ophthalmology, University of Leicester, Clinical Sciences Building, Leicester LE2 7LX.

LUCY K. SMITH, Research Associate

JOHN R. THOMPSON, Senior Lecturer in Ophthalmic Epidemiology
GEOFFREY WOODRUFF, Senior Lecturer/Consultant in Ophthalmology

FIONA HISCOX, Research Orthoptist

Address correspondence to Lucy Smith.

© Oxford University Press 1994

TABLE 1 Mean age at presentation in years (with numbers of children given in parentheses) for types of amblyopia and quintile of Townsend deprivation score within clinic

Quintile of Townsend score	Age (years)		
	Anisometropic amblyopia	Strabismic amblyopia	Mixed amblyopia
1 (affluent)	5.3 (34)	3.5 (96)	4.5 (49)
2	5.4 (25)	3.2 (94)	4.1 (59)
3	6.2 (34)	3.7 (95)	4.1 (45)
4	6.1 (26)	3.4 (110)	5.1 (46)
5 (deprived)	6.4 (32)	3.3 (98)	4.5 (53)
All	5.9 (151)	3.4 (493)	4.4 (252)*

* Missing data for one patient.

discharged and then re-referred before starting treatment, and one child who was originally referred for strabismus and found to be normal but developed anisometropic amblyopia over a year later, were excluded from the analysis as unlikely to have amblyopia at the time of their first visit. This left 897 children who were classified as having strabismic amblyopia, anisometropic amblyopia or mixed strabismic and anisometropic amblyopia. Anisometropia was defined as a difference of at least one dioptre of either sphere or cylinder between the two eyes, and strabismus was defined as the presence of manifest squint on cover testing. Children with strabismus were classified as having microtropia (less than 5° of misalignment), exotropia (divergent squint) or esotropia (convergent squint).

The degree of social deprivation of each child was estimated by the Townsend deprivation score,⁸ a composite score based on electoral ward data from the census. A 1981 post-code directory linked the patient's post-code to electoral wards, and the Townsend score was calculated by linking these wards to 1981 Census data. The post-code directory contained co-ordinates for each post-code, allowing the distance from the orthoptic centre to be calculated.

Regression analysis was used to investigate relationships between the age at presentation and the following variables and their interactions: referral visual acuity, type of squint, sex, race (Asian or non-Asian forename or surname), distance from orthoptic clinic, Townsend deprivation score, and orthoptic clinic attended.

Results

Children with strabismic amblyopia presented significantly younger (mean 3.4 years, $n = 493$) than

TABLE 2 Mean age at presentation in years (with numbers of children given in parentheses) for types of amblyopia and clinic

Clinic	Age (years)		
	Anisometropic amblyopia	Strabismic amblyopia	Mixed amblyopia
A	5.6 (19)	3.2 (71)	4.7 (29)
B	5.1 (27)	3.8 (51)	5.1 (43)
C	5.5 (4)	3.2 (33)	3.5 (15)
D	6.8 (46)	3.4 (134)	4.4 (57)
E	5.9 (23)	3.7 (97)	4.2 (67)
F	4.8 (23)	3.0 (84)	4.4 (31)
G	6.6 (9)	3.4 (23)	3.7 (11)
All	5.9 (151)	3.4 (493)	4.4 (253)

children with pure anisometropic amblyopia (mean 5.9 years, $n = 151$) or children with mixed amblyopia (mean 4.4 years, $n = 253$) ($p = 0.0001$). Children at each clinic were classified by the Townsend deprivation score of the ward in which they lived into quintiles. These quintiles ranged from the 20 per cent of children living in the most affluent areas to the 20 per cent living in the most deprived areas. An association between quintile of social deprivation and age at presentation was evident for children with pure anisometropic amblyopia but not for those with strabismus (Table 1). Table 2 compares the seven clinics and shows a range of 22 months in the mean age at presentation for children with anisometropic amblyopia compared with nine months for those with strabismic amblyopia and 20 months for those with mixed strabismic and anisometropic amblyopia.

Regression analysis of logarithm of age at presentation confirmed that there were significant differences in the age at presentation to the clinics for children with anisometropic amblyopia ($p = 0.0001$) and for those with strabismic amblyopia ($p = 0.04$) but not for those with mixed amblyopia ($p = 0.6$). The Townsend deprivation score was significantly associated with the age at presentation for children with anisometropic amblyopia ($p = 0.01$, regression coefficient 0.017, standard error 0.007). Thus an increase of four units in Townsend score (approximately 1 SD) led to an average delay in presentation of six months. No similar relationship was observed in children with strabismic amblyopia ($p = 0.9$) or mixed amblyopia ($p = 0.3$). Age at presentation was related to type of squint in children with strabismic and mixed amblyopia ($p = 0.0001$ in both cases), showing that children with microtropia presented up to two years later than those with exotropia, esotropia or other squint.

Discussion

Wang *et al.*³ showed that parents are often the first to notice visual problems in their children. Squints are often readily apparent, but poor vision is harder to identify. We hypothesized that patients from more deprived areas would have amblyopia detected later than children from less deprived areas. Our results show that this is true of anisometropic amblyopia, a condition which is generally asymptomatic, but not true of strabismic amblyopia.

We measured social deprivation using the Townsend deprivation score⁸ based on census data from electoral ward areas, which has been shown to correlate highly with a range of health measures.⁹ However, there may be a stronger underlying relationship between individual social deprivation and age at presentation than we have been able to show by using this relatively insensitive measure.

The relationship between age at presentation and social deprivation is best illustrated by the most extreme wards in the study. The child with the lowest Townsend score (-7.65) came from a ward where 85 per cent of households had a car, no households were overcrowded, 84 per cent of households were owner-occupied and there was only 4 per cent unemployment. In contrast, the child with the highest deprivation score (+7.65) came from a ward where 73 per cent of households had no car, 8 per cent of households were overcrowded, 2 per cent of households were owner-occupied and there was 29 per cent unemployment. Our regression equation indicates that the most deprived children with anisometropic amblyopia present on average 22 months later than the least deprived.

We also found there was a much wider difference in age at presentation between centres for children with anisometropic amblyopia than for those with strabismus. This may be because factors associated with local screening and health care provision play a more important role in the referral of children with anisometropic amblyopia because it is usually asymptomatic.

This cohort represents one of the largest series of children with amblyopia ever studied. The children were recruited in 1983 and their treatment was followed through to 1991. The referral patterns thus relate to a period about 10 years ago, and it is possible that the importance of deprivation has either increased or decreased since that time. Missing data are usually a problem with retrospective studies, although in this study this factor was minimal. Unfortunately, information is not available on the source of referral of the children in our study; for example, we do not know whether they were referred by a screening service. The

sources of referral for the seven clinics are known to have varied greatly in 1983, with some clinics running extensive screening programmes whereas others had none. Despite this, we find no evidence of an interaction between clinic and social deprivation, with the effect of social deprivation on age at presentation being of the same magnitude in each centre despite their different sources of referral.

Shaw *et al.*⁴ found that Asians presented later than Caucasians. In our study there was no significant relationship in any of the diagnostic groups between age at presentation and Asian parentage even before adjusting for the Townsend score. However, we had relatively few Asians in our study (5 per cent). Shaw *et al.* also found that boys presented later than girls, but no similar relationship between sex and age at presentation was observed in this study.

Campbell and Charney⁵ found age at diagnosis depended on degree of parental suspicion and level of maternal education for all types of amblyopia. Our study could not measure this specifically, but the Townsend score is thought to reflect these factors.

Most districts of the United Kingdom implement extensive visual screening of children with the early detection of asymptomatic amblyopia as one of the main objectives. Our results suggest that, to be most effective, these screening programmes should attempt to direct their resources towards children from the more deprived parts of the community.

Acknowledgements

We thank the staff of the orthoptic clinics at Birmingham Children's Hospital, Bristol Eye Hospital, Leeds General Infirmary, Leicester Royal Infirmary, Nottingham University Hospital, Sunderland Eye Infirmary and Worthing Hospital. We would also like to thank the staff in the department of General Practice at St Mary's College, London, for access to the 1981 post-code directory. Material from Crown-copyright records has been made available through the Post Office and the ESRC Data Archive. We gratefully acknowledge the financial support of the British Council for Prevention of Blindness, the Iris fund, the British Orthoptic Society and the Anne Allerton Fund.

References

- 1 Von Noorden GK. *Binocular vision and ocular motility: theory and management of strabismus*, 4th edn. St. Louis, MO: Mosby, 1985: 208-209.
- 2 Thompson JR, Woodruff G, Hiscox F, Strong N, Minshull C. The incidence and prevalence of amblyopia detected in childhood. *Public Hlth* 1991; **105**: 455-463.

- ³ Wang Y, Thompson JR, Goulstine DB, Rosenthal AR. A survey of the initial referral of children to an ophthalmology department. *Br J Ophthalmol* 1990; **74**: 650-653.
- ⁴ Shaw DE, Fielder AR, Minshull C, Rosenthal AR. Amblyopia - factors influencing age of presentation. *Lancet* 1988; **ii**: 207-209.
- ⁵ Campbell LR, Charney E. Factors associated with delay in diagnosis of childhood amblyopia. *Pediatrics* 1991; **75**: 111-116.
- ⁶ Department of Health and Social Security. *Inequalities in health* (Black report). London: DHSS, 1980.
- ⁷ Reading R, Jarvis S, Openshaw S. Measurement of social inequalities in health and use of health services among children in Northumberland. *Arch Dis Child* 1993; **68**: 626-631.
- ⁸ Townsend P, Phillimore P, Beattie A. *Health and deprivation: inequality and the North*. London: Croom Helm, 1988.
- ⁹ Morris R, Carstairs V. Which deprivation? A comparison of selected deprivation indexes. *J Publ Hlth Med* 1991; **13**: 318-326.

Accepted on 15 March 1994

BEST COPY NOTE

**THE FOLLOWING PAGES
ARE STUCK IN SUCH A
MANNER THAT
FILMING IS IMPEDED**

APPENDIX C

THE PRESENTATION OF CHILDREN WITH AMBLYOPIA

G. WOODRUFF, F. HISCOX, J. R. THOMPSON and L. K. SMITH

Leicester

SUMMARY

This study reports the presentation of 961 children who underwent amblyopia treatment at seven orthoptic centres in the United Kingdom. We confirmed previous authors' findings of a small but significant increased incidence of left-sided compared with right-sided amblyopia overall. For pure anisometropic amblyopia this difference was very marked and a possible pathophysiological mechanism is proposed. The mean age of presentation for anisometropic, strabismic and mixed amblyopia was 5.6, 3.3 and 4.4 years, respectively. Neither sex nor race affected the age of presentation. Despite their older age, children with pure anisometropic amblyopia had the best initial visual acuity, with 25% of anisometropes having an initial visual acuity of less than 6/18 compared with 39% of strabismics and 50% of mixed amblyopes. The ages and initial acuities of the strabismic patients in this series are at least as favourable as those of patients reported from outside the UK. There were variations in the age and proportion of patients presenting with anisometropic amblyopia at the different centres, suggesting a failure in the referral of anisometropic amblyopia of importance in interpreting epidemiological studies.

The selection of patients and their clinical presentation are major factors responsible for the variable results of amblyopia treatment reported in the literature, with success rates ranging from 30% to over 90%.¹⁻⁴ There is very little information about the presentation of children with amblyopia in the United Kingdom so that the relevance to British practice of the different results of treatment claimed in the literature is not clear. In addition, previous epidemiological studies based on clinic data have relied upon data from a single centre. We report on the presentation of a cohort of 961 children started on treatment for amblyopia at seven English centres. The outcome of this treatment is reported separately.⁵

PATIENTS AND METHODS

A research orthoptist and research assistant visited each of

Correspondence to: Geoffrey Woodruff, FRCOphth, Department of Ophthalmology, University of Leicester, Clinical Sciences Building, Leicester Royal Infirmary, PO Box 65, Leicester LE2 7LX, UK.

seven English orthoptic centres which had agreed to participate in an audit of amblyopia treatment. The age at presentation, initial visual acuity and sex were recorded for every child who had had a first appointment in the orthoptic department in 1983, who had not had any previous treatment, and who was prescribed occlusion treatment for anisometropic, strabismic or mixed amblyopia at any time either at the first appointment or subsequently. Anisometropia was defined as a difference of 1 dioptre or more of either sphere or cylinder between the two eyes, and strabismus was defined as manifest strabismus on cover testing. The same research orthoptist determined the diagnosis from the information in the orthoptic records in the seven different centres. Children with an Asian surname or forename were classified as of Asian origin.⁶ Categorical data were analysed using the chi-squared (χ^2) test. Proportions were analysed using the Normal approximation to the Binomial distribution. Continuous data were analysed with regression analysis using SAS⁷.

RESULTS

Overall, 535 children had strabismic amblyopia (55%), 164 had pure anisometropic amblyopia (17%) and 262 had mixed anisometropic and strabismic amblyopia (27%) (Table I). There was a wide range in the proportion of patients with each type of amblyopia at the different centres ($\chi^2 = 21.1, p = 0.05$), with only 7% of the patients presenting with anisometropic amblyopia at centre B compared with 24% at centre C (Table I).

Left anisometropic amblyopia was more common (105 cases) than right anisometropic amblyopia (59 cases). For pure anisometropia this was so at each centre studied and was highly significant ($p = 0.0003$). For mixed anisometropic and strabismic amblyopia there was a less marked but still significant preponderance of left amblyopia (115 right, 147 left; $p = 0.048$), while for pure strabismic amblyopia there was no significant difference (252 right, 283 left; $p = 0.18$).

The mean age of first attending the orthoptic department for strabismic amblyopia was 3.3 years. Mixed strabismic and anisometropic amblyopia presented over 1 year later at 4.4 years. Pure anisometropes presented latest

Table I. Number of patients with each type of amblyopia at seven centres

Centre	Type of hospital	No. of children with each type of amblyopia			
		Strabismic	Anisometropic	Mixed	Total
A	District eye hospital	88	23	35	146
B	Teaching hospital	36	4	16	56
C	University eye hospital	57	31	43	131
D	Children's teaching hospital	77	20	31	128
E	District general hospital	23	10	11	44
F	Teaching hospital	113	29	68	210
G	Teaching hospital	141	47	58	246
Total		535	164	262	961

at 5.6 years (Table II). There was no significant difference in the age of presentation for children with strabismus (pure strabismic or mixed amblyopia) at the different centres (linear regression, $p = 0.14$ and $p = 0.13$ respectively). For pure anisometropic amblyopia the age of presentation was very significantly associated with the centre attended ($p = 0.0001$), with patients presenting to centres E and G 2 years or more later than to centres A and C. The mean age of presentation of patients with an Asian forename or surname was 4.7 years compared with 4.0 years in the non-Asians, but this was not significantly different after adjusting for centre and type of amblyopia (linear regression, $p = 0.87$). There was no significant difference in the age of presentation of girls (4.0 years) compared with boys (4.0 years), even after adjusting for centre and type of amblyopia (linear regression, $p = 0.91$).

In 80% of patients a measure of visual acuity was made before commencing treatment. The method used to determine visual acuity before commencing treatment varied according to age (Fig. 1). Except in two centres where Catford drum or mounted ball testing was used, the visual acuity of children in the 0–2 age group was not measured. The level of visual acuity prior to starting treatment recorded by Snellen or matching system is given in Fig. 2. Overall, of children tested using one of these methods, 38% had a visual acuity (VA) less than 6/18 and 62% had 6/18 or better. The mean visual acuity $\exp[(\Sigma \log VA)/n]$ was 6/20.3 ($n = 341$) for strabismic, 6/16.6 ($n = 162$) for pure anisometropic and 6/25.8 ($n = 208$) for mixed amblyopia. The range of visual acuities and the numbers of patients in each group were not large enough for differences between centres to be discernible.

DISCUSSION

A preponderance of left anisometropic amblyopia has not

been widely recognised. However, review of the literature shows that Massie,⁸ Cole,⁹ Burian¹⁰ and Lippmann¹¹ each found a slight overall preponderance of left amblyopia but did not analyse the contribution of different types of amblyopia to this observation. A consistent error in diagnosis due to a preference to test the acuity of the right eye first or due to a preference for refracting children from the right side rather than the left, could conceivably explain these results. However, we believe, in view of the previous literature, that our finding of an increased incidence of anisometropic amblyopia affecting the left eye is likely to be of pathophysiological significance. Right-handedness and right eye dominance are both more common than left dominance.¹² Fabian¹³ termed the progressive reduction in the range of refractive errors in children emmetropisation. It is not clear how this process is controlled but it appears to be a vision-dependent phenomenon.^{14,15} We postulate that when random variations in refraction occur in infancy, the vision of the more dominant eye may be more effective at driving emmetropisation, thus making significant refractive errors and anisometropic amblyopia less common in dominant eyes.

Shaw *et al.*¹⁶ noted an association between age of presentation and female sex and Asian origin. We found no evidence of this in our study. Pre-school vision screening has been introduced in an effort to promote an earlier approach to children with amblyopia. Children seen in the three centres (A, C, G) which received referrals from pre-school vision screening programmes were not significantly younger than children seen at other centres. However, we did not have sufficient data about the route of referral of individual patients or sufficient details of local screening programmes for any firm conclusions to be drawn about the effect of pre-school vision screening.

Table II. Age of presentation to seven orthoptic centres of patients with each type of amblyopia

Centre	<i>n</i>	Age of presentation (yr)			
		Anisometropic	Strabismic	Mixed	Total
A	146	4.8	2.9	4.4	3.6
B	56	5.5	3.0	3.6	3.4
C	131	4.8	3.7	5.1	4.4
D	128	5.4	3.1	4.6	3.8
E	44	7.1	3.4	3.7	4.3
F	210	5.0	3.6	4.1	3.9
G	246	6.7	3.3	4.4	4.2
Total	961	5.6	3.3	4.4	4.0

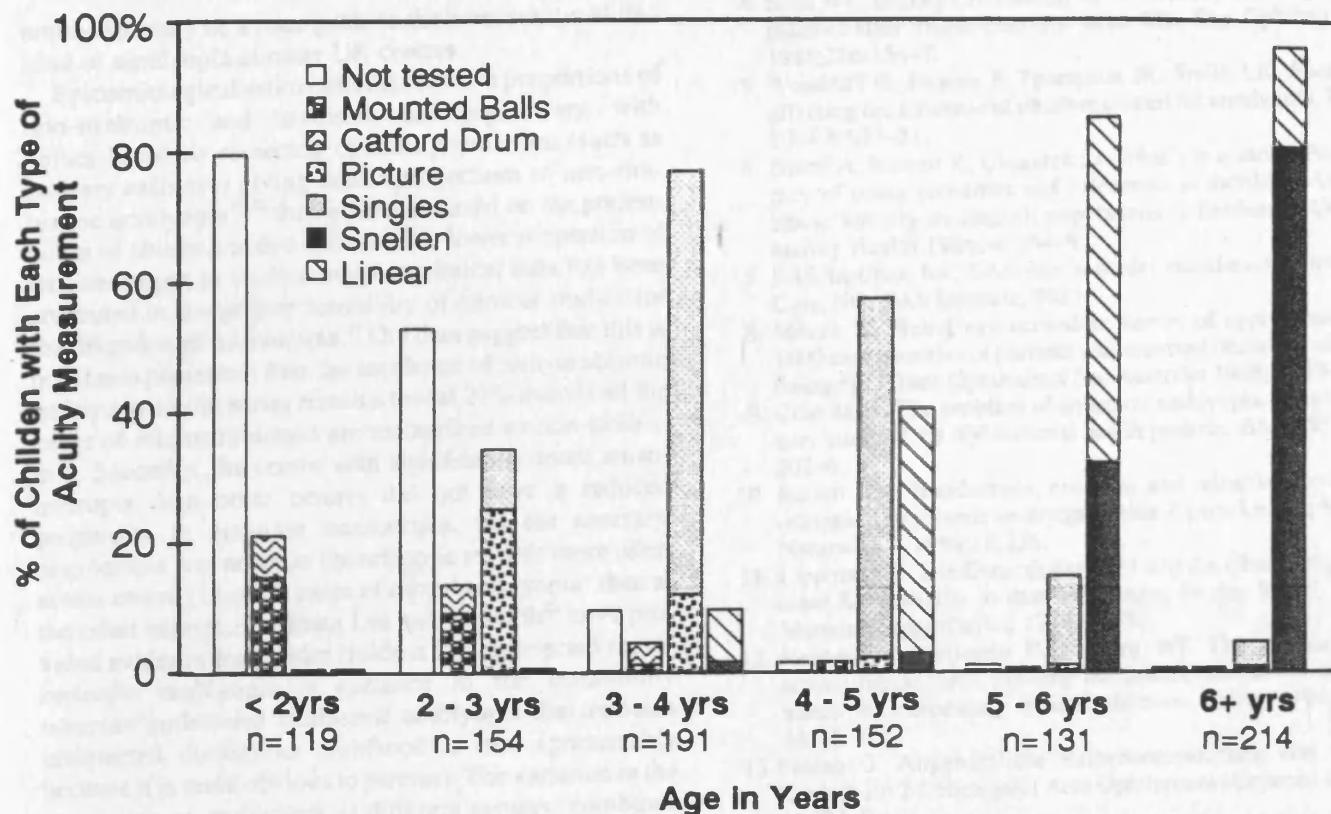


Fig. 1. Method of measuring the visual acuity of amblyopic patients presenting at different ages.

Sixty-two per cent of our patients had a visual acuity equal to or better than 6/18. While the Sheridan Gardiner method of testing may underestimate the degree of amblyopia in some patients it is clear that only a minority of children starting treatment in our centres had severe amblyopia. Lithander and Sjöstrand¹⁷ reported an overall age of children starting amblyopia treatment of 52 months – similar to our patients. However, only 27% of their patients had an acuity of 6/18 or better before starting treatment. Fulton and Mayer¹⁸ reported a median age at first clinic visit of approximately 50 months in their

patients with strabismic amblyopia compared with a mean age of 38 months in our series. The mean initial visual acuity ratio of their patients was approximately 0.34, suggesting a higher proportion of patients with an acuity of less than 6/18 than at any of our centres. Because of the relative consistency of our findings for strabismic amblyopia at the different centres, it is possible that the presentation of strabismic amblyopia follows a similar pattern in other centres in the UK. In contrast, because of the differences in age of presentation of the anisometropic patients at our different centres, our findings for anisometropic

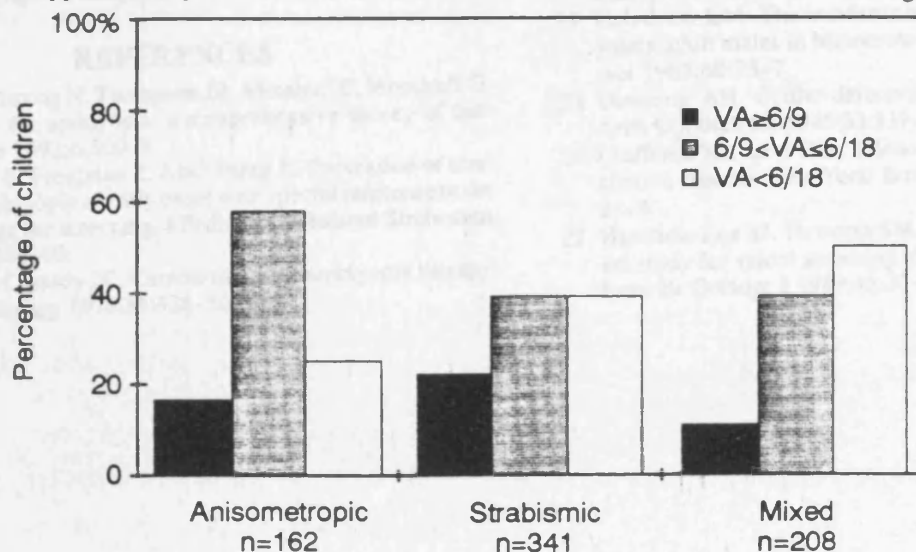


Fig. 2. Percentage of children with different levels of visual acuity recorded by Snellen or matching methods.

amblyopia may be a poor guide to the presentation of this kind of amblyopia at other UK centres.

Epidemiological estimates of the relative proportions of non-strabismic and strabismic amblyopia vary, with values based on screening of adult populations (such as military enlistees) giving higher proportions of non-strabismic amblyopia^{19,20} than estimates based on the presentation of children to eye clinics.³ The lower proportion of anisometropes in studies based on clinical data has been attributed to the greater sensitivity of clinical studies for the diagnosis of microtropia.²¹ Our data suggest that this is not the explanation: first the incidence of non-strabismic amblyopia in our series remains low at 23% even if all the cases of microstrabismus are reclassified as non-strabismic. Secondly, the centre with significantly more anisometropia than other centres did not have a reduced propensity to diagnose microtropia. On the contrary, microtropia was noted in the orthoptic records more often at this centre (14 of 42 cases of mixed amblyopia) than at the other centres. Hardman Lea and Haworth²² have provided evidence from older children that undetected anisometropic amblyopia is common in the community whereas undetected strabismic amblyopia that remains undetected throughout childhood is rare (presumably because it is more obvious to parents). The variation in the proportion of amblyopes at different centres, combined with the variation in the age of presentation of anisometropic amblyopia, suggests to us a failure in the referral of anisometropic amblyopia that may be important in interpreting epidemiological studies.

We thank Jackie Nolan, Liz Newcombe, Rowena McNamara, Susan Viney, Jenny Elmer-Monaghan, Sheelagh Baker and the staff of the Orthoptic Clinics at Birmingham Children's Hospital, Bristol Eye Hospital, Leeds General Infirmary, Leicester Royal Infirmary, Queen's Medical Centre, Nottingham, Sunderland Eye Infirmary and Worthing Hospital. We gratefully acknowledge the financial support of the British Council for the Prevention of Blindness, The Iris Fund for Prevention of Blindness, The British Orthoptic Society and the Anne Allerton Fund.

Key words: Age, Amblyopia, Anisometropia, Epidemiology, Presentation.

REFERENCES

1. Hiscox F, Strong N, Thompson JR, Minshull C, Woodruff G. Occlusion for amblyopia: a comprehensive survey of outcome. *Eye* 1992;6:300-4.
2. Neumann E, Freidman Z, Abel-Peleg B. Prevention of strabismic amblyopia of early onset with special reference to the optimal age for screening. *J Pediatr Ophthalmol Strabismus* 1987;24:106-10.
3. Flynn JT, Cassady JC. Current trends in amblyopia therapy. *Ophthalmology* 1978;85:428-50.
4. Scott WE, Dickey CF. Stability of visual acuity in amblyopic patients after visual maturity. *Arch Clin Exp Ophthalmol* 1988;226:154-7.
5. Woodruff G, Hiscox F, Thompson JR, Smith LK. Factors affecting the outcome of children treated for amblyopia. *Eye* 1994;8:627-31.
6. Nicoll A, Bassett K, Ulijaszek SJ. What's in a name? Accuracy of using surnames and forenames in ascribing Asian ethnic identity in English populations. *J Epidemiol Community Health* 1986;40:364-8.
7. SAS Institute Inc. SAS user's guide: statistics, version 5. Cary, NC: SAS Institute, 1985.
8. Massie H. Fixing eye occlusion: survey of approximately 1000 case histories of patients who received occlusion of the fixing eye. *Trans Ophthalmol Soc Australia* 1965;24:39-46.
9. Cole RBW. The problem of unilateral amblyopia: a preliminary study of 10 000 national health patients. *BMJ* 1959;1:202-6.
10. Burian HM. Handedness, eyedness and refractive error in relation to strabismic amblyopia. *Wiss Z Univ Leipzig Math Naturwiss R* 1969;18:236.
11. Lippmann O. Die Entartungszeichen und die Überwertigkeit einer Körperhälfte in ihrer Bedeutung für das Sehen. *Klin Monatsbl Augenheilkd* 1934;92:370.
12. Hebben N, Benjamin D, Milberg WP. The relationship among handedness, sighting dominance and acuity dominance in elementary school children. *Cortex* 1981;17:441-6.
13. Fabian G. Augenärztliche Reihenuntersuchung von 1200 Kindern im 2 Lebensjahr. *Acta Ophthalmol (Copenh)* 1966;44:473-9.
14. Raviola E, Wiesel T. The mechanism of lid-suture myopia. *Acta Ophthalmol (Copenh)* 1988;185:91-2.
15. Rabin J, Van Sluyters RC, Malach R. Emmetropization: a vision-dependent phenomenon. *Invest Ophthalmol Vis Sci* 1981;20:561-4.
16. Shaw DE, Fielder AR, Minshull C, Rosenthal AR. Amblyopia: factors influencing age of presentation. *Lancet* 1988;2:207-9.
17. Lithander J, Sjöstrand J. Anisometropic and strabismic amblyopia in the age group 2 years and above: a prospective study of the results of treatment. *Br J Ophthalmol* 1991;75:111-6.
18. Fulton AB, Mayer DL. Esotropic children with amblyopia: effects of patching on acuity. *Graefes Arch Clin Exp Ophthalmol* 1988;226:309-12.
19. Helveston EM. The incidence of amblyopia ex anopsia in young adult males in Minnesota in 1962-63. *Am J Ophthalmol* 1965;60:75-7.
20. Downing AH. Ocular defects in sixty thousand selectees. *Arch Ophthalmol* 1945;33:137-43.
21. Ciuffreda KJ, Levi DM, Selenow A. Amblyopia: basic and clinical aspects. New York: Butterworth Heinemann, 1990: 21-3.
22. Hardman Lea SJ, Haworth SM. An assessment of the present study for visual screening of school children in Nottingham. *Br Orthopt J* 1989;46:20-5.

APPENDIX D

FACTORS AFFECTING THE OUTCOME OF CHILDREN TREATED FOR AMBLYOPIA

G. WOODRUFF, F. HISCOX, J. R. THOMPSON and L. K. SMITH

Leicester

SUMMARY

The outcome of treatment for amblyopia and the factors that affect this are not well understood. A major reason for this has been the exclusion from previous large studies of a sometimes unknown number of patients because of failure to comply with treatment. This paper analyses the outcome of amblyopia treatment in a retrospective review of the orthoptic records of a cohort of 961 children treated for amblyopia at seven centres who first attended in 1983. The final visual acuity was recorded by Snellen or matching methods in 894 children (93%). Of these, 48% achieved 6/9 or better, 35% less than 6/9 but better than or equal to 6/18, and 17% achieved less than 6/18. The outcome was best for pure anisometropic amblyopia, intermediate for pure strabismic amblyopia and least good for mixed strabismic and anisometropic amblyopia with a final visual acuity of 6/10.2, 6/12.8 and 6/14.8 respectively. While the age at start of treatment did not correlate with final visual acuity both poor initial visual acuity and poor compliance were associated with poor outcome. The main factor affecting the outcome of amblyopia treatment is the initial visual acuity. Comparison with the literature suggests that the results of treatment in this country may be falling far short of what would be possible in ideal circumstances with unlimited resources.

There is a wide range in the reported results of treatments of amblyopia with success rates ranging from 30% to 92%.^{1,2} Reasons for this include the selection of patients and the exclusion from some studies of those not completing treatment.³ There have been few studies with large numbers of patients which include information on children who do not complete treatment and there have been no multicentre studies. This paper analyses factors associated with the outcome of amblyopia treatment in a multicentre study of 961 children treated for amblyopia and followed for up to 10 years.

Correspondence to: Geoffrey Woodruff, FRCOphth, Department of Ophthalmology, University of Leicester, Clinical Sciences Building, Leicester Royal Infirmary, PO Box 65, Leicester LE2 7LX, UK.

PATIENTS AND METHODS

Data were collected on all 961 patients first seen at each of seven English Orthoptic Centres in 1983 who were prescribed occlusion for anisometropic, strabismic, or mixed amblyopia at any time either at the first appointment or up to 10 years subsequently. The visual acuity (including the method of measurement) and the treatment prescribed at each visit were entered into a database from the orthoptic records. The same orthoptist supervised the collection of data at each centre. Details of the presentation of these children are given in an accompanying paper.⁴ The degree of social deprivation of each child was estimated by the Townsend deprivation score⁵ using electoral ward data from the 1981 census linked to the postcode of each patient. Statistical analysis of social deprivation was based on these raw values although for tabulation the deprivation scores are grouped into quintiles.

RESULTS

Outcome for Different Types of Amblyopia

The final visual acuity, i.e. the best visual acuity within three visits of the cessation of treatment, was recorded by either matching (49%) or Snellen methods (51%) in 894 (93%) of the children. Of these, 48% achieved 6/9 or better, 35% 6/18 or 6/12 and 17% achieved less than 6/18. The outcome was better for pure anisometropes, intermediate for pure strabismic patients and least good for mixed strabismic and anisometropic patients (Fig. 1), with a mean final visual acuity (VA) $\exp[(\sum \log VA)/n]$ of 6/10.2 ($n = 163$), 6/12.8 ($n = 477$) and 6/14.8 ($n = 254$), respectively. This was statistically significant ($p < 0.0001$).

There was a significant relationship between difference in spherical equivalent between the two eyes and final visual acuity amongst those children with anisometropia (linear regression: $n = 157$, $p < 0.0001$ for pure anisometropes; $n = 248$, $p < 0.0001$ for mixed amblyopes), with worse final visual acuity associated with higher degrees of anisometropia.

Age at Start of Treatment

Neither the age at presentation nor the age at the start of

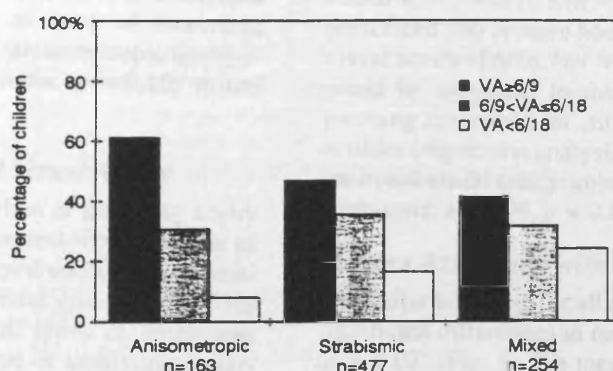


Fig. 1. Percentage of children with different levels of final visual acuity for each type of amblyopia.

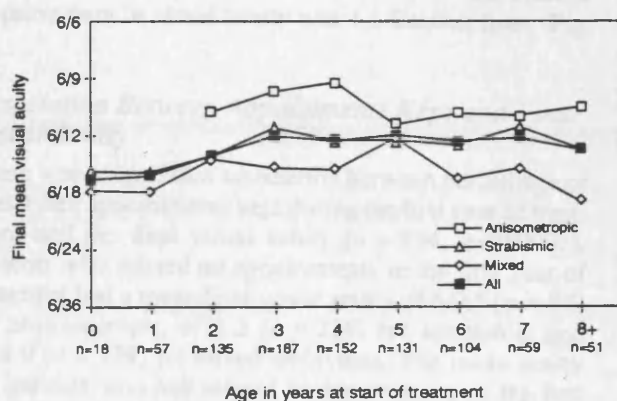


Fig. 2. Mean final visual acuity and age at starting treatment for each type of amblyopia and for all types combined. There was no significant association between age at start of treatment and final visual acuity.

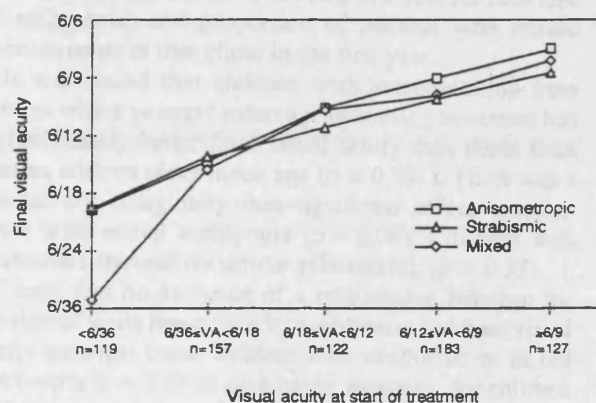


Fig. 3. Mean final visual acuity and visual acuity at start of treatment. The final visual acuity was significantly associated with the visual acuity at the start of the treatment.

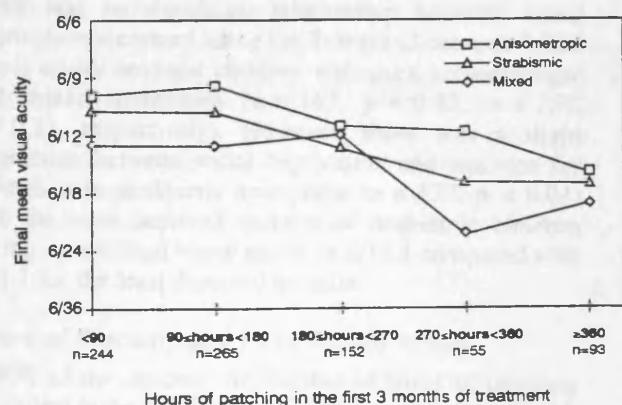


Fig. 4. Mean final visual acuity and hours of patching in the first 3 months of treatment.

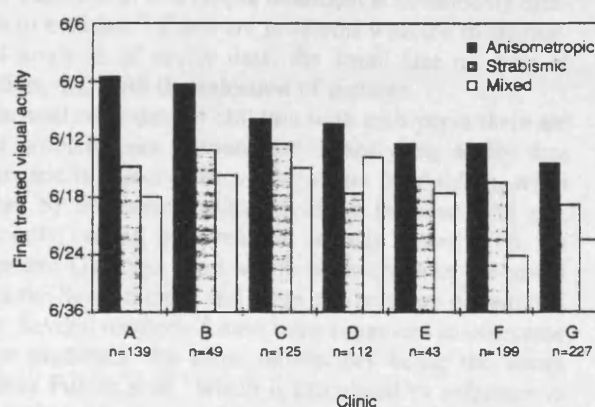


Fig. 5. Outcome of amblyopia treatment at seven centres (A-G). Regression analysis showed that even after allowing for differences in initial acuity there were significant differences between centres.

treatment (Fig. 2) was associated with the final outcome. This was true whether children with all types of amblyopia were considered together (age at start of treatment, $p = 0.08$) or analysed separately (anisometropic amblyopia: $p = 0.48$; strabismic amblyopia: $p = 0.10$; mixed amblyopia: $p = 0.64$).

Final Visual Acuity and Initial Visual Acuity

A total of 708 children had a Snellen or matching acuity measurement prior to starting treatment. For each type of amblyopia the visual acuity at referral and at start of treatment correlated closely with the final visual acuity. This was statistically significant for all types of amblyopia ($n = 708$, $p < 0.0001$) and each type of amblyopia separately (anisometropic amblyopia: $n = 162$, $p < 0.0001$; strabismic amblyopia: $n = 338$, $p < 0.0001$; mixed amblyopia: $n = 208$, $p < 0.0001$). For children whose acuity at start of treatment was less than 6/9 the average improvement in visual acuity was 1.8 Snellen lines (Fig. 3).

Association Between Appointments Kept and Final Visual Acuity

There was a significant association between percentage of prescribed appointments kept during the first year of treatment and the final visual acuity ($n = 894$, $p < 0.0001$). Patients who missed no appointments in the first year of treatment had a mean final visual acuity of 6/9.5 ($n = 94$) for anisometropic, 6/11.3 ($n = 225$) for strabismic and 6/14.0 ($n = 139$) for mixed amblyopia. The mean acuity for patients who had missed an appointment in the first year of treatment was 6/11.1 ($n = 69$), 6/14.3 ($n = 252$) and 6/15.9 ($n = 115$), respectively.

Association Between Social Deprivation and Outcome

There was no significant relationship between social deprivation measured using the Townsend score and final visual acuity amongst children with pure anisometropic and mixed amblyopia ($n = 163$, $p = 0.45$; $n = 253$, $p = 0.33$, respectively). However, there was a slight association between social deprivation and outcome for patients with strabismic amblyopia ($n = 477$, $p = 0.04$) with the most deprived quintile of strabismic children having a mean final visual acuity of 6/14.4 compared with 6/11.7 for the least deprived quintile.

Hours of Patching and Final Visual Acuity

In 80% of the children the number of hours of patching prescribed in the first 3 months of treatment was recorded. There was a highly significant relationship between hours of patching prescribed in the first 3 months and final visual acuity for all types of amblyopia ($n = 809$, $p = 0.0001$) and for each type of amblyopia analysed separately (Fig. 4) (anisometropic: $n = 152$, $p = 0.008$; strabismic: $n = 423$, $p = 0.007$; mixed: $n = 234$, $p = 0.004$). On average children who were prescribed less than 90 hours of

patching in the first 3 months of treatment had a final visual acuity of 6/10.8 ($n = 244$), while children who were prescribed 360 or more hours of patching had a mean final visual acuity of 6/16.7 ($n = 93$). However, this association could be attributed to the greater number of hours of patching prescribed for children with poorer initial visual acuities (regression analysis of final visual acuity adjusted for initial visual acuity: anisometropic: $n = 152$, $p = 0.31$; strabismic: $n = 299$, $p = 0.81$; mixed: $n = 193$, $p = 0.90$).

Factors Associated with Clinic

Even after adjusting for all other variables there remained significant differences in outcome at the different centres ($p = 0.02$) (Fig. 5). We therefore analysed the characteristics of different centres to see whether these related to outcome. The following factors were considered: average age of presentation of each type of amblyopia to each centre; mean number of visits in the first year for each type of amblyopia; and proportion of patients who missed appointments at that clinic in the first year.

It was found that children with anisometropia from centres with a younger mean age of starting treatment had a significantly better final visual acuity than those from centres with an older mean age ($p = 0.001$). There was a similar but marginally non-significant effect amongst those with mixed amblyopia ($p = 0.06$). Children with strabismus showed no similar relationship ($p = 0.33$).

There was no evidence of a relationship between the number of visits prescribed by each centre and final visual acuity amongst those children with strabismic or mixed amblyopia ($p = 0.68$ in each case). However, for children with anisometropic amblyopia centres which prescribed, on average, more visits to their patients had better results ($p = 0.003$).

DISCUSSION

The outcome of amblyopia treatment is notoriously difficult to evaluate.⁶ There are problems with the mathematical analysis of acuity data, the small size of most of studies, and with the selection of patients.

In analysing data of children with amblyopia there are two problems not encountered in analysing acuity data from adults. Firstly, the visual acuity of children, when tested by the same method, tends to improve with age; secondly, young children are usually tested with the Sheridan Gardiner chart while older children are tested with the Snellen chart and these two tests are not equivalent. Several methods^{7,8} have been suggested to overcome these problems, the most satisfactory being the acuity ratio of Fulton *et al.*⁹ which is calculated by reference to the acuity of the non-amblyopic eye. In our series a visual acuity of better than 6/6 was almost never recorded and thus we were not confident that the minimum angle of resolution of the non-amblyopic eye had been accurately recorded. Better analysis of our data would have been possible if more accurate testing of the acuity in the better eye had been done.

In the largest series of amblyopic patients to date,

Massie¹⁰ reported on 949 patients treated at one centre. However, the method of selection of patients, particularly with regard to patients who failed to attend, is not clear. Since more than two thirds of the patients were aged 7 years or more when treatment started, it is unlikely that his series is representative of patients treated at most clinics in North America and Britain today. Similarly, Bremner¹¹ reported on 240 patients treated with the Cam stimulator, but the results of only 42% of the patients initially started on treatment are recorded in her study and nearly half of these were more than 8 years of age. More satisfactory studies of amblyopia outcome such as those by Kutschke *et al.*,¹² Fulton and Mayer,⁹ Neumann *et al.*⁸ and Lithander and Sjöstrand¹³ have reported on smaller numbers of patients and, with the exception of Lithander and Sjöstrand,¹³ these authors have all specifically excluded patients who failed to comply with treatment.

Overall 48% of the patients in this study achieved 6/9 or better visual acuity. This compares poorly with the more than 80% of the patients of Fulton & Mayer⁹ and of Ching *et al.*¹⁴ who achieved 6/9 or better and the 83% of the patients of Kutschke *et al.*¹² with anisometropic amblyopia who achieved 20/40 or better. At least part of this difference can be attributed to the exclusion of non-compliant patients in these other series. The results of treatment of our fully compliant patients were better, with 52% of the fully compliant patients achieving 6/9, compared with 44% of those who missed appointments. However, Lithander and Sjöstrand¹³ have reported the results of treatment of 44 consecutive children treated for amblyopia in the context of private practice and weekly follow-up. In this series no patients were excluded because of failure to attend. Lithander and Sjöstrand's¹³ results suggest that with ideal treatment and generous resources a success rate of nearly 100% can be aimed for rather than the 50% recorded from the centres we have studied in the United Kingdom.

Despite the large numbers in our series we, like previous authors, could find no significant association between young age at presentation and better outcome. In fact there was a tendency for the reverse to be true. The time of onset of strabismus, and therefore of strabismic amblyopia at least, varies widely between children. We found a marked correlation between initial visual acuity and outcome, showing that poor initial visual acuity indicates severe (i.e. difficult to treat) amblyopia. We suspect that later onset amblyopia tends to be less severe than earlier onset amblyopia, and also that during any delay between onset and start of treatment there is a progressive deterioration towards more severe amblyopia. These two factors could result in the finding of no overall correlation between age and outcome.

Like Lithander and Sjöstrand¹³ we found that good compliance was associated with better outcome. Since social deprivation did not make much difference to outcome, we suspect that their system of intensive follow-up contributed to better compliance in their series and was the main factor responsible for the difference in overall out-

come in the two series. Patients prescribed more hours of patching in the first year of treatment had a worse outcome than children prescribed fewer hours. For all children we demonstrated that this association could be attributed to the fact that more hours of patching were prescribed for children with worse initial visual acuities.

Overall we conclude that the main factor affecting the outcome of amblyopia treatment is the initial visual acuity. If differences in the severity of amblyopia at the start of treatment are taken into account the mean outcome for children with strabismic amblyopia at the different centres in our study was similar. However, there remained differences in outcome of anisometropic and mixed amblyopia and those centres having less good results with these types of amblyopia were those where patients with anisometropia presented late. It is not clear whether this is because the centres with the better results treated patients sooner after the onset of amblyopia than the centres with the less good results or whether the treatment itself was indeed better at these centres. For better analysis of treatment in future studies the visual acuity of the good eye should be recorded accurately. However, it appears that the outcome of amblyopia treatment in many centres in this country is worse than it could be in ideal circumstances with unlimited resources.

We thank Jackie Nolan, Liz Newcombe, Rowena McNamara, Susan Viney, Jenny Elmer-Monaghan, Sheelagh Baker and the staff of the Orthoptic Clinics at Birmingham Children's Hospital, Bristol Eye Hospital, Leeds General Infirmary, Leicester Royal Infirmary, Queen's Medical Centre Nottingham, Sunderland Eye Infirmary and Worthing Hospital. We gratefully acknowledge the financial support of the British Council for the Prevention of Blindness, The Iris Fund for Prevention of Blindness, The British Orthoptic Society and the Anne Allerton Fund.

REFERENCES

1. Flynn JT, Cassady JC. Current trends in amblyopia therapy. *Ophthalmology* 1978;85:428-50.
2. Scott WE, Dickey CF. Stability of visual acuity in amblyopic patients after visual maturity. *Graefes Arch Clin Exp Ophthalmol* 1988;226:154-7.
3. Hiscox FA, Strong N, Thompson JR, Minshull C, Woodruff G. Occlusion for amblyopia: a comprehensive survey of outcome. *Eye* 1992;6:300-4.
4. Woodruff G, Hiscox F, Thompson JR, Smith LK. The presentation of children with amblyopia. *Eye* 1994;8:623-6.
5. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the North. London: Croom Helm, 1988.
6. Romano PE. Needed: better outcome evaluation of treatment of amblyopia and strabismus. [editorial]. *Binocular Vis Q* 1991;6(2):71-2.
7. Meyer E, Mizrahi E, Perlman I. Amblyopia Success Index: a new method of quantitative assessment of treatment efficiency; application in a study of 473 anisometropic amblyopic patients. *Binocular Vis Q* 1991;6(2):83-90.
8. Neumann R, Oliver M, Gottesmann N, Shimshoni M. Prognosis for occlusive therapy for strabismic and anisometropic amblyopia and for different initial depths of amblyopia. *Chibret Int J Ophthalmol* 1989;6(3):22-6.
9. Fulton AB, Mayer DL. Esotropic children with amblyopia: effects of patching on acuity. *Graefes Arch Clin Exp Ophthalmol* 1988;226:309-12.
10. Massie H. Fixing eye occlusion: survey of approximately 1000 case histories of patients who received occlusion of the fixing eye. *Trans Ophthalmol Soc Aust* 1965;24:39-46.
11. Bremner MH. Visual acuity in the primary school child

OUTCOME OF AMBLYOPIA TREATMENT

631

- aged four to twelve years: a review of amblyopia treatment in this age group at Princess Margaret Hospital. *Aust J Ophthalmol* 1984;12:395-9.
12. Kutschke PJ, Scott WE, Keech RV. Anisometropic amblyopia. *Ophthalmology* 1991;98:258-63.
 13. Lithander J, Sjöstrand J. Anisometropic and strabismic amblyopia in the age group 2 years and above: a prospective study of the results of treatment. *Br J Ophthalmol* 1991;75:111-6.
 14. Ching FC, Parks MM, Friendly DS. Practical management of amblyopia. *J Pediatr Ophthalmol Strabismus* 1986;23:12-6.

APPENDIX E

J Pediatr Ophthalmol Strabismus 1995;32:98-101.

Factors Affecting Treatment Compliance in Amblyopia

Lucy K. Smith, MSc, John R. Thompson, PhD, Geoffrey Woodruff, FCOphth, and Fiona Hiscox, DBO

ABSTRACT

Amblyopia is the most common form of visual disability in children. Successful treatment by patching depends on compliance, but evidence of factors affecting compliance is limited and contradictory. Because there is a well established relationship between social deprivation and access to health care, we hypothesized that social deprivation might be associated with noncompliance. Data from a historical cohort of 961 children from seven English orthoptic clinics starting treatment for amblyopia in 1983 were used to study factors affecting compliance with amblyopia treatment. Children were classified as noncompliant if they failed to attend all appointments prescribed during the first year of treatment. There was a significant difference in compliance between centers ($P=.0001$). Overall, children with anisometropic amblyopia were more compliant than those with strabismus but this varied significantly between centers. A relationship between social deprivation and compliance was also found ($P=.00001$). Only 41% of children from the most deprived wards were compliant compared with 61% in the least deprived wards. Compliance was not found to be related to age at starting treatment.

From the Department of Ophthalmology, University of Leicester, Clinical Sciences Building, Leicester LE2 7LX, UK.

The authors acknowledge the staff of the orthoptic clinics at Birmingham Children's Hospital, Bristol Eye Hospital, Leeds General Infirmary, Leicester Royal Infirmary, Nottingham University Hospital, Sunderland Eye Infirmary, and Worthing Hospital. They also acknowledge the staff in the department of General Practice at St Mary's College, London, for access to the 1981 frozen postcode directory; and the financial support of the British Council for Prevention of Blindness, the Iris fund, the British Orthoptic Society, and the Anne Allerton Fund.

Material from Crown-copyright records have been made available through the Post Office and the ESRC Data Archive.

Reprint requests should be addressed to Lucy K. Smith, MSc, University of Leicester, School of Medicine, Clinical Sciences Building, Leicester Royal Infirmary, PO Box 65, Leicester LE2 7LX, UK.

INTRODUCTION

Amblyopia is the most common visual disability in children, with a prevalence of between 2% and 5%.^{1,2} In amblyopia, normal vision fails to develop because of a difference in vision between the two eyes in early life. It is usually associated with either misalignment of the eyes (strabismus), or a refractive error that is greater in one eye than the other (anisometropia), or a combination of these factors. Treatment usually involves patching the unaffected eye to stimulate use of the amblyopic eye.

Noncompliance is known to reduce the improvement a child could achieve from treatment,³ but there has been little investigation of this problem. Children do not like having their eyes occluded,⁴ and thus, in previous studies of the outcome of amblyopia treatment, 30%,⁵ 47%,⁶ even 59%⁷ of children have been excluded from analysis because of noncompliance.

Compliance with patching treatment is difficult to assess, so several workers have used clinic attendance as a surrogate measure but they have found conflicting results. Nucci et al⁸ classified patients as compliant if they attended all of their appointments during the first 6 months of treatment. They found that age, initial visual acuity, and refractive error affected compliance, with an increase in compliance with age. Oliver et al⁹ classified patients as compliant if they attended all of their appointments in the first year of treatment, but they found a decrease in compliance with age.

While a relationship between social deprivation and access to health care is well established,¹⁰ the relationship between social deprivation and compliance with prescribed treatment has been investigated relatively little. Gadjosik and Campbell¹¹ found that socioeconomic status was associated with therapists' estimation of the compliance of parents with a home exercise program for disabled children.

Using data from a multi-center study of patients treated for amblyopia, we showed that outcome of amblyopia treatment is related to compliance.³ We now investigate the factors affecting compliance using clinic attendance as a surrogate measure and taking data from a large multi-

TABLE 1
Percentage of Children (With Total Number of Children in Parenthesis)
Who Were Compliant for Each Type of Amblyopia and at Each Center

Clinic	Mean Visits	Type of Amblyopia			
		Anisometropic	Strabismic	Mixed	All
A	6.9	80% (10)	65% (23)	91% (11)	75% (44)
B	9.5	50% (4)	67% (36)	81% (16)	70% (56)
C	8.0	69% (29)	59% (113)	46% (68)	56% (210)
D	7.9	61% (23)	45% (88)	54% (35)	50% (146)
E	7.6	30% (20)	45% (77)	68% (31)	48% (128)
F	6.7	71% (31)	35% (57)	47% (43)	47% (131)
G	7.2	47% (47)	39% (141)	52% (58)	44% (246)
All	7.6	57% (164)	48% (535)	55% (262)	51% (961)

center study. We looked at age at starting treatment, social deprivation, orthoptic clinic attended, visual acuity at start of treatment, and other possible explanatory variables.

PATIENTS AND METHODS

Data were obtained from a follow-up study of amblyopia that included 961 children starting treatment in 1983 and followed up through 1992 in seven volunteering English orthoptic clinics. Consistency of information was achieved by using one research orthoptist to visit all seven clinics. Details of the study are given in Woodruff et al.¹² Strabismus was diagnosed in 535 (56%) cases, combined anisometropia and strabismus in 262 (27%) cases, and anisometropia in 164 (17%) cases.

We classified children as compliant if they attended all of their prescribed appointments in the first year of treatment. In this survey, 98% of patients were prescribed part-time occlusion in the first year. The relationship between compliance and the following variables was studied: type of amblyopia, age at starting treatment, visual acuity at start of treatment, sex, ethnicity, distance from orthoptic clinic, degree of social deprivation, refractive error, and orthoptic clinic attended.

The degree of social deprivation of each child was estimated by the Townsend deprivation score,¹³ a composite score based on ward level data from the census. This is calculated using the percentage of households with no car, the percentage of households with more than one person per room, the percentage of households not owner-occupied, and the percentage of economically active people who are unemployed. Children at each clinic were grouped into quintiles based on the Townsend deprivation score. This means that, for each center separately, the children were arranged in five ordered groups, ranging from the 20% who lived in the most affluent wards to the 20% who lived in the most deprived wards. In the population studied, the most common ethnic group as recorded by the census was white, with South Asian being the second largest group. Ethnic group was, therefore, assessed by

identifying children with a South Asian forename or surname as described by Nicoll et al.¹⁴

Categorical data were analyzed using the χ^2 test and ordered categorical data were analyzed using the χ^2 test for trend. Unconditional logistic regression was then performed to investigate the multivariate relationship between the factors affecting compliance. The principle conclusions arising from this model were the same as for the univariate analysis, and we, therefore, concentrate on the univariate analysis.

RESULTS

The 961 children were prescribed an average of 7.6 visits in the first year of treatment for amblyopia. Fifty-one percent of patients attended all of their prescribed appointments in the first year and hence were classified as compliant.

Table 1 compares the compliance rates for the different types of amblyopia at the seven centers. There is a significant difference in the proportion of compliant children at each clinic, ranging from 43% to 75% ($P = .0001$). Clinics with higher compliance rates were generally those which prescribed more visits in the first year. There is some evidence to show that children with strabismic amblyopia had slightly lower rates of compliance than those with anisometropia or mixed amblyopia ($P = .04$). However, this varied significantly between centers ($P = .01$).

When compliance is tabulated by quintile of Townsend deprivation score within clinic (Table 2), it can be seen that there is a significant decrease in compliance with increased deprivation ($P < .0001$), with compliance rates over 50% better in the least deprived areas compared with the most deprived areas.

Compliance was not seen to be significantly related to sex ($P = .43$), visual acuity at start of treatment ($P = .14$), difference in refractive error between the two eyes ($P = .6$) or ethnic group ($P = .11$). No other factors studied were seen to be significantly related to compliance. Table 3 shows compliance and age at start of treatment. It can be

TREATMENT COMPLIANCE IN AMBLYOPIA

TABLE 2
Percentage of Children Who Were Compliant by Quintile of Townsend Deprivation Score Within Clinic

Quintile of Townsend Deprivation Score	Total Number of Children	Number of Compliant Children	% of Children Who Were Compliant
0% to 20% Affluent	192	119	62%
20% to 40%	189	112	59%
40% to 60%	187	89	48%
60% to 80%	200	97	48%
80% to 100% Deprived	191	76	40%
Total	*959	493	51%
$\chi^2 = 25.3$, χ^2 test for trend = 23.2,	4° freedom, 1° freedom,	$P < .0001$ $P < .0001$	

* Data missing for two patients.

TABLE 3
Percentage of Children Who Were Compliant (Total Number of Children in Parenthesis)
at Different Ages of Starting Treatment for Each Type of Amblyopia

Age	Type of Amblyopia			
	Anisometropia	Strabismus	Mixed	All
0 to 1 year		51% (97)	73% (22)	55% (119)
2 to 3 years	61% (23)	50% (232)	54% (90)	52% (345)
4 to 5 years	53% (58)	44% (142)	53% (83)	49% (283)
6 to 7 years	53% (60)	46% (54)	47% (49)	49% (163)
8+ years	74% (23)	65% (10)	66% (18)	65% (51)
Total	57% (164)	48% (535)	55% (262)	51% (961)
χ^2 (4 degrees freedom)	3.45	3.85	5.2	5.07
P-value	0.327	0.427	0.267	0.280

seen that for each type of amblyopia, the youngest and oldest referrals were the most compliant, and those between 4 and 7 years old were the least compliant. These patterns are not statistically significant (anisometropic amblyopia: $P = .33$; strabismic amblyopia: $P = .43$; mixed amblyopia: $P = .27$).

Two further analyses were performed on the data using different criteria for compliance. Patients were classified as compliant if they attended all but one of their appointments in the first year of treatment and attended all of their appointments in the first 6 months of treatment. The conclusions drawn from these two analyses were the same as those for the original method of classification.

DISCUSSION

Compliance in amblyopia is difficult to measure but clinic attendance is a simple and frequently used measure. This measure of compliance has also been shown to be significantly related to the results of treatment.³ Parents may be reluctant to bring their child into the clinic if they are failing to apply prescribed treatment at home, so this measure not

only reflects the ability to attend the clinic but also psychological factors affecting attendance. However, further study is necessary to understand the degree of association between compliance to treatment and attendance at clinic.

Rates of compliance vary greatly between studies, although this is partly explained by differences in the definitions of compliance. Our overall rate of 51% compliance seems to compare poorly with that shown by Oliver et al,⁹ who used the same criteria as we did. Only the two smallest of our seven centers had compliance rates that exceeded the 65% of Oliver et al and compliance in our worst center was only 44%. Trying to explain this variability, we found that within our study those centers that arranged more visits in the first year of treatment had higher rates of compliance. This perhaps shows that these centers reinforce the importance of treatment compliance by inviting the patients into the clinic more frequently. There would certainly seem to be scope for improving compliance in some UK centers.

Social deprivation as measured using the Townsend deprivation score¹¹ has been shown to correlate highly with a range of health measures.¹⁵ In this study, the proportion of

compliant children in affluent areas was 50% higher than in the deprived areas. As deprivation is studied within each clinic the effect is not due to a difference in deprivation between the clinics. There was no evidence of a relationship between distance from orthoptic clinic and compliance. This suggests that compliance is not related to the actual distance travelled, but may be due to factors associated with deprivation such as the ability to attend the clinic, access to transportation, family support, and motivation.

Oliver et al⁹ and Nucci et al⁸ give conflicting evidence as to the relationship between age at starting treatment and compliance. Our results were similar to those of Nucci et al in that we found a higher rate of compliance among the youngest and oldest children and a decrease in compliance among the 3- to 5- year olds. However, the effect was not statistically significant.

Compliance is of great importance to the outcome of amblyopia treatment but is difficult to measure. At present, there is no accurate quantitative method for measuring compliance, and assessments can only be based on attendance at clinic during treatment or the opinions of the orthoptist when the patient is seen at the clinic. Here we have used a surrogate measure for compliance, but a further study to look at an actual measure is needed to validate our findings.

REFERENCES

1. von Noorden GK. *Binocular Vision and Ocular Motility: Theory and Management of Strabismus*, 4th ed. St Louis, Mo: Mosby; 1985:208-209.
2. Thompson JR, Woodruff G, Hiscox F, Strong N, Minshull C. The incidence and prevalence of amblyopia detected in childhood. *Public Health*. 1991;105:455-463.
3. Woodruff G, Hiscox F, Thompson JR, Smith LK. Factors affecting outcome of amblyopia treatment. *Eye*. 1994;8:627-631.
4. Kushner BJ. Amblyopia. In: Nelson LB, Calhoun JH, Harley RD, eds. *Pediatric Ophthalmology*, 3rd ed. WB Saunders, USA, 1991;107-121.
5. Neumann R, Oliver M, Gottesmann N, Shimshoni M. Prognosis for occlusion therapy for strabismic and anisometropic amblyopia and for different initial depths of amblyopia. *Chibret International Journal of Ophthalmology*. 1989;6:22-26.
6. Kutschke PJ, Scott WE, Keech RV. Anisometropic amblyopia. *Ophthalmology*. 1991;75:111-116.
7. Bremner MH. Visual acuity in the primary school child aged four to twelve years: a review of amblyopia treatment in this age group at Princess Margaret Hospital. *Aust J Ophthalmol*. 1984;12:395-399.
8. Nucci P, Alfarano R, Piantanida A, Brancato R. Compliance in antiamblyopia occlusion therapy. *Acta Ophthalmol*. 1992;70:128-131.
9. Oliver M, Neumann R, Chaimovich Y, Gotesman N, Shimshoni M. Compliance and results of treatment for amblyopia in children more than 8 years old. *Am J Ophthalmol*. 1986;102:340-345.
10. Reading R, Jarvis S, Openshaw S. Measurement of social inequalities in health and use of health services among children in Northumberland. *Arch Dis Child*. 1993;68:626-631.
11. Gadjosik, Campbell. Effects of weekly review, socio-economic status, and maternal belief on mothers' compliance with their disabled children's home exercise program. *Physical and Occupational Therapy in Pediatrics*. 1991;11:47-65.
12. Woodruff G, Hiscox F, Thompson JR, Smith LK. The presentation of children with amblyopia. *Eye*. 1994;8:623-626.
13. Townsend P, Phillimore P, Beattie A. *Health and Deprivation: Inequality and the North*. London: Croom Helm; 1988.
14. Nicoll A, Bassett K, Ulijaszek SJ. What's in a name? Accuracy of using surnames and forenames in ascribing Asian ethnic identity in English populations. *J Epidemiol Community Health*. 1986;40:364-368.
15. Morris R, Carstairs V. Which deprivation? A comparison of selected deprivation indexes. *J Public Health Med*. 1991;13:318-326.

APPENDIX F

Copy of questionnaire used in the 1992 amblyopia study

To be completed by the parent or guardian of «FirstName» «LastName»

How are you related to «FirstName»? ☐ Mother ☐ Father ☐ Other -----

How was «FirstName» normally brought to hospital?

☐ Own car ☐ Bus ☐ Train ☐ Lift from friend
☐ Bicycle ☐ Taxi ☐ On foot

Did someone have to take time off work to bring «FirstName»? ☐ Yes ☐ No

What problem was first noticed in «FirstName»?

☐ Squint ☐ Poor vision ☐ Unknown ☐ Other _____

How old was «FirstName» when a problem was noticed? ☐ Years ☐☐ Months

When was the problem first noticed by an ophthalmologist or orthoptist?

☐ Years ☐☐ Months

Who first noticed the problem? Who did «FirstName» see next?

e.g. parent → GP → Leicester Royal Infirmary
or Health visitor → orthoptic screening → Ophthalmologist

_____ → _____ → _____ →

Did «FirstName» ever have their vision checked without any problem being noticed?

☐ No ☐ Yes at age : ☐ Years ☐☐ Months

Had «FirstName» ever attended hospital or received any other specialist treatment ?

☐ No ☐ Yes Reason _____

Did / does «FirstName» go to school/nursery?

☐ No ☐ Yes since age ☐ Years ☐☐ Months

Is there a family history of amblyopia or lazy eye in «FirstName»'s family?

☐ No ☐ Yes Relation _____

How difficult was it to get «FirstName» to wear the patch ? ☐

- | | | |
|-----------------------------|---------------------|----------------------|
| 1. Extremely difficult | 2. Fairly difficult | 3. Some difficulties |
| 4. Very occasional problems | 5. No problems | |

Did «FirstName» wear the patch for the prescribed time ?

- | | | | |
|--------------------|---------------------|---------------------|--------------------------|
| 1. All of the time | 2. Most of the time | 3. Some of the time | <input type="checkbox"/> |
| 4. Occasionally | 5. Never | | |

When did «FirstName» usually wear the patch ? _____
(e.g. mornings, at school)

Did «FirstName» wear the patch for one continuous period or several short periods

☐ Continuous ☐ Different times

When did «FirstName» have most problems? _____

How many children are there in your family ?
(including «FirstName»)

☐ children

Which number child was «FirstName» ? i.e. 1st , 2nd etc.

☐ child

Have you moved since «FirstName» was born ? ☐ Yes ☐ No

What sort of accommodation do you live in ?

☐ House ☐ Flat ☐ Hostel/B&B ☐ Caravan

Do you own your own home ? (with or without mortgage) ☐ Yes ☐ No

How many adults usually live in your home ? adults

How many children usually live in your home ? children

How many rooms does your home have for its own use ?
(excluding kitchens and bathrooms) rooms

How many cars or vans are normally available for use by you or members of your home ?
 cars / vans

Is English the main language spoken in your home ? ☐ Yes ☐ No

Are you bringing up «FirstName» on your own ?

☐ No (with partner) ☐ No (with parents) ☐ Yes

(If yes)

Have you been alone since «FirstName» was born ? ☐ Yes ☐ No

Please answer the following questions about yourself and your partner

	Yourself	Your partner
How old are you ?	<input type="text"/> <input type="text"/> years	<input type="text"/> <input type="text"/> years
How would you assess your ethnic group ?		
1. White	<input type="checkbox"/>	<input type="checkbox"/>
2. Black African, Black Caribbean, Black Other	<input type="checkbox"/>	<input type="checkbox"/>
3. Indian, Pakistani, Bangladeshi	<input type="checkbox"/>	<input type="checkbox"/>
4. Chinese,	<input type="checkbox"/>	<input type="checkbox"/>
5. Other ethnic group	<input type="checkbox"/>	<input type="checkbox"/>
Which country were you born in ?	<input type="checkbox"/>	<input type="checkbox"/>
1. UK		
2. Ireland		
3. EEC country		
4. Africa		
5. India/Pakistan/Bangladesh		
6. West Indies/Caribbean		
7. Hong Kong/East Asia		
8. Other (specify)		
(If not UK) How long have you lived in the UK	<input type="text"/> <input type="text"/> years	<input type="text"/> <input type="text"/> years
At what age did you leave full-time education ?	<input type="text"/> <input type="text"/> years	<input type="text"/> <input type="text"/> years

Yourself

Your partner

Which of these things were you doing last week ? ☐

☐

1. Full time work

7. School/full time education

2. Part time work (less than 30 hours)

8. Unemployed and looking for job

3. Self employed with employees

9. Long term illness/disability

4. Self employed without employees.

10. Retired from paid work

5. On a Training Scheme

11. Looking after the home or family

6. Waiting to start a job

12. Other

.Please give the title of your present or last job (within the last ten years) and describe the main things done in that job.

Yourself

Your Partner

Full job

Title

Main

Duties

Thank you very much for your time .

You have been extremely helpful in completing this questionnaire

Please contact us if you have any further queries

APPENDIX G

Letter to parents asking them to participate in the amblyopia study

Parent or guardian of (child's name)

Address

Postcode

Reference:

Dear Sir / Madam,

SURVEY OF CHILDREN ATTENDING THE LEICESTER EYE DEPARTMENT

Your child, (child's name), attended the eye department as an outpatient during 1992. We are now doing a survey of referral patterns and looking at every child who first attended in that year. We are asking all these families to help us in our study and we would be grateful if you would answer some questions relating to you and your child.

We have been unable to contact you by telephone and so we would appreciate it if you could complete a short questionnaire relating to information about you and your child similar to that collected in the census, and return it to us in the prepaid envelope. All your answers will be treated in the strictest confidence. If you have any queries or you would prefer to answer the questions by telephone then please indicate this on the form at the bottom of the letter and return it to us or ring Mrs Bharti Patel on 0533 586624. If however you do not wish to have any further correspondence from us please tick the box at the bottom and return the letter to us in the prepaid envelope.

Thank you in anticipation of your help.

Yours faithfully,

Lucy Smith

Research Associate, Eye Department.

I would prefer to answer the questions by telephone ☐

Your telephone number _____

Please give times you are available _____

I do not wish to have any further correspondence from you ☐

Children's vision screening: impact on inequalities in central England

Lucy K Smith, John R Thompson, Geoffrey Woodruff

Abstract

Study objective - To investigate the relationship between age at presentation of amblyopia and social deprivation before and after the introduction of changes to a vision screening service.

Design - Two cohorts of children treated for amblyopia in 1983 and 1992.

Setting - The orthoptic department of Leicester Royal Infirmary.

Participants - The 209 patients treated for amblyopia who first attended the orthoptic department in 1983, and 203 who first attended in 1992.

Measurements - Age at presentation to the orthoptic department was the main outcome measure. Social deprivation was measured by Townsend deprivation score for the electoral ward in which the child lived, using 1981 and 1991 census data.

Main results - After the introduction of changes in the screening programme, the mean age at presentation of amblyopia associated with microtropia or no strabismus was reduced from 6.6 years to 5.0 years. In 1983 there was a significant relationship between deprivation and age at presentation ($p=0.0001$), with those from more deprived areas presenting later. No similar association was found in children referred in 1992 ($p=0.17$). There was no change in the mean age of presentation of amblyopia associated with a large angle of strabismus (3.3 years in 1983 and 1992) and no relationship between deprivation and age at presentation 1983 or 1992 ($p=0.24$ and $p=0.39$ respectively).

Conclusions - Since the introduction of changes to vision screening, the relationship between social deprivation and the age at presentation of asymptomatic amblyopia seems to have disappeared. Children are now referred earlier and those from deprived areas are not being overlooked.

Large angle strabismus in a child is usually noticed by their parents¹ and when they seek medical help any associated amblyopia is likely to be detected. In contrast, amblyopia associated with a very small angle of strabismus (microtropia) or with anisometropia offers no obvious outward signs and is usually only detected by a vision screening test.

Over the last 20 years a system has been in place in Leicestershire for children to be screened for amblyopia and strabismus. Screening takes place at 6 weeks, 7½ months, 18 months (since 1991), and 3½ years. At first, health visitors were responsible for most of the screening and children thought to have strabismus or amblyopia were referred to a consultant ophthalmologist via their GP. Between 1988 and 1991 radical changes were made to the county's vision screening. A major development has been the introduction of a secondary orthoptic screening service which allows a much more prompt and readily available referral service for children suspected of having amblyopia or strabismus at the initial screening. Although most of the initial screening is still carried out by health visitors, responsibility for child health surveillance has been transferred to GPs who are required to make a return for every child screened.

The Hall report² reviewed all screening services for preschool children and highlighted the fact that many new screening programmes were introduced before their benefit had been established. Although the report noted the widespread practice of preschool vision screening, it found no evidence of health gain to support this practice and questioned the continuation of screening tests for amblyopia and strabismus. Orthoptic based preschool screening has been shown to be more effective³ but there has been little research to investigate the effectiveness of secondary orthoptic screening.

In a previously reported multicentre study we showed that the age at presentation of children with amblyopia but no strabismus was related to social deprivation⁴ but that there was no similar relationship among children with strabismus. We suggested that this may be because deprivation does not affect the detection of large angle strabismus but does affect a child's access to the screening necessary to detect amblyopia associated with anisometropia or microtropia. This paper investigates the changes in the age of detection of amblyopia over a nine year period during which there have been major changes in the screening services. We hypothesised that among children with microtropia or no strabismus, the introduction of improved referral

Department of Ophthalmology, University of Leicester, Clinical Sciences Building, Leicester LE2 7LX
L K Smith
J R Thompson
G Woodruff

Correspondence to:
Ms L K Smith.

Accepted for publication
May 1995

(J Epidemiol Community Health 1995;49:606-609)

Amblyopia is the most common visual disability in children, with a prevalence estimated to be between 2 and 5%.^{1,2} It is characterised by poor visual acuity in an otherwise healthy eye caused by failure to experience normal vision during early childhood. Misalignment of the eyes (strabismus), and differences in refractive error between the two eyes (anisometropia) may cause a child to use one eye less than the other and so cause amblyopia.

Mean age at presentation (numbers in parenthesis) in relation to the quintile of Townsend deprivation score, year, and type of amblyopia

Quintile of Townsend deprivation score	Type of amblyopia			
	Microtropia or no strabismus		Large angle of strabismus	
	1983	1992	1983	1992
1 Least deprived	5.4 (10)	4.5 (17)	2.8 (19)	3.8 (14)
2	5.4 (8)	4.8 (20)	3.2 (28)	3.3 (18)
3	6.5 (16)	5.4 (21)	3.3 (28)	3.3 (25)
4	7.3 (17)	5.3 (20)	3.4 (42)	2.9 (26)
5 Most deprived	7.1 (19)	4.7 (14)	3.5 (22)	3.2 (28)
All	6.6 (70)	5.0 (92)	3.3 (130)	3.3 (111)
Slope (deprivation)	0.204	0.069	0.051	-0.041
95% confidence interval	0.10, 0.31	-0.03, 0.17	-0.03, 0.12	-0.13, 0.05
p value	0.0001	0.17	0.24	0.39

processes would lessen or reduce the effect of social deprivation on the age at presentation. We expected no change among children with large angle strabismus.

Methods

Data were collected on two cohorts of children defined as all children who first attended Leicester Royal Infirmary orthoptic clinic in 1983 or 1992 and who were subsequently treated for amblyopia. Comprehensive information was collected from the orthoptic notes for each patient. This included the type of strabismus, ethnic origin (based on Asian or non-Asian forename⁶), postcode, age at presentation, and diagnosis. Large angle strabismus was defined as manifest strabismus on cover testing of more than 5° and microtropia as strabismus of 5° or less. Anisometropia was defined as at least one dioptre difference in refraction in either sphere or cylinder between the two eyes. Children who were first treated for amblyopia more than two years after first attending the orthoptic clinic were excluded as being unlikely to have had amblyopia at the time of presentation. To avoid the confounding effect of children attending outside clinics, only children from within a 10 km radius of the hospital are included in this analysis.

Each child's social deprivation was estimated using the Townsend deprivation score⁹ for the electoral ward in which they lived. This score is calculated using the percentage of households with no car; the percentage of households with more than one person per room; the percentage of households not owner occupied and the percentage of economically active people who are unemployed, as recorded in the census. A 1981 and 1991 computerised postcode directory linked the patient's postcode to the ward in which they lived. Data from the 1981 census were used to calculate the Townsend scores for the 1983 cohort and data from the 1991 census were used for the 1992 cohort. To calculate quintiles of deprivation for each census year we took all the wards in the 10 km radius and ranked them by their Townsend score. Wards were then placed into five groups containing approximately equal numbers of children, on the basis of the Townsend score, ranging from the least deprived group of wards to the most deprived group of wards.

Linear regression was performed separately for each year and type of strabismus, including

the deprivation score as a continuous variable. A model was fitted to the combined data for both years in order to test for an interaction between year and deprivation.

Results

In 1983, 209 children were treated for amblyopia, of whom 139 had amblyopia associated with a large angle of strabismus and 70 had a microtropia or did not have strabismus at all. In 1992, 203 children were treated, 111 with a large angle of strabismus and 92 others. There has thus been a significant change in the proportion of children treated for amblyopia with microtropia or no strabismus from 33% to 45% (a change of 12%, 95% confidence interval (2.5%, 21.2%)).

AMBLYOPIA ASSOCIATED WITH MICROTROPIA OR NO STRABISMUS

The table shows the mean age at presentation in relation to deprivation, year, and type of amblyopia. Among children with amblyopia associated with microtropia or no strabismus, the mean age of presentation was reduced by 19 months from 6.6 years in 1983 to 5.0 years in 1992 ($p=0.0001$ (95% CI 12.8, 26 months)). The table also shows that in 1983 there was a trend of increasing age of presentation with increasing deprivation but that in 1992 there is no similar pattern. This is illustrated in figures 1 and 2. Fitting a regression model to the data for 1983 shows a significant effect of deprivation on age at presentation ($p=0.0001$). The model indicates that on average children from the most deprived areas present two and a half years later than those from the least deprived areas. At first sight there seems to be a difference between the age at presentation of Asian children compared with non-Asian children ($p=0.09$) within the 1983 cohort but this difference is completely explained by adjusting for social deprivation ($p=0.98$). The same regression fitted to the data for 1992 shows the effects of ethnic origin and deprivation are both non-significant ($p=0.99$ and $p=0.17$ respectively).

The data for the two cohorts were combined in a single regression to test for an interaction between year and deprivation. This interaction was found to approach formal significance ($p=0.066$) supporting our prior hypothesis of a reduction in the effect of deprivation on age at presentation.

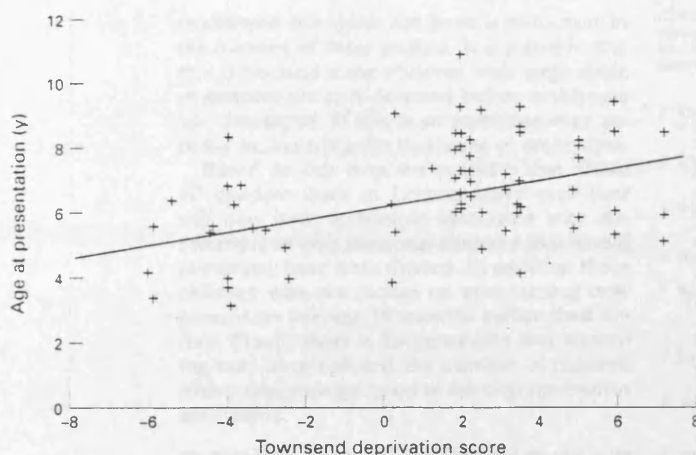


Figure 1 Age at presentation and Townsend deprivation score for children with amblyopia associated with microtropia or no strabismus in 1983

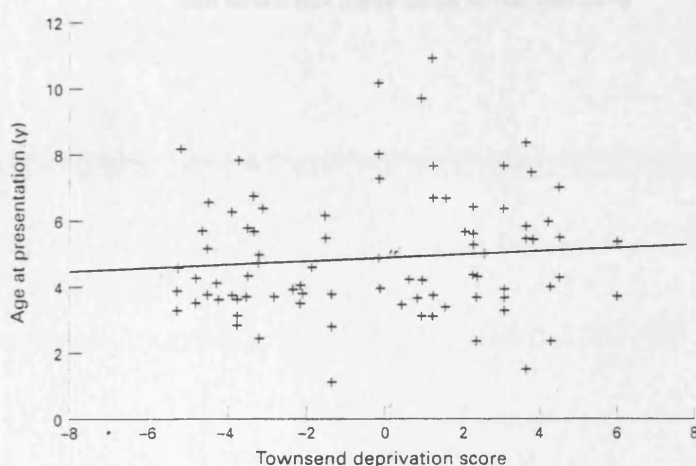


Figure 2 Age at presentation and Townsend deprivation score for children with amblyopia associated with microtropia or no strabismus in 1992.

AMBLYOPIA ASSOCIATED WITH A LARGE ANGLE OF STRABISMUS

The table shows there has been no apparent change in the mean age at presentation of children with a large angle of strabismus. In neither 1983 nor 1992 is there evidence of an effect of deprivation on age at presentation. Regression models showed no significant effect of deprivation or ethnic origin for 1983 or 1992.

Discussion

Between 1988 and 1991 the organisation of vision screening in Leicestershire for preschool children has been radically changed with the introduction of secondary orthoptic screening, an emphasis on GPs managing screening and formal monitoring of children as part of the child health surveillance programme. Over this period of change we have demonstrated a dramatic alteration in the referral pattern of children with amblyopia which we attribute to these improvements in screening.

There have been two changes in the pattern of presentation of amblyopia that is not associated with a large angle of strabismus; the average age has been reduced by 19 months and there is no longer any link between social deprivation and age at presentation.

It is important to consider whether these results could be partly due to the design of our study. By taking children who started treatment in a given year we do not have pure birth cohorts and it is likely that the experience of screening would have varied within the two groups. In particular the changes to child surveillance that were introduced in 1988 were phased in over several years so that some children in the 1992 group may not have experienced the full benefit of the changes. However, this effect would have acted to reduce the apparent impact of the changes to child surveillance. The other problem with the design is that we do not have details of how each child was detected and in particular whether this was through screening or self referral. Thus, we are working with population level information and it is possible, if unlikely, that some external factor may have acted to improve the self referral of children from poorer homes. However, we do know that few children with anisometropic or small angle strabismic amblyopia are detected other than by screening⁷ and we feel confident that any changes to the detection of these conditions would be attributable to changes in screening.

Before 1988 health visitors were required to refer children suspected of a vision problem to their GP who could refer them on to an ophthalmologist. This process offered the opportunity for delay, drop out, and error. The current system whereby children are referred directly from primary screening to the secondary orthoptic screen reduces delay, cuts down the possibility of a child dropping out of the system, and offers a trained assessment of the child's problem.

We believe that improved organisation of child health surveillance is the most likely reason for the removal of the relationship between social deprivation and age at presentation. The percentage of the population of children in Leicestershire screened at $3\frac{1}{2}$ years of age has increased from 80% in 1983 to 88% in 1992 and similarly those screened at $7\frac{1}{2}$ months has increased from 91% in 1983 to 95% in 1992. An improvement in coverage is likely to have particularly benefited children from poorer areas.

There were a third more cases of amblyopia without a large angle of strabismus treated in 1992 than in 1983, with no change in the size of the population. It is possible that in 1983 a smaller proportion of those detected with amblyopia were treated, or that there has been a rise in the incidence of amblyopia associated with microtropia or anisometropia but we think this is unlikely. Rather we believe that there has been an increase in detection with children being treated today who would previously have gone undetected.

There are no apparent changes in the age at presentation of children with large angle

strabismus but there has been a reduction in the number of cases treated. It is possible that this is because some children with large angle strabismus are now detected before amblyopia has developed. If this is so screening may actually be lowering the incidence of amblyopia.

Based on our data we estimate that about 40 children born in Leicestershire each year will now have amblyopia associated with microtropia or no strabismus detected that would previously have been missed. In addition those children who are picked up at screening now present on average 19 months earlier than before. Finally there is the possibility that screening may have reduced the number of children with strabismus going on to develop strabismus amblyopia.

We thank Fiona Hiscox and Diane Oliver and the staff of the orthoptic clinic at Leicester Royal Infirmary for their help in this study. We also thank the staff in the Department of General Practice at St Mary's College, London for access to the 1981 frozen postcode directory. Material from Crown-copyright records has been made available through the Post Office and the

ESRC Data Archive. We gratefully acknowledge the financial support of the British Council for Prevention of Blindness, the Iris fund for Prevention of Blindness, the British Orthoptic Society and the Anne Allerton Fund.

- 1 Von Noorden GK. Binocular vision and ocular motility: theory and management of strabismus. 4th ed. St. Louis: Mosby, 1985:208-9.
- 2 Thompson JR, Woodruff G, Hiscox F, Strong N, Minshull C. The incidence and prevalence of amblyopia detected in childhood. *Public Health* 1991;105:455-63.
- 3 Wang Yu-dong, Thompson JR, Goulstine DB, Rosenthal AR. A survey of the initial referral of children to an ophthalmology department. *Br J Ophthalmol* 1990;74:650-3.
- 4 Hall DMB ed. *Health for all children*. Oxford: Oxford University Press, 1989.
- 5 Bolger PG, Stewart-Brown SL, Newcombe E, Starbuck A. Vision screening in pre-school children: comparison of orthoptists and clinical medical officers as primary screeners. *BMJ* 1991;303:1291-4.
- 6 Jarvis SN, Tamhne RC, Thompson L, Francis PM, Anderson J, Colver AF. Preschool vision screening. *Arch Dis Child* 1990;65:288-94.
- 7 Smith LK, Thompson JR, Woodruff G, Hiscox F. Social deprivation and age at presentation of amblyopia. *J Public Health Med* 1994;16:348-51.
- 8 Nicoll A, Bassett K, Uliaszek SJ. What's in a name? Accuracy of using surnames and forenames in ascribing Asian ethnic identity in English populations. *J Epidemiol Community Health* 1986;40:364-8.
- 9 Townsend P, Phillimore P, Beattie A. *Health and deprivation: inequality and the North*. London: Croom Helm, 1988.

APPENDIX I

SAS program for fitting Espeland and Hui's log-linear model

```
proc iml;
/* set up the data and the design matrix*/
s={1 0 0 0 0 0 0 0,
   0 1 0 0 0 0 0 0,
   1 0 0 0 0 0 0 0,
   0 1 0 0 0 0 0 0,
   0 0 1 0 0 0 0 0,
   0 0 0 1 0 0 0 0,
   0 0 1 0 0 0 0 0,
   0 0 0 1 0 0 0 0,
   0 0 0 0 1 0 0 0,
   0 0 0 0 1 0 0 0,
   0 0 0 0 0 1 0 0,
   0 0 0 0 0 1 0 0,
   0 0 0 0 0 0 1 0,
   0 0 0 0 0 0 1 0,
   0 0 0 0 0 0 0 1,
   0 0 0 0 0 0 0 1};

m={1 1 1 1 1 1 1 1,
   1 -1 1 1 1 1 -1 1,
   1 1 -1 1 1 -1 1 1,
   1 -1 -1 1 1 -1 -1 1,
   1 1 1 -1 1 -1 -1 -1,
   1 -1 1 -1 1 -1 1 -1,
   1 1 -1 -1 1 1 -1 -1,
   1 -1 -1 -1 1 1 1 -1,
   1 1 1 1 -1 1 1 -1,
   1 -1 1 1 -1 1 -1 -1,
   1 1 -1 1 -1 -1 1 -1,
   1 -1 -1 1 -1 -1 -1 -1,
   1 1 1 -1 -1 -1 -1 1,
   1 -1 1 -1 -1 -1 1 1,
   1 1 -1 -1 -1 1 -1 1,
   1 -1 -1 -1 -1 1 1 1};

/* data for males aged 16-24 */
y={447,16307,689,28189,9384,7370,11404,17474};

/* Make an initial guess (this has been done by OLS on z */

z=s*inv(t(s)*s)*y;
b=inv(t(m)*m)*t(m)*log(z);

/* perform a few iterations to get a solution */
do i=1 to 10 by 1;
mu=exp(m*b);
d=diag(mu);
p=s*inv(t(s)*d*s)*t(s)*d;
x=t(p)*z;
t=t(p*m)*d*p*m;
delta=inv(t)*t(p*m)*(mu-z);
b=b-delta;
print delta,b;
end;
/* output the estimates of  $x_{ijkl}$  */
print x;
quit;
```

APPENDIX J

Glim program for calculation of relative risks using method of Ekholm et al

GLIM macros

\$mac eta *User supplied macro to define link function for each cell*

```
$scal %b=p(4)/%exp(p(5))
$scal %l(1)=45632*(p(1)*p(2)*P(4)+(1-P(1))*p(3)*%b)
$scal %l(2)=45632*(p(1)*(1-p(2))*p(4)+(1-p(1))*(1-p(3))*%b)
$scal %l(3)=45632*(p(1)*p(2)*(1-p(4))+(1-p(1))*p(3)*(1-%b))
$scal %l(4)=45632*(p(1)*(1-p(2))*(1-p(4))+(1-p(1))*(1-p(3))*(1-%b))
$scal %l(5)=45632*p(1)*p(2)
$scal %l(6)=45632*p(1)*(1-p(2))
$scal %l(7)=45632*(1-p(1))*p(3)
$scal %l(8)=45632*(1-p(1))*(1-p(3))
```

\$\$endmac

\$subfile NLIN4 *Supplied macros for fitting nonlinear models*

\$mac NLIN !

\$scal %z6=%len(P) \$var %z6 IN \$cal IN=1 \$!

\$var %z6 DLT_ \$cal DLT_=0.001 \$

\$init FIRST !

\$method * M1 \$!

\$scal %z4=%z6 \$whi %z4 DER \$!

\$endmac!

!

\$mac VAR \$endm!

!

\$mac FIRST \$cal %lp=0 : %z4=%z6 \$whi %z4 LP \$\$endm!

!

\$mac LP \$cal %z5=1+%z6-%z4 \$pr(store=VAR) ' p' *i %z5 \$!

\$scal %z7=IN(%z5) \$switch %z7 NDER \$!

\$scal %lp=%lp+P(%z5)*#VAR : %z4=%z4-1 \$\$endm!

!

\$mac M1 \$ext %pe \$cal P(IN*%cu(1))=%pe(IN*%cu(IN)) \$!

\$use ETA %eta \$use ASSIGN \$cal %z4=%z6 \$whi %z4 DER \$!

\$endm!

!

\$mac DER !

\$scal %z5=1+%z6-%z4 \$pr(store=VAR) ' P' *i %z5 \$!

\$scal %z7=IN(%z5) \$switch %z7 NDER \$!

\$scal %z4=%z4-1 \$\$endm!

!

\$mac NDER !

\$scal %z1=DLT_(%z5) : %z2=2*%z1 \$!

\$scal P(%z5)=P(%z5)+%z1 \$use ETA ETAU_ \$!

\$scal P(%z5)=P(%z5)-%z2 \$use ETA ETAL_ \$!

\$scal P(%z5)=P(%z5)+%z1 \$!

\$scal #VAR=(ETAU_-ETAL_)/%z2 \$!

\$scal R_=%abs((ETAU_-ETAL_)/(ETAU_+ETAL_)) \$!

\$scal R_=%if(R_>1,0.002,R_) \$!

\$stab the R_ m into %z9 \$!

\$scal DLT_(%z5)=DLT_(%z5)*0.001/%z9 \$!

\$endm!

\$return

GLIM program to execute macros with data

```
$unit 8$                8 observations per strata
$data y$
$read                  B* by A table, B* by B table, males 16-24
447 689 16307 28189    P(B*+A+), P(B*-A+), P(B*+,A-), P(B*-A-)
9384 11404 7370 17474 P(B*+B+), P(B*-B+), P(B*+,B-), P(B*-B-)
$yvar y$
$serr p$
$link i$
$ass p=0.5,0.45,0.3,0.05,0.1$ Sets initial estimates P(1)-P(5)
$use nlin$
$mac assign            Puts each vector of P into matrix TM
$calc %m=8$
$calc %n=5$
$arr m %m,%n$
$ass m=p1,p2,p3,p4,p5$
$calc %l=%m*%n$
$vari %l index$
$calc index=(%gl(%n,1)-1)*%m+%gl(%m,%n)$
$calc tm=m(index)$
$arr tm %m,%n$
$endmac

$use assign
$cycle 20 2 1.0e-8$
$set noconst$
$fit tm$              Fits model and displays estimates
$dis e$
$return $
```

BIBLIOGRAPHY

- Abrahamsson, M., Fabian, G., Andersson, A. K., & Sjostrand, J. (1990). A longitudinal study of a population based sample of astigmatic children. 1. Refraction and amblyopia. Acta Ophthalmologica, 68, 428-34.
- Agresti, A. (1988). A model for agreement between ratings on an ordinal scale. Biometrics, 44, 539-548.
- Aitkin, M. (1987). Modelling variance heterogeneity in Normal regression using GLIM. Applied Statistics, 36, 332-9.
- Altman, D. G. (1991). Practical statistics for medical research. London: Chapman and Hall.
- Armstrong, B. K., White, E., & Saracci, R. (1994). Principles of exposure measurement in epidemiology. Oxford: OUP.
- Atri, J., Falshaw, M., Livingstone, A., & Robson, J. (1996). Fair shares in health care - ethnic and socio-economic influences on recording of preventive care in selected inner London general practices. British Medical Journal, 312, 614-617.
- Bartley, M., & Blane, D. (1994). Appropriateness of deprivation indices must be ensured. British Medical Journal, 309, 1479.
- Beardsell, R. (1989). Orthoptic visual screening at 3½ years by Huntingdon Health Authority. British Orthoptic Journal, 46, 7-13.
- Benzeval, M., Judge, K., & Whitehead, M. (1995). Unfinished business. M. Benzeval, K. Judge, & M. Whitehead (Editors), Tackling inequalities in health: an agenda for action (pp. 122-140). London: King's Fund.

- Blane, D. (1991). Inequality and social class. G. Scambler (Editor), Sociology as applied to medicine (3 ed., pp. 109-128). London: Bailliere Tindall.
- Blaxter, M. (1990). Health and Lifestyles. London: Routledge.
- Blaxter, M. (1981). The health of the children: a review of research on the place of health in cycles of disadvantage. London: Heinemann Educational Books Limited.
- Bowman, R. J. C., Bennett, H. G. B., Houston, C. A., Aitchison, T. C., & Dutton, G. N. (1996). Waiting times for and attendance at paediatric ophthalmology outpatient appointments. British Medical Journal, 313, 1244.
- Bross, I. (1954). Misclassification in 2 X 2 tables. Biometrics, 10, 478-486.
- Butler, J. (1989). Child Health Surveillance in primary care: a critical review. London: Department of Health .
- Campbell, D. A., Radford, J. M. C., & Burton, P. (1991). Unemployment rates: an alternative to the Jarman index ? British Medical Journal, 303, 750-5.
- Campbell, F. W., Hess, R. F., Watson, P. G., & Banks, R. (1978). Preliminary results of a physiologically based treatment of amblyopia. British Journal of Ophthalmology, 62, 748-755.
- Campbell, L. R., & Charney, E. (1991). Factors associated with delay in diagnosis of childhood amblyopia. Pediatrics, 75, 111-116.
- Carr-Hill, R. (1994). Efficiency and equity implications of the health care reforms. Social Science and Medicine, 39, 1189-1201.
- Carr-Hill, R. A. (1993). Underenumeration in 1991 census: Census inaccurate for many ethnic groups. British Medical Journal, 307, 1563-4.

- Carr-Hill, R. A., Sheldon, T. A., & Thunhurst, C. (1992). The politics of deprivation indices and health. N. Spencer, & H. Janes (Editors), Uses and abuses of deprivation indices (pp. 24-32). Warwick: University of Warwick.
- Carr-Hill, R., & Rice, N. (1995). Is enumeration district level an improvement on ward level analysis in studies of deprivation and health. Journal of Epidemiology and Community Health, 49(Supplement 2), s28-s29.
- Carruthers, A. W., & Waugh, T. C. (1992). GIMMS reference manual. Edition 5. Edinburgh: GIMMS.
- Carstairs, V. (1995). Deprivation indices: their interpretation and use in relation to health. Journal of Epidemiology and Community Health, 49(Supplement 2), s3-s8.
- Carstairs, V., & Morris, R. (1989a). Deprivation and mortality: an alternative to social class. Community Medicine, 11, 210-219.
- Carstairs, V., & Morris, R. (1989b). Deprivation: explaining differences in mortality between Scotland and England and Wales. British Medical Journal, 299, 886-889.
- Clayton, D., & Hills, M. (1993). Statistical models in epidemiology. Oxford: Oxford University Press.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37-46.
- Davey Smith, G., Bartley, M., & Blane, D. (1990). The Black report on socio-economic inequalities in health 10 years on. British Medical Journal, 301, 373-377.
- Delamothe, T. (1991). Social inequalities in health. British Medical Journal, 303, 1046-1050.

- Denham, C., & Rhind, D. (1983). The 1981 Census and Its Results. D. Rhind (Editor), A Census User's Handbook (pp. 17-88). London: Methuen and Co. Ltd.
- Department of Health . (1991). The health of the nation: A consultative document for health in England. London: HMSO.
- Department of Health. (1992). The health of the nation: A strategy for health in England. London: HMSO.
- Department of Health. (1995). Variations in health: what can the Department of Health and the NHS do? London: Department of Health.
- DHSS. (1980). Inequalities in health: report of a research working group (The Black Report). London: DHSS.
- Diamond, E. L., & Lilienfeld, A. M. (1962). Effects of errors in classification and diagnosis in various types of epidemiological studies. American Journal of Public Health, 52, 1137-1144.
- DOE. (1983). Urban deprivation (Information note no. 2). London: Inner Cities Directorate, Department of Environment.
- Dolan. S.A., Jarman, B., Bajekal, M., Davies, P. M., & Hart, D. (1995). Measuring disadvantage: changes in the underprivileged area, Townsend, Townsend and Carstairs scores 1981-1991. Journal of Epidemiology and Community Health, 49(Suppl 2), S30-S33.
- Dorling, D. (1995). Visualizing the 1991 Census. S. Openshaw (Editor), Census Users' Handbook (pp. 167-211). Cambridge: GeoInformation International.
- Dosemeci, M., Wacholder, S., & Lubin, J. H. (1990). Does nondifferential misclassification of exposure always bias a true effect toward the null value . American Journal of Epidemiology, 132, 746-748.

- Drever, F., & Whitehead, W. (1995). Mortality in regions and local authority districts in the 1990s: exploring the relationship with deprivation. Population Trends, 82, 19-26.
- Duguid, G., & Grant, R. (1983). Areas of special need in Scotland. Edinburgh: Central Research Unit, Scottish Office.
- Ecob, R. (1996). A multi-level modelling approach to examining the effects of area of residence on health and functioning. Journal of the Royal Statistical Society Series A, 159, 61-76.
- Edwards, R. S., Whitelaw, A. J., & Abbott, A. G. (1989). Orthoptists as pre-school screeners: a 2-year study. British Orthoptic Journal, 46, 14-19.
- Eibschitz, N., Friedman, Z., & Neumann, E. (1978). Comparative results of amblyopia treatment. Metabolic Ophthalmology, 2, 111-2.
- Ekholm, A. (1991). Algorithms versus models for analyzing data that contain misclassification errors. Biometrics, 47, 1171-1182.
- Ekholm, A., & Green, M. (1995). Fitting nonlinear models in GLIM 4 using numerical derivatives. GLIM Newsletter, 23, 12-20.
- Ekholm, A., Green, M., & Palmgren, J. (1986). Fitting exponential family nonlinear models in GLIM 3.77. GLIM Newsletter, 13, 4-13.
- Ekholm, A., & Palmgren, J. (1987). Correction for misclassification using doubly sampled data. Journal of Official Statistics, 3, 419-429.
- Epelbaum, M., Milleret, C., Buisseret, P., & Dufier, J. L. (1993). The sensitive period for strabismic amblyopia in humans. Ophthalmology, 100, 323-327.
- Espeland, M. A., & Hui, S. L. (1987). A general approach to analyzing epidemiologic data that contain misclassification errors. Biometrics, 43, 1001-1012.

- Espeland, M. A., & Odoroff, C. L. (1985). Log-linear models for doubly sampled categorical data fitted by the EM algorithm. Journal of the American Statistical Association, 80, 663-670.
- Essen, J., & Wedge, P. (1982). Continuities in childhood disadvantage. London: Heinemann Educational Books Limited.
- Feinstein, J. S. (1993). The relationship between socio-economic status and health: A review of the literature. The Milbank Quarterly, 71, 279-322.
- Flegal, K. M., Brownie, C., & Haas, J. D. (1986). The effects of exposure misclassification on estimates of relative risk. American Journal of Epidemiology, 123, 736-751.
- Flegal, K. M., Keyl, P. M., & Nieto, F. J. (1991). Differential misclassification arising from nondifferential errors in exposure measurement. American Journal of Epidemiology, 134, 1233-1244.
- Fleiss, J. L. (1981). Statistical methods for rates and proportions. New York: John Wiley and Sons.
- Flynn, J. T., & Cassady, J. C. (1978). Current trends in amblyopia therapy. Ophthalmology, 85, 428-436.
- Folwell, F. (1995). Single measures of deprivation. Journal of Epidemiology and Community Health, 49(Suppl 2), S51-S56.
- Fox, A. J., Goldblatt, P. O., & Jones, D. R. (1985). Social class mortality differentials: artefact, selection or life circumstances ? Journal of Epidemiology and Community Health, 39, 1-8.
- Gilthorpe, M. S. (1995). The importance of normalisation in the construction of deprivation indices. Journal of Epidemiology and Community Health, 49(Suppl 2), S45-S50.
- Glover, G. R. (1993). Sex ratio errors in census data. British Medical Journal, 307, 506.

- Goldblatt, P. (1989). Mortality by social class, 1971-1985. Population Trends, 56, 6-15.
- Goldstein, H. (1995). Multilevel Statistical Models. London: Edward Arnold.
- Grant, J. (1994). The state of the world's children 1994. New York: Unicef.
- Hall, D. M. B. ed. (1989). Health for all children. Oxford: Oxford University Press.
- Hardman Lea, S. J., Loades, J., & Rubinstein, M. P. (1989). The sensitive period for anisometropia amblyopia. Eye, 3, 783-790.
- Haynes, R., Gale, S., Lovett, A., & Bentham, G. (1996). Unemployment rate as an updatable health needs indicator for small areas. Journal of Public Health Medicine, 18, 27-32.
- Hillis, A. (1986). Amblyopia: prevalent, curable, neglected. Public Health Reviews, 14, 213-235.
- Hillis, A., Flynn, J. T., & Hawkins, B. S. (1983). The evolving concept of amblyopia: a challenge to epidemiologists. American Journal of Epidemiology, 118, 192-205.
- Hiscox, F., Strong, N., Thompson, J. R., Minshull, C., & Woodruff, G. (1992). Occlusion for amblyopia: a comprehensive survey of outcome. Eye, 6, 300-304.
- Hopton, J. L., Howie, J. G. R., & Porter, A. M. D. (1992). Social indicators of health needs for general practice: a simpler approach. British Journal of General Practice, 42, 236-240.
- Illsley, R., & Le Grand, J. (1993). Regional inequalities in mortality. Journal of Epidemiology and Community Health, 47, 444-449.
- Ingram, R. M., Holland, W. W., Walker, C., Wilson, J. M., Arnold, P. E., & Dally, S. (1986). Screening for visual defects in preschool children. British Journal of Ophthalmology, 70, 16-21.

- Ingram, R. M., Walker, C., Wilson, J. M., Arnold, P. E., & Dally, S. (1986). Prediction of amblyopia and squint by means of refraction at age 1 year. British Journal of Ophthalmology, 70, 12-15.
- Irving, D. (1983). Identification of underprivileged areas (Technical report). London: School of Economics.
- Jarman, B. (1983). Identification of underprivileged areas. British Medical Journal, 286, 1705-1709.
- Jarman, B. (1990). Social deprivation and health service funding. Papers in Science, Technology, Medicine and Public Policy No. 22. London: Imperial College .
- Jarman, B. (1984). Underprivileged areas: validation and distribution of scores. British Medical Journal, 289, 1587-1592.
- Jarvis, S. N., Tamhne, R. C., Thompson, L., Francis, P. M., Anderson, J., & Colver, A. F. (1990). Preschool vision screening. Archives of Disease in Childhood, 65, 288-294.
- Jones, I. G., & Cameron, D. (1984). Social class analysis: An embarrassment to epidemiology. Community Medicine, 6, 37-46.
- Judge, K., & Benzeval, M. (1993). Health inequalities: new concerns about the children of single mothers. British Medical Journal, 306, 677-680.
- Kammerling, R. M., & O'Connor, S. (1993). Unemployment rate as predictor of rate of psychiatric admission. British Medical Journal, 307, 1536-1539.
- Kutschke, P. J., Scott, W. E., & Keech, R. V. (1991). Anisometropic amblyopia. Ophthalmology, 98, 258-63.

- Labour Party. (1995). Renewing the NHS: Labour's agenda for a healthier Britain. London: The Labour Party.
- Laurance, J. (1997). Inquiry launched into public health. The Independent, p. June 11.
- Leese, M., Loftus, L., & Thornicroft, G. (1995). Adjusting for underenumeration in the 1991 census (letter). British Medical Journal, 311, 394.
- Leese, R., & Fox, J. (1977). Registrar General's social classes: origins and uses. Population Trends, 8, 1-7.
- Levartovsky, S., Gottesman, N., Shimshoni, M., & Oliver, M. (1992). Factors affecting longterm results of successfully treated amblyopia: age at beginning of treatment and age at cessation of monitoring. Journal of Pediatric Ophthalmology and Strabismus, 29, 219-23.
- Li, P., & Logan, S. (1996). The current state of screening in general practice. Journal of Public Health Medicine, 18, 350-6.
- Liberatos, P., Link, B. G., & Kelsey, J. L. (1988). The measurement of social class in epidemiology. Epidemiologic Reviews, 10, 87-121.
- Lithander, J., & Sjostrand, J. (1991). Anisometropic and strabismic amblyopia in the age group 2 years and above: a prospective study of the results of treatment. British Journal of Ophthalmology, 75, 111-6.
- London Research Centre. (1992). SASPAC user manual. London: London Research Centre.
- Lynch, M. (1995). Effect of practice and patient population characteristics on the uptake of childhood immunizations. British Journal of General Practice, 45, 205-8.
- Macintyre, S. (1989). The role of health services in relation to inequalities in health in Europe. J. Fox (Editor), Health inequalities in European countries (pp. 317-32). Aldershot: Gower.

- Mackenbach, J. P. (1994). Socioeconomic inequalities in health in the Netherlands: impact of a five year research programme. British Medical Journal, 309, 1487-1491.
- MacLellan, A. V., & Harker, P. (1979). Mobile orthoptic service for primary screening of visual disorder in young children. British Medical Journal, 1, 994-995.
- MacRae, K. (1994). Socioeconomic deprivation and health and the ecological fallacy. British Medical Journal, 309, 1478-1479.
- Majeed, F. A., Cook, D. G., Poloniecki, J., & Martin, D. (1995). Using data from the 1991 census. British Medical Journal, 310, 1511-1514.
- Majeed, F. A., Martin, D., & Crayford, T. (1996). Deprivation payments to general practitioners: limitations of census data. British Medical Journal, 313, 669-670.
- Marmot, M. G., Davey Smith, G., Stansfield, S., Patel, C., North, F., Head, J., White, I., Brunner, E., & Feeney, A. (1991). Health inequalities among British civil servants: the Whitehall II study. The Lancet, 337, 1387-1393.
- Marsh, G. N., & Channing, D. M. (1986). Deprivation and health in one general practice. British Medical Journal, 292, 1173-6.
- McCarron, P. G., Davey Smith, G., & Womersley, J. J. (1994). Deprivation and mortality in Glasgow: changes from 1980 to 1992. British Medical Journal, 309, 1481-1482.
- McCullagh, P., & Nelder, J. A. (1988). Generalised linear models. London: Chapman and Hall.
- Mertens, T. E. (1993). Estimating the effects of misclassification. The Lancet, 342, 418-421.
- Meyer, E., Mizrahi, E., & Perlman, I. (1991). Amblyopia Success Index: a new method of quantitative assessment of treatment efficiency; application in a study of 473 anisometropic amblyopic patients. Binocular Vision Quarterly, 6, 83-90.

- Morgan, M., & Chinn, S. (1983). ACORN group, social class, and child health. Journal of Epidemiology and Community Health, 37, 196-203.
- Morganstern, H. (1982). Uses of ecologic analysis in epidemiologic research. American Journal of Public Health, 72, 1336-44.
- Morris, J. K., Cook, D. G., & Shaper, A. G. (1994). Loss of employment and mortality. British Medical Journal, 308, 1135-1139.
- Morris, J. N. (1990). Inequalities in health: ten years and little further on. The Lancet, 336, 491-493.
- Morris, R., & Carstairs, V. (1991). Which deprivation ? A comparison of selected deprivation indexes. Journal of Public Health Medicine, 13, 318-326.
- Neumann, E., Friedman, Z., & Abel-Peleg, B. (1987). Prevention of strabismic amblyopia of early onset with special reference to the optimal age for screening. Journal of Pediatric Ophthalmology and Strabismus, 24, 106-110.
- Newman, D. K., Hitchcock, A., McCarthy, H., Keast-Butler, J., & Moore, A. T. (1996). Preschool vision screening: outcome of children referred to the hospital eye service. British Journal of Ophthalmology, 80, 1077-82.
- NHS Centre for Reviews and Dissemination. (1995). Review of the Research on the effectiveness of health service interventions to reduce variations in health. York: NHS Centre for Reviews and Disseminations.
- Nicholl, A., Bassett, K., & Ulijasek, S. J. (1986). What's in a name ? Accuracy of using surnames and forenames in ascribing Asian ethnic identity in English populations. Journal of Epidemiology and Community Health, 40, 364-8.

- Nucci, P., Alfarano, R., Piantanida, A., & Brancato, R. (1992). Compliance in antiamblyopia occlusion therapy. Acta Ophthalmologica, 70, 128-31.
- Oliver, M., Neumann, R., Chaimovich, Y., Gotesman, N., & Shimshoni, M. (1986). Compliance and results of treatment for amblyopia in children more than 8 years old. American Journal of Ophthalmology, 102, 340-5.
- OPCS. (1978). Occupational mortality 1970-72. London: HMSO.
- Openshaw, S., & Blake, M. (1995). Geodemographic segmentation systems for screening health data. Journal of Epidemiology and Community Health, 49(Suppl 2), S34-S38.
- Orchard, C. (1994). Comparing health outcomes. British Medical Journal, 308, 1493-6.
- Palmgren, J., & Ekholm, A. (1987). Exponential family non-linear models for categorical data with errors of observation. Applied Stochastic Models and Data Analysis, 3, 111-124.
- Paton, C. R. (1992). The economics of screening. C. R. Hart, & P. Burke (Editors), Screening and surveillance in general practice (pp. 25-33). Edinburgh: Churchill Livingstone.
- Piantadosi, S. (1994). Ecologic biases. American Journal of Epidemiology, 139, 761-771.
- Piantadosi, S., Byar, D. P., & Green, S. B. (1988). The ecological fallacy. American Journal of Epidemiology, 127, 893-904.
- Quick, A., & Wilkinson, R. G. (1991). Income and Health. London: Socialist Health Association.
- Radical Statistics Health Group. (1991). Missing: a strategy for the health of the nation. R. Smith (Editor), Health of the Nation: the BMJ view. London: British Medical Journal.
- Reade-Christopher, S. J., & Kupper, L. L. (1991). Effects of exposure misclassification on regression analyses of epidemiologic follow-up study data. Biometrics, 47, 535-548.

- Reading, R. (1997). Poverty and the health of children and adolescents. Archives of Diseases in Childhood, 76, 463-7.
- Reading, R., Colver, A., Openshaw, S., & Jarvis, S. (1994). Do interventions that improve immunisation uptake also reduce social inequalities in uptake ? British Medical Journal, 308, 1142-1144.
- Reading, R., Jarvis, S., & Openshaw, S. (1993). Measurement of social inequalities in health and use of health services among children in Northumberland. Archives of Disease in Childhood, 68, 626-631.
- Reading, R., & Openshaw, S. (1993). Do inaccuracies in small area deprivation analyses matter ? Journal of Epidemiology and Community Health, 47, 238-241.
- Reading, R., Openshaw, S., & Jarvis, S. (1994). Are multidimensional social classifications of areas useful in UK health services research ? Journal of Epidemiology and Community Health, 48, 192-200.
- Rhind, D. (1983). Mapping Census Data. D. Rhind (Editor), A Census User's Handbook . London: Methuen.
- Rhind, D., & Tannenbaum, E. (1983). Linking census and other data. D. Rhind (Editor), A census user's handbook (pp. 287-300). London: Methuen.
- Roberts, H. (1997). Children, inequalities and health. British Medical Journal, 314, 1122-5.
- Robinson, W. S. (1950). Ecological correlations and the behaviour of individuals. American Sociological Review, 15, 351-7.
- Romano, P. E. (1991). Needed: better outcome evaluation of treatment of amblyopia and strabismus. Binocular Vision Quarterly, 6, 71-2.

- Royal College of Ophthalmologists and the British Paediatric Association. (1994). Ophthalmic services for children. Report of a joint working party. London: The Royal College of Ophthalmologists and the British Paediatric Society.
- SAS Institute Inc. (1992). SAS technical report P-229, version 6.07. Cary, NC.: SAS Institute.
- Scott, W. E., & Dickey, C. F. (1988). Stability of visual acuity in amblyopic patients after visual maturity. Graefe's Archive for Clinical and Experimental Ophthalmology, 226, 154-157.
- Shaw, D. E., Fielder, A. R., Minshull, C., & Rosenthal, A. R. (1988). Amblyopia: factors affecting age of presentation. Lancet, ii, 207-209.
- Shouls, S., Congdon, P., & Curtis, S. (1996). Modelling inequality in reported long term illness in the UK: combining individual and area characteristics. Journal of Epidemiology and Community Health, 50 , 366-376.
- Sitthi-amorn, C., & Poshychinda, V. (1993). Bias. Lancet, 342, 286-8.
- Sjostrand, J., & Abrahamsson, M. (1990). Risk factors in amblyopia. Eye, 4, 787-93.
- Sloggett, A., & Joshi, H. (1994). Higher mortality in deprived areas: community or personal disadvantage. British Medical Journal, 309, 1470-4.
- Smaje, C. (1995). Health, 'race' and ethnicity: making sense of the evidence. London: King's Fund Institute.
- Smith, L. K., Thompson, J. R., & Woodruff, G. (1995). Children's vision screening: impact on inequalities in central England. Journal of Epidemiology and Community Health, 49, 606-609.

- Smith, L. K., Thompson, J. R., & Woodruff, G. (1994). Social deprivation and age at presentation of anisometropia: the effect of introducing a screening service. Journal of Epidemiology and Community Health, 48, 507.
- Smith, L. K., Thompson, J. R., Woodruff, G., & Hiscox, F. (1995). Factors affecting treatment compliance in amblyopia. Journal of Pediatric Ophthalmology and Strabismus, 32, 98-101.
- Smith, L. K., Thompson, J. R., Woodruff, G., & Hiscox, F. (1994). Social deprivation and age at presentation in amblyopia. Journal of Public Health Medicine, 16, 348-351.
- Spencer, N. (1996). Poverty and child health. Oxford: Radcliffe Medical Press Limited.
- Spencer, N., Lewis, M., & Logan, S. (1993). Diagnostic and sociodemographic changes in multiple hospital admission in children under 2 over a 5 year period. Journal of Public Health Medicine, 15, 332-6.
- Stewart-Brown, S. L., Haslum, M. N., & Howlett, B. (1988). Preschool vision screening: a service in need of rationalisation. Archives of Disease in Childhood, 63, 356-359.
- Taylor, D. (1985). Screening ? Transactions of the Ophthalmological Societies of the United Kingdom, 104, 637-640.
- Taylor, D. (1987). Screening for squint and poor vision. Archives of Disease in Childhood, 62, 982-983.
- Thompson, E. J. (1995). The 1991 census of population in England and Wales. Journal of the Royal Statistical Society Series A, 158, 203-241.
- Thompson, J. R., Woodruff, G., Hiscox, F. A., Strong, N., & Minshull, C. (1991). The incidence and prevalence of amblyopia detected in childhood. Public Health, 105, 455-462.

- Thunhurst, C. (1985). The analysis of small area statistics and planning for health. The Statistician, 34, 93-106.
- Thunhurst, C., & MacFarlane, A. (1992). Monitoring the health of urban populations: what statistics do we need ? Journal of the Royal Statistical Association Series A, 155, 317-352.
- Townsend, P. (1993). The politics of poverty and health (Letter). British Medical Journal, 306, 337.
- Townsend, P. (1979). Poverty in the United Kingdom. London: Penguin.
- Townsend, P., Phillimore, P., & Beattie, A. (1988). Health and deprivation: inequality and the North. London: Croom Helm.
- Vereecken, E. P., & Brabant, P. (1984). Prognosis for vision in amblyopia after loss of the good eye. Archives of Ophthalmology, 102, 220-4.
- Von Noorden, G. K. (1985). Binocular vision and ocular motility: theory and management of strabismus. 4th Edition. St. Louis, MO.: Mosby.
- Wadsworth, M. (1988). Inequalities in child health. Archives of Disease in Childhood, 63, 353-5.
- Waller, D., Agass, M., Mant, D., Coulter, A., Fuller, A., & Jones, L. (1990). Health checks in General Practice: another example of inverse care? British Medical Journal, 300, 1115-8.
- Wang, Y. D., Thompson, J. R., Goulstine, D. B., & Rosenthal, A. R. (1990). A survey of the initial referral of children to an ophthalmology department. British Journal of Ophthalmology, 74, 650-653.
- Watson, P. G., Banks, R. V., Campbell, F. W., & Hess, R. F. (1978). Clinical assessment of a new treatment for amblyopia. Transactions of Ophthalmological Societies in the United Kingdom, 98, 201-208.

- Whitehead, M. (1988). Inequalities in health: The health divide. Harmondsworth: Penguin.
- Whitehead, M., & Dahlgren, G. (1991). What can be done about inequalities in health ? The Lancet, 338, 1059-1063.
- WHO. (1985). Targets for health for all: Targets in support of the European regional strategy for health for all by the year 2000. Copenhagen: World Health Organisation Regional Office for Europe.
- Wick, B., Wingard, M., Cotter, S., & Scheiman, M. (1992). Anisometropic amblyopia: is the patients ever too old to treat ? Optometry and Visual Science, 69, 866-878.
- Williamson, T. H., Andrews, R., Dutton, G. N., Murray, G., & Graham, N. (1995). Assessment of an inner city visual screening programme for preschool children. British Journal of Ophthalmology, 79, 1068-73.
- Wilson, J. M. G., & Junger, G. (1968). Principles and Practice. Public Health Papers No. 34. Geneva: WHO.
- Woodruff, G., Hiscox, F., Thompson, J.R., & Smith, L. K. (1994a). The presentation of children with amblyopia. Eye, 8, 623-626.
- Woodruff, G., Hiscox, F., Thompson, J. R., & Smith, L. K. (1994b). Factors affecting the outcome of children treated for amblyopia. Eye, 8, 627-631.