Explaining variations in mortality rates at general practice level

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

by

Catherine Eliza Honeyford BSc MSc
Department of Health Sciences
University of Leicester.

2014
Abstract

Explaining variations in mortality rates at general practice level.

Catherine E Honeyford.

This thesis aims to identify features of primary care associated with improved health outcomes using premature coronary heart disease (CHD) mortality as an example. Impacts of different modelling approaches are also explored.

A cross-sectional study of 229 general practices in the East Midlands was undertaken. The main outcome measure was numbers of premature CHD deaths in patients registered at the practices (April 2006 to March 2009). Publicly available data describing both population characteristics and aspects of primary care were utilised. A novel method of estimating smoking prevalence in practice populations was described and differing methods of describing the performance of primary care in detecting hypertension were evaluated.

Population characteristics and markers of quality of primary care were associated with variations in premature CHD mortality. Increases in: the percentage of practice populations on practice diabetes registers, the proportion who were over 65, the proportion who were male, and the estimated smoking prevalence in patients with chronic conditions were all associated with increasing levels of premature CHD mortality. Control of serum cholesterol levels in those with CHD and the percentage of patients recalling access to their preferred general practitioner, a measure of continuity of care, were both associated with decreased counts of premature CHD mortality. Increasing levels of undiagnosed hypertension prevalence were associated with increased levels of premature CHD mortality. Similar results were found for all-age mortality; there is less evidence that continuity of care is associated with all-age CHD mortality.

High-quality primary care, including aspects of access to and continuity of care, disease detection and management, appear to be associated with reduced CHD mortality. Data gathered as part of the Health Checks initiative has the potential to improve studies of this type, particularly if published by age group. Determining the most useful measures of quality of primary care needs further consideration.
Acknowledgements

Funding from NIHR CLAHRC LNR is gratefully acknowledged. The views expressed in this thesis do not necessarily reflect those of the Department of Health or the NIHR. I would like to thank the three PCTs involved for giving me access to the mortality data on which this research is based.

I would like to express my heartfelt thanks to my three supervisors for sharing their knowledge and expertise so generously. I have learnt so much during my studies - about statistics, health care, health services research and semi-colons. In particular I would like to thank Professor David Jones for helping me navigate my way through the many dead ends and blind corners of research; Professor Richard Baker for keeping the research grounded in the real world of primary care and Dr John Bankart for his unwavering attention to detail.

Thank you to the students and research associates who have normalised the daily concerns of a PhD student. A special mention to Dr Mark Rutherford and Dr Michael Crowther for their apparently infinite patience as they have answered my many questions about Stata and LaTeX. Many thanks to my parents and sister who have supported my mathematical development, improved my writing and interpreted many LaTeX error messages.

I would not have been able to do this without the unwavering support of my husband, Gareth, and my daughter, Beth. They gave me the courage to start, the faith to keep going and the motivation to finish. Thank you.
## Contents

I  Context  3

1 Introduction  4
   1.1 Introduction .................................................. 4
   1.2 Primary Care .................................................. 5
   1.3 Health outcomes ............................................. 7
   1.4 Multiple determinants of health ........................... 9
   1.5 Primary care and health outcomes ....................... 10
   1.6 Conclusions .................................................. 14

2 Primary care factors associated with improved health outcomes  15
   2.1 Introduction .................................................. 15
   2.2 Prevention, detection and disease management ........ 16
   2.3 Premature mortality .......................................... 17
   2.4 Access and continuity of care ............................. 17
   2.5 Conclusions .................................................. 20

3 Modelling Mortality  21
   3.1 Introduction .................................................. 21
   3.2 Measuring mortality ......................................... 22
   3.3 Multiple linear regression ................................ 23
   3.4 Count Models ................................................ 26
   3.5 Poisson models .............................................. 26
   3.6 Overdispersion .............................................. 30
   3.7 Models for overdispersed count data .................... 32
   3.8 Modelling approach - conclusions ...................... 34
   3.9 Variable selection .......................................... 35
   3.10 Conclusions ............................................... 39
7.8 Clinical management of CHD ........................................ 97
7.9 Previous research ...................................................... 98
7.10 Exception reporting .................................................. 101
7.11 Other factors describing quality of care ......................... 101
7.12 Conclusions ............................................................ 102

8 Indicators of quality of primary care - access and continuity of care.103
8.1 Introduction ............................................................ 103
8.2 Measure of access ..................................................... 103
8.3 The GP patient survey (GPPS) ..................................... 104
8.4 Response rates and satisfaction ................................... 110
8.5 Selection of key access and continuity of care responses .... 121
8.6 Conclusions ............................................................. 122

9 Variable selection ......................................................... 124
9.1 Introduction ............................................................. 124
9.2 Variable selection in this data set ................................. 124
9.3 Approach to be used in this model ............................... 127
9.4 Conclusions ............................................................. 128

IV Initial Model - Approach and Results ............................ 129

10 Modelling CHD mortality using count models .................. 130
10.1 Introduction ........................................................... 130
10.2 Summary of explanatory variables .............................. 130
10.3 Poisson model .......................................................... 132
10.4 Over-dispersed Poisson Model .................................... 134
10.5 Negative Binomial Model ............................................ 135
10.6 Conclusions ............................................................. 145

11 Results of initial model ............................................... 146
11.1 Associations between quality of primary care and premature CHD mor-
tality ................................................................. 146
11.2 Comparison between premature and all-age mortality ........ 149
11.3 Effect of changes in explanatory variables ..................... 151
11.4 Conclusions ............................................................. 154
V Developing and extending the initial model

12 Sensitivity Analyses
   12.1 Introduction .................................................. 156
   12.2 Age categories in count models ............................ 156
   12.3 Sex covariate in model ...................................... 160
   12.4 Missing data ................................................... 161
   12.5 Conclusions .................................................... 164

13 Multiple linear regression of standardised mortality rates
   13.1 Introduction .................................................. 165
   13.2 Unweighted regression ........................................ 165
   13.3 Weighted least squares regression ......................... 168
   13.4 Use of log of SMR ............................................ 174
   13.5 Conclusions .................................................... 175

14 Multilevel Modelling Approach
   14.1 Introduction .................................................. 177
   14.2 Linear regression ............................................. 178
   14.3 Poisson regression ............................................ 185
   14.4 Conclusions .................................................... 190

15 Estimating smoking prevalence in general practices: an evaluation
   of data from the Quality and Outcomes Framework (QOF). 191
   15.1 Introduction .................................................. 191
   15.2 Summary of QOF smoking indicators ....................... 192
   15.3 Estimating smoking prevalence from QOF data ............ 194
   15.4 Estimates of smoking prevalence in practice populations. 197
   15.5 Usefulness of estimates ...................................... 197
   15.6 Including smoking prevalence estimates in models of mortality 199
   15.7 Discussion .................................................... 200
   15.8 Conclusions .................................................... 201

16 Consideration of alternative explanatory variables 203
   16.1 Obesity ....................................................... 203
   16.2 Hypertension .................................................. 204
   16.3 Access and continuity of care variables .................... 207
   16.4 Conclusions .................................................... 208
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>Developing and extending the initial model - results</td>
<td>209</td>
</tr>
<tr>
<td>17.1</td>
<td>Introduction</td>
<td>209</td>
</tr>
<tr>
<td>17.2</td>
<td>Results</td>
<td>210</td>
</tr>
<tr>
<td>VI</td>
<td>Discussion, recommendations and conclusions</td>
<td>217</td>
</tr>
<tr>
<td>18</td>
<td>Discussion, recommendations and conclusions</td>
<td>218</td>
</tr>
<tr>
<td>18.1</td>
<td>Introduction</td>
<td>218</td>
</tr>
<tr>
<td>18.2</td>
<td>Principal findings</td>
<td>221</td>
</tr>
<tr>
<td>18.3</td>
<td>Approaches to modelling</td>
<td>221</td>
</tr>
<tr>
<td>18.4</td>
<td>Population characteristics</td>
<td>224</td>
</tr>
<tr>
<td>18.5</td>
<td>Primary care - prevention, detection and clinical management</td>
<td>229</td>
</tr>
<tr>
<td>18.6</td>
<td>Primary care - access and continuity of care</td>
<td>237</td>
</tr>
<tr>
<td>18.7</td>
<td>Comparisons with all-age mortality</td>
<td>239</td>
</tr>
<tr>
<td>18.8</td>
<td>Critical review - summary</td>
<td>242</td>
</tr>
<tr>
<td>18.9</td>
<td>Recommendations - summary</td>
<td>244</td>
</tr>
<tr>
<td>18.10</td>
<td>Conclusions</td>
<td>246</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Summary of known risk factors for CHD. ........................................ 10
3.1 Hilbe’s Table to compare unscaled AIC measures. .......................... 29
3.2 Using absolute differences in BIC to compare relative performance of nonnested models (Raftery, 1995). ................................................ 30
4.1 Percentage of population in key ethnic groups in each PCT, and East Midlands and England for comparison. Mid year estimates for 2009 (ONS). . 43
4.2 Summary of deprivation in various geographical regions in the area of interest (based on Index of Deprivation, 2007). ...................................... 44
4.3 Summary of health indicators included in CVD profiles compiled by SEPHO (2009/10); significance as given by SEPHO (2013). .......................... 46
4.4 Summary of indirectly standardised mortality rates and confidence intervals for CVD, CHD and stroke, 2006-08 (pooled) (NCHOD - The NHS Information Centre for health and social care, 2009). ................................. 47
5.1 Numbers of deaths by key categories of CVD, including percentage of those deaths which are under 75. .................................................. 51
5.2 Numbers of premature deaths caused by key categories of CVD, for total population, male and female. ................................................. 51
5.3 Numbers of deaths caused by key categories of CVD, for total population and male. .......................................................... 52
5.4 Patterns of CVD deaths in the area of study - percentage of all and premature CVD deaths which have underlying cause CHD and stroke. ........ 53
6.1 Super Output Areas ................................................................. 64
6.2 Summary of key sources of data. .................................................. 65
6.3 Summary statistics for diabetes prevalence (%) measured by QOF (practices for which there is data for all three years n = 228). .......................... 69
6.4 Weightings of each domain in the final index .................................. 79
6.5 Summary statistics for practice IMD scores based on the ‘gold standard’ or the practice postcode method (practices for which there is data for all four measures \( n = 222 \)).

6.6 Indicators included in the health domain of IMD-2007 (Noble et al., 2008).

7.1 Percentage of practices in LNR achieving maximum points in six QOF clinical domains.

7.2 Underlying achievement of recording and target achievement indicators, 2006/07 (\( n=229 \)).

7.3 Summary of indicators within the blood pressure area indicators; 2006/07 (\( n=229 \)).

7.4 Median underlying achievement (%) and interquartile range for blood pressure indicators for four disease areas (\( n=229 \)).

7.5 Median underlying achievement (%) and interquartile range for cholesterol indicators for three disease areas (\( n=229 \)).

7.6 Summary of smoking indicators; 2006/07 (\( n=229 \)).

7.7 CHD prevalence, detection and undiagnosed prevalence, practices with all measures (\( n = 221 \)).

7.8 Underlying achievement for potential indicators (\( n = 229 \)).

8.1 Five key themes in the GPPS.

8.2 2007/08 GPPS survey sections.

8.3 Questions associated with five key themes in GPPS - 2008/09.

8.4 Links between the GPPS and QOF.

8.5 GPPS annual response rates, based on technical reports.

8.6 National response rates for different groups.

8.7 Correlations between practice ethnic groups and response rates in LNR (*\( p < 0.001 \)).

8.8 National and local positive response in five key areas.

8.9 Median satisfaction and positive experiences in practices in the area of study based on their urban classification (2006/07).

9.1 Summary of population characteristics which have been considered and availability of data.

9.2 Summary of variables included in two models considered in this thesis.

10.1 Summary statistics for explanatory variables used.

10.2 Pairwise correlations \( R_S \) between key explanatory variables (\( p \leq 0.0001 \) in all cases).
10.3 IRR and 95% confidence intervals for premature CHD mortality - after fitting a Poisson model. 132
10.4 IRR and 95% confidence intervals for premature mortality - after fitting a Negative Binomial Model. 135
10.5 Impact of exclusion outliers on over-dispersion. 142
10.6 Impact of deleting two ‘unusual’ outlying practices on overdispersion. 143
10.7 Impact of deleting two ‘unusual’ outlying practices on incident rate ratios. 144

11.1 IRRs, 95% confidence intervals and \( p \)-values for premature mortality - inclusion of CHD achievement score (after fitting a Negative Binomial Model). 148
11.2 IRRs, 95% confidence intervals and \( p \)-values for premature mortality and all-age mortality (after fitting a Negative Binomial Model). 150
11.3 Effects on premature CHD mortality count of a one unit increase in the value of the explanatory variables and the impact of an improvement in primary care or decrease in population burden. Percentage change in premature count (95% confidence interval) \( (n = 228) \) (after fitting a Negative Binomial Model). 152

12.1 IRR and 95% confidence intervals for age covariates based on total population and population under 75 (model includes 11 other explanatory variables). 157
12.2 Impact of decrease in percentage aged 65 and over on premature mortality account, based on total population and population under 75 (model includes 11 other explanatory variables). 158
12.3 IRR and 95% confidence intervals for different age groups in a model which includes 11 other explanatory variables. 158
12.4 IRR and 95% confidence intervals for diabetes prevalence and hypertension detection when different upper age boundaries are used, (in a model which includes other specified explanatory variables). 159
12.5 Correlations between different age groups in practice populations \( (p < 0.0001 \) in all cases except \( ^{1} \) where \( p = 0.80) \). 159
12.6 Deviance of models with additional age groups in the model. 160
12.7 Concordance and 95% limits of agreement between actual count of deaths and predicted count of models with additional age groups in the model. 160
12.8 Impact of different imputation approaches for missing data for GPs per 1000 patients. 162
12.9 Imputed response to the GPPS question related to being able to see a preferred GP for practices with missing data from three alternative methods. 164
13.1 Estimated regression coefficients for unweighted linear regression showing the association between population and service characteristics and indirectly standardised premature (U75) CHD mortality rate ($R^2 : 0.38$). .......................... 166

13.2 Estimated regression coefficients for weighted linear regression showing the association between two CHD indicators and indirectly standardised premature (U75) CHD mortality rate. ................................................................. 170

13.3 Estimated regression coefficients for unweighted linear regression of $\log(SMR)$, $n = 227$ ($R^2 : 0.35$). ................................................................. 174

14.1 Key results from different two level models. ................................. 181

14.2 Key results from different two level models - inclusion of level 2 smoking variable. ................................................................. 182

14.3 Estimate of district level slope for diabetes. ..................................... 184

14.4 Estimate of ICC for three practices. .................................................. 187

14.5 Estimate of ICC for three practices for three Poisson models. ............ 188

15.1 Summary of smoking indicators for which underlying achievement is published; QOF guidance is available from www.nhsemployers.org (2014). .................................................. 193

15.2 Example of QOF data from 2012/13, showing how it can be used to calculate smoking prevalence for individual practices. All data available from HSCIC (HSCIC, 2014e) ................................................................. 196

15.3 Smoking prevalence (%) in those with chronic conditions ($n = 215$). ........ 197

15.4 IRRs, 95% confidence intervals and $p$–values for premature mortality - inclusion of smoking prevalence estimate (after fitting a Negative Binomial Model model). .................................................. 202

16.1 IRRs, 95% confidence intervals and $p$–values for premature mortality - inclusion of hypertension detection and undiagnosed hypertension ($n = 228$) (after fitting a Negative Binomial Model model). .................................................. 206

16.2 Comparison of access and continuity of care variables ($n = 225$). ........... 207

17.1 IRRs, 95% confidence intervals and $p$–values for premature mortality - inclusion of estimated smoking prevalence and undiagnosed hypertension and individual CHD indicators ($n = 228$) (after fitting a Negative Binomial Model). 212

17.2 IRRs, 95% confidence intervals and $p$–values for premature mortality - inclusion of estimated smoking prevalence and undiagnosed hypertension ($n = 228$) (after fitting a Negative Binomial Model). .................................................. 213
17.3 Effects on premature CHD mortality count of a one unit increase in the value of the explanatory variables and the impact of an improvement in primary care or decrease in population burden. Percentage change in premature count (95% confidence interval) \( n = 228 \) (after fitting a Negative Binomial Model).

17.4 IRRs or \( \beta \) coefficients (and 95% confidence intervals) for all models considered in this thesis. Premature CHD mortality is the dependent variable throughout. Numbers in **bold** indicate where confidence intervals do not include one (IRR) or zero (\( \beta \) coefficients). (*NBM: Negative Binomial Model*)

18.1 Summary of population characteristics which have been considered and availability of data.

18.2 Recommendations for policy makers, practitioners and further research.
List of Figures

1.1 Model to describe role of primary care in improving health outcomes in populations - Levene et al (2012). .......................... 12
1.2 Model to describe role of primary care in improving health outcomes in populations - Honeyford. .......................... 13

5.1 All CVD deaths - 2007/08 .......................... 54
5.2 Premature CVD deaths - 2007/08. .......................... 54
5.3 All CHD deaths - 2007/08 .......................... 54
5.4 Premature CHD deaths - 2007/08. .......................... 54
5.5 All stroke deaths - 2007/08 .......................... 54
5.6 Premature stroke deaths - 2007/08. .......................... 54
5.7 All CHD deaths - three years combined .......................... 55
5.8 Premature CHD deaths - three years combined. .......................... 55
5.9 All stroke deaths - three years combined .......................... 55
5.10 Premature stroke deaths - three years combined. .......................... 55
5.11 Funnel plot of crude rate of CHD deaths. .......................... 57
5.12 Funnel plot of crude rate of premature CHD deaths. .......................... 57
5.13 Scatter plot of premature and all-age crude death rates. .......................... 58
5.14 Bar plot to show variation in all-age CHD SMR. .......................... 58
5.15 Bar plot to show variation in all-age CHD SMR - outlier removed. .......................... 59
5.16 Histogram of premature SMRs - outlier removed. .......................... 59
5.17 Scatter plot of premature and all-age SMRs - outlier removed. .......................... 60
5.18 Funnel plot of SMR of CHD deaths - outlier removed. .......................... 61
5.19 Funnel plot of SMR of premature CHD deaths - outlier removed. .......................... 61

6.1 Bland-Altman Plot showing agreement between QOF based estimates and estimates derived from the APHO model for 25 practices. △ Corby • South Northamptonshire ○ Wellingborough. .......................... 70
6.2 Boxplots showing distribution of obesity prevalence as measured by estimates (APHO, 2010) and QOF registers. .......................... 72
6.3 Association between hypertension prevalence indicated by QOF and synthetic estimates from APHO model, (practices in LNR n = 229). 74
6.4 Bland-Altman Plot showing agreement between hypertension prevalence indicated by QOF and synthetic estimates from APHO model, (practices in LNR n = 229). 75
6.5 Correlation between ‘gold standard’ IMD scores for practices with postcodes from two different sources matching and those which don’t (practices in LNR for which both measures are present n = 222). 81
6.6 Correlation plots for domains of the Index of Multiple Deprivation, (practices in LNR n = 230). 83
6.7 Relationship between IMD score with and without health domain for LSOAs in LNR (n = 1013). 84
7.1 Distribution of points achieved in six clinical domains for practices in LNR, 2006/07 (n = 229). 90
7.2 Distribution of mean underlying achievement achieved in six clinical domains for practices in LNR, 2006/07 (n = 229). 92
8.1 Association between response rate and levels of socio-economic deprivation, as measured by the Index of Multiple Deprivation (2007). 112
8.2 Association between response rate and percentage white in the practice population. 113
8.3 Association between response rate and classification of practices into urban, town and village. 114
8.4 Association between response rate and practice population size. 114
8.5 Association between response rate and the percentage of the practice population over 65. 115
8.6 Satisfaction and positive experience in the five key areas across practices in the area of study. 116
8.7 Association between satisfaction and deprivation for the five key questions in LNR - 2006/07. 118
8.8 Association between satisfaction and positive experiences in the five key questions - 2006/07. 122
9.1 QOF clinical indicators: process of variable selection. 1 and 2 discussed in previous chapters. 126
10.1 Plot of variance against mean - practices divided into 10 groups (Poisson model). 133
10.2 Plot of variance against mean - practices divided into 10 groups (quasi-Poisson model). ................................. 134
10.3 Plot of variance against mean - observations divided into 10 groups. .... 136
10.4 Standardised deviance and Anscombe residuals plotted against transformed fitted values. ........................................... 137
10.5 Absolute residuals plotted against transformed fitted values. ............. 137
10.6 Standardised deviance residuals plotted against explanatory values. A: Population characteristics. ........................................... 138
10.7 Standardised deviance residuals plotted against explanatory values. B: Service characteristics. ........................................... 138
10.8 Added variable plot. A: Population characteristics. 1: Percentage white 2: Deprivation (IMD) 3: Diabetes prevalence (%) 4: Percentage 65 and over 5: Percentage male. ........................................... 139
10.9 Added variable plot. B: Service characteristics. 6: Hypertension detection 7: GPs per 1000 patients 8: Preferred GP - patients recall 9: SM02 10. CHD08 11. CHD09 ......................................................................................... 139
10.10 Residuals against transformed fitted values - identification of outliers. ... 140
10.11 Standardised hat values - index plot. ........................................ 141
10.12 Cook’s distances - index plot. .................................................. 141
11.1 Effects on premature CHD mortality count of a decrease in population burden. Percentage change in premature count (95% confidence interval). ........ 153
11.2 Effects on premature CHD mortality count of an improvement in primary care. Percentage change in premature count (95% confidence interval). .... 153
12.1 Age standardised death rates varying with age groups (ONS). ............. 157
13.1 Studentised residuals for unweighted regression. ................................ 167
13.2 Expected number of deaths in individual practices. ............................ 168
13.3 Absolute residuals of unweighted and weighted regression. ................. 172
13.4 Absolute residuals of unweighted and weighted regression, y-axes limited to clarify patterns. ........................................... 172
13.5 Absolute residuals for unweighted linear regression of log(SMR), n = 227. 175
13.6 Observed SMR compared to fitted values; comparing weighted least squares and OLS of log(SMR). ................................. 176
15.1 Relationship between aggregated QOF estimates and IHS estimates for local authority districts (QOF estimates based on 2012/13 data; IHS estimates based on 2011/12 survey). ........................................... 198
17.1 Effects on premature CHD mortality count of a decrease in population burden. Percentage change in premature count (95% confidence interval). . . . . . . 215
17.2 Effects on premature CHD mortality count of an improvement in primary care. Percentage change in premature count (95% confidence interval). . . . 215

18.1 Model to describe role of primary care in improving health outcomes in populations. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
List of abbreviations
A&E  Accident and Emergency
AF   Atrial Fibrillation
AIC  Akaike information criterion
AMI  Acute Myocardial Infarction
APHO Association of Public Health Observatories
BIC  Bayesian information criterion
CCGs Clinical Commissioning Groups
CHD  Coronary Heart Disease
CKD  Chronic Kidney Disease
COPD Chronic Obstructive Pulmonary Disorder
CVD  Cardiovascular disease
DPM  Diabetes Prevalence Model
GLM  Generalised Linear Models
GP   General Practitioner
GPPS GP Patient Survey
HES  Hospital Episode Statistics
HSCIC Health and Social Care Information Centre
HSE  Health Survey for England
ICC  Intraclass Correlation Coefficient
IHS  Integrated Household Survey
IMD  Index of Multiple Deprivation
IQR  Inter Quartile Range
IRLS Iteratively Reweighted Least Squares
IRR  Incident Rate Ratio
LAD  Local Authority District
L&R  Leicestershire and Rutland
LNR  Leicestershire (including Leicester City), Northamptonshire and Rutland
LRT  Likelihood Ratio Statistic
LSOA Lower Super Output Area
MSOA Medium Super Output Area
NCHOD National Clinical Health Outcomes Data
NHS  National Health Service
OLS  Ordinary Least Squares
PAD  Peripheral Arterial Disease
PCMD Primary Care Mortality Database
PCT  Primary Care Trust
QALYs Quality Adjusted Life Years
QOF  Quality and Outcomes Framework
SEPHO South East Public Health Authority
SHA  Strategic Health Authority
SMR  Standardised Mortality Ratio
UA   Unitary Authority
VPC  Variance Partition Coefficient
WHO  World Health Organisation
YHPHO Yorkshire and Humber Public Health Observatory
Part I

Context
Chapter 1

Introduction

1.1 Introduction

Developing a sound understanding of the impact of health care on health outcomes is necessary in formulating health system policies, particularly in a country with universal access to health care funded by the government. It is important to identify which aspects of primary care can play a role in reducing inequalities in health and improving health for all. This thesis focusses on the associations between aspects of primary care and improved health outcomes; premature coronary heart disease (CHD) mortality is the primary health outcome considered. Results of analysis of premature CHD mortality are compared with all-age CHD mortality. The methods would be applicable to other chronic diseases. In order to quantify these associations, there is a need to identify and evaluate the approaches to modelling which are available and this forms a key part of the thesis. In addition, the availability of data to describe the patient lists associated with general practice are explored in detail. A thorough understanding of modelling approaches and available data allow the substantive results to be discussed in context and areas for further research identified.

Initially, in Part I, the context of the research is described. In this chapter the concept of primary care and its potential to impact on health outcomes is outlined and risk factors associated with CHD are described, setting the context for the remainder of the thesis. This is not an in-depth review of primary care, but an introductory overview concentrating on aspects relevant to subsequent sections of the thesis. Chapter 2 summarises current research with a focus on the role of primary care in improving health outcomes related to cardiovascular diseases (CVD). Possible approaches to modelling
mortality data are summarised in Chapter 3. Part II describes the geography and populations of the area of study (Chapter 4) and the patterns of mortality in the area of study (Chapter 5).

Part III provides a detailed description and evaluation of data which are available and useful in exploring the relationship between premature CHD mortality and primary care. Chapter 6 considers the range of datasets available to describe key aspects of practice populations in terms of the known risk factors for CHD (Table 1.1 in Chapter 1). Chapters 7 and 8 describe the data which are available to assess the different aspects of primary care which may be important in modifying the predictive nature of practice populations. Approaches to variable selection in the context of the available data are described in Chapter 9.

An initial approach to modelling both premature and all-age mortality based on measures of quality of primary care suggested by previous research (Kiran et al., 2010; Levene et al., 2012) are described in Part IV. Part V develops the model in various ways, including a comparison with linear regression (Chapter 13), a consideration of a novel method of estimating smoking prevalence (Chapter 15) and consideration of measures of risk factor prevalence and quality of primary care different to those included in the initial model (Chapter 16). Part V concludes with the results of the model identified as the most useful (Chapter 17).

Part VI discusses the principle findings, strengths and limitations and implications. Recommendations for further research, practitioners and policy makers are identified.

1.2 Primary Care

Why primary care?

Health care is often broken down into various levels of care: primary - generally considered the first point of contact for all patients; secondary - specialist care generally provided within hospitals; and tertiary - more specialized care.

The importance of primary care was enshrined in 1978 by the World Health Organisation in the Alma-Ata Declaration (1978), in which primary care was considered key to achieving the target that all peoples of the world would attain a level of health permitting them to lead a socially and economically productive life in the year 2000. Since then extensive work by Starfield has shown that primary care services reduce
morbidity and mortality, again emphasising its importance (Starfield et al., 2005a). Primary care is a key area of research into health care because of the level of service it provides to the general population. In England primary care only accounts for 7.5% of the total expenditure, but over 90% of patient contacts within the NHS take place within primary care (NHS England (London Region), 2013).

The worldwide acknowledgement that primary care is vital to the improvement of the health of populations and its importance given the scope of service it provides means that primary care is an important area of research. This is particularly true in times when health services are under increasing pressure and clinicians, policy makers and patients are discussing the future role of the National Health Service in England (see (NHS England (London Region), 2013), for example).

What is primary care?

Primary care has been variously defined, with the definition provided by the Institute of Medicine (1978) being commonly cited: *Primary care is the provision of integrated, accessible healthcare services by clinicians who are accountable for addressing a large majority of personal health care needs, developing sustained partnerships with patients, and practicing in the context of family and community.*

This definition provides an international understanding of what is meant by primary care, but it is also useful to consider the features of good primary care and the services it provides: Starfield (2005a) suggests the four main features of primary care are:

- First contact for each new need;
- Long-term person (not disease) focused care;
- Comprehensive care for most health needs;
- Co-ordinated care (when multiple agencies are involved).

The Kings Fund suggest the following characteristics of high-quality primary care (Smith et al., 2013):

- Comprehensive;
- Person-centred;
- Population oriented;
In England, primary care is generally considered to be made up of general practices, dental practices, community pharmacies and high street optometrists, all of which have a role to play in improving health outcomes. This thesis focusses particularly on the associations between general practices and CHD mortality and consequently on the role of primary care in reducing CHD mortality. Clearly there are many other parts of health care in England which have a key role to play. For example optometrists can play a key role in diagnosing diabetes, an important risk factor for CHD. General practices in England can take various forms, from practices serving less than 1000 patients with one sole practitioner to very large practices, serving over 20,000 patients with 20 general practitioners and 15 practice nurses (Latham House Medical Practice, 2014). Larger practices often offer a wider range of services including family planning, and chronic condition clinics covering healthy heart, asthma or hypertension. During the period of study covered in this thesis, April 2006 to March 2009, general practices were members of Primary Care Trusts (PCTs), of which there were 152 in England, responsible for the provision of hospital, community health and some aspects of public health, assessing the local health needs and services which needed to be provided. Whilst the NHS undergoes regular changes, between April 2006 and March 2009 there were no major organisational changes. In April 2013 PCTs were abolished as Clinical Commissioning Groups (CCGs) were created.

One thing all general practices in England have in common is that each practice holds a registered list of patients, allowing practices and their over-arching organisational bodies to take an active population based approach to planning and delivery of care for registered patients (Thorlby, 2013). The recognized need for primary care to play a key role in improving the health of populations makes population based studies particularly appropriate.

### 1.3 Health outcomes

When considering the role of primary health care in improving health outcomes we need to consider what is meant by health outcomes and why some may be more suitable for study than others. Various outcomes have been suggested:
- Mortality;
- Morbidity;
- Disability;
- Health status;
- Quality of life;
- Health related quality of life;
- Quality Adjusted Life Years (QALYs);

The global burden of disease study, commissioned by the World Health Organisation (WHO), considers years of life lost; years lived with disability; disability adjusted life years; and healthy life expectancy (WHO, 2014a). Some reports also include mortality rates (Murray et al., 2013).

In addition accident and emergency (A&E) attendances and emergency admissions are commonly studied as a health outcome, as they create a considerable financial burden on health care services (Tian et al., 2012). Ideally for ambulatory care-sensitive conditions both visits and admissions would be prevented by provision of other care (Purdy et al., 2011).

Associations between aspects of primary care and changes in health behaviour and risk factor prevalence may also be studied, as well as the rate of certain processes in secondary care, for example attendance at cervical screening programs (Bang et al., 2012), the uptake of health checks (Artac et al., 2013) and rates of angiography (Jones et al., 2004).

Researchers may have a clear rationale for considering particular health outcomes, including current priorities, for example the current interest in reducing A&E visits (NHS England, 2013). The ease of access to data and the completeness and reliability of that data will also play a key role in the selection of health outcomes. In England Hospital Episode Statistics processes data on all admissions, outpatient appointments and A&E attendances at NHS hospitals in England (HSCIC, 2014a) and provides a useful source of data for research into associations between primary care and emergency admissions. The Primary Care Mortality Database includes mortality data as provided at the time of registration of death and links mortality data to an individuals registered practice, allowing research into associations between primary care and mortality rates (HSCIC, 2014c).
These datasets allow relatively straightforward calculations of mortality and emergency admissions rates, but do not allow determination of outcomes such as QALYS as they do not allow easy linkage to information on disability and quality of life. In addition, the ease of analysis may be a factor in deciding on the health outcome, modelling counts of death and emergency admissions is relatively straightforward, even if the choice of modelling is complex.

### 1.4 Multiple determinants of health

There is a general acceptance that there are multiple determinants of health. These are presented and categorised in a variety of ways. McGinnis (2002) describes five domains which influence our health prospects: genetic and gestational endowments; social circumstances; environmental conditions; behavioural choices; and medical care. Starfield (2005b), having accepted the importance of the underlying genotype of individuals, divides determinants into cultural and behavioural characteristics of populations; the social and physical environment; and medical practice. In addition, there is considerable research into psychosocial factors and their potential impact on physical health. These factors include stress, job control and depression (Macleod and Davey Smith, 2003). These broad determinants contribute to the existing risk caused by biological factors. Within these broad headings, for any given disease there will be multiple risk factors which increase the probability of an individual becoming ill or dying.

#### 1.4.1 Cardiovascular disease (CVD)

CVD includes all diseases of the heart and circulatory system; including CHD, peripheral vascular disease and stroke. The date of disease initiation for diseases of this type is unlikely to be known, but progression of the disease is relatively well understood. Fatty streaks are deposited in the endothelium of blood vessels (20-30 years old), these may progress into fibrous plaques (30-40 years old), occlusive atherosclerotic plaques, plaque ruptures and thrombosis and lead progressively on to myocardial infarction, coronary death, stroke or peripheral arterial disease (adapted from (Greenland et al., 2002)). Risk factors for CHD are relatively well known and summarised in Table 1.1. It is not completely understood which risk factors are important in the progression between stages and it is thought that some risk factors will be associated with more than one stage. Many of the risk factors described in this study are associated with all types
of cardiovascular diseases, but they do not influence them equally (Newschaffer et al., 2010). For example, CHD is rarely found in populations without elevated cholesterol; in contrast stroke is most strongly associated with high blood pressure.

<table>
<thead>
<tr>
<th>Unmodifiable Biological risk factors</th>
<th>Biological risk factors</th>
<th>Key health behaviours</th>
<th>Psychosocial wellbeing</th>
<th>Environmental conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increasing age</td>
<td>High blood pressure</td>
<td>Poor nutrition</td>
<td>Depression</td>
<td>Socioeconomic deprivation</td>
</tr>
<tr>
<td>Gender</td>
<td>High serum cholesterol</td>
<td>Physical inactivity</td>
<td>Anxiety</td>
<td>including education, income</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Diabetes</td>
<td>Smoking</td>
<td>Hostility</td>
<td>and the built environment</td>
</tr>
<tr>
<td>Family history</td>
<td>Obesity</td>
<td>Alcohol</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genetic make up</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.1: Summary of known risk factors for CHD.

There are complex associations between the risk factors above and the progression of CHD. For example, within the category poor nutrition, links have been identified between diets high in fruit, vegetables and soluble fibre leading to lower blood cholesterol (Threapleton et al., 2013), diets high in salt linked to high blood pressure (Frisoli et al., 2012), diets high in saturated fat leading to high blood cholesterol and generally poor nutrition leading to obesity which is associated with increased risk of diabetes (Whitney and Rolfs, 2012). The identification of independent risk factors is also complex. For example, people living in poverty are known to have higher levels of depression, poorer nutrition and higher rates of smoking. However, research suggests that even when associated risk factors are taken into account there is an association between socioeconomic factors and increased risk of CHD (Kivimäki et al., 2007). The main health behaviour risk factors of poor nutrition, physical inactivity and smoking have been identified as explaining 75% of new cases of CHD, and it is suggested that these should be the main focus of public health care (Magnus and Beaglehole, 2001).

For primary care to have an impact on risk factors, the factors need to be modifiable. Some factors such as age and genetic make up are not modifiable; some factors are modifiable and are within the scope of primary care, for example high blood pressure and smoking behaviour; and some factors are potentially modifiable but lie outside the scope of primary care, for example housing conditions.

1.5 Primary care and health outcomes

Starfield (2005b) has proposed that there are six mechanisms, acting alone and in combination, which may account for the beneficial impact of primary care on population
health:

1. Greater access to needed services;
2. Better quality of care;
3. A greater focus on prevention;
4. Early management of health problems;
5. Cumulative effect of the main primary care delivery characteristics;
6. Role of primary care in reducing unnecessary and potentially harmful specialist care.

Levene et al (2012), drawing on Starfield’s mechanisms and a previous finding that lower detection of hypertension was associated with higher CHD mortality, devised a conceptual model to...

...explain how the effect of population characteristics upon variations in mortality of long-term conditions might be modified by variations in the delivery of primary care that incorporates whole population coverage (greater access to services) and offers sustained relationships with patients (the cumulative effect of primary care delivery characteristics, including continuity and comprehensiveness).

Levene et al (2012) proposed that primary care can play a key role in altering the predictive effect of population characteristics on the proportion of any given population transitioning from healthy or morbid to deceased. These effects may either act directly on the progression to death by appropriate clinical management or by affecting a modifiable risk factor. Appropriate interventions, targeting both healthy and morbid populations, include early detection, prevention and clinical management. This can be seen pictorially in Figure 1.1.

However, this representation of the modifying effect of primary care on the transition from healthy or morbid to deceased omits various important feedback loops and may limit our interpretation of the process; Figure 1.2 starts to explore some of these. For
example, non-modifiable risk factors for CHD may also affect access to primary care and development of sustained relationships, as well as the effectiveness of prevention, early detection and clinical management, including referral to secondary care. Having established that the relationships between CHD mortality and population characteristics, primary care and other potentially modifying factors are extremely complex, it is acknowledged that it is difficult to develop a statistical model which allows all these relationships to be considered, but when the model is interpreted the complexity of the relationships should be considered.
Figure 1.2: Model to describe role of primary care in improving health outcomes in populations - Honeyford.
1.6 Conclusions

Primary care has a key role to play in ensuring that people have a level of health which allows them to lead a socially and economically productive life. The scope of service it provides and the way in which it delivers this service are both important. This chapter has briefly described the role of primary care in improving health outcomes; the variety of health outcomes which can be studied and the multiple determinants of health, with cardiovascular disease as an example. The chapter concludes by aiming to conceptualise the many aspects of population characteristics and primary care that can impact on population health.

The following chapter reviews published research which has investigated the associations between population characteristics, primary care and cardiovascular disease health outcomes.
Chapter 2

Primary care factors associated with improved health outcomes

2.1 Introduction

Primary care has an important role to play in improving a variety of health outcomes, as described in Chapter 1. There are many studies which have analysed individual patient records and focussed on the treatment individual patients have had and the subsequent impact on reducing mortality. This is not the aim of this study; the focus here is on the aspects of primary care associated with population level health outcomes. Given the wealth of knowledge concerning individual level risk factors and the effectiveness of treatment, there are at least two valid approaches to understanding variations in mortality between practices. Detailed analysis of what is happening to individuals is useful; but understanding the way in which practices are functioning in terms of overall quality of care is also important. The role of primary care organisations in reducing and managing overall risk in their patients should not be underestimated.

This thesis specifically analyses associations between primary care and premature CHD mortality. That is, mortality in those under the age of 75, as adopted by Public Health England in their ‘Longer Lives’ project (Public Health England, 2014a). This specific health outcome has not been studied previously in relation to English primary health care. The review of current literature therefore focusses on similar studies. Health outcomes which are reviewed in this chapter are all-age mortality, A&E visits and emergency admissions, with a focus on CVD related health outcomes.
The introduction of the Quality and Outcomes Framework (QOF) in 2005 (described in more detail in Chapter 7) means that there is comprehensive information on the levels of clinical management for groups of patients diagnosed with a range of chronic conditions, including CHD. These readily available data have resulted in a multitude of studies at practice level, with each study using a slightly different approach to selection of relevant indicators or scores. This chapter reviews studies analysing the role of primary care factors in reducing unplanned hospital admissions and mortality.

2.2 Prevention, detection and disease management

Many studies have considered the associations between primary care and emergency admissions and mortality; two studies focus particularly on CHD mortality.

Levene et al (2010; 2012) have published two papers which consider the role of primary care in reducing CHD mortality at PCT level, including all PCTs in England, using slightly different methodology. In 2010, using multiple linear regression and a stepwise approach to variable selection, after the inclusion of ethnicity, socioeconomic deprivation, diabetes and smoking prevalence, the only aspect of primary care found to be associated with CHD mortality was hypertension detection (Levene et al., 2010). A similar result was found using negative binomial regression (Levene et al., 2012); control of cholesterol in CHD patients and treatment with aspirin were not found to be associated with lower levels of CHD mortality.

Work by Kiran et al (2010) on CHD mortality in 1531 general practices in London found that a devised ‘CHD achievement score’, based on 12 QOF indicators, was associated with lower levels of CHD mortality. The CHD achievement score included control of cholesterol and blood pressure in high risk patients, as well as treatment with aspirin and β-blockers.

Studies of practice characteristics and emergency admissions have shown that various aspects of primary care are associated with decreased rates of admissions. Increased control of blood pressure and cholesterol in stroke patients and control of blood pressure in patients with hypertension were all associated with reduced stroke admissions (Soljak et al., 2011). However, when practice characteristics including deprivation and smoking prevalence were included in the model, only control of cholesterol was significantly associated with emergency stroke admissions. Overall clinical performance, as measured through an overall QOF clinical score, was found to be associated with
admissions for angina, with admissions being 16% lower for each quartile increase in score; however specific CHD management factors were not significantly associated when adjusted for deprivation and CHD prevalence (Purdy et al., 2011). Higher QOF performance protected against unplanned admissions for cancer (Bottle et al., 2012) and chronic obstructive pulmonary disorder (COPD) admissions (Calderon-Larranaga et al., 2011). In contrast, a PCT-level study found no significant association between the quality of CHD care, as measured by a mean of QOF CHD indicators and control of blood pressure and cholesterol separately, and rates of either elective or emergency admissions for CHD (Bottle et al., 2008). Similarly practice level studies of all emergency admissions found no significant association with total clinical QOF points (Bankart et al., 2011).

2.3 Premature mortality

Few studies have considered the associations between quality of primary health care, as measured by QOF, and premature mortality. In a study of small area mortality rates, not linked to health care, Allender et al (2012) demonstrated a stronger association between socio-economic deprivation and premature CHD mortality (under 75), compared to all-age CHD mortality. In a study of CHD hospitalisation and primary care, patients were divided into 45-74 and 75 and over; socio-economic status was found to be more important in the younger age group, but no difference was found in the association of quality of care and health outcomes in the two different age groups (Bottle et al., 2008).

2.4 Access and continuity of care

Access may appear a simple idea – ‘is it possible for someone to see a medical professional?’ – but in reality it is a complex concept. Gulliford et al (2002) distinguish between ‘having access’ and ‘gaining access’. Having access can be considered the potential to use a service if required and is related to the supply of health services and is traditionally measured in primary care using indicators such as doctors or other health professionals per capita. Starfield et al (2005b) summarise a range of studies from the 1990s which have shown that U.S. states with higher numbers of primary care physicians per 100,000 population had better health outcomes. More recently physicians per capita was a significant predictor of cardiovascular mortality in the central Appalachian
states (Esch and Hendryx, 2011) and improved health outcomes in chronic conditions have been shown to be associated with physicians per capita in Korea, which has a system of universal health care coverage (Lee et al., 2010).

Gulliford suggests that ‘proof of access is use of service, not simply the presence of a facility’. Barriers to use of service include ‘personal barriers’ related to a patient’s perception of their needs, their attitudes and experiences. ‘Financial barriers’ which, in a country of universal eligibility, free at the point of delivery, include costs associated with travel, time and aspects of health care which do have cost, such as prescriptions. ‘Organisational barriers’ also affect use of primary care, such as apportionment systems, phone access and waiting lists for secondary care which may affect a patient’s decision to access primary care.

Having gained access to primary care, ‘continuity of care’ is another key dimension of primary care which has been shown to be associated with some improved health outcomes. Whilst the definitions of continuity of care vary, the majority involve an emphasis on an ongoing relationship between a patient and medical professionals; the emphasis in this research concerns the relationship between general practitioners and patients; the role of other medical professionals and auxiliary staff should not be underestimated. Continuity of care has been divided into relationship continuity; information; and management continuity (Haggerty et al., 2003; Freeman and Hughes, 2010). The varied definitions available have led Gulliford (2006) to identify two core aspirations of continuity of care as a ‘continuous caring relationship’ and as a ‘seamless service’.

Whilst the reasons for the positive impact of access are relatively clear, (if a patient cannot gain access to a medical professional they are unlikely to receive medical help), the reasons for the impact of continuity of care is not clear. A possible mechanism is that relationship continuity may result in increased trust between the patient and doctor, which may result in patients being more likely to accept medical advice and to comply with medication (Ettlinger and Freeman, 1981; Chen et al., 2013), and improved diagnosis both in terms of ‘wait and see’ and detailed knowledge of the ‘norm’ for that patient (Freeman and Hughes, 2010; Round et al., 2013).

Studies of mortality rates and GP supply in England do not support studies in U.S. states linking physician supply with improved health outcomes. Apparent associations between mortality and GP supply were largely explained by confounding variables such as ethnicity and social class (Gulliford, 2002) and no association between CHD mortality and GPs per capita was found in an analysis of PCT level mortality (Levene
et al., 2010).

Studies of emergency admission rates in England suggest the relationships between GPs per capita and admission rates is not clear cut. Soljak et al (2011) found a positive relationship between stroke admissions, although the incident rate ratio (IRR) was close to unity; a negative association was found in studies of heart failure admissions (Kiran et al., 2010; Purdy et al., 2011; Brettell et al., 2013) and no significant associations were found between GPs per capita and diabetes based admissions (Dusheiko et al., 2011).

Other research has sought to address the impact on health of ‘gaining access’, rather than simply ‘having access’. For example, Baker’s team in Leicester have used questions on the GP patient survey (GPPS), discussed in more detail in Chapter 8, to measure patients’ perception of access, but did not find associations with emergency admissions when other factors were included in the model (Bankart et al., 2011). Others have found that being able to book an appointment within 48 hours and/or in advance were both associated with lower admissions for heart failure (Brettell et al., 2013; Cowling et al., 2013). Being able to book an appointment in advance was found to be negatively associated with stroke admissions (Soljak et al., 2011).

Continuity of care has also been shown to be associated with improved health outcomes. Patient level analysis of continuity of care has shown that high levels of continuity is associated with reduced hospitalisation in patients aged 67 or over (Menec et al., 2006). The measure of continuity in this study was that at least 75% of visits were with the same family physician. The evidence for an association between continuity of care and reduced hospitalisation was stronger for ambulatory care-sensitive conditions. Practice and PCT level studies using measures such as patients’ perceptions of being able to book an appointment in advance and with a preferred GP have been shown to be associated with improved health outcomes. This has included lower COPD and cancer mortality (Levene et al., 2012), reduced emergency admissions (Bankart et al., 2011; Gunther et al., 2013) and improved cholesterol in patients with diabetes (Suleman et al., 2011). However, Anwar et al (2012) found that being able to book an appointment in advance was associated with lower levels of recording of chronic disease.

Although only emergency admissions and mortality have been highlighted here, there is increasing evidence from recent studies that both ‘gaining access’ and continuity of care are related to a variety of improved health outcomes. This is supported by a European wide study which strongly suggests that better physician-patient relationships rather than the organisation of CHD care are related to higher levels of health related quality.
of life (Ose et al., 2012).

2.5 Conclusions

This chapter has summarised recent research focussing on the associations between primary care and health outcomes for groups of patients, with a focus on studies based in England and CVD health outcomes.

Studies have shown that primary care is associated with improved health outcomes. The research is all of an observational nature and demonstrates a range of associations and does not confirm causation. The range of associations which have been found may be related to the range of health outcome studied, the availability of data, the approach to modelling and the selection of explanatory variables included in statistical models. The last three are discussed in the following chapters.
Chapter 3

Modelling Mortality

3.1 Introduction

Research summarised in Chapter 2 suggests that the evidence that there is a strong association between cardiovascular health outcomes and aspects of primary care is not conclusive. One possible reason for this is the statistical approaches adopted to model the data. Multiple linear regression of standardised mortality (or hospitalisation) rates as the dependent variable is relatively common and has been used in several similar studies (Bottle et al., 2008; Muller, 2002; Levene et al., 2010; Lee et al., 2010; Esch and Hendryx, 2011; Tunstall et al., 2012). In addition, the linear regression may be weighted (Kiran et al., 2010) or the dependent variable may be log(SMR) (Pocock et al., 1980). However, Rosenbaum and Rubin (1984) have shown that when SMRs are the dependent variable in multiple linear regression and the explanatory variables are not age adjusted, risk estimates can be biased. An alternative approach, which is arguably more natural and appropriate, is the use of Poisson models of count data. Age and sex can then be included in the model as explanatory variables, thus overcoming the lack of age standardised explanatory variables. As the counts are often overdispersed, that is the variance is greater than the mean, a negative binomial model may be more appropriate than a Poisson model (Hilbe, 2011). Models based on count data, for example the numbers of deaths or hospital admissions, can also be found in the literature (Alter et al., 2008; Freemantle et al., 2009; Bankart et al., 2011; Chauhan et al., 2012; Levene et al., 2012).

This chapter initially summarises approaches to standardisation and then describes the two main approaches which have been adopted for modelling health outcomes, in par-
ticular mortality: multiple linear regression of mortality which has been standardised in some way; and models of counts of deaths. Both Poisson and Negative Binomial Regression are described, the latter in more detail as this is a less familiar model. The chapter concludes with another important aspect of model building - variable selection.

3.2 Measuring mortality

Mortality as a health outcome has been an interest for health researchers for centuries. Different modelling approaches utilise different measures of mortality. Different approaches to standardisation are briefly described here; with the details of the approach used described in Chapter 5.

3.2.1 Standardisation

Age adjusted rates were developed in 1841 (Curtin and Klein, 1995). At this time the impact of study populations which had very different age structures were noted, with populations with a higher proportion of older people often having higher crude death rates than comparison populations even though the age specific death rate of composite age groups were lower throughout the population. Age standardisation remains important when comparing practices today (Webb and Esmail, 2002). Two approaches are discussed below.

Direct Standardisation

In direct standardisation the age specific death rates of a study population are compared to a standard population; a weighted-average of the age specific death rates. Details of the methodology can be found in standard textbooks, (for example see Donaldson and Scally, 2009, Chapter 1). Directly standardised mortality rates based on small numbers of deaths will exhibit a large amount of random variation; as a rough guideline at least 25 total deaths are needed for direct standardisation to be appropriate (Curtin and Klein, 1995).
Indirect standardisation

Standardised mortality ratios (SMRs) are produced by a process of indirect standardisation. The SMR is the ratio of the observed number of deaths in a study population to the expected number of deaths in the same population. A standard set of age-specific death rates is applied to the age-specific local area population to produce an expected number of deaths.

Indirect standardisation is useful because it can be used when age specific death rates of a population are not known, but is also considered more appropriate when the number of deaths is small (Goldman and Brender, 2000). An SMR of 100 indicates that the study population has the same mortality risk as that of the reference population; higher values suggest an increased risk of mortality in the study population.

SMRs are commonly used when studying differences between health institutions but there are concerns about their use which arise when the age distributions of populations differ considerably (Goldman and Brender, 2000; Julious et al., 2001; Evans et al., 2013). Examples show that even when the SMR is lower in all age groups for one population, compared to another, the aggregated SMR can be higher; an example of Simpson’s Paradox (Julious and Mullee, 1994). Goldman and Brender (2000) argue that some studies which criticise SMRs contrasted hypothetical populations with extremely different age distributions, but when ‘more realistic’ populations were compared there were less concerns with SMRs. However, Evans et al (2013) found significant differences in SMRs when comparing neonatal units and suggest that in certain cases this could lead to misleading conclusions.

3.3 Multiple linear regression

Multiple linear regression is a commonly used approach to describe the relationship between a dependent variable, $Y$, and a set of $p$ explanatory variables. The model has the general form:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_p x_{pi} + e_i$$

where $x_{1i}, x_{2i}, \ldots, x_{pi}$ are as set of explanatory variables, $\beta_1, \beta_2, \ldots, \beta_p$ are regression coefficients and $e_i$ is an error term. Ordinary least squares (OLS) is used to estimate
coefficients, this is based on minimising the sum of the squares of the residuals (the differences between the observed value of the dependent variable \( y \) and the predicted value \( \hat{y} \)). In multiple linear regression the effect of each explanatory variable on the dependent variable is estimated in the presence of the other explanatory variables in the model. Assumptions of the multiple linear regression model are similar to those of the simple linear regression model (Dobson and Barnett, 2008):

- It is assumed that errors are independent and normally distributed; \( e_i \sim N(0, \sigma^2) \).
- The variance is assumed to be constant, an assumption termed homoscedasticity.
- It is assumed that there is a linear association between the dependent variable and the explanatory variables.

In addition, in multiple regression, it is important that the explanatory variables are independent. Associations between explanatory variables is termed collinearity. Various approaches are available to check that assumptions have been met; (see, for example, Chatterjee and Simonoff, 2012).

### 3.3.1 Weighted least squares

Weighted least squares is a possible method to both resolve and to take into account heteroscedasticity. Assigning different weights to different observations because they are more variable, or simply to reduce the importance of remote data points, has a long history (Willett and Singer, 1989). This approach has been used when data are grouped in some way; for example geographical groupings (Pocock et al., 1981). One challenge of using weighted least squares is that the underlying theory assumes that the variance of each data point is known, which is not the case. Therefore, estimates of weights need to be used. One approach is to use the reciprocal of the variance of the observation, in this case the variance of the estimate of the SMR. If weighted least squares has been used model diagnostics should utilise weighted residuals (PennState, 2014).

### 3.3.2 Use of log of SMR

With multiple linear regression the final model is not constrained to predicting dependent values which are greater than zero; a model which predicts SMRs below zero.
is not useful. Multiple linear regression of log(SMR) overcomes this problem by con-
straining the SMR to be greater than zero. In addition, a log transformation can be
used to reduce the level of skewness in highly skewed distributions and to homogenise
variances. Similarities between multiple linear regression of log(SMR) and the Poisson
model are discussed in Section 3.5.

3.3.3 Limitations of approach

A key limitation of using an SMR as a dependent variable is that standardised explana-
tory variables are rarely available. Regression of age-adjusted rates on crude rates, for
example smoking prevalence, may result in biased parameter estimates (Rosenbaum
and Rubin, 1984). Rosenbaum and Rubin (1984) suggest possible approaches which
can lead to unbiased estimates:

1. Regression of the responses of individuals on the age of individuals and explana-
tory variables;
2. Weighted regression of the age-specific response rates (mortality) on age and on
the age-specific explanatory variable averages;
3. Weighted regression of the crude response rates (mortality) on age and the crude
explanatory variable averages;
4. Weighted regression of age-adjusted rates on the age-adjusted explanatory vari-
ables.

However, it is rare for candidate explanatory variables to be published in terms of age
groups, this is certainly the case for data coming from QOF.

An additional approach is to use weighted regression of age-adjusted rates on age and
crude explanatory variables, which can result in unbiased estimates in certain condi-
tions. When this approach is used results of modelling can be sensitive to different
methods of adjusting for age. For example, income inequality was found to be sig-
ificantly positively associated with age-adjusted mortality rates (Milyo and Mellor,
2003). When crude mortality rates were modelled with age groups included as pre-
dictor variables the association remained positive but was not significant and when
mean age was used in place of age groups the relationship was reversed. Associations
between mortality and certain socioeconomic factors may be extremely sensitive to
different age-adjustment methods.
3.4 Count Models

An alternative approach to using rates, either standardised or crude, as a measure of mortality, is to consider the underlying data, the counts of deaths, as the key outcome variable. If the underlying populations are not the same, in terms of sex, age and size, these need to be included in the model in some way. Standard linear regression is inappropriate for modelling count data; counts cannot be negative and count data often exhibit heteroscedasticity, a larger variance accompanying larger means (Chatterjee and Simonoff, 2012). The Poisson regression model, an example of a generalised linear model (GLM), is a commonly used example of a count model, alternative models, such as the negative binomial model, may be appropriate if certain assumptions of the Poisson model are not met.

3.5 Poisson models

The Poisson regression model is an example of a GLM of the form:

\[ E(Y_i) = \mu_i = n_i e^{X_i \beta}; \quad Y_i \sim \text{Poisson}(\mu_i). \]

The link function is the logarithmic function:

\[ \log \mu_i = X_i \beta \]

(Dobson and Barnett, 2008).

The Poisson distribution is dependent on its mean (\( \mu \)) as its variance is equal to the mean; a property known as equidispersion.

Maximum likelihood estimation is used to estimate the unknown parameters; an iteratively reweighted least squares (IRLS) algorithm is applied to find the solution.

**Interpretation**

Poisson regression models the log of the expected count; the \( \beta \) coefficients describe a change in the log of the expected count. It is common to express effects associated with each variable as an incident rate ratio (IRR). For every one unit increase in the
predictor, the predicted count changes by:

\[(\text{IRR} - 1) \times 100\%\].

Consequently, an IRR of 1.015 would mean that a one unit change in the explanatory variable would be associated with a 1.5% increase in the expected count; conversely with an IRR of 0.985 one unit change would be associated with a 1.5% decrease in the expected count.

**Exposure**

Poisson regression can be used to model ‘rate data’, when events occur over time or space; this is termed the ‘exposure’. In this case the ‘space’ is the number of patients on a practice list, this will affect the number of deaths per practice. It is, therefore, important to take this into account when modelling counts of deaths. The \(\log\) of the exposure is included in the model on the right hand side and the coefficient is required to have a value of one (Dobson and Barnett, 2008, p 152). The \(\log\) of the exposure is termed the offset.

The link function with an offset included is shown below:

\[
\log \mu_i = \log n_i + X_i \beta;
\]

where \(\log n_i\) is the offset.

**Assessing model fit and model checking**

Residuals are important in assessing both overall model fit and model checking. In linear models residuals are the difference between observed data and fitted data and if the model is appropriate the residuals follow a Normal distribution with mean zero. In Poisson models the raw residuals are both heteroscedastic and asymmetric (Cameron and Trivedi, 1998, p142). Pearson, deviance and Anscombe residuals are all recommended for measures of model fit and model diagnostics of Poisson models (Hilbe, 2011, p41).
Goodness of fit

Both the Pearson and deviance statistics are commonly used measures of goodness of fit in GLMs (Cameron and Trivedi, 1998, p152). The Pearson statistic is a weighted sum of the of the residuals and can be compared to the number of observations \( n \) minus the number of parameters \( k \). Cameron and Trivedi (Cameron and Trivedi, 1998, p152) suggest that when the Pearson statistic does not equal \( n - k \) there is evidence of model mis-specification. The deviance statistic is twice the difference between the log-likelihood achievable and the log-likelihood of the fitted model.

\( R^2 \) is commonly used in linear regression to measure the global fit of the model and is specifically used to summarise the percentage of variance explained. Either Efron’s pseudo-\( R^2 \) or McFadden’s likelihood-ratio index can be used to assess percentage variance explained in count models (Hardin and Hilbe, 2007, p59) in comparison to other models of the same type. These should not be compared to \( R^2 \) values from linear regression estimated through OLS. Pseudo-\( R^2 \) values are generally much lower than those arising from linear regression. Values of 0.2-0.4 are considered to be indicative of extremely good model fit, possibly approximate to 0.7-0.9 from linear regression (Louviere et al., 2000). Pseudo-\( R^2 \) values are most useful for comparing models based on the same data.

Model checking

McCullagh and Nelder (1989) divide model checking into two key areas: i) checks for systematic departures from the model and ii) checks for isolated departures from the model.

Three plots are recommended to identify potential systematic departures from the model (McCullagh and Nelder, 1989, p399):

Plots of residuals against some function of the fitted values. Fitted values can be transformed to the constant information scale of the error distribution.

Plots of residuals against an explanatory variable.

Added-variable plot. This plot gives an indication of whether an omitted covariate should be included in the linear predictor.

Patterns in the first two residual plots may arise as a result of various causes, including
the wrong choice of link function, wrong choice of scale of one or more covariates, or omission of a quadratic term in a covariate.

Checks for isolated departures from the model are also important. Individual observations which may need further investigation, either because they are outliers or are particularly influential, need to be identified. Plots of residuals against transformed fitted values, as above, are useful for identifying outliers. The diagonal elements of the ‘hat’ matrix are a useful measure of leverage in both linear regression and GLMs (McCullagh and Nelder, 1989, p405). A limit of \( h > 2p/n \) has been suggested has a measure of high leverage; this expression can be rearranged to produce standardised hat values - consequently hat values greater than 2 are considered high. An approximation for Cook’s distances is available to estimate the influence of individual data points; values greater than \( 4/n \) are associated with high influence (Hardin and Hilbe, 2007, p49). Observations with Cook’s distances greater than \( 4/n - p - 1 \) may be problematic.

**Comparing models**

Information criteria are available to allow different models which have been fitted to the same data to be compared. They are likelihood-based measures that include a penalty function (Cameron and Trivedi, 1998, p150). The penalty can be a function of the number of parameters in the model or the number of observations. The Akaike information criterion (AIC) and the Bayesian information criteria (BIC) are commonly used in linear regression. The theory underlying these criteria is not directly applicable to Poisson regression, but they have proved useful in practice (Chatterjee and Simonoff, 2012).

Generally models with the smaller value for either AIC or BIC are considered to represent a better fit. There is no specific statistical test from which a \( p \)-value may be computed to determine whether a decrease in AIC is representative of a better model. Hilbe (2011, p. 70) has devised a subjective table based on simulation studies that can be used to make such a decision for comparing unscaled AIC measures, shown in Table 3.1.

<table>
<thead>
<tr>
<th>Difference: Model A and B</th>
<th>Decision (assuming ( A &lt; B ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00 &lt;diff( \leq 2.50 )</td>
<td>No difference in models</td>
</tr>
<tr>
<td>2.50 &lt;diff( \leq 6.00 )</td>
<td>Prefer A if ( n &gt; 256 )</td>
</tr>
<tr>
<td>6.00 &lt;diff( \leq 9.00 )</td>
<td>Prefer A if ( n &gt; 64 )</td>
</tr>
<tr>
<td>10.00 &lt; difference</td>
<td>Prefer A</td>
</tr>
</tbody>
</table>

Table 3.1: Hilbe’s Table to compare unscaled AIC measures.
Absolute differences in BIC statistics can be used to compare non-nested models. If $BIC_A - BIC_B < 0$ then model A is preferred. If $BIC_A - BIC_B > 0$ then model B is preferred. The degree of preference is shown in Table 3.2.

<table>
<thead>
<tr>
<th>Difference</th>
<th>Degree of preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-2</td>
<td>Weak</td>
</tr>
<tr>
<td>2-6</td>
<td>Positive</td>
</tr>
<tr>
<td>6-10</td>
<td>Strong</td>
</tr>
<tr>
<td>&gt; 10</td>
<td>Very strong</td>
</tr>
</tbody>
</table>

Table 3.2: Using absolute differences in BIC to compare relative performance of nonnested models (Raftery, 1995).

### 3.5.1 Comparing the Poisson to multiple linear regression

If the log SMR is modelled using multiple linear regression this can be written as:

$$\log \frac{O}{E} = x_i \beta + \epsilon_i$$

$O$: observed number of deaths; $E$: expected number of deaths.

This in turn can be seen to be:

$$\log(O) = \log(E) + x_i \beta + \epsilon_i.$$  

This is similar to a Poisson model where the expected number of deaths is the offset. However, the Poisson regression model does not contain an error term and therefore the log-likelihood is derived from the distribution of the dependent variable, which is assumed to be Poisson.

### 3.6 Overdispersion

A central assumption of the Poisson distribution is that the variance is equal to the mean; this is often not the case. Overdispersion occurs simply when the variance is greater than the mean. When Poisson models are applied to overdispersed count data
the estimates of coefficients are likely to be reliable, but the standard errors tend to be conservative. This leads to wide confidence intervals and impacts on interpretation.

Overdispersion in count data is common, McCullagh and Nelder (1989, p125) point out that “overdispersion is the norm in practice and nominal dispersion is the exception”. Hilbe (2011, p140) agrees, stating that the assumption of equidispersion is “rarely met with real data” and that “overdispersion is in fact the norm”.

Two approaches to overdispersion are most commonly seen. A dispersion parameter can be introduced to a standard Poisson model (quasi-Poisson), or a different probability distribution can be used, such as the negative binomial (Ver Hoef and Boveng, 2007). These approaches are described below.

Assessing overdispersion

It is common to assess overdispersion by first applying a Poisson model to the count data and then applying a goodness of fit test. Whilst this does not indicate that overdispersion is necessarily present, it does indicate that a Poisson model may not be suitable for the data.

In addition, Hilbe (2011, p175) suggests two tests; a score test and the Lagrange multiplier test, both of which test the hypothesis that the Poisson model is overdispersed.

Score Test

\[ Z_i = \frac{(y_i - \mu_i)^2 - y_i}{\mu_i \sqrt{2}}. \]

Lagrange Multiplier Test

\[ \chi^2 = \frac{(\Sigma_i \mu_i^2 - ny)^2}{2 \Sigma_i \mu_i^2}. \]

Within most statistical packages when data are modelled using a negative binomial approach, included in the output is a test of whether the dispersion parameter included in the negative binomial model is zero, and an estimate of the value of the dispersion parameter.

Czardo and Sikora (2002) argue that merely testing whether the dispersion parameter is zero is insufficient if one is to determine whether or not to switch from the Poisson model to the negative binomial model, which involves the estimate of an additional parameter.
and is therefore less parsimonious. A discrepancy measure has been developed in an attempt to answer whether overdispersion in the data is low enough to ignore it, or whether the effort in switching to the negative binomial model is justified. Zuur et al (2010) argue that if the parameters of a Poisson GLM are highly significant, then a small correction of the standard errors due to overdispersion is not going to make any differences to the conclusions, if those conclusions are primarily based on significance.

3.7 Models for overdispersed count data

3.7.1 Quasi-Poisson

In the presence of overdispersion standard errors tend to be conservative. In a quasi-Poisson model a scale parameter is used to correct the standard errors; it is assumed that the variance is proportional to the mean:

$$\text{Var}(Y) = \phi E(Y) = \phi \mu$$

and the scale parameter $\phi$ can be based on the Pearson $\chi^2$:

$$\hat{\phi} = \frac{\chi^2_p}{n - p}.$$

This approach assumes that the model is relatively well fitted, as it assumes that $\chi^2_p$ is due to extra-Poisson variation rather than lack of fit (Rodríguez, 2013).

3.7.2 Negative Binomial models

There are several distinct negative binomial models (Hilbe, 2011), but the most common is derived from a Poisson-gamma mixture distribution, symbolised as NB2. If $Y$ follows a Poisson distribution then $Y_i \sim \text{Poisson}(\mu_i)$, where $\mu_i$ is a fixed term, which is a function of the explanatory variables. In an NB2 model it is assumed that $\mu_i$ is a random variable following a Gamma distribution. The first two moments of the NB2 model are:

$$E(y_i) = \mu_i$$
\[ V(y_i) = \mu_i(1 + \alpha \mu_i^2), \]

if \( \alpha = 0 \) this reduces the NB2 model to a Poisson model. In contrast to the quasi-Poisson model described above the variance is a function of the square of the mean. Maximum likelihood is used to estimate the parameters (Chatterjee and Simonoff, 2012).

Similar model diagnostics as those used for Poisson models are suitable for NB2 models, standardised deviance residuals and Anscombe residuals are recommended by Hilbe (2011, p. 37).

**Assessing the appropriateness of a Negative Binomial Model**

It is common for statistical packages, for example Stata, to include a likelihood ratio test of \( \alpha = 0 \) when a negative binomial regression is used to model data (StataCorp, 2013, p.1136). McCullagh and Nelder (1989) recommend both checking the variance and the link function. The variance function can be checked by plotting absolute values of Studentized deviance residuals against the fitted values transformed to the constant information scale. If the variance function is satisfactory there will be no pattern; if there is a positive trend there is evidence that the current variance function is increasing too slowly with the mean. The link function can be tested by plotting studentized deviance residuals against fitted values on the transformed to the constant information scale (Pokropp et al., 2006).

**3.7.3 Selecting between quasi-Poisson and Negative Binomial regression**

Both models take overdispersion into account, both models have two parameters, one more than Poisson models, and the majority of statistical packages include both regression models. It is sensible to consider how an informed choice between models can be made. AIC and BIC can be used to choose between the models, but these models depend on a distributional form, which is not the case for the quasi-Poisson model. Although quasi-information approaches have been developed there is some concern that this approach has not been shown to be valid (Ver Hoef and Boveng, 2007). The main advantage of using negative binomial regression is that it is based on a distribution in its own right and does not involve the adoption of quasi-likelihood methods. However,
it has been shown that the difference in the relationship between the variance and the mean can have important implications. Through example it has been shown that negative binomial regression gives smaller units more weight relative to the quasi-Poisson (Ver Hoef and Boveng, 2007). The possible implications of this should be considered when data are modelled.

**Apparent or actual dispersion**

The choice of a negative binomial or quasi-Possion over a Poisson model is based on the presence of overdispersion. It is important to be sure the data are actually overdispersed. Hilbe (2011) distinguishes between apparent and actual overdispersion. Reasons for apparent overdispersion are as follows:

- The model omits important explanatory variables.
- The data includes outliers.
- The model fails to include sufficient interaction terms.
- A predictor needs to be transformed.
- The link is misspecified.

These fundamentally describe mis-specification of the model which should be able to be modified. Zuur et al (2010) argue that only once these have been ruled out can we conclude there is real overdispersion. Actual dispersion is caused by positive correlation between responses or by an excess variation between response probabilities or counts.

### 3.8 Modelling approach - conclusions

Given the concerns outlined in Section 3.3.3, modelling SMRs using multiple linear regression will not be the primary modelling approach. Initially a Poisson model will be used to model counts of deaths; with the practice size as the exposure. If there is evidence of overdispersion both the quasi-Poisson and the negative binomial model will be considered. The negative binomial model is the preferred option to take overdispersion into account due to the known distributional form. In addition, multiple linear regression of indirectly standardised rates will be completed and the results compared to the appropriate model of count data; this is described in Chapter 13.
3.9 Variable selection

Variable selection has been described as the hardest part of model building (Ratner, 2010). The aim is to identify the necessary variables and delete irrelevant and redundant variables (those not affecting and those not adding to the dependent variable respectively). A process is needed which is sensitive, including necessary variables, and specific, excluding unnecessary variables (Greenland, 1989). The general problem of variable selection and possible approaches are discussed below.

3.9.1 The general problem

Observational studies often have a large number of potential variables which may explain the observed response. The large number of potential predictors means it is necessary to select a subset of potential predictors in the final model to avoid overfitting. ‘Overfitting’ means using more variables than necessary in the model and can lead to estimations of $R^2$ being unrealistically high and a belief that the model explains a large proportion of variance. There is also an increased likelihood of collinearity which leads to large standard errors, leading to misinterpretation of parameter estimates. In addition, including a large number of variables means the model is less likely to be generalisable to other situations, such as periods of time or geographical locations. Parsimony in model building is therefore important, although the need to further reduce the number of predictors may depend on the overall purpose of the model. If the main aim of the model is prediction it is important to include the smallest subset of variables which leads to unbiased prediction. However, if the main purpose is explanation, a fuller model with predictors which are not significant may be more appropriate (Whittingham et al., 2006). As a general rule, Harrell (2001, p.61) recommends that to avoid overfitting, the number of candidate variables which should be considered is $m/10$ or $m/20$ where $M$ is the limiting sample size.

The aim is to develop a useful model, which adequately describes the relationship, is generalisable to other situations and does not contain nuisance variables which do not add useful information.
3.9.2 Approaches to variable selection

Stepwise selection approaches

Stepwise procedures are approaches to variable selection which are based on using $p$-values to select or eliminate variables from a model. Since the use of statistical packages it has become common to use automated approaches based on statistical tests, statistical criteria and statistical stopping rules. Methods which eliminate or include variables based on statistical significance, such as ‘stepwise regression’ have become particularly popular and are still commonly used despite it being condemned in many textbooks (Greenland and Pearce, 2014).

Backward elimination procedures begin with all the potential predictors in the model, and the predictor with the highest $p$-value above a critical, user defined value, is ‘eliminated’ from the model and the process is repeated until all $p$-values are less than the critical value. Forward selection starts with an empty model and all predictors are included in the model in turn; the variable with lowest $p$-value is added to the model. This process is repeated until no additional predictor is under the critical value. Stepwise regression is a combination of both forwards and backward selection and is commonly available in statistical software packages.

Although it is now easy to find criticisms of stepwise approaches to variable selection it remains a commonly used approach in research of this type, for example in a study of heart failure admission rates stepwise selection was used (Brettell et al., 2013):

Bivariate analysis was performed initially, followed by multivariate analysis. Covariates were selected using backwards stepwise selection, and non-significant factors were removed using likelihood ratio tests.

Stepwise selection has also been used in a variety of related studies; see Baker et al., 2009; Soljak et al., 2011; Ashworth et al., 2007 as examples. The approach is also common in research in other areas, for example, 57% of studies in three leading ecological journals in 2004 used stepwise selection (Whittingham et al., 2006).

Harrell (2001, p.56) suggests seven problems with using stepwise approaches to variable selection; including yielding $R^2$ values that are too high, standard errors of regression coefficients that are biased low and $p$-values that are too small. Possibly most importantly it ‘allows us to not think about the problem’.

Stepwise regression may have an important role to play in model building. For example
when other approaches have reduced the number of variables, it may still be advantageous or necessary to reduce the number of variables; in this case stepwise approaches may support the researcher in developing a useful model. This author recommends that the use of stepwise selection methods should be noted in research papers and the potential impact of stepwise variable selection should be taken into when the results of the model are interpreted.

**Best subsets regression**

This approach considers all the possible subsets of variables, and interaction terms if appropriate, and the subset which does the best at meeting predetermined criteria is selected. Models may be selected based on maximising $R^2$ (or adjusted $R^2$), minimising the residual sum of squares and other goodness of fit measures such as information criteria (AIC and BIC) or Mallows’ $C_p$, although Mallows himself considered this was not a satisfactory approach (George, 2000).

This approach depends on a researcher-based definition of ‘best’ and may result in models which are difficult to interpret; the approach can also be sensitive to small changes in data (George, 2000).

**Bayesian model averaging**

Bayesian model averaging does not aim to select one model, but computes a weighted average for each parameter based on all possible models. Prior distributions are required for parameters within models and prior weights are needed for each model. Advantages of this approach include avoiding multiple hypothesis testing and produce more realistic measures of uncertainty for IRRs (Clyde, 2000).

**The full model**

One approach is to include all potential predictors in the model; this has the advantage of meaning there is a single set of hypotheses on a single model and the properties of the model are well understood. This is a useful model when the aim is to determine which predictors are associated with the response variable. The full model may not be useful if we wish to make predictions, as the inclusion of non-significant predictors which may be nuisance variables, will influence the parameters of predictors which may be more important. The full model needs careful interpretation; the reliance on
‘$p$-values’ to determine if a variable is significant should be avoided and it should be noted that the null hypothesis of ‘no association’ may not be appropriate as variables would only have been included if there was prior knowledge to suggest an association (Whittingham et al., 2006).

However, a full model is difficult to define in all cases, particularly where data are available from a variety of sources rather than having been collected for the specific purpose of the study. Selection of variables to include in the full model is subject to researcher bias and as only variables that the researcher considers important are likely to be included in the model; whilst this may be based on previous research the selection of previous research may also be subject to bias (Rentsch et al., 2014).

### 3.9.3 Collinearity

Collinearity occurs when explanatory variables are included in a statistical model which are correlated with other explanatory variables and this may have an impact on which explanatory variables can be included in a model (Harrell, 2001, p65). Collinearity increases the standard errors of the coefficients; this may result in explanatory variables not considered ‘significant’. Variance inflation factors are often used to quantify collinearity (Chatterjee and Simonoff, 2012, p28). The variance inflation factor (VIF) for each explanatory variable is defined as:

$$VIF = \frac{1}{1 - R^2_j},$$

where $R^2_j$ is the $R^2$ of the regression of the variable $x_j$ on the explanatory variables.

Everitt and Rabe-Hesketh (2003, p63) suggest that VIF values of less than 10 should not cause concern. A second measure is the mean of the VIFs for all explanatory variables; a mean score considerably larger than one suggests collinearity.

### 3.9.4 A note on $p$-values

One important issue with both variable selection methods and interpretation of final models is the frequent emphasis on $p$-values. The $p$-value is the probability of obtaining a particular result if the null hypothesis is correct. The smaller the $p$-value the stronger the evidence is against the null hypothesis. It is common to divide $p$-values into significant, if they are below 0.05, and if they are above 0.05 as ‘not significant’. 
This arbitrary distinction is not recommended, and it has been suggested that this cut-off should not be given any special importance (Sterne, 2001). In addition the null hypothesis in multiple linear regression is that there is no association between the dependent variable and the explanatory variable, which may not be a realistic hypothesis given the explanatory variable may have been selected based on prior evidence of association (Whittingham et al., 2006). The decision on whether to include an explanatory variable or interpret the variable based on a \( p \)-value should always be treated with caution, especially if the commonly used 0.05 criterion is enforced without discretion. As recently discussed in Nature (Nuzzo, 2014) \( p \)-values should be treated with caution; they do not adequately describe or determine the reality of the situation. Multiple regression involves multiple hypothesis testing, which can lead to an increase in Type 1 errors (Gordon, 2012). (Approaches to counter this are available, for example the Bonferroni correction, which is considered conservative; the false discovery rate has gained in popularity in comparing multiple health care providers (Jones et al., 2008).) Confidence intervals are likely to be more use in interpretation, as they give information about both statistical significance and the direction and strength of the association (Prel et al., 2009). Some methods of variable selection rely on \( p \)-values, for example stepwise methods. In this case, choice of \( p \)-values needs to be carefully considered, although 0.05 is a common choice, this can lead to relatively poor model performance and the interpretation of \( p \)-values in the final model should take into account the multiple hypothesis testing inherent in variable selection methods of this type (Steyerberg et al., 2000).

### 3.10 Conclusions

This chapter has briefly described approaches to standardisation of mortality data and approaches to modelling data. The challenges of a commonly used method, linear regression of SMRs are summarised. As there is strong evidence that estimates will be biased if this approach is used, this thesis will primarily use counts of deaths as the dependent variable. Appropriate count models will be utilised, depending on the level of dispersion. This is described in Chapter 10. The impact of modelling standardised rates using multiple linear regression on results and their interpretation will be discussed in Chapter 13. Various approaches to variable selection are available (Section 3.9); a pragmatic approach based on clinical research, availability of data and previous research will be adopted as the primary approach in this thesis. The process is described in Chapter 9.
In the next chapters the data which are the focus of the thesis are described. Initially
the context of the data, in terms of geography and population, is described (Chapter
4). Chapter 5 describes patterns in mortality in the local area. Part III identifies
and evaluates the data which are available to measure factors which are of interest as
candidate explanatory variables.
Part II

The local context
Chapter 4

Geography, demographics and health of the area of study

4.1 Introduction

The area of study covers general practices in three Primary Care Trusts (PCTs) located within the East Midlands of England.

The East Midlands is one of 10 regions of England. It is the fourth largest region in England, covering 12% of England; 6% of the UK. It includes a diverse landscape including large conurbations and a National Park. There is considerable variation in the socio-economic make-up of the area, as well as the proportion of the population who are from different ethnic groups. The East Midlands is made up of five counties, and in 2006/07 nine PCTs, covered by a single Strategic Health Authority (SHA). This chapter provides an overview of the area in terms of age, ethnicity and deprivation. Relevant aspects of the health of the resident population, according to published data, are also described.

4.2 Demographic characteristics of the specific area

This research covers general practices in three PCTs: Leicestershire and Rutland (L&R); Leicester City; and Northamptonshire (Northants). These PCTs collaborated as part of the CLAHRC (Collaboration for Leadership in Applied Health Research and Care) project (LNR-CLAHRC, 2014).
The three PCTs cover the same geographical area as two county councils: Leicestershire and Northamptonshire; and two unitary authorities: Leicester City and Rutland. The total population served by the three PCTs was estimated at 1.67 million by the middle of 2009.

230 general practices were operating for the complete period of study; a three year period, April 2006 to March 2009. 84 practices were in Leicestershire and Rutland (L&R); 83 in Northamptonshire; and 63 in Leicester City.

### 4.2.1 Age

L&R and Northamptonshire have age profiles in line with England as a whole; Leicester has a younger population. The mean age in Leicester 35.5 years which compares to 38.7 for England and Wales. Leicester has a slightly higher proportion of people in the 0-15 age group (Leicester City Council and Leicester PCT, 2007). Leicester has a higher than average proportion of people in the 20-24 age group which may be explained by the large student population and the large number of migrants in Leicester (Krausova and Vargas-Silva, 2013).

### 4.2.2 Ethnicity

There is considerable variation in the ethnic breakdown of the resident population of the three PCTs; Leicester City PCT has a high proportion of Asian/Asian other when compare to the other PCTs and the East Midlands as a whole, 71.6% of this group in Leicester are of Indian descent.

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Leicester</th>
<th>L&amp;R</th>
<th>Northants</th>
<th>East Midlands</th>
<th>England</th>
</tr>
</thead>
<tbody>
<tr>
<td>White British/White other</td>
<td>64.1%</td>
<td>90.8%</td>
<td>90.8%</td>
<td>90.1%</td>
<td>87.4%</td>
</tr>
<tr>
<td>Mixed</td>
<td>2.9%</td>
<td>1.3%</td>
<td>1.8%</td>
<td>1.6%</td>
<td>1.8%</td>
</tr>
<tr>
<td>Asian/Asian other</td>
<td>26.1%</td>
<td>5.6%</td>
<td>4.0%</td>
<td>5.4%</td>
<td>6.1%</td>
</tr>
<tr>
<td>Black/Black other</td>
<td>3.8%</td>
<td>0.1%</td>
<td>1.9%</td>
<td>1.6%</td>
<td>2.9%</td>
</tr>
<tr>
<td>Chinese/Chinese other</td>
<td>3.1%</td>
<td>1.2%</td>
<td>1.5%</td>
<td>1.3%</td>
<td>1.3%</td>
</tr>
</tbody>
</table>

Table 4.1: Percentage of population in key ethnic groups in each PCT, and East Midlands and England for comparison. Mid year estimates for 2009 (ONS).
4.2.3 Deprivation

Using the Index of Deprivation 2007, described in more detail in Chapter 6, it can be seen that the three PCTs are quite different in terms of levels of deprivation. Leicester is ranked as the 20th most deprived local authority (out of 354) and nearly half of Leicester’s population live in areas classed as the fifth most deprived in England. This compares to Leicestershire which was ranked 138th.

PCT boundaries and populations are similar to local authorities but not exactly the same, for example Leicestershire and Rutland are combined for PCT data. A summary of the score and rank is shown in Table 4.2. The overall scores for the PCTs can be seen to exhibit considerable variation.

<table>
<thead>
<tr>
<th>Area code</th>
<th>Area</th>
<th>IMD score*</th>
<th>rank**</th>
</tr>
</thead>
<tbody>
<tr>
<td>5PC</td>
<td>Leicester</td>
<td>34.68</td>
<td>20</td>
</tr>
<tr>
<td>5PA</td>
<td>Leicestershire and Rutland</td>
<td>10.61</td>
<td>146</td>
</tr>
<tr>
<td>5PD</td>
<td>Northamptonshire</td>
<td>16.05</td>
<td>116</td>
</tr>
<tr>
<td>00FN</td>
<td>Leicester Unitary Authority***</td>
<td>34.68</td>
<td>20</td>
</tr>
<tr>
<td>00FP</td>
<td>Rutland Unitary Authority***</td>
<td>7.49</td>
<td>334</td>
</tr>
<tr>
<td>31UB</td>
<td>Blaby</td>
<td>8.41</td>
<td>326</td>
</tr>
<tr>
<td>31UC</td>
<td>Charnwood</td>
<td>11.95</td>
<td>264</td>
</tr>
<tr>
<td>31UD</td>
<td>Harborough</td>
<td>7.08</td>
<td>344</td>
</tr>
<tr>
<td>31UE</td>
<td>Hinckley and Bosworth</td>
<td>10.90</td>
<td>283</td>
</tr>
<tr>
<td>31UG</td>
<td>Melton</td>
<td>10.43</td>
<td>294</td>
</tr>
<tr>
<td>31UH</td>
<td>North West Leicestershire</td>
<td>14.73</td>
<td>219</td>
</tr>
<tr>
<td>31UJ</td>
<td>Oadby and Wigston</td>
<td>10.51</td>
<td>293</td>
</tr>
<tr>
<td>34UB</td>
<td>Corby</td>
<td>26.16</td>
<td>75</td>
</tr>
<tr>
<td>34UC</td>
<td>Daventry</td>
<td>10.61</td>
<td>292</td>
</tr>
<tr>
<td>34UD</td>
<td>East Northamptonshire</td>
<td>11.78</td>
<td>266</td>
</tr>
<tr>
<td>34UE</td>
<td>Kettering</td>
<td>15.09</td>
<td>214</td>
</tr>
<tr>
<td>34UF</td>
<td>Northampton</td>
<td>21.15</td>
<td>129</td>
</tr>
<tr>
<td>34UG</td>
<td>South Northamptonshire</td>
<td>6.46</td>
<td>351</td>
</tr>
<tr>
<td>34UH</td>
<td>Wellingborough</td>
<td>17.79</td>
<td>168</td>
</tr>
</tbody>
</table>

* a high score indicates a more deprived area.
** the most deprived area is given a rank of one.
*** Leicester and Rutland are not subdivided because they are unitary authorities.

Table 4.2: Summary of deprivation in various geographical regions in the area of interest (based on Index of Deprivation, 2007).
4.2.4 Other features

There are four universities in the area of study, two in Leicester City, one in Leicestershire and one in Northamptonshire. All four universities have practices closely associated with the student population, although the practice servicing the University of Northampton has two branches, one serving a local village population and one the university, meaning that it has a more mixed patient population.

The area is made up of a mix of large urban, other urban and rural districts (ONS, 2013).

4.2.5 Conclusions - demography of the local area

The demography of the area of interest is basically similar to England as a whole, in that there is considerable variation between districts. Some areas show a high percentage of people from Asian/Asian other ethnic groups; some districts have a high percentage living in postcodes associated with high levels of socioeconomic deprivation and some have a distinctive age profile. There are clear differences between the three PCTs and between local authority districts; however, people are free to access primary care where they choose, generally close to where they live, but not restricted by district or PCT boundaries. It is therefore important to consider the impact of practice populations and less important to consider district and PCT population demographics; while acknowledging that local area environments may have an impact on individuals and their health.

4.3 Health of the local area

Life expectancy in the East Midlands was similar to England as a whole in 2005-07, 77.6 years for males and 81.6 for female (77.5 and 81.7 for England respectively). There is considerable variation in the area; male life expectancy in Rutland is 2.5 years longer than England as a whole. In contrast, Leicester City has a life expectancy is 2.2 years less than England. By 2009 life expectancy for females was lower than England as a whole (Department of Health & APHO, 2010).

The South East Public Health Observatory (SEPHO) has published CVD profiles for each of the three PCTs (SEPHO, 2013); these were first published in 2009/10, which
is after the period of study but are still a useful reference point for comparing the health of the three PCTs; these can be seen in Table 4.3. Premature CVD mortality in Leicester is significantly worse than in England as a whole; this is also true for premature AMI (acute myocardial infarction) and stroke mortality. In comparison, in Northamptonshire and Leicestershire and Rutland premature mortality is significantly better than England. A similar pattern can be seen for CHD admissions, although there is less difference for stroke admissions, with only Leicestershire and Rutland being significantly better than in England as a whole. Risk factors are given as part of the CVD profiles (SEPHO, 2013); both Northamptonshire and Leicester have higher than average smoking prevalence, whereas in Leicestershire and Rutland it is significantly lower than England. The CVD profiles also show that observed CHD prevalence divided by expected CHD prevalence is similar across the three PCTs, although slightly lower in Leicester. A similar pattern is seen for observed divided by expected hypertension prevalence. In addition diabetes profiles produced by the Yorkshire and Humber Public Health Observatory (YHPHO) showed that diabetes prevalence in Leicester City was 5.1% in 2008/09 compared to 4.0% in L&R and 3.9% in Northamptonshire (YHPHO, 2013).

<table>
<thead>
<tr>
<th>Indicator</th>
<th>5PD</th>
<th>5PA</th>
<th>5PC</th>
<th>England Average</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mortality and admissions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVD mortality (under 75)</td>
<td>65.4+</td>
<td>55.3+</td>
<td>111.5+</td>
<td>70.4</td>
</tr>
<tr>
<td>AMI mortality (under 75)</td>
<td>13.4+</td>
<td>11.6+</td>
<td>24.4+</td>
<td>16.3</td>
</tr>
<tr>
<td>Stroke mortality (under 75)</td>
<td>10.9+</td>
<td>10.1+</td>
<td>18.6+</td>
<td>12.8</td>
</tr>
<tr>
<td>CHD emergency admissions</td>
<td>178.5+</td>
<td>154.7+</td>
<td>264.8+</td>
<td>205.3</td>
</tr>
<tr>
<td>Stroke emergency admissions</td>
<td>101.7</td>
<td>86.3+</td>
<td>101.3</td>
<td>104.2</td>
</tr>
<tr>
<td><strong>Lifestyle factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated % smokers (16+)</td>
<td>23.3−</td>
<td>20.3+</td>
<td>31.3+</td>
<td>22.2</td>
</tr>
<tr>
<td>Estimated % obese (16+)</td>
<td>25.1−</td>
<td>24.3</td>
<td>24.6</td>
<td>24.2</td>
</tr>
<tr>
<td>4 week quitters per smokers (%)</td>
<td>4.1</td>
<td>4.4</td>
<td>3.3</td>
<td>4.0</td>
</tr>
<tr>
<td><strong>Detection of key risk factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs/Exp CHD prevalence</td>
<td>0.62</td>
<td>0.66</td>
<td>0.55</td>
<td>0.61</td>
</tr>
<tr>
<td>Obs/Exp Hypertension prevalence</td>
<td>0.44</td>
<td>0.48</td>
<td>0.42</td>
<td>0.44</td>
</tr>
</tbody>
</table>

+ significantly better than England as a whole.
− significantly worse than England as a whole.

Table 4.3: Summary of health indicators included in CVD profiles compiled by SEPHO (2009/10); significance as given by SEPHO (2013).

Standardised mortality rates published by the National Centre for Health Outcomes Development (NCHOD) are shown in Table 4.4 (NCHOD - The NHS Information Centre for health and social care, 2009). These clearly show that mortality rates for Leicester are higher than expected for all cardiovascular diseases and CHD and stroke as specific categories. This is the case for all-age and premature mortality and for
males and females. All-age female stroke deaths is the only category of deaths which is not significantly above 100. L&R has SMRs lower than 100 for all categories of death. Northamptonshire has mortality rates lower than 100 in most categories; the confidence intervals overlap 100 for most female deaths and all premature stroke deaths.

<table>
<thead>
<tr>
<th>PCT</th>
<th>SMR (95% CI)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
<td>All persons</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CVD - all-age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L&amp;R</td>
<td>91 (88, 94)</td>
<td>89 (86, 92)</td>
<td>90 (88, 92)</td>
<td></td>
</tr>
<tr>
<td>Leicester</td>
<td>125 (119, 132)</td>
<td>116 (110, 123)</td>
<td>121 (116, 125)</td>
<td></td>
</tr>
<tr>
<td>Northants</td>
<td>96 (92, 99)</td>
<td>96 (92, 100)</td>
<td>96 (93, 98)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CVD - U75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L&amp;R</td>
<td>80 (75, 86)</td>
<td>80 (73, 88)</td>
<td>80 (76, 85)</td>
<td></td>
</tr>
<tr>
<td>Leicester</td>
<td>147 (135, 160)</td>
<td>154 (137, 174)</td>
<td>149 (139, 160)</td>
<td></td>
</tr>
<tr>
<td>Northants</td>
<td>89 (84, 95)</td>
<td>101 (92, 110)</td>
<td>93 (88, 98)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHD - all-age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L&amp;R</td>
<td>94 (89, 99)</td>
<td>91 (86, 97)</td>
<td>93 (89, 96)</td>
<td></td>
</tr>
<tr>
<td>Leicester</td>
<td>137 (127, 147)</td>
<td>129 (118, 140)</td>
<td>133 (126, 141)</td>
<td></td>
</tr>
<tr>
<td>Northants</td>
<td>94 (90, 99)</td>
<td>90 (85, 96)</td>
<td>93 (89, 96)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHD - U75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L&amp;R</td>
<td>86 (80, 93)</td>
<td>84 (73, 96)</td>
<td>85 (80, 91)</td>
<td></td>
</tr>
<tr>
<td>Leicester</td>
<td>162 (146, 180)</td>
<td>183 (155, 216)</td>
<td>168 (154, 183)</td>
<td></td>
</tr>
<tr>
<td>Northants</td>
<td>87 (80, 94)</td>
<td>91 (79, 105)</td>
<td>88 (82, 94)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stroke - all-age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L&amp;R</td>
<td>92 (85, 99)</td>
<td>92 (86, 98)</td>
<td>92 (88, 97)</td>
<td></td>
</tr>
<tr>
<td>Leicester</td>
<td>115 (101, 130)</td>
<td>108 (97, 119)</td>
<td>110 (102, 119)</td>
<td></td>
</tr>
<tr>
<td>Northants</td>
<td>94 (87, 102)</td>
<td>93 (87, 99)</td>
<td>93 (89, 98)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stroke - U75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L&amp;R</td>
<td>78 (65, 92)</td>
<td>77 (63, 93)</td>
<td>77 (68, 88)</td>
<td></td>
</tr>
<tr>
<td>Leicester</td>
<td>143 (113, 178)</td>
<td>131 (100, 168)</td>
<td>137 (115, 162)</td>
<td></td>
</tr>
<tr>
<td>Northants</td>
<td>87 (74, 103)</td>
<td>105 (88, 124)</td>
<td>95 (84, 107)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4: Summary of indirectly standardised mortality rates and confidence intervals for CVD, CHD and stroke, 2006-08 (pooled) (NCHOD - The NHS Information Centre for health and social care, 2009).

4.4 Conclusions

The population covered by the area of study is determined by the patients registered at each practice included in the study. The information presented in this chapter describes the resident population of the three PCTs included in the study, or the geographical areas with similar boundaries where PCT based information is not available.

The area is a diverse area; when taken as a whole many characteristics are similar to those of England, but within the area there are extreme differences. This is likely to be reflected in the characteristics of practice populations, which can be expected
to be heterogeneous in terms of age, ethnicity and socioeconomic status. There is evidence that localities within the area of study have higher than average CVD and CHD mortality. There is also evidence of variation in the prevalence of some risk factors.

The heterogenous nature of the population makes the area of study of particular interest as it means the practice populations reflect the range of populations which might be seen across England. It also demonstrates how important it is to take population characteristics into account when comparing practices and studying associations between primary care and health outcomes.

In the next chapter patterns of mortality are described in more detail.
Chapter 5

Patterns of mortality in the area of study

5.1 Introduction

In the previous chapter information about the health of the area of study was summarised using information published at PCT or local authority level. In this chapter practice level mortality data provided directly for this research have been analysed and presented, for the whole sample, aggregated to PCTs and at individual practice level.

Initially the numbers of deaths across the whole area and across the individual PCTs are described. In Section 5.4 practice level data are described, including standardised mortality rates. Throughout the chapter only deaths caused by diseases classified as cardiovascular are described.

5.2 Processing of data

Deaths of interest were those included in the extract of the Primary Care Mortality Database (PCMD), April 2006 to March 2009. Only deaths matched to general practices within the three PCTs were considered. In total 40,998 deaths were included in the PCMD extract, eight deaths were excluded due to no underlying cause of death. Practices where there was concern that deaths did not represent the complete three year period or where population data were not available were investigated. There was evidence that for seven practices they either opened or closed during the three year
period. These were excluded from the study, meaning 35 deaths, 11 with CVD as underlying cause, were excluded from analysis.

Underlying cause of death based on the International Classification of Disease revision 10 (ICD-10) (WHO, 2014b) was used to classify deaths as CVD (Chapter IX - Diseases of the circulatory system) and further subdivided into coronary heart disease (CHD) (ICD-10 I20-I25) and stroke (I60-I69).

5.2.1 Practice with restricted patient list

Practice Z was identified as being an ‘unusual’ practice when preliminary work including descriptive statistics and initial linear regression analysis was carried out. Practice Z is an extremely small practice, with 830 or fewer patients registered in each year, this is less than 60% of the size of the next smallest practice. Over 80% of patients are male. Initial linear regression analysis clearly showed that this practice was having an extremely high influence on analysis. Having identified the practice as an unusual practice further investigation showed that the practice offers a service to a restricted patient list, specialising in care for homeless people. For this reason it was decided that this practice should be excluded from all analysis as including it would potentially distort the main message and would not add to the body of knowledge about ‘general practice’.

5.3 Area level summary data

In this section the numbers and percentages of deaths are described. Key differences between genders are discussed. The nature of the available data means it is not possible to discuss deaths in terms of ethnicity.

5.3.1 Cause of death

All summary data shown below is based on 40924 deaths from 229 practices across three PCTs, from April 2006 to March 2009.

Of the 40924 deaths in the sample, 13424 had an underlying cause of death within Chapter IX - diseases of the circulatory system (I00-I99). This accounts for 32.8%
of all deaths, similar to data for England (33.0% - 2008). Table 5.1 shows overall cardiovascular deaths and two sub-categories: CHD and stroke.

<table>
<thead>
<tr>
<th>Underlying cause of death</th>
<th>No. of deaths (% of CVD deaths) [England]</th>
<th>No. of deaths under 75 (% of all-age) [England]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVD (I00-I99)</td>
<td>13424 (100%) [n/a]</td>
<td>3621 (27.0%) [25.4]</td>
</tr>
<tr>
<td>CHD (I20-I25)</td>
<td>6443 (48.0%) [45.7]</td>
<td>2148 (33.3%) [31.4]</td>
</tr>
<tr>
<td>Stroke (I60-I69)</td>
<td>3556 (26.5%) [21.3]</td>
<td>622 (17.5%) [21.7]</td>
</tr>
</tbody>
</table>

Table 5.1: Numbers of deaths by key categories of CVD, including percentage of those deaths which are under 75.

CHD is the single biggest cause of death within CVD, accounting for 48.0% of the deaths; 33.3% of these deaths affecting people aged under 75. These data are similar to England (Table 5.1), although in England a slightly lower proportion of deaths are caused by CHD and a slightly lower proportion are under 75. A higher proportion of CHD deaths are premature when compared to stroke deaths (33.3% compared to 17.5%), again similar to England. A lower proportion of stroke deaths are premature in the area of study compared to England as a whole (17.5% compared to 21.7% in England).

59.3% of premature (under 75) CVD deaths are caused by CHD, this compares to 56.5% in England as a whole. Table 5.2 shows that approximately two thirds of premature CVD deaths in males are caused by CHD. CHD is clearly an important category of CVD to study if premature deaths is an area of interest.

<table>
<thead>
<tr>
<th>Underlying cause of death</th>
<th>No. of deaths under 75 (% of premature (under 75) CVD deaths)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
</tr>
<tr>
<td>CVD (I00-I99)</td>
<td>3621 deaths</td>
</tr>
<tr>
<td>CHD (I20-I25)</td>
<td>2148 deaths (59.3%)</td>
</tr>
<tr>
<td>Stroke (I60-I69)</td>
<td>622 deaths (17.2%)</td>
</tr>
</tbody>
</table>

Table 5.2: Numbers of premature deaths caused by key categories of CVD, for total population, male and female.

Overall, when cause of death is considered, the area of study is similar to England.

**Differences across genders**

The patterns of deaths, both all-age and premature, for males and females demonstrate some interesting patterns, shown in Table 5.3. Approximately half of all deaths caused
by CVDs are male, slightly higher than in England as a whole, this increases to 67.5% when premature deaths are considered. When CHD is the underlying cause of death 59.2% are male, this increases to 75.3% for premature deaths. This has implications for the interpretation of statistical modelling and associations, as the majority of premature deaths are male, any associations may be more important for males than females. When stroke is the underlying cause of premature death, the balance of deaths in males and females is more even, although still higher in males than females (55% males compared to 45% females). In addition, 42.4% of male CHD deaths are premature, compared to 20.2% of female deaths. A similar, although less extreme, pattern is found for stroke - 24.4% of male stroke deaths are premature compared to 13.0% of female stroke deaths.

Table 5.2 emphasizes other key differences; over half of male premature CVD deaths are caused by CHD (66.1%); this is higher than for England (62.0%), compared to 45.2% of female premature deaths, similar to England. When all-ages are considered 60.8% of CHD deaths are male, whereas 40.8% of stroke deaths are male, similar to England as a whole.

<table>
<thead>
<tr>
<th>Cause of death</th>
<th>Age</th>
<th>All</th>
<th>Male (% of total) [England]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVD (I00-I99)</td>
<td>all</td>
<td>13424</td>
<td>6705 (49.9%) [48.1%]</td>
</tr>
<tr>
<td></td>
<td>under 75</td>
<td>3621</td>
<td>2445 (67.5%) [67.6%]</td>
</tr>
<tr>
<td>CHD (I20-I25)</td>
<td>all</td>
<td>6443</td>
<td>3815 (59.2%) [56.5%]</td>
</tr>
<tr>
<td></td>
<td>under 75</td>
<td>2148</td>
<td>1617 (75.3%) [74.2%]</td>
</tr>
<tr>
<td>Stroke (I60-I69)</td>
<td>all</td>
<td>3556</td>
<td>1404 (39.4%) [49.9%]</td>
</tr>
<tr>
<td></td>
<td>under 75</td>
<td>622</td>
<td>342 (55.0%) [55.9%]</td>
</tr>
</tbody>
</table>

Table 5.3: Numbers of deaths caused by key categories of CVD, for total population and male.

Generally, patterns in the area of study are similar to those of England, although the proportion of all-age stroke deaths affecting males is lower than in England.

**Differences across PCTs**

The proportion of CVD deaths caused by CHD is higher in Leicester City than in L&R and Northamptonshire, and the proportion caused by stroke is lower in Leicester City than the other two PCTs (shown in Table 5.4). The proportion of deaths which are premature is higher in Leicester City for CVD, CHD and stroke than in L&R and Northamptonshire.

There are considerable differences in the population characteristics of the three PCTs.
Table 5.4: Patterns of CVD deaths in the area of study - percentage of all and premature CVD deaths which have underlying cause CHD and stroke.

and this may explain these differences; this was discussed in Chapter 4.

5.4 Practice level summary data

5.4.1 Numbers of deaths

Figures 5.1 to 5.6 show the relatively small numbers of deaths connected with practices on an annual basis. The numbers of all cardiovascular deaths varies from 0 to 96; 30% of practices had less than 10 deaths in 2007/08. The numbers of premature CHD is naturally lower; 29 practices had no deaths and only three practices have 10 or more deaths in 2007/08. For this reason the data for the three year period had been combined. When the three years are combined the number of premature deaths caused by CHD varies from 0 to 39 and only two practices had zero deaths (see Figures 5.7 and 5.8 for more detail). The median number of deaths caused by stroke was eight (IQR: (4, 13)), the median number of premature deaths is 2 (IQR: (1,4)) and 36 practices had zero deaths (see Figures 5.9 and 5.10).
Figure 5.1: All CVD deaths - 2007/08

Figure 5.2: Premature CVD deaths - 2007/08.

Figure 5.3: All CHD deaths - 2007/08

Figure 5.4: Premature CHD deaths - 2007/08.

Figure 5.5: All stroke deaths - 2007/08

Figure 5.6: Premature stroke deaths - 2007/08.
Figure 5.7: All CHD deaths - three years combined

Figure 5.8: Premature CHD deaths - three years combined.

Figure 5.9: All stroke deaths - three years combined

Figure 5.10: Premature stroke deaths - three years combined.
5.4.2 Mortality rates

Calculations of mortality rates

Crude rates were calculated based on the aggregated number of deaths in each of the three years divided by the aggregated practice population over the three years. For premature deaths the practice population aged under 75 was used. Indirect standardisation was chosen as an appropriate method of standardisation due to the low numbers of deaths, even when aggregated over the three years. 99% of practices had less than 25 premature CHD deaths over the three years and over 50% had less than 10.

Both SMRs and 95% confidence intervals were calculated based on the guidance given in the Compendium of Clinical Health Indicators User Guide (Lakhani et al., 2011, p656–8). Age specific death rates were calculated from mortality and population data published by the Office for National Statistics (2014). Data for the calendar year which represented the majority of the relevant financial year were used with the practice populations for each financial year to determine the expected numbers of deaths for each practice for each sex in each year. The expected numbers of deaths for each year were then summed for each sex; the practice total expected deaths is the sum of the male and female expected numbers. Expected numbers of deaths were then divided by the total observed deaths for each practice. This was done separately for all-age and premature deaths.

When PCT level SMRs were compared to externally published data, the estimate for one PCT lay outside the confidence interval of one external source, National Clinical Health Outcomes Database (NCHOD) (2009). This discrepancy was investigated. When the resident populations used by NCHOD were compared with estimated populations of the local authorities with concurrent boundaries to the PCTs it was evident that NCHOD use the resident populations within the PCT boundaries, whereas the populations used in this study were based on the populations of the practices within the PCTs. This difference in resident population is sufficient to explain differences in SMRs.

Funnel plots

Funnel Plots have been used to present mortality rates. Funnel plots are a type of control chart where the indicator of interest is plotted against the denominator, in this case the SMR is plotted against expected deaths. Included in the plot is the target value
and control limits. Funnel plots help to show whether the variation from the target value is ‘common cause variation’ which could be expected or ‘special-cause’ variation which highlights observations which are worthy of further investigation (Flowers, 2007). Various tools are available to enable easy production of funnel plots, for example APHO have published a ‘tool’ in the form of a spreadsheet (APHO, 2008). Although funnel plots are now considered preferable to other methods for comparing institutions (Spiegelhalter, 2005), they should be interpreted with caution as different approaches to calculating control limits and different underlying populations can result in different observations being considered outliers (Seaton and Manktelow, 2012; Evans et al., 2013; Seaton et al., 2013). Funnel plots presented here were created using Stata, with control limits based on Wald confidence intervals (Seaton and Manktelow, 2012) and a target value of 100, the expected SMR if each practice had the same age specific death rate as England and Wales.

**Crude rates**

Crude rates, deaths per 10,000, vary considerably. The mean all-age crude death rate for CHD is 12.90, compared to 4.82 for premature death. The funnel plots shown in Figures 5.11 and 5.12 show that the majority of practices are within control limits, although there are clear outliers. Three practices are below the control limits for both all-age and premature death rates.

Figure 5.11: Funnel plot of crude rate of CHD deaths.  
Figure 5.12: Funnel plot of crude rate of premature CHD deaths.

There is medium positive correlation between all-age and premature crude death rates (Fig. 5.13).
Standardised mortality ratios

There is considerable variation in indirectly standardised mortality ratios (SMRs) for both premature and all-age CHD mortality. SMRs for all-age CHD mortality vary from zero to 402.44, with a median rate of 96.02 (IQR: (78.32, 119.92)), showing that over 50% of practices have mortality rates below that expected. SMRs for premature CHD mortality vary from zero to 765.69, with a median rate of 93.07 (IQR: (64.92, 135.62). Figure 5.14 shows that the practice with a rate of 765.69 is a distinct outlier, the next highest SMR is 334. This outlier has been removed from subsequent graphs as this allows easier interpretation (see Fig. 5.15).
Figure 5.15: Bar plot to show variation in all-age CHD SMR - outlier removed.

A histogram of premature CHD SMR suggests that the premature mortality rates are not following a normal distribution but are positively skewed (Fig. 5.16).

Figure 5.16: Histogram of premature SMRs - outlier removed.

There is a strong positive correlation between premature and all-age CHD SMRs, practices with high premature CHD mortality rates also have high all-age CHD mortality rates. Figure 5.17 suggests that premature SMRs increase more rapidly, in comparison to expected rates, than all-age mortality rates.
Figs. 5.18 and 5.19 are funnel plots for all-age and premature CHD mortality rates. Five practices lie outside the 99.8% control limits for all-age mortality, compared to six for premature mortality. 35 practices lie outside the 95% control limits for all-age mortality, 16 above and 19 below, suggesting that higher numbers of practices than expected have both low and high SMR for all-age CHD. 30 practices lie outside the lower control limits for premature mortality, with 18 practices having a higher than expected SMR. For the 99.8% control limit, again there is a balance of practices above and below the limit for all-age mortality, (three practices are above and two practices are below). However for premature mortality, five practices are above the 99.8% control limit and only one below.

Of the two practices above the 99.8% control limit for all-age mortality one is also above the upper control limit for all-age mortality. Of the 21 practices which are above the 95% control limit, six are above the control limit for both all-age and premature mortality, seven are above the control limit for premature mortality and eight are above the limit for all-age mortality.

5.5 Conclusions

In this chapter the numbers of deaths and standardised mortality rates based on the Primary Care Mortality Database, provided specifically for this research. The numbers of premature deaths in each practice with CHD as the underlying cause is relatively low; deaths have therefore been aggregated over the three year period. Although the variation in standardised rates is wide, only small numbers of practices lie outside the
99.8% control limits.

In the following chapters data which describes both the characteristics of practice populations and quality of primary care are evaluated.
Part III

Data available on CHD risk factors
Chapter 6

Population characteristics - available data

6.1 Introduction

Although there are many surveys which aim to collect information on characteristics of the population of England, very few of them are useful for describing practice populations. For information describing practice populations, data collected by the practice itself is the most reliable. QOF data are routinely published and include measures of the prevalence of certain conditions as well as indicators summarising the delivery of aspects of primary care. In addition, information which is associated with postcodes can be linked to patient postcodes and weighted averages can be calculated. Some data are now published at local area level and these can be matched to practice postcodes, although this will not necessarily reflect where patients live. This chapter will consider the range of datasets available to describe key aspects of practice populations in terms of the known risk factors for CHD (Table 1.1 in Chapter 1).

This chapter is the first stage in the identification and evaluation of data, focussing on population characteristics. The following chapters (7 and 8) focus on data which describe aspects of primary care.
6.2 Key sources of data

Some data are available which specifically describe practice populations, but many datasets describe local area populations. Local areas include a range of administrative and geographical areas including PCTs, local authorities (a general term for any level of local government in the UK) and lower tier local authorities (local authority districts (LADs)). Within the study area, Leicestershire and Northamptonshire are each made up of seven LADs; Rutland and Leicester City are both unitary authorities (UAs), but for convenience are included within the term LAD. There is considerable variation between LADs; for example Leicester City Unitary Authority has a population of over 300,000 and has two hospitals, whereas Rutland UA has a population of under 40,000 and no hospital (Rogers, 2010).

In addition smaller areas, Lower Super Output Areas (LSOAs), designed to have social homogeneity based on the type of dwelling and the nature of tenure, have been devised by the Office of National Statistics. LSOAs are aggregated into Medium Super Output Areas (MSOAs), which were designed to fit into Local Authority boundaries. LSOAs and MSOAs are a consistent size across England, (see Table 6.1 for more information). Some data are available for LADs, MSOAs and/or LSOAs, but not readily available for practice populations.

<table>
<thead>
<tr>
<th></th>
<th>LSOA</th>
<th>MSOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum population</td>
<td>1000</td>
<td>5000</td>
</tr>
<tr>
<td>Median population</td>
<td>1500</td>
<td>7200</td>
</tr>
<tr>
<td>Avg. no of households</td>
<td>400</td>
<td>2000</td>
</tr>
<tr>
<td>Number in England</td>
<td>32482</td>
<td>6780</td>
</tr>
<tr>
<td>Number in area of study</td>
<td>1013</td>
<td>215</td>
</tr>
</tbody>
</table>

Table 6.1: Super Output Areas

Sources of data of potential interest are briefly summarised in Table 6.2.
<table>
<thead>
<tr>
<th>Information sources</th>
<th>Details</th>
<th>Population</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice register lists</td>
<td>Personal details of registered patients; including date of birth, address and sex. Medical and QOF information is also recorded for individual patients. Personal details are collated and reported to PCTs and other relevant organisations and was not generally publicly available. QOF information is collated and published. Other medical information is not generally collated nor publicly available.</td>
<td>Practice level.</td>
<td>Collated QOF data is publicly available; the majority of other information is not publicly available</td>
</tr>
<tr>
<td>Hospital episode statistics (HES)</td>
<td>HES has details of all NHS admitted patient care, outpatient appointments and accident and emergency (A&amp;E) attendances, excluding mental health and maternity admissions. Information about patients, their diagnoses and treatments and administrative information, for example waiting times, is recorded.</td>
<td>Practice level.</td>
<td>The majority of this information is not routinely, publicly available. Ethnicity of practice populations is available, this is discussed below.</td>
</tr>
<tr>
<td>Health Survey for England (HSE)</td>
<td>A cross-sectional survey which has been carried out on an annual basis since 1991 (National Centre of Social Research, 2012). Core topics which are covered each year include general health; smoking, drinking and fruit and vegetable consumption; height; weight; blood pressure measurements and blood and saliva samples.</td>
<td>National and regional.</td>
<td>Publicly available.</td>
</tr>
<tr>
<td>Modelled estimates</td>
<td>APHO (Association of Public Health Observatories; now part of Public Health England) have used the HSE to develop a range of models to provide estimates of underlying disease prevalence. Estimates are available for a range of diseases and healthy lifestyle behaviours.</td>
<td>Some estimates are available at practice level. Behaviour estimates are available at MSOA level.</td>
<td>Publicly available.</td>
</tr>
<tr>
<td>QOF prevalence registers</td>
<td>Available for a range of diseases.</td>
<td>Practice level.</td>
<td>Publicly available.</td>
</tr>
<tr>
<td>Other sources</td>
<td>Including Census data which are often used in composite measures, specific surveys and information from various Government departments.</td>
<td>LSOAs to regional.</td>
<td>Publicly available.</td>
</tr>
</tbody>
</table>

Table 6.2: Summary of key sources of data.
6.3 Unmodifiable biological risk factors

The underlying genetic make up of the population is likely to be a principal predictor of CHD; information relating to this is currently unavailable. Sex, ethnicity and age are the key unmodifiable biological risk factors about which there is available information. Age and sex are either used to standardise health outcomes or are used directly in count models.

Age and sex

For the purposes of this research, the PCTs involved supplied information regarding the age profile, by sex, of the practices for the years 2005/06 to 2009/10. These are based on practice prescribing populations and the registered patients who live in the PCTs. The data are collated on one specific day in each financial year and are therefore a snapshot of the practice population. Some practices will experience more mobility than others; that is the movement of patients onto and off the register throughout the year. This information was not available for this study, but may be an area for further research.

Ethnicity

APHO produce estimates of the ethnic breakdown of the registered patients derived from HES. HES was considered to be the only routine dataset which included both practice information and patient ethnicity (Walford et al., 2011). The GP patient survey (GPPS), described in Chapter 8 is now used to estimate proportions of black minority ethnic populations for National General Practice Profiles (Public Health England, 2014b). This information is not available for the period of study. Although the HES and GPPS collect ethnicity information in 16 detailed categories, the ethnic makeup of practice populations published by APHO, and used in the National General Practice Profiles, are broken down into six aggregated ethnic groupings: • Asian/Asian Other • Black/Black Other • Irish White • Mixed • Other • White (Healthcare Commission, 2008).

Ethnicity recorded in hospital records has been shown to be relatively accurate in comparison with self-reported ethnicity for white/non-white differences (Saunders et al., 2013). However, the same study found that between 20% and 35% of patients who self-report that they belong to one of the four major ethnic groups will be miscoded by
HES. This suggests that percentage white ethnicity may be the most reliable measure. When HES data are used to estimate practice ethnicity information it is assumed that the hospital admissions reflect the true ethnic population of the practice and there is no systematic bias. Unfortunately the same ethnic distribution is applied across all age bands as there are insufficient hospital admissions to reliably calculate the distribution of ethnic groups by age and sex for practices.

Estimates of ethnic breakdown based on the HES will be used in this analysis. It is acknowledged that this will not adequately describe certain ethnic groups in some practice populations, which will need to be considered as part of the discussion.

### 6.4 Modifiable biological risk factors

The key modifiable biological risk factors for CHD which have been identified as relevant for inclusion in this study are: diabetes; obesity; hypertension and serum cholesterol. Although it is common to include a range of biological risk factors in individual, patient level studies when studying the impact of interventions or increasing understanding of risk factors; it is relatively rare to include modifiable biological risk factors, such as diabetes and obesity prevalence in population based studies of the associations of primary care and health outcomes. Diabetes prevalence has been included as an explanatory variable in studies of associations between primary care and CHD mortality (Levene et al., 2010; Levene et al., 2012; Honeyford et al., 2013). In a study of unplanned admissions and diabetes management COPD, CHD, mental health and obesity prevalences were all included as potential explanatory variables (Dusheiko et al., 2011). A possible reason for the sparsity of studies including modifiable biological risk factors is that there are limited routinely available data describing these factors.

#### QOF prevalence registers

QOF registers are measures of detection - as only people diagnosed with the condition are included in the register. However, they may reflect prevalence if detection rates are high. Using QOF measures as measures of detection was an approach used by Levene et al (2010; 2012) for hypertension and for a range of chronic diseases in a study of detection and access (Anwar et al., 2012). Registers have also been used as measures of prevalence, for example the use of diabetes registers as measures of prevalence (Levene et al., 2012). Through QOF practices are awarded points for achieving particular
targets; each register is allocated six points. More details of QOF are given in Chapter 7.

The decision as to whether QOF registers should be used as measures of prevalence or detection has been influenced by studies comparing QOF registers with modelled estimates of prevalence. Comparisons of modelled disease prevalence and QOF registers at LAD level suggest that while QOF detects 92% of diabetes cases predicted by the prevalence model, this figure is reduced to 69% for hypertension (Martin and Wright, 2009), confirming earlier work (Standing et al., 2005). The potential to use QOF registers as measures of prevalence is discussed in the relevant sections.

6.4.1 Diabetes mellitus

Diabetes is a chronic condition in which the blood sugar levels are high. Diabetes can be caused by too little insulin (a hormone which controls blood sugar levels), resistance to insulin or both. Type 1 diabetes is associated with low production of insulin and Type 2 diabetes, in which the body’s cells become resistant to insulin, which is much more common (NHS, 2012). Diabetes is a known risk factor for CHD, although the precise pathophysiology of the link is complex (Paneni et al., 2013) and there is debate about the precise quantification of the risk (Beckman et al., 2013).

Diabetes QOF prevalence register

All practices in the study have received maximum points for producing a register of patients aged 17 and over with diabetes; QOF prevalence registers include patients with Type 1 and Type 2 diabetes without distinction. QOF does not specify the method of diagnosis but practices are encouraged to use a ‘systematic approach’ (BMA NHS Employers, 2006, p. 98).

Summary statistics for the three year period (see Table 6.3) suggest a gradual upward trend in diabetes prevalence. This may be due to increased prevalence, improved diagnosis or more complete recording.

Modelled estimates of diabetes prevalence

APHO have developed the APHO Diabetes Prevalence Model (DPM); this has been used to provide estimates diabetes prevalence in people aged 16 years and older for
<table>
<thead>
<tr>
<th>QOF 06/07</th>
<th>Median (IQR)</th>
<th>(Min, Max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>QOF 07/08</td>
<td>3.9 (3.5, 4.6)</td>
<td>(0.4, 10.3)</td>
</tr>
<tr>
<td>QOF 08/09</td>
<td>5.2 (4.6, 5.9)</td>
<td>(0.5, 13.1)</td>
</tr>
</tbody>
</table>

Table 6.3: Summary statistics for diabetes prevalence (%) measured by QOF (practices for which there is data for all three years $n = 228$).

England, SHAs and PCTs (Holman, 2012). The models take age, sex, ethnicity, deprivation and obesity. Synthetic estimates are not published at practice level but a tool, in the form of an Excel spreadsheet, is published by APHO which enables users to enter information about populations and estimates are then calculated. The model is based on the population aged 16 and over and has been developed to include undiagnosed cases, estimated to be 26.4% in men and 21.7% in women (Holman, 2012).

Comparing QOF registers and modelled estimates of diabetes prevalence

The estimate of diabetes prevalence, based on the DPM, for the East Midlands SHA in 2009 was 7.3% (YHPHO, 2010); this compares to 5.4% based on aggregated QOF registers for 2008/09. Dividing the number of people on general practice diabetes registers by the number of people estimated to have diabetes suggests that 74% of estimated diabetes cases are diagnosed and therefore on QOF registers in the East Midlands. This equates to 26% of diabetes cases not included on general practice diabetes registers. This is similar to estimates of 27.1% for England as a whole in a similar analysis (Holman et al., 2011).

Practice level estimates can be estimated from the diabetes prevalence model (DPM), although in small populations there is a high level of uncertainty. Information regarding age, sex and ethnicity of the practice population and the socio-economic deprivation of the local area need to be entered into the tool. Socio-economic data are in the format of the population in each local authority district living within each quintile of IMD-2010, which is not routinely published, but can be calculated based on the LSOAs within each LAD. In addition, the tool requires ethnicity in 10 year age groups; this information was not available. The lack of patient level socioeconomic deprivation and ethnicity by age group means that using the tool to estimate diabetes prevalence in practice populations is subject to error over and above error in the model itself. Relevant data were entered into the APHO tool for a sample of 25 practices in three districts in Northamptonshire, which provided estimates for the 25 practices.
The agreement between estimates provided by the APHO tool and QOF based prevalence are shown in a Bland-Altman Plot (Fig. 6.1). Bland-Altman plots allow two measurement approaches to be compared to determine the level of agreement between the two measures. The plots are a plot of the difference between the two methods against their mean; the mean level of agreement and the 95% limits of agreement are included to help interpret the difference (Bland and Altman, 1986).

![Bland-Altman Plot](image)

Figure 6.1: Bland-Altman Plot showing agreement between QOF based estimates and estimates derived from the APHO model for 25 practices. △ Corby ● South Northamptonshire ○ Wellingborough.

The difference between the estimated prevalence and the QOF based prevalence varies from 1.4% to 4.8%, with a mean difference of 3.1%. This suggests that between 42.5% and 79.5% of estimated cases are accounted for by QOF registers. There is no evidence that the difference is associated with the mean. There is some evidence that in South Northamptonshire the difference between the modelled estimates and QOF prevalence is lower than the average and the difference is higher in Corby.

**Conclusion**

Whilst there are some concerns that QOF prevalence registers are not reflecting the actual diabetes prevalence, the lack of an alternative, valid measure of diabetes prevalence for practice populations means that QOF diabetes prevalence registers are a useful source of data.
6.4.2 Obesity

Obesity is known to be associated with CHD (Lavie et al., 2009). Body Mass Index (BMI) is the most commonly used measure of obesity, calculated by dividing body weight in kg by the height in metres squared. People with BMIs between 25 and 29.9 kg/m\(^2\) are considered overweight, and those over 30.0 kg/m\(^2\) are obese (NHS, 2014b), although there is evidence that BMI has to be over 35 before an association with increased risk of CHD is seen. Waist circumference and waist to hip circumference are also collected and used to measure ‘central obesity’ (Janssen et al., 2004).

**Obesity QOF prevalence register**

All practices in the area of study achieved maximum points for producing a register of patients aged 16 and over who are considered obese (BMI 30 and over), based on QOF guidelines. QOF obesity prevalence has increased from 7.38% in 2006/07 to 9.52% in 2008/09 and further increased to 10.0% in 2011/12. This is likely to reflect improved case seeking rather than an increase in underlying prevalence as it is still considerably lower than obesity prevalence based on the HSE (HSCIC, 2013).

**Modelled estimates of obesity prevalence**

Modelled estimates of obesity prevalence in MSOAs in England are available, based on data from the HSE 2003-05 and 2006-08. Relationships seen at national level between obesity prevalence and ethnicity (percentage black origin), percentage male and aspects of socio-economic deprivation were used to develop the model (APHO, 2010).

**Comparing QOF registers and modelled estimates of obesity prevalence**

Nationally the HSE 2007 suggested that 24% of the English population aged over 16 have a BMI of 30 or over, compared to 7.4% indicated by the 2007/08 QOF registers.

It is possible to match individual practices with MSOAs using the postcode of the practice. However, the population of an individual MSOA will not be the same as the practice population. QOF registers indicate a much lower prevalence than estimates based on the postcode of the practice, median of 7.33% compared to 25.10% (95% limits of agreement are 9.93 to 23.63). Figure 6.2 clearly shows the difference in distribution.
and levels of obesity prevalence indicated by QOF registers and estimates based on the HSE.

Figure 6.2: Boxplots showing distribution of obesity prevalence as measured by estimates (APHO, 2010) and QOF registers.

Conclusion

There is considerable evidence that QOF registers under-estimate the levels of obesity in practice populations. Modelled estimates are only available for MSOAs and cannot be regarded as valid measures of obesity prevalence in practice populations. In addition, a BMI of over 30 may not be the most appropriate measure of obesity to include in studies of CHD mortality (Janssen et al., 2004; Lavie et al., 2009).

Including estimates based on MSOAs in modelling CHD mortality may be possible using multilevel models which are considered in Chapter 14.

6.4.3 Hypertension

Hypertension is sustained high blood pressure. QOF guidelines for 2006/07 suggest that elevated blood pressure readings of greater than 140/90 on three separate occasions are generally taken to confirm sustained high blood pressure (BMA NHS Employers, 2006). (140/90 is a lower systolic pressure than the target described in later chapters).
Hypertension is a known risk factor for CHD as it leads to damage of the lining of blood vessels (Remington et al., 2010).

**QOF hypertension register**

All practices in the study have received maximum points for producing a register of patients with hypertension. QOF based hypertension prevalence is based on the percentage of the total practice population who have hypertension and not merely the adult population. QOF prevalence registers suggest that 13.4% of people across the three PCTs have diagnosed hypertension. The median practice based prevalence is 12.3% (IQR: 11.04% to 14.69%).

**Modelled estimates of hypertension prevalence**

Practice level, modelled estimates of hypertension prevalence are available from APHO (APHO, 2009). The model, based on the HSE 2003 and 2004, takes into account the following factors:

- **Age** Hypertension prevalence is associated with increasing age;
- **Gender** Men generally have a higher risk;
- **Interaction term - age and gender** Hypertension prevalence in females is higher than males as age increases;
- **Socio-economic deprivation** Increasing levels of deprivation, as measured through the quintiles of IMD-2004 is associated with increased risk;
- **Ethnicity** People of Black or Black British ethnicity show a higher risk (Walford et al., 2008).

**Comparing QOF registers and modelled estimates of hypertension prevalence**

The estimate of prevalence across the three PCTs is 23.2%, just under 10% higher than diagnosed hypertension based on QOF registers, suggesting that only 58% of estimated hypertension cases are diagnosed and included on QOF registers. Fig 6.3 illustrates the differences between prevalence indicated by QOF registers in 2007/08.
and the synthetic estimates for the adult population. Lin’s concordance correlation coefficient is 0.13 indicating poor agreement between the two measures (Lin, 1989). The 95% limits of agreement are 10.74 to 21.25 with a mean difference of 16.0%. The Bland-Altman plot (Fig.6.4) suggests that there may be some pattern in the differences, with, perhaps, a slight increase in difference as the average increases.

![Bland-Altman plot](image)

Figure 6.3: Association between hypertension prevalence indicated by QOF and synthetic estimates from APHO model, (practices in LNR $n = 229$).

Differences between QOF registers and estimates may be due to the method of diagnosis. The HSE relies on a one-off measure of blood pressure and is not a measure of sustained high blood pressure, which would need to be present before GPs would record patients as having hypertension. This could lead to estimates of hypertension prevalence being higher than QOF registers. Conversely, the APHO technical briefing (Walford et al., 2011) suggests that the syndrome of ‘white coat hypertension’ can cause patients to have high blood pressure measurements in a general practice setting, leading to QOF registers suggesting higher prevalence than alternative methods of diagnosis.

Walford (2011) points out that discrepancies between modelled estimates and other sources of data may be due to local variations not captured by the model and that if
Figure 6.4: Bland-Altman Plot showing agreement between hypertension prevalence indicated by QOF and synthetic estimates from APHO model, (practices in LNR \( n = 229 \)).

A practice population differs significantly from a ‘typical’ population the assumptions of the model may not apply. For example, one assumption is that smoking prevalence is the same across all ethnic groups, but if this does not hold for a particular practice population the synthetic estimates are not likely to be accurate.

**Conclusion**

The model used to estimate hypertension includes key variables which are risk factors for CHD and will be considered in their own right, for example age and deprivation, this suggests that including estimates of hypertension prevalence may not add to any analysis. QOF hypertension registers reflect detection of hypertension and are not an accurate source of information describing the underlying prevalence in the population.

**6.4.4 Serum cholesterol**

Serum cholesterol and other lipid serum measurements are risk factors for CHD, directly increasing the risk of atherosclerosis (Stamler et al., 1986), and including these measure measurement in patient level studies is relatively common. However, there are no QOF indicators which can be used to estimate the underlying prevalence within
practice populations, as only patients who have been diagnosed with conditions such as diabetes and CHD have their serum cholesterol measured as part of QOF. There are no models designed to estimate serum cholesterol within small populations. The introduction of the health check programme in 2009 means large numbers of people have had their serum cholesterol checked - over three million people since 2011 (NHS, 2014a). However, there is no clear evidence that data from health checks is being collated and used to estimate the prevalence of high levels of serum cholesterol in England. In addition to serum cholesterol there are other biomarkers, such as C-reactive protein and fibrinogen, which are associated with increased risk of cardiovascular events (Kaptoge et al., 2012). There are no practice level measures of these biomarkers.

6.4.5 CHD prevalence

A risk factor which has been included in some studies, but is not being considered here, is CHD prevalence. In studies of unplanned admissions, CHD prevalence was considered as a key explanatory factor (Bottle et al., 2008), heart failure and CHD prevalence were considered important in explaining unplanned admissions for heart failure (Brettell et al., 2013) and stroke prevalence for unplanned stroke admissions (Soljak et al., 2011). Bankart et al. (2011) considered the use of CHD prevalence as a measure of overall morbidity in a study considering unplanned hospital admissions, but found it was too highly correlated with other variables in the model (mainly the percentage aged 65 and over) to include.

This thesis focusses on the role primary care can play in modifying the predictive effect of population characteristics on CHD mortality, with a focus on primary prevention. CHD prevalence is directly associated with CHD mortality, regardless of the effectiveness of primary care. Therefore, CHD prevalence would be a strong predictor of CHD mortality, without improving our understanding of the role of primary care in reducing CHD mortality. In addition, CHD prevalence is strongly associated with population characteristics which introduces issues of collinearity into the model. In conclusion CHD prevalence is not considered a useful predictor to include in this analysis.

6.5 Key health behaviours

Key health behaviours identified in Chapter 1 (Table 1.1) for inclusion in the study are smoking; poor nutrition; physical (in)activity; and alcohol consumption. These
factors have been identified as important in contributing to the risk of developing a range of chronic diseases, including various cancers, COPD, cardiovascular diseases and diabetes. Despite the demonstrated importance of these risk factors, there are relatively few targets within QOF relating to these areas.

6.5.1 Smoking

Smoking is a known risk factor for the development of a range of diseases, including CHD, and is considered one of the key causes of premature death (McCarthy, 2014). Nationally the Smoking Toolkit Survey is a valuable source of information on national smoking prevalence and trends. The HSE collects data on smoking which has allowed estimates of smoking prevalence in MSOAs. In addition, from 2009, the Integrated Household Survey (IHS) includes questions on smoking habits. These statistics are designated as experimental, in a testing phase and not yet fully developed (ONS, 2011), but estimates are publicly available which describe smoking prevalence in MSOAs and information is available for LSOAs for approved research. None of these surveys aims to establish the smoking prevalence within practice populations. Some researchers have matched patient postcodes to MSOAs and used the smoking prevalence in each LSOA and the proportion of patients in the practice population resident in each LSOA to calculate a weighted average to estimate smoking prevalence in the practice population (Soljak, 2013). Individual patient postcodes were not available for this study and this method could therefore not be implemented.

During the course of this analysis QOF data for 2012/13 were published, this enabled smoking prevalence estimates to be determined for practice populations, this is discussed in more detail in Chapter 15.

Conclusion

Initial models will not include a measure of smoking prevalence; the potential of using local area estimates is discussed in Chapter 14. The use of QOF smoking data from 2012/13 is described in Chapter 15.
6.5.2 Poor nutrition, physical (in)activity and alcohol consumption

Although there are various surveys which aim to measure important lifestyle factors such as nutrition, physical (in)activity and alcohol consumption, none are available which describe these factors at practice level. The HSE includes questions about the consumption of fruit and vegetables, physical (in)activity and alcohol consumption and modelled estimates are published at MSOA level (HSCIC, 2013). The Active People Survey (APS), commissioned by Sport England, is a large telephone survey, covering over 150,000 people in 2012. Data are available for LADs, including the percentage of people who participate in activity for less than 30 minutes a week, 30-89, 90-149 and 150 minutes or more, although in earlier years only participation in at least 30 minutes is reported (Sport England, 2014). In the future, health checks may provide self-reported information on these important factors.

6.6 Environmental condition

It is common for health outcomes to be examined in relation to socioeconomic status; in the UK this is generally thought of in terms of social class as revealed in employment status and working conditions. Research has also shown that the role of the neighbourhood environment is also important in understanding inequalities in health (Sundquist et al., 2004). Indices including multiple markers of socioeconomic deprivation have been developed over the last 20 years. The Townsend Index was one of the first examples, and included car ownership, unemployment, overcrowded housing and housing tenure. An alternative measure, the Carstairs Index, was developed for Scotland (Smith, 2003). Since the development of the Index of Multiple Deprivation by the Department for Communities and Local Government in 2000 it has become the most commonly used marker of deprivation in UK studies.

In addition to inclusion of an indicator of socioeconomic deprivation some studies (for example Erskine et al. (2010)) include a classification to distinguish between urban and rural areas.
6.6.1 The 2007 Index of Multiple Deprivation

The Index of Multiple Deprivation (IMD) is a measure of multiple deprivation, made up of 38 indicators across seven domains, and is published at various geographical levels, the smallest level being LSOAs (Communities and Local Government, 2011). Information used to compile the IMD comes from the most recent census, from various government departments and from NHS data. A single Index of Multiple Deprivation was first published in 2000 and most recently in 2010; the underlying principles behind the index have remained the same over the period of publication, although there have been modifications to the smallest geography and the domains.

The mortality data being analysed in this study cover the period 2006/07 to 2008/09. Using IMD-2007 as an indicator of socioeconomic deprivation for the general practices gives an indication of the deprivation experienced by patients in the period leading up to death, and is more likely to reflect the deprivation of patients being treated by practices during that period. However, it is acknowledged that the relationship between deprivation and health is not a simple one. Identifying which periods of a patient’s life are most sensitive to the impact of deprivation would be a valuable area of research, but beyond the scope of this thesis.

IMD 2007 is based on seven domains combined with specific weightings to produce the final Index of Multiple Deprivation (Table 6.4).

<table>
<thead>
<tr>
<th>Domain</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income deprivation</td>
<td>22.5</td>
</tr>
<tr>
<td>Employment deprivation</td>
<td>22.5</td>
</tr>
<tr>
<td>Health deprivation and disability</td>
<td>13.5</td>
</tr>
<tr>
<td>Education, skills and training deprivation</td>
<td>13.5</td>
</tr>
<tr>
<td>Barriers to housing and services</td>
<td>9.3</td>
</tr>
<tr>
<td>Crime</td>
<td>9.3</td>
</tr>
<tr>
<td>Living environment deprivation</td>
<td>9.3</td>
</tr>
</tbody>
</table>

Table 6.4: Weightings of each domain in the final index

6.6.2 Summarising the index score and rank for a practice population.

General practices do not serve particular geographical areas. Patients within one LSOA are able to choose between different general practices. There are two commonly used
approaches for determining an IMD score for a general practice.

- Individual registered patients’ residential postcodes can be matched to an LSOA and a weighted average of the LSOA linked deprivation scores calculated. This approach is termed the ‘gold standard’ by Strong et al (2006).
- Identify the LSOA in which the practice is located and use the IMD score and rank for that LSOA as a measure for the practice population.

Indices based on the second approach have been shown to be highly correlated with the ‘gold standard’ (Strong et al., 2006). Griffin et al (2010) suggest that matching the postcode of the practice to the MSOA gives results closer to the ‘gold standard’. However, there is evidence that approaches based on the practice postcode may result in the relationship between deprivation and ill health being under-estimated (McLean et al., 2008). In addition Strong et al (Strong et al., 2007) have developed a method using Geographical Information Systems (GIS) which they consider to be the most reliable measure if the ‘gold standard’ is not available.

Methods and Results

‘Gold standard’ deprivation scores, based on the postcodes of individual patients have been provided by the PCTs for each practice. In order to determine the correlation between practice postcode based scores and ‘gold standard’ scores in this sample of practices, practice postcodes were matched to LSOA and relevant IMD scores (2007 and 2010) were derived for each practice. The scores based on different approaches were compared; summary statistics are shown in Table 6.5. There is a strong correlation between the ‘gold standard’ and the practice postcode IMD scores ($R_p = 0.82$, 95% CI [0.77, 0.86]); a slightly stronger correlation than found by Strong et al (2006) ($R_p = 0.74$, 95% CI [0.54, 0.85]).

<table>
<thead>
<tr>
<th></th>
<th>Median (IQR)</th>
<th>(Min, Max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Gold’ 2007</td>
<td>16.17 (9.99, 27.01)</td>
<td>(5.10, 60.28)</td>
</tr>
<tr>
<td>Practice postcode 2007</td>
<td>17.69 (9.53, 32.24)</td>
<td>(2.33, 71.74)</td>
</tr>
</tbody>
</table>

Table 6.5: Summary statistics for practice IMD scores based on the ‘gold standard’ or the practice postcode method (practices for which there is data for all four measures $n = 222$).
Conclusions

As ‘gold standard’ scores are available for this study, these will be used.

An issue with the IMD-07 scores

The ‘gold standard’ scores were supplied by the PCTs with additional practice information, including the postcode of the practice. Careful examination of the postcodes led to the identification of 62 practices for which the postcodes given in this spreadsheet were different from the practice postcodes given by Connecting for Health (NHSBSA, 2012). The latter postcodes were the same as postcodes given on practice websites.

The concordance between the ‘gold standard’ scores for 2007 and 2010 is high, \(\rho_c = 0.99, p < 0.0001\) and the average difference is -0.48 (95% limits of agreement: -3.81, 2.85). A visual inspection of the correlation between the scores for 2007 and 2010 for practices with and without matching postcodes does not suggest any concern (Fig. 6.5); mismatch is likely to have arisen after the scores had been allocated to practices. Hence, PCT-provided ‘gold standard’ IMD scores based on the IMD-2007 will be used. Information including scores and ranks for all individual domains is available for all practices for which there is mortality information.

![Figure 6.5: Correlation between ‘gold standard’ IMD scores for practices with postcodes from two different sources matching and those which don’t (practices in LNR for which both measures are present \(n = 222\)).](image-url)
6.6.3 Is the IMD score a valid measure of deprivation for health studies?

One of the seven domains of the Index was developed to identify areas with relatively high rates of people who die prematurely or whose quality of life is impaired by poor health or who are disabled. This domain, the ‘health deprivation and disability’ domain, contributes 13.5\% to the overall index and is summarised in Table 6.6. A strong correlation between IMD and mortality could, in part, be due to measures of mortality being part of the index. A summary of an investigation into the relationship between the seven domains and the impact of excluding health from the overall score is presented below.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Source</th>
<th>Calendar period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of Potential Life Lost (YPLL)</td>
<td>Office for National Statistics</td>
<td>2001 to 2005</td>
</tr>
<tr>
<td>Comparative Illness and Disability Ratio (CIDR)</td>
<td>Department for Work and Pensions</td>
<td>2005</td>
</tr>
<tr>
<td>Measures of acute morbidity</td>
<td>Derived from Hospital Episode Statistics</td>
<td>2004 to 2005</td>
</tr>
<tr>
<td>The proportion of adults under 60 suffering from mood or anxiety disorders based on prescribing</td>
<td>Sources include Prescribing Pricing Authority, Hospital Episode Statistics and Incapacity Benefit data</td>
<td>2004-2005</td>
</tr>
</tbody>
</table>

Table 6.6: Indicators included in the health domain of IMD-2007 (Noble et al., 2008).

**Methods and Results**

The correlation plots of all domain scores in practices across LNR, shown in Figure 6.6, shows a very strong positive correlation between the overall deprivation and individual domains scores for practices. The correlation between the score based on the health domain and the overall domain is high (Pearson $R_p = 0.88, p < 0.0001$, increasing to $R_s = 0.98, p < 0.0001$ if rank of the practices is used). The housing domain is the only domain which does not show a strong correlation with the overall score.
Indices with and without health were calculated, ensuring that overall weightings of each of the six domains, without health, summed to 100, as recommended by Rose (2012). There is very high agreement between ‘indices’ with and without health (Bland Altman’s 95% limits of agreement (-1.932, 1.129)); Figure 6.7 shows high concordance between the estimates. Hence, using the complete IMD score here, as in the majority of related research, does not cause concern.

### 6.7 Psychosocial wellbeing

There is a growing body of evidence that psychosocial wellbeing, both long term and short term, is associated with adverse health (Mostofsky et al., 2014; Toren et al., 2014)

Whilst these may be important factors, with primary care potentially playing a role in modifying the predictive events, reliable data to describe these factors at population level are extremely limited. For example, although QOF registers allow the determination of the prevalence of patients with mental health conditions, these focus on schizophrenia, bipolar disorder and other psychoses (2006/07) and so exclude depres-
The depression domain includes two indicators which requests that case finding for depression has been undertaken for people with CHD and/or diabetes, not the population as a whole. In addition patients with a new diagnosis should have further assessment. These indicators do not allow the prevalence of depression to be determined. Whilst acknowledging that psychosocial wellbeing may be an important modifiable predictor of CHD it is not included in this analysis.

6.8 The ‘nursing home effect’

Special local factors may be important for particular practices, for example the proportion of registered patients who are resident in nursing homes, which is described here. Various research has shown that practices with a high proportion of ‘nursing home’ patients have higher than expected mortality rates. Mortality rates have been shown to be higher for those living in nursing homes than for people of the same age living at home (Nimmo et al., 2006) and this may have an impact on practice level mortality rates. This has been termed the ‘nursing home effect’. Mohammed et al (2004) investigated general practitioners in the West Midlands who were associated with an unacceptable high mortality of patients over an eight year period. A high
correlation was found between excess mortality and deaths in nursing homes and when one practice relinquished a nursing home from their patient list there was a subsequent decrease in mortality. A similar study of five GPs in West Sussex who had been identified as having excessively high mortality rates found that there was a high correlation of deaths occurring in nursing homes and the SMR for each GP’s patient population (Billett et al., 2005). A study of life expectancy in local areas also found a significant inverse relation between life expectancy and the proportion of nursing home deaths (Williams et al., 2004). In addition, the association between deprivation and mortality was found to be less clear when there were high proportions of nursing home residents. Information about the percentage of patients resident in nursing home was not available for this study. However, it is a potential area of interest in studies of mortality.

6.9 Conclusions

This chapter has identified and evaluated data which could be used to measure practice population characteristics which are known risk factors for CHD. There is no information available describing the underlying genetic make up of practice populations or family history of CHD. Information about age, sex and ethnicity is reliable, although information about specific ethnic minority groups may not be. Determining whether QOF prevalence registers are useful measures of prevalence or ‘detection’ is problematic as true underlying prevalence is not known. There is evidence to suggest that QOF diabetes registers are a useful measure of prevalence, but that this is not the case for hypertension. This is discussed in more detail in Chapters 7 and 16. There is considerable evidence that QOF obesity registers are not reflecting obesity prevalence in practice populations. There are no data available to describe the percentage of patients who have high levels of serum cholesterol. Collation of data from the Health Checks program may provide this in future years.

Reliable information about smoking rates, alcohol consumption, physical (in)activity and nutrition within practice populations is lacking. Various estimates are available, but many of these are modelled estimates for geographical areas, not practice populations, based on levels of deprivation, ethnicity and age. The publication of new QOF smoking indicators in October 2013 has led to a novel approach of estimating smoking prevalence in practice populations (see Chapter 15 and (Honeyford et al., 2014) for more details).
The Index of Multiple Deprivation is a useful measure of socio-economic deprivation and data supplied for this study are of a high quality. Other factors which may be of interest but for which there are no useful data available for this study include psychosocial wellbeing and the number of patients who are resident at nursing homes. None of the data are available by age group which would be of interest in this analysis as premature mortality is the key health outcome.

Chapter 9 summarises the candidate explanatory variables, available data and the process of variable selection used in this thesis. In the following chapters the data available to describe primary care are identified and evaluated.
Chapter 7

Indicators of quality of primary care - prevention, detection and clinical management.

7.1 Introduction

In the previous chapter (Chapter 6) data describing the characteristics of practice populations have been identified and evaluated. In this chapter, data describing the quality of primary care in terms of prevention, detection and clinical management are identified and evaluated. Measuring the quality of primary care through the use of routinely available data is not straightforward, in that the data may not reflect the experience of the patient or be the most useful measure. Prescribing data, the presence of health promotion clinics and the provision of specialists have been used as measures of primary care (Barnett et al., 2000; Saxena et al., 2006; Millett et al., 2009). Since the introduction of QOF, it has become common practice to use information from QOF as a way of assessing the quality of primary care delivered by general practices, or aggregated to give information about Primary Care Trusts (PCTs) (Bottle et al., 2008; Dusheiko et al., 2011; Bankart et al., 2011).
7.2 QOF - an introduction

QOF is made up of four domains (clinical; organisational; patient experience and additional), and is a rich source of publicly available information about the delivery of primary care in England. Within the clinical domain are several disease areas, each made up of individual indicators. For each indicator, points are awarded to practices based on the percentage of patients for whom the target has been achieved, this is known as the underlying achievement. Both the score and the underlying achievement for each indicator are publicly available. In 2006/07 there were 80 indicators within the clinical domain covering 19 disease areas and worth 655 points out of the 1000 point total.

The way in which QOF has been used in health services research varies from study to study. Total scores for the clinical domain have been used in studies of emergency admissions (Purdy et al., 2011; Bankart et al., 2011) and elective admissions (Chauhan et al., 2012). Combinations of scores for an individual disease area as well as individual indicators (Bottle et al., 2008) were used in a study of hospitalisation due to CHD. Levene et al (2012) used the underlying achievement of individual indicators whilst Kiran et al (2010) used combinations of individual indicators underlying achievement to give a devised CHD quality achievement score. In addition, the underlying achievement of individual indicators within the patient experience domain have been used as measures of access and sustained relationships (Levene et al., 2012; Chauhan et al., 2012; Bankart et al., 2011).

QOF prevalence registers have also been used as a measure of quality of primary care. The number of people on hypertension registers held by general practices as part of QOF has been used as a measure of hypertension detection and been shown to be associated with decreases in CHD mortality (Levene et al., 2012).

Another variation lies in the choice of year for recorded health outcomes and QOF data. It is most common for the period of health outcomes to be contemporaneous with the QOF data. However, Kiran et al (2010) included QOF information from 2006/07 in a study of mortality between 2004 and 2007, meaning not all patients experienced the primary care directly assessed by the QOF scores.

The different ways in which QOF data have been used may start to explain some of the differences in results found in different studies. For example, Kiran et al (2010) found that the overall CHD quality achievement score was significantly negatively associated with CHD mortality. A higher overall clinical QOF score was found to be negatively
associated with the risk of admission for angina (Purdy et al., 2011). However, other studies have shown no association between individual indicators and CHD mortality (Levene et al., 2012) and CHD admissions (Bottle et al., 2008) or overall scores and emergency and elective admissions (Bankart et al., 2011; Chauhan et al., 2012).

### 7.3 Selection of indicators

Given the multitude of indicators available to select from, it is important to have a clear strategy for prioritising indicators which may be relevant for the health outcome of interest, even if some of these indicators are not included in a statistical model for other reasons, such as collinearity. In this thesis, the role of primary care in reducing CHD mortality has been divided into aspects relating to a) access and relationships and b) clinical care. Within clinical care the multitude of services and roles primary care has are divided into i) prevention, ii) detection and iii) clinical management. It is therefore relevant to include indicators from these three areas as well as indicators representing the relevant risk factors for CHD. Evaluating indicators on these grounds provides a clear strategy for identifying relevant individual indicators, and using previous research adds to this strategy.

### 7.4 Which QOF data are the most useful in describing the quality of primary care?

Within the clinical domain, the total points allocated to each practice has been determined for six disease areas. The percentage of practices achieving maximum points ranges from 42.0% for the diabetes domain in 2006/7 to 92.5% for the heart failure domain in 2008/09. Over 60% of practices achieved maximum points in CHD, stroke, blood pressure, heart failure and atrial fibrillation (AF). For diabetes the percentage of practices achieving maximum points has increased from 42.0% in 2006/07 to 48.7% in 2008/09 (see Table 7.1 and Fig 7.1).
<table>
<thead>
<tr>
<th>Domain</th>
<th>2006/07</th>
<th>2007/08</th>
<th>2008/09</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD</td>
<td>73.6</td>
<td>73.7</td>
<td>74.8</td>
</tr>
<tr>
<td>Diabetes</td>
<td>42.0</td>
<td>47.0</td>
<td>48.7</td>
</tr>
<tr>
<td>Stroke</td>
<td>65.8</td>
<td>73.7</td>
<td>72.8</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>80.1</td>
<td>78.7</td>
<td>78.9</td>
</tr>
<tr>
<td>Heart failure</td>
<td>88.7</td>
<td>86.5</td>
<td>91.2</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>90.0</td>
<td>90.0</td>
<td>92.5</td>
</tr>
</tbody>
</table>

Table 7.1: Percentage of practices in LNR achieving maximum points in six QOF clinical domains.

Figure 7.1: Distribution of points achieved in six clinical domains for practices in LNR, 2006/07 (n = 229).
In addition to the publication of points the underlying achievement for each indicator is published. That is, the actual percentage of patients in the practice (or on the relevant register) who have met the target for each indicator is published and/or the data used to calculate the percentage. For example, for an indicator relating to the number of patients on the CHD register whose last blood pressure reading is at or below the target of 150/90, both the number of people with CHD (the denominator) and the number who have had a blood pressure reading of 150/90 or less is published for each practice. For 2008/09 the ‘underlying achievement’ is published as a percentage, but for 2006/07 and 2007/08 the numerator and denominator are published and the ‘underlying achievement’ can be calculated. Either underlying achievement for each individual indicator or mean underlying achievement for domains can be included in analysis. If a practice has achieved the target for all relevant patients for all indicators in one domain they would receive an overall underlying achievement of 100%.

Mean underlying achievement was calculated for each practice for each of the six domains included in Table 7.1. For CHD, diabetes, stroke and blood pressure 1% or fewer of the practices achieved 100%. For atrial fibrillation and heart failure the percentage was much higher. Fig 7.2 shows the distribution of mean underlying achievement. Underlying achievement more closely describes the clinical care received by patients and shows more variation between practices and is therefore a more useful measure of quality of care than simply taking points achieved.
Figure 7.2: Distribution of mean underlying achievement achieved in six clinical domains for practices in LNR, 2006/07 ($n = 229$).
7.5 General form of indicators

Although indicators are varied, some indicators relate to recording and some relate to the achievement of targets for patients. For example, within the clinical disease area there is an indicator which relates to all patients on the CHD register having a record of their blood pressure (CHD05) and a second which sets a target for blood pressure for those patients (CHD06). Achievement is generally higher for indicators relating to measurement than recording; the median achievement for recording blood pressure is 99% but for achieving the blood pressure target the median achievement is 90%. Table 7.2 has more detail and a second example.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Indicator details</th>
<th>Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD 05</td>
<td>The percentage of patients with coronary heart disease whose notes have a record of blood pressure in the previous 15 months.</td>
<td>98.73 (97.25, 99.54)</td>
</tr>
<tr>
<td>CHD 06</td>
<td>The percentage of patients with coronary heart disease in whom the last blood pressure reading (measured in the previous 15 months) is 150/90 or less.</td>
<td>89.87 (85.22, 93.22)</td>
</tr>
<tr>
<td>CHD07</td>
<td>The percentage of patients with coronary heart disease whose notes have a record of total cholesterol in the previous 15 months.</td>
<td>95.21 (92.96, 97.44)</td>
</tr>
<tr>
<td>CHD08</td>
<td>The percentage of patients with coronary heart disease whose notes have a last measured cholesterol (measured in the previous 15 months) is 5 mmol/l or less.</td>
<td>82.45 (77.59, 86.58)</td>
</tr>
</tbody>
</table>

Table 7.2: Underlying achievement of recording and target achievement indicators, 2006/07 (n=229).

7.6 Prevention of CHD

This section considers the role of primary care in reducing the prevalence of modifiable biological risk factors and modifiable health behaviours, as these are associated with CHD development.

7.6.1 Modifiable biological risk factors

Relevant modifiable biological risk factors have been previously identified as hypertension, serum cholesterol, diabetes and obesity. In 2006/07-2008/09 there were no QOF indicators within the clinical domain relating to the prevention of these risk factors. QOF indicators are nearly all based on achieving management for people who have already been diagnosed with chronic conditions.
Detection of biological risk factors has been shown to be associated with lower levels of CHD mortality (Levene et al., 2010; Levene et al., 2012); however measurement of detection is not straightforward. QOF prevalence registers have been used as measures of detection (Anwar et al., 2012; Levene et al., 2012). QOF based detection will depend on both the performance of primary care and the actual underlying prevalence in the practice population. An alternative measure is to consider undiagnosed prevalence which would reflect both detection and underlying prevalence. This approach has been used in studies of stroke and CHD outcomes (Purdy et al., 2011; Soljak et al., 2011). True underlying prevalence is unavailable; therefore modelled estimates of prevalence have to be used, which may not accurately reflect prevalence for all practices, particularly when the practice population differs significantly from a ‘typical’ population (Walford et al., 2011).

Modelled estimates for practice populations were only available for hypertension for the relevant period of time; modelled estimates for obesity are available at local area level. These have been discussed in Chapter 6; hypertension will be discussed in more detail in Chapter 16. QOF includes a range of indicators considering the clinical management of hypertension, diabetes and serum cholesterol, and these are described below.

**Hypertension**

Hypertension indicators occur within the blood pressure area within the clinical domain; summarised in Table 7.3.

<table>
<thead>
<tr>
<th>Indicator code</th>
<th>Indicator details</th>
<th>Median underlying achievement (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP 01</td>
<td>The practice can produce a register of patients with established hypertension.</td>
<td>n/a</td>
</tr>
<tr>
<td>BP 04</td>
<td>The percentage of patients with hypertension in whom there is a record of the blood pressure in the previous 9 months.</td>
<td>92.8 (90.7, 94.9)</td>
</tr>
<tr>
<td>BP 05</td>
<td>The percentage of patients with hypertension in whom the last blood pressure (measured in the previous 9 months) is 150/90 or less.</td>
<td>77.5 (73.1, 82.8)</td>
</tr>
</tbody>
</table>

Table 7.3: Summary of indicators within the blood pressure area indicators; 2006/07 (n=229).

Similar indicators exist applying to those on the CHD, stroke, diabetes and chronic kidney disease (CKD) registers, although the target blood pressure is 145/85 for CKD and
diabetes. Underlying achievement for recording of blood pressure is high, particularly for patients diagnosed with chronic conditions; underlying achievement for achieving blood pressure targets is lower, particularly for those diagnosed with hypertension and diabetes. This is summarised in Table 7.4.

<table>
<thead>
<tr>
<th>Indicator type:</th>
<th>Patients with...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Blood pressure</td>
</tr>
<tr>
<td>Record of blood pressure</td>
<td>93 (91, 95)</td>
</tr>
<tr>
<td>Achieving target</td>
<td>77 (73, 83)</td>
</tr>
</tbody>
</table>

Table 7.4: Median underlying achievement (%) and interquartile range for blood pressure indicators for four disease areas (n=229)

In addition there are indicators within the organisational domain which require practices to have blood pressure recordings for patients over the age of 45. Underlying achievement for these indicators is published; it is therefore possible to determine the percentage of patients for whom there is blood pressure recording in the previous five years, which may be a useful measure of a practice’s approach to hypertension detection.

**Serum cholesterol**

Indicators relating to the control of serum cholesterol follow a similar pattern and appear in three disease areas: stroke, diabetes and CHD. Patients should have a record of serum cholesterol in the previous 15 months and the last measurement should be 5 mmol/l or less. Again the achievement of recording is higher than achievement of target, which is lowest for patients diagnosed with stroke (Table 7.5).

<table>
<thead>
<tr>
<th>Indicator type:</th>
<th>Patients with...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CHD</td>
</tr>
<tr>
<td>Record of cholesterol</td>
<td>95 (93, 97)</td>
</tr>
<tr>
<td>Achieving target</td>
<td>82 (78, 87)</td>
</tr>
</tbody>
</table>

Table 7.5: Median underlying achievement (%) and interquartile range for cholesterol indicators for three disease areas (n=229).

**Diabetes**

There are 15 ‘management’ indicators within the diabetes disease area. Nine of these relate to the percentage of patients with diabetes who have a record of various tests
or referrals having been completed. Two indicators concern the control of $HbA_{1c}$ (glycated haemoglobin levels). These indicators were considered measures of overall diabetes control by Kiran et al (2010) in their ‘CHD achievement score’ and Dusheiko et al (2011) in a study of emergency admissions for diabetes. A further approach would be to take a mean diabetes achievement score.

**Obesity**

Although there is an obesity indicator, this does not include any requirement regarding management. In addition patients with diabetes should have a record of their BMI (median underlying achievement: 96% (93, 98)).

### 7.6.2 Key health behaviours

Key health behaviours which could be targets for primary care are smoking, poor nutrition, physical (in)activity and alcohol consumption. Despite the recognised importance of the key health behaviours only smoking was included in QOF indicators 2006/07-2008/09. A new indicator under the heading ‘Cardiovascular disease - primary prevention’ in the clinical domain was introduced in 2009/10. The indicator requests that patients with a diagnosis of hypertension are ‘given lifestyle advice in the last 15 months for: increasing physical activity, smoking cessation, safe alcohol consumption and healthy diet’. This has not been considered for this analysis as it relates to the year after the period of study. Smoking indicators for 2006/07 are summarised in Table 7.6.

<table>
<thead>
<tr>
<th>Indicator code</th>
<th>Indicator details</th>
<th>Median underlying achievement (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMOKING 01</td>
<td>The percentage of patients with any or any combination of specific conditions(^1) whose notes record smoking status in the previous 15 months.</td>
<td>96.0 (94.1, 97.6)</td>
</tr>
<tr>
<td>SMOKING 02</td>
<td>The percentage of patients with any or any combination of specific conditions(^1) who smoke whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered within the previous 15 months.</td>
<td>93.6 (91.5, 96.0)</td>
</tr>
</tbody>
</table>

\(^1\) coronary heart disease, stroke or TIA, hypertension, diabetes, COPD, CKD or asthma

Table 7.6: Summary of smoking indicators; 2006/07 (n=229).
SM02 was included in a ‘CHD achievement score’ used by Kiran et al (2010) as a measure of prevention of this key lifestyle factor, although it only applies to patients who have already started smoking and have been diagnosed with a chronic condition. The validity of this indicator has been challenged by Coleman (2010) who suggests that higher levels of achievement may be indicative of better recording of advice rather than differences in advice. Another study has shown that increases in recording are not reflected in patients’ recall of having been offered advice (Szatkowski et al., 2011). Researchers have used referrals to smoking cessation clinics (Levene et al., 2012), published by the NHS Health and Social Care Information Centre (HSCIC) however these are not published at geographical levels relevant for this study (HSCIC, 2009).

### 7.7 Detection of CHD

Modelled prevalence of CHD at practice level are available; termed ‘expected prevalence’ in the the CVD practice profiles (SEPHO, 2013). This has been used to estimate undiagnosed prevalence in studies of CHD admissions (Purdy et al., 2011); a similar approach has been used in a study of stroke admissions (Soljak et al., 2011). Prevalence, detection and undiagnosed prevalence are summarised in Table 7.7. The median undiagnosed prevalence is 1.12%. In contrast, Martin and Wright (2009) found that CHD prevalence as estimated by QOF was higher than modelled prevalence, although the model used to estimate prevalence was different to the one used to produce these estimates.

<table>
<thead>
<tr>
<th></th>
<th>Median (IQR)</th>
<th>(min, max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated CHD prevalence (%)</td>
<td>4.36 (3.88, 4.83)</td>
<td>(0.10, 6.84)</td>
</tr>
<tr>
<td>CHD detected prevalence (QOF registers) (%)</td>
<td>3.26 (2.78, 3.64)</td>
<td>(0.01, 5.08)</td>
</tr>
<tr>
<td>Undiagnosed CHD prevalence (%)</td>
<td>1.12 (0.98, 1.48)</td>
<td>(-0.41, 2.98)</td>
</tr>
</tbody>
</table>

Table 7.7: CHD prevalence, detection and undiagnosed prevalence, practices with all measures (n = 221). %

### 7.8 Clinical management of CHD

In 2006/07 there were eight indicators concerned with the ongoing management of patients with CHD, two relate to recording. These include control of blood pressure,
serum cholesterol, medication with aspirin, beta-blockers and ACE inhibitor (or an angiotensin II antagonist) and influenza immunisation. A mean CHD achievement score or individual indicators based on their role in managing the CHD or the patients to whom they apply could be used.

7.9 Previous research

A variety of studies have been published considering the impact of QOF on various health outcomes, including those linked to cardiovascular outcomes. Levene et al (2010) analysed the association of various features of primary health care with CHD mortality for PCTs in England. Included in their analysis were the underlying achievement for three QOF indicators:

**BP 5** The percentage of patients with hypertension in whom the last blood pressure (measured in the previous 9 months) is 150/90 or less.

**CHD 06** The percentage of patients with CHD in whom the last blood pressure reading (measured in the previous 15 months) is 150/90 or less.

**CHD 08** The percentage of patients with CHD whose last measured total cholesterol (measured in the previous 15 months) is 5mmol/l or less.

In addition Levene et al (2010) included the proportion of total QOF points achieved for CHD, hypertension and diabetes. Clinical performance as reflected in the QOF indicator scores or underlying achievement did not predict mortality in any of the three years that were analysed. In further work Levene et al (2012) selected CHD08 and CHD 09 (the percentage of CHD patients who are taking aspirin (or equivalent)), and both were negatively associated with CHD mortality, although the evidence of an association was weak.

Kiran et al (2010) also analysed the association between quality of primary care and cardiovascular outcomes, with a focus on CHD mortality and morbidity for 1531 general practices in London. Quality indicators were chosen based on ‘four modifiable cardiovascular risk factors’:

- high blood pressure;
- high blood cholesterol level;
• diabetes;
• smoking.

They included five CHD indicators: two stroke indicators; one blood pressure indicator; three diabetes indicators and one smoking indicator. Indicators describing blood pressure and cholesterol were included if they described the proportion of patients for whom a target has been achieved, rather than ‘recorded’. One general diabetes indicator was included (DM20). In addition three CHD indicators relating to medication were included. The mean achievement for each practice was calculated using the 12 quality indicators, with each weighted equally. Kiran et al (2010) also included a smoking indicator describing the percentage of patients with any or any combination of specific conditions who have a record that smoking cessation advice has been offered, although they comment that ‘the (QOF) incentive scheme does not include a quality indicator that reflects objective reduction in smoking’. These indicators are summarised in Table 7.8. There are challenges with the ‘combined’ indicator approach, although Kiran et al (2010) found that higher levels of CHD care were associated with improved health outcomes it is difficult for clinicians to determine which indicators may be most important. In addition, many patients will appear on more than one register and this will affect some indicators more than others. It may be estimated that 75% of the people on the CHD register will also be on the hypertension registers, so including both control of blood pressure in patients on the hypertension register and the CHD register may be counting the same achievement twice.

In an analysis of associations between primary care factors and rates of hospital admissions for CHD, between PCTs, Bottle et al (2008) used total CHD points and points for CHD 06 and CHD 09 (described in Table 7.8). They found no association between QOF points and elective or unplanned hospital admissions.

In a study of emergency stroke admissions, the percentage of stroke patients whose cholesterol and blood pressure were included, as well as influenza immunisation rates, antiplatelet medication and blood pressure control in those on the hypertension register (Soljak et al., 2011). When stroke mortality was considered, control of cholesterol and aspirin medication in those on the stroke register as well as hypertension detection and overall influenza vaccination rates were included as key measures of primary care.

Fleetcroft et al (2010) determined that the clinical indicators with the greatest potential for mortality reduction were primary prevention for hypertension (BP05) and influenza immunisation.
<table>
<thead>
<tr>
<th>Indicator code</th>
<th>Detail of indicator</th>
<th>Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD06</td>
<td>The percentage of patients with CHD in whom the last blood pressure reading (measured in the previous 15 months) is 150/90 mm Hg.</td>
<td>90 (85, 93)</td>
</tr>
<tr>
<td>CHD08</td>
<td>The percentage of patients with CHD whose last measured total cholesterol level (measured in the previous 15 months) is 5 mmol/l.</td>
<td>82 (78, 87)</td>
</tr>
<tr>
<td>CHD09</td>
<td>The percentage of patients with CHD with a record in the previous 15 months that aspirin, an alternative antiplatelet therapy, or an anti-coagulant is being taken (unless a contraindication or adverse effects are recorded).</td>
<td>95 (93, 97)</td>
</tr>
<tr>
<td>CHD10</td>
<td>The percentage of patients with CHD who are currently treated with a β-blocker (unless a contraindication or adverse effects are recorded).</td>
<td>73 (66, 81)</td>
</tr>
<tr>
<td>CHD11</td>
<td>The percentage of patients with a history of myocardial infarction (diagnosed after 1 April 2003) who are currently treated with an angiotensin-converting enzyme inhibitor or angiotensin II antagonist.</td>
<td>92 (88, 97)</td>
</tr>
<tr>
<td>STROKE06</td>
<td>The percentage of patients with TIA or stroke in whom the last blood pressure reading (measured in the previous 15 months) is 150/90 mm Hg.</td>
<td>88 (82, 92)</td>
</tr>
<tr>
<td>STROKE08</td>
<td>The percentage of patients with TIA or stroke whose last measured total cholesterol level (measured in the previous 15 months) is 5 mmol/l.</td>
<td>76 (70, 83)</td>
</tr>
<tr>
<td>BP05</td>
<td>The percentage of patients hypertension in whom the last blood pressure reading (measured in the previous 9 months) is 150/90 mm Hg.</td>
<td>77 (73, 83)</td>
</tr>
<tr>
<td>DM12</td>
<td>The percentage of patients with diabetes in whom the last blood pressure reading is 145/85 mm Hg.</td>
<td>80 (74, 85)</td>
</tr>
<tr>
<td>DM17</td>
<td>The percentage of patients with diabetes whose last measured total cholesterol level (measured in the previous 15 months) is 5 mmol/l.</td>
<td>83 (77, 87)</td>
</tr>
<tr>
<td>DM20</td>
<td>The percentage of patients with diabetes in whom the last HbA1C is 7.5 (or equivalent test/reference range depending on laboratory) in the previous 15 months.</td>
<td>67 (61, 73)</td>
</tr>
<tr>
<td>SMOKE02</td>
<td>The percentage of patients with any or any combination of the following condition: CHD, stroke or TIA, hypertension, diabetes, COPD or asthma, who smoke and whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered within the previous 15 months.</td>
<td>94 (92, 96)</td>
</tr>
<tr>
<td>Mean of the above 12 indicators (Kiran CHD achievement score)</td>
<td>83 (80, 86)</td>
<td></td>
</tr>
<tr>
<td>CHD12</td>
<td>The percentage of patients with coronary heart disease who have a record of influenza immunisation in the preceding 1 September to 31 March</td>
<td>94 (91, 96)</td>
</tr>
</tbody>
</table>

Table 7.8: Underlying achievement for potential indicators ($n = 229$).
7.10 Exception reporting

Exception reporting allows general practices to exclude some patients from achievement calculations for QOF indicators. This may be due to clinical reasons such as contraindications for treatments or logistical reasons such as patients who have been recently registered. In addition, patient factors such as ‘patient informed dissent’ (not agreeing to the investigation or treatment) and patients who have refused to attend review despite having been invited on at least three occasions during the preceding twelve months can allow exception reporting. Exception reporting means that practices do not need to provide required care for all eligible patients. In English practices the median rate of exception reporting was 2.7% (IQR: 1.9-3.9%) (Doran et al., 2012).

7.11 Other factors describing quality of care

Various other factors have been included in studies of associations between primary care and health outcomes, for example physician age, the percentage of physicians who qualified in the UK and percentage female physicians (Kiran et al., 2010; Dusheiko et al., 2011). When hospital admissions is the health outcome of interest, it is common for the distance from general practices to hospitals to be included in studies (Bankart et al., 2011; Purdy et al., 2011). Kiran et al (2010) created relevant categorical indicators to describe these aspects of the practice and found higher percentages of GPs trained outside of the UK was associated with poorer CHD outcomes, in terms of both mortality and admissions, and practices with older GPs were associated with higher CHD admissions, but not mortality.

Various aspects of practice staff and organisation have been shown to be associated with QOF achievement. For example, increased percentage of GPs aged 50 and over and decreased percentage of GPs educated in the UK were associated with lower overall QOF achievement in the first year of QOF (Doran et al., 2006). Practices which remained low scoring for four years were analysed by Ashworth et al (2011) and were found to be more likely to be singlehanded, non-training practices and GPs were more likely to be 65 and over and male. In contrast to other studies, these practices were more likely to have GPs qualified in the UK. Lower achievement in QOF by single handed and smaller practices has been found to be explained by lower achievement in the organisational domain only (Wang et al., 2006). Factors associated with lower achievement have been found to be correlated. For example, singlehanded practices
are more likely to be male, older and have qualified in South Asian countries and are more likely to be in areas of high socio-economic deprivation (Wang et al., 2006).

The association between quality of primary care and various features of practice staffing is not clear cut. The aim of this research is to determine whether the quality of primary care can modify the effects of population risk factors on CHD mortality. Whilst there may be associations between various practice features and quality of primary care, measured through QOF, including these factors in a study of health outcomes may increase our understanding of the organisation of primary care in affecting quality of delivery rather than increasing our understanding of the role of primary care in improving health outcomes.

7.12 Conclusions

In this chapter the data describing the quality of primary care in terms prevention, detection and clinical management of CHD are identified and evaluated. The importance of data from QOF is discussed and the differing ways in which the data can be used in studies of this type considered. The advantage of using the ‘underlying achievement’ of individual indicators, the percentage of patients for whom the target has been met, has been demonstrated. Disease areas relevant to CHD have been identified and possible approaches to selecting relevant indicators described; the approach to indicator selection is discussed in Chapter 9.
Chapter 8

Indicators of quality of primary care - access and continuity of care.

8.1 Introduction

Access and continuity of care have been shown to be important aspects of primary care in explaining a variety of health outcomes, as discussed in Chapter 2. Concepts of access and continuity of care were discussed in Section 2.4. In addition, the potential to use responses to the GP Patient Survey (GPPS) as useful measures of these aspects of primary care was described. In this chapter the potential use and validity of the GPPS in the areas of study is discussed in detail, as well as practice level features associated with high levels of satisfaction with access and continuity.

8.2 Measure of access

The number of full-time-equivalent GPs has been accessed from the NHS Information Centre and is based on the year 2008/09, although it is a snapshot at a fixed point in the year. This has been used with the practice population information to estimate the number of GPs per 1000 patients. The numbers of other practice staff were not available for this study.
8.3 The GP patient survey (GPPS)

The GPPS has developed from the Department of Health’s (1997) aim to include patients’ views on quality of care and to use their views to inform the improvement of local services. The first iteration of the Quality and Outcomes Framework (QOF) in 2004 included incentives to ask patients their views using one of two approved questionnaires. In 2006/07 a national survey was introduced and in 2008/09 the results of two key questions were linked to indicators included in QOF within the patient experience domain (see Section 8.3.1 for more details).

8.3.1 Questions in the GPPS

The GPPS has existed in its current format, a postal survey administered by Ipos MORI on behalf of the Department of Health, since 2006/07, although there was extensive modification for 2008/09 (Roland et al., 2009). In 2006/07, in addition to three questions characterising the respondent, the survey contained 10 questions asking about experiences of making an appointment. Five key questions have been identified as being relevant; other questions ascertain whether the questions are relevant to the respondent. These five questions relate to five key areas relating to satisfaction with opening hours and experiences of attempting to obtain an appointment, see, for example, Addink et al (2011) and Baker et al (2011). These are shown in Table 8.1. In 2007/08 the survey was split into three sections and questions were added, shown in Table 8.2.

In 2008/09 the survey was considerably modified and additional questions were included. The style of the majority of response options changed from yes/no to ‘scales’ of satisfaction and frequency of positive experiences. The survey was divided into 12 sections. Questions similar to the five key questions were included and are shown in Table 8.3. In 2008/09 respondents are asked to complete a scale for questions relating to ‘phone’, ‘preferred’ and ‘open’. The published data tables combine the ‘positive’ response for these questions into an overall ‘satisfaction’/‘positive experience’ percentage. For example, for Section B, Question 5, part (a):‘Getting through on the phone’, patients who responded with ‘very easy’ and ‘fairly easy’ are combined into ‘easy’. The summary percentages can be used as an equivalent to the satisfaction/positive
experience score in the earlier two years. It can be expected that there will be differences between 2007/08 and 2008/09 because of the increased options for patients to complete.
<table>
<thead>
<tr>
<th>Theme</th>
<th>Abbreviation</th>
<th>Question number</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiences of phone access.</td>
<td>phone</td>
<td>Q2</td>
<td>Are you satisfied with how easy it is to get through on the phone?</td>
</tr>
<tr>
<td>Experiences of being able to get an</td>
<td>quick</td>
<td>Q3</td>
<td>Were you able to get an appointment fairly quickly? (On the same day or on the next 2 days the surgery was open)</td>
</tr>
<tr>
<td>appointment fairly quickly.</td>
<td>ahead</td>
<td>Q6</td>
<td>Were you able to get an appointment more than two full days in advance?</td>
</tr>
<tr>
<td>Experiences of being able to make an</td>
<td>preferred</td>
<td>Q8</td>
<td>Were you able to make an appointment with a particular Doctor?</td>
</tr>
<tr>
<td>appointment in advance.</td>
<td>open</td>
<td>Q9a</td>
<td>Were you satisfied with the hours your surgery was open?</td>
</tr>
</tbody>
</table>

Table 8.1: Five key themes in the GPPS.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Section - number of questions</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Getting to see a doctor</td>
<td>A - 10</td>
<td>Identical questions to 2006/07.</td>
</tr>
<tr>
<td>Referrals to hospital</td>
<td>B - 2</td>
<td>Relates to whether doctors discussed hospital choice with patients.</td>
</tr>
<tr>
<td>Some questions about you</td>
<td>C - 12</td>
<td>Including more detailed questions relating to individuals completing the questionnaire (for example, long-standing medical conditions) and questions about individuals which might inform the context of satisfaction with opening hours (eg distance travelled to work and hours of employment).</td>
</tr>
</tbody>
</table>

Table 8.2: 2007/08 GPPS survey sections.
<table>
<thead>
<tr>
<th>Theme</th>
<th>Section &amp; question number</th>
<th>Question</th>
<th>Response options</th>
</tr>
</thead>
<tbody>
<tr>
<td>phone</td>
<td>Section B Q5</td>
<td>In the past six months, how easy have you found the following: a) Getting through on the phone. b) Speaking to a doctor on the phone. c) Speaking to a nurse on the phone. d) Getting test results on the phone.</td>
<td>a) Haven’t tried. b) Very easy. c) Fairly easy. d) Not very easy. e) Not at all easy. f) Don’t know.</td>
</tr>
<tr>
<td>quick</td>
<td>Section C Q7</td>
<td>Were you able to get the appointment on the same day or on the next 2 days the GP surgery or health centre was open?</td>
<td>Yes/No/Can’t remember.</td>
</tr>
<tr>
<td>ahead</td>
<td>Section C Q10</td>
<td>Were you able to get an appointment more than two full days in advance?</td>
<td>Yes/No</td>
</tr>
<tr>
<td>preferred</td>
<td>Section E Q15</td>
<td>Is there a particular doctor you prefer to see at your GP surgery or health centre?</td>
<td>Yes/No/Can’t remember</td>
</tr>
<tr>
<td></td>
<td>Section E Q16</td>
<td>How often do you see the doctor you prefer to see?</td>
<td>a) Always or almost always. b) A lot of the time. c) Some of the time. d) Never or almost never. e) Not tried at this GP surgery or health centre.</td>
</tr>
<tr>
<td>open</td>
<td>Section F Q17</td>
<td>How satisfied are you with the hours that your surgery or health centre is open?</td>
<td>a) Very satisfied. b) Fairly satisfied. c) Neither satisfied nor dissatisfied. d) Fairly dissatisfied. e) Very dissatisfied. f) I’m not sure when my GP surgery or health centre is open.</td>
</tr>
</tbody>
</table>

Table 8.3: Questions associated with five key themes in GPPS - 2008/09.
The GPPS and QOF

The patient experience (PE) domain has been a domain within QOF since its first year. The domain is divided into ‘length of consultations’ and ‘patient surveys’; points allocated to the PE domain and requirements of practices are shown in Table 8.4. In 2008/09 the link between the survey and QOF became more explicit; 23.5 points were specifically allocated to the patients being able to obtain a consultation with a GP within two working days (PE7 - links directly to the response to Section C Q7 [quick]) and 35 points allocated to patients being able to book an appointment with a GP more than two days ahead (PE8 - links directly to section C Q10 [ahead]). By 2009 £68 million of general practitioners pay was tied to patient-reported experiences of access to care, in particular patients being able to book ahead and being able to book urgent appointments. This equates to £13700 for a practice with 10000 patients achieving maximum thresholds in patient response (Addink et al., 2011).

<table>
<thead>
<tr>
<th>Year</th>
<th>Points allocated to PE domain</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004/05</td>
<td>100</td>
<td>Practices were expected to undertake a patient survey, reflect on the results and propose changes.</td>
</tr>
<tr>
<td>2005/06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006/07</td>
<td>108</td>
<td>Practices were expected to have undertaken ‘an approved patient survey’ and produce an action plan.</td>
</tr>
<tr>
<td>2007/08</td>
<td>146.5</td>
<td>Practices were expected to have undertaken ‘an approved patient survey’ and produce an action plan.</td>
</tr>
</tbody>
</table>

Table 8.4: Links between the GPPS and QOF.

8.3.2 Are questions from the GPPS useful measures of access and continuity of care.

Baker’s team at Leicester have studied the relationship between a range of health outcomes, access and relationship continuity using answers to the GPPS. Being satisfied with opening hours (open) and positive experiences of phone access (phone) and being able to get an appointment within 48 hours (quick) have been identified as being measures of ‘access’, whereas it is suggested that being able to make an appointment ahead (ahead) and with a preferred GP (preferred) reflect personal continuity (Addink et al., 2011; Bankart et al., 2011; Suleman et al., 2011; Anwar et al., 2012; Levene et al., 2012; Gunther et al., 2013).
Campbell et al (2009) clearly consider being able to make an appointment within 48 hours and in advance (quick and ahead) as measures of access, not of continuity of care. In their own questionnaire they use the question ‘how often do you see your usual doctor?’ to assess continuity of care. This is similar to the GPPS question regarding preferred doctor (preferred). Similar question were used as measures of continuity in a survey to establish which patients value interpersonality (Baker et al., 2007). Although there are slight differences in the 2009/10 GPPS, questions relating to preference for and success in seeing a preferred doctor were used to assess continuity of care (Aboulghate et al., 2012). Brettell et al (2013) use both quick and ahead as measures of access from patients perspective in addition to including a measure of GP:patient ratio.

The advantage of using the GPPS is that it considers both access and continuity of care from the patients’ perspective, which is considered key by Haggarty et al (2003) and allows the consideration of ‘gaining access’ rather than simply ‘having access’ (Gulliford et al., 2002).

### 8.3.3 Summary - consideration of key questions

The five ‘key’ questions are reasonably consistent over the three years, although there are important differences between 2007/08 and 2008/09, including the introduction of scale options for answers and pay incentives related to QOF. This means that comparing results over this three year period will not give an accurate picture of how patient experience of general practices has changed over time. However, for individual years the five key questions can be used as measures of similar, if not identical, aspects of patient experience.

Despite the complicated nature of the concepts of access and continuity of care, the consistency of the five ‘key questions’, their use in previous research and the fact that they arise from the patients’ perspective suggests that there is a justification for including them in this analysis.

For the purposes of this research, open, phone and quick are considered measures of patient perceived access, whereas preferred and ahead are considered to measure aspects of patient perceived relationship continuity.
8.4 Response rates and satisfaction

8.4.1 Response rate

The response rate for the GPPS is relatively low, although possibly in line with other surveys of this type in the UK (Roland et al., 2009), decreasing from 46.1% in 2006/07 to 39.7% in 2008/09. The increase in the number of questions from 24 to 47 between 2007/08 to 2008/09 may in part explain the decrease in response rate, as may the deliberate over sampling of practices with low response rates. Roland et al (2009) describe an ‘adjusted response rate’ which takes this over-sampling into account and suggest an adjusted response rate for 2008/09 of 42.3%.

<table>
<thead>
<tr>
<th>Year</th>
<th>Sample size</th>
<th>Responses</th>
<th>Response rate</th>
<th>Undeliverable</th>
<th>Adjusted for undeliverable surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006/07</td>
<td>5,220,482</td>
<td>2,295,587</td>
<td>44.0%</td>
<td>239,697</td>
<td>46.1%</td>
</tr>
<tr>
<td>2007/08</td>
<td>4,922,080</td>
<td>1,999,523</td>
<td>40.6%</td>
<td>272,132</td>
<td>43.0%</td>
</tr>
<tr>
<td>2008/09</td>
<td>5,660,217</td>
<td>2,163,456</td>
<td>38.2%</td>
<td>216,162</td>
<td>39.7%</td>
</tr>
</tbody>
</table>

Table 8.5: GPPS annual response rates, based on technical reports.

Patients eligible for inclusion in the 2006/07 survey were those who had been registered at the GP practice both on 16/07/06 and 15/10/06 and aged 18 and over.

Patterns in response rate

The technical reports published each year to accompany the results of the GPPS do not include a detailed breakdown of responses but do provide information on response rates by gender, age group and regions. These show that females are more likely to respond than males and response rate increases with age, shown in Table 8.6. Response rates dropped in all groups over the three years; suggesting that the reasons for reductions in response rate are general.

Detailed analysis of the 2008/09 survey by Roland et al (2009) suggests these patterns have continued. Young people were the least likely to respond, with those aged 50-59 2.54 times as likely to respond as those aged 18-29. Patients aged 70-79 were 5.54 times as likely to respond as those in the youngest age group (18-29). The odds of responding declined approximately linearly with increasing deprivation. The odds of responding were 41% lower for men than for women, after allowing for the effects of
Table 8.6: National response rates for different groups.

<table>
<thead>
<tr>
<th></th>
<th>2006/07</th>
<th>2007/08</th>
<th>2008/09</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>48.7</td>
<td>46.7</td>
<td>45.4</td>
</tr>
<tr>
<td>Male</td>
<td>36.1</td>
<td>34.7</td>
<td>31.1</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-34</td>
<td>23.5</td>
<td>21.6</td>
<td>19.7</td>
</tr>
<tr>
<td>35-44</td>
<td>33.6</td>
<td>31.6</td>
<td>29.0</td>
</tr>
<tr>
<td>45-54</td>
<td>44.9</td>
<td>42.5</td>
<td>38.5</td>
</tr>
<tr>
<td>55-64</td>
<td>59.8</td>
<td>57.8</td>
<td>53.2</td>
</tr>
<tr>
<td>65 plus</td>
<td>67.4</td>
<td>65.9</td>
<td>60.4</td>
</tr>
</tbody>
</table>

In comparison with the general population, therefore, there is a clear under representation of men, younger people and people from more deprived areas in those that responded to the survey. Analysing the representation of ethnic groups is more complex. The proportion of the survey respondents who described themselves as white British is similar to the general population. There is under-representation of Asian/Asian British, Black/Black British, Chinese and mixed, but the proportion describing themselves as ‘other’ is 3.8%, rather than 0.8% in the general population (Roland et al., 2009).

**Response rate - the local context.**

The response rate in the sample of practices analysed in this thesis is similar to that across the East Midlands, 47% in 2006/07, 43% in 2007/08 and 40% in 2008/09. The variation in response rates between practices is large, with the maximum response rate 67% and the lowest response rate 7%. The interquartile range in 2006/07 was 43% to 57%.

Patient level data could not be easily obtained for the practices included in this analysis, but the associations between features of the practices and the response rate in the patient survey have been analysed.

There is moderate to strong positive correlation between socio-economic deprivation and response rate. As levels of deprivation increase (shown by an increase in deprivation score), response rate decreases, varying from -0.66 in 2008/09 to -0.71 in 2007/08 ($p < 0.0001$ for all three years). Fig 8.1 shows a reasonably linear relationship between
deprivation and response rate, with some clear outliers.

![Graph showing association between response rate and deprivation score]

Figure 8.1: Association between response rate and levels of socio-economic deprivation, as measured by the Index of Multiple Deprivation (2007).

There is a moderate positive correlation between the percentage of the practice population who are white (according to hospital episode statistics) and response rate. Figure 8.2 suggests that the relationship is not linear, with a distinct group of five practices with low response rates.

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>2006/07</th>
<th>2007/08</th>
<th>2008/09</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>0.53*</td>
<td>0.56*</td>
<td>0.51*</td>
</tr>
<tr>
<td>Asian and Asian other</td>
<td>-0.41*</td>
<td>-0.43*</td>
<td>-0.38</td>
</tr>
<tr>
<td>Black and Black other</td>
<td>-0.618</td>
<td>0.62*</td>
<td>-0.61*</td>
</tr>
<tr>
<td>Other</td>
<td>-0.53*</td>
<td>-0.53</td>
<td>-0.51*</td>
</tr>
<tr>
<td>Irish white</td>
<td>-0.26</td>
<td>-0.27</td>
<td>-0.32</td>
</tr>
<tr>
<td>Mixed</td>
<td>-0.11</td>
<td>-0.11</td>
<td>-0.08</td>
</tr>
</tbody>
</table>

Table 8.7: Correlations between practice ethnic groups and response rates in LNR (*p < 0.001).
There is a clear pattern in the response rate and the classification of practices into urban, town or village. Practices within an urban setting have the lowest median response rate across all the years, 40% for 2008/09, compared with 50% for town and fringe practices and 53.5% for village based practices (Fig. 8.3). There is a weak, positive correlation between the size of the practice list and response rates, the correlation increases over the three years (Fig. 8.4). There is a strong, positive correlation between the percentage of the practice list aged over 65 and response rate, with $r > 0.7$ for all years. Figure 8.5 shows that the five practices with low response rates also have a low percentage of the practice population aged 65 and over; this includes three practices which serve mainly student populations.

These results suggest a similar pattern in response rate to those found by Roland et al (2009). There are clear links between features of general practices across LNR and response rates, practices with higher levels of socio-economic deprivation and a higher proportion of patients aged over 65 have higher response rates.
8.4.2 Reporting of positive experiences and satisfaction

Results of the national survey suggest that the majority of patients report positive experiences and satisfaction with their practices with regards to both type of appointment and opening hours and this is also true in the area of study (Table 8.8). The drop in positive response for phone and preferred may reflect the introduction of the scale option, although this may be due to the lack of financial incentive related to this
Figure 8.5: Association between response rate and the percentage of the practice population over 65.

question resulting in practices prioritising other aspects of appointment management. 2008/09 was the first year that QOF included specific indicators relating to responses on a patient survey, specifically quick and ahead. Research which included use of an additional survey in 1998, 2003, 2005 and 2007 in 42 practices suggests that continuity of care as measured by the question ‘how often do you see your usual doctor’, decreased after the introduction of QOF (Campbell et al., 2009) and this decline may have continued as QOF incentives became more specific.

<table>
<thead>
<tr>
<th></th>
<th>2006/07</th>
<th>2007/08</th>
<th>2008/09</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>National</td>
<td>Local</td>
<td>National</td>
</tr>
<tr>
<td>open</td>
<td>84%</td>
<td>83%</td>
<td>82%</td>
</tr>
<tr>
<td>phone</td>
<td>86%</td>
<td>84%</td>
<td>87%</td>
</tr>
<tr>
<td>quick</td>
<td>86%</td>
<td>86%</td>
<td>87%</td>
</tr>
<tr>
<td>ahead</td>
<td>75%</td>
<td>68%</td>
<td>72%</td>
</tr>
<tr>
<td>preferred</td>
<td>88%</td>
<td>85%</td>
<td>88%</td>
</tr>
</tbody>
</table>

Table 8.8: National and local positive response in five key areas.

Practice level results show similar patterns, with decreases in median levels of positive response for phone and preferred in 2008/09. There is considerable variation in the percentage of patients recalling positive experiences at practice level. The lowest level of positive response in any practice was 13%, in response to being able to book an appointment more than 48 hours ahead (ahead) in 2006/07. The maximum was 100% recall of positive experience for quick in all three years.
Figure 8.6: Satisfaction and positive experience in the five key areas across practices in the area of study.

Possible practice level predictors of positive response and satisfaction.

In 2006/07 the technical report clearly states that ‘older patients, those registered with smaller practices, patients living in rural areas and White British respondents were more likely to report that they were able to get the types of appointments they wanted’. Kontonpantelis et al (2010) considered both patient and practice level indicators of patient experience for the 2007/08 patient survey. Satisfaction and positive experience were associated with increasing age; in contrast, satisfaction and positive experience
was lower amongst those working full time and all non-white ethnic groups, most notably patients who described their ethnicity as Asian or mixed Asian. Females were slightly less likely to report positive experiences than males in all areas; the effect was strongest for being able to book an appointment in advance.

Possible practice level indicators were also examined (Kontopantelis et al., 2010). Practice size was found to be a strong practice-level predictor, larger practices receiving poorer ratings in all areas. Practices with a higher GP-to-patient ratio were reported as being better in terms of phone access (phone) and the availability of appointments within two days (quick). Practices serving more deprived areas receive lower satisfaction scores in all areas, except being able to make an appointment with a preferred doctor (preferred).

A particularly interesting result found by Kontopantelis et al (2010) was the effect of the ethnic make up of the practice list on satisfaction and experience over and above a patient’s own ethnic identity. They found that the larger the non-white community the more likely that both white and non-white patents were to give lower ratings on all domains of access, after controlling for other patient and area characteristics.

**Local area practice level features associated with satisfaction and positive experience**

To determine if there is any similarity in patterns of satisfaction and positive experience in practices in the sample a simple correlation analysis has been completed on various practice level predictors. Satisfaction rates and levels of positive experience are all positively skewed, so associations between practice level features and satisfaction are assessed using Spearman’s rank correlation coefficient.

**Socioeconomic deprivation.**

Socio-economic deprivation is moderately negatively correlated with satisfaction and positive experience; the strongest association is with being able to see a preferred Doctor (preferred), ($R_S = -0.43$ (06/07); $R_S = -0.41$ (07/08) and $R_S = -0.37$ (08/09)). Fig. 8.7 shows the relationship between deprivation and satisfaction for the five key questions for 2006/07. As with national results, lower levels of positive experience are associated with practices with a more deprived practice population.
Figure 8.7: Association between satisfaction and deprivation for the five key questions in LNR - 2006/07.

**Practice size and GPs per capita.**

As with national patterns, larger practices received poorer satisfaction ratings and lower levels of positive experiences. There are weak to moderate negative correlations between four of the five key questions and practice size in each of the three years. Opening hours (open) shows no association with practice size. The strongest association is between positive experience of being able to get through on the phone (phone) and size of practice ($R_S = -0.34$ (06/07), $R_S = -0.30$ (07/08) and $R_S = -0.30$ (08/09), $p < 0.001$ for all three years). The association between being able to book an appointment in advance (ahead) also shows a moderate negative correlation for all three years ($R_S = -0.24$ (06/07), $R_S = -0.22$ (07/08) and $R_s = -0.25$ (08/09), $p < 0.001$ for all three years). Being able to see a preferred Doctor shows a moderate negative correlation in 2008/09 ($R_S = -0.29$, $p < 0.001$), although the correlation is much weaker in 2006/07 and 2008/09 ($R_S = -0.16$ and $p < 0.001$ in both years).

In this sample of practices no associations were found between GPs per 1000 patients and satisfaction and positive experience.
Urban/Rural indicator

Practices defined as being in a village location generally had higher levels of satisfaction and positive experience than those in town and fringe or urban settings. This remained even when deprivation was taken into account; results not shown here.

<table>
<thead>
<tr>
<th>Classification</th>
<th>phone</th>
<th>quick</th>
<th>ahead</th>
<th>preferred</th>
<th>open</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban (171)</td>
<td>0.85</td>
<td>0.88</td>
<td>0.70</td>
<td>0.86</td>
<td>0.84</td>
</tr>
<tr>
<td>Town and Fringe (41)</td>
<td>0.93</td>
<td>0.92</td>
<td>0.75</td>
<td>0.90</td>
<td>0.85</td>
</tr>
<tr>
<td>Village (12)</td>
<td>0.99</td>
<td>0.95</td>
<td>0.96</td>
<td>0.97</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Table 8.9: Median satisfaction and positive experiences in practices in the area of study based on their urban classification (2006/07).

Age of practice population

There was a weak to moderate association between being able to see a preferred Doctor (preferred) and the percentage of the practice population aged over 65. There was a weak association between satisfaction with opening hours (open) and phone access (phone) in 2006/07, but the association was weaker in 2007/08 and 2008/09. There was no evidence of an association between age being able to book an appointment within 48 hours (quick) or in advance (ahead).

Percentage female in the practice population

There was little evidence of an association between the percentage of the practice population who were female and positive experience of with access in four of the key areas. There was a weak positive association between the percentage of the practice who were female and being able to make an appointment with a preferred Doctor (preferred).

8.4.3 Response rate and satisfaction.

The relatively low levels of response across all practices, with extremely low response rates in some practices, lead to questions about the reliability of the response as an indicator of satisfaction with opening hours and positive experiences of aspects of
access and continuity of care. In all three years there is a positive correlation between response rate and satisfaction or levels of positive experiences in all five 'key' areas. The relationship is strongest for the association between being able to see a preferred Doctor (preferred) and overall response rate ($R_s = 0.41$ (06/07), $R_S = 0.41$ (07/08) and $R_s = 0.38$ (08/09), $p < 0.001$ for all three years). The relationship is also reasonably strong between being able to make an appointment within 48 hours (quick) and response rate ($R_s = 0.38$ (06/07), $R_s = 0.30$ (07/08) and $R_s = 0.41$ (08/09), $p < 0.001$ for all three years). The relationship between being able to book ahead (ahead) and response rate is the weakest ($R_s < 0.2$ and $p > 0.1$ for all three years).

Whilst the correlation is low to moderate and the pattern is unclear, there is some evidence that practices with higher response rates are achieving higher levels of satisfaction and positive experiences in particular areas and the underlying reasons for this are not known. This may be due to confounding variables, such as deprivation, which are associated with both response rate and satisfaction or levels of positive experience. When practice level variables and response rate are included in a model considering the associations between mean levels of positive experience across the five key areas rate in 2006/07, percentage of the practice population who are white (positive association), size of practice population and deprivation levels (both negative associations) are the only explanatory variables which are significant in explaining mean ‘positive experience’. However, only 8.8% of the variation in mean satisfaction is explained by the model, suggesting other factors such as the actual service received by patients may be important in explaining variation in satisfaction. This finding is supported by detailed analysis of the 2008/09 survey (Roland et al., 2009) which found significant, positive correlation between quick and ahead and response rate ($R_S = 0.34$ and $R_S = 0.18$ respectively; $p < 0.001$ for both), but that these correlations were reduced to 0.04 and -0.07 respectively when demographic features of practices were taken into account. No systematic variation between response rates and questionnaire scores was found for quick and ahead and there were higher levels of practice-level reliability.

Kontopantelis et al. (2010) are also not concerned that the low response rate could mean that results are affected by response bias. They argue that any over-representation of females will not introduce bias as the size of the effect of gender was estimated to be very small. They do suggest that satisfaction and positive experience with access might have been overestimated as a result of the lower response rate in younger patients, who tend to be more negative in their responses. In addition, a ‘mystery shopper’ based study provides evidence that the GPPS is a valid measure of access to care and appointment
availability (Campbell et al., 2013). A researcher simulated a patient and contacted practices to request appointments. Results showed that practices with lower scores in the GPPS had poorer access as experienced by the ‘mystery shopper’.

8.5 Selection of key access and continuity of care responses

Although the five key questions are all potentially useful in studies of health outcomes; for methodological reasons it may be necessary to select one or two measures. If variables which are highly correlated are included in a statistical model this may result in large standard errors; a problem termed collinearity (Harrell, 2001, p.64). In addition, the number of explanatory variables included in a model is limited by the number of observations; this was discussed in Section 3.9. Whilst variable selection approaches will be discussed in more detail in Chapter 9, the possibility of combining the measures is discussed here.

A simple mean or total ‘positive experience’ measure would be easy to determine, however, specific information about patient satisfaction with, and positive experiences of, access and continuity of care may be lost.

A consideration of the correlations shows that there are moderate to strong correlations between positive experiences in all five questions. This is shown clearly in Figure 8.8. More detailed analysis shows that the relationships between the five questions are not the same for all years. However, for all years, the lowest correlation is between being able to make an appointment in less than 48 hours (quick) and being able to book an appointment in advance (ahead). \( R_s = 0.26, 0.29 \) and \( 0.34 \) for 2006/07, 2007/08 and 2008/09 respectively, \( p < 0.0001 \) for all three years. This may reflect the difficulty of arranging practice appointment systems to meet both demands. For 2006/07 and 2007/08 the three strongest pairwise correlations were between being able to see preferred doctor (preferred), ahead and phone (\( R_s \) ranges from 0.68 to 0.73 over the three pairs and three years). For 2008/09 the associations show a slightly different pattern, the highest association is between quick and ahead (\( R_s = 0.65 \)), all other correlations are between 0.34 and 0.55. This may be related to the introduction of new response options in 2008/09.
Principal component analysis allows further investigation. Component 1 accounts for 59% to 62% of the variance; each key question achieves approximately equal loadings, this can be interpreted as an overall ‘satisfaction’. Component 2 accounts for between 15% and 17% of the variance and shows a contrast between two aspects of access: quick and open and aspects more related to continuity of care: ahead and preferred, although preferred is not important in 2008/09. In 2006/07 and 2007/08 component 3 accounts for 11% of the variance and shows a contrast between open and quick; however the in 2008/09 component 3 is a contrast between preferred and open and phone.

Consideration of individual practice scores for each of these two components does not identify any clear groups of practices.

The high correlation between ahead and preferred and their grouping in the PCA suggests that it may be possible to use one of these measures in the analysis, rather than the need to include both, the same may be true for quick and open.

8.6 Conclusions

Although access and continuity of care are important concepts, their definition and measurement is not straightforward, as discussed in Chapter 2. The increasingly com-
mon use of responses to key questions on the GPPS and their interpretation in terms of patient perceived access and continuity of care means they are a useful measure of this aspect of primary care. Whilst much of this work has been based in Baker’s team in Leicester (for examples, see Baker et al., 2009; Addink et al., 2011; Anwar et al., 2012), other researchers have described these measures in terms of access and continuity of care.

The GPPS is the most comprehensive measure of patient experiences of access and continuity; despite low response rates, there is evidence of reliability (Roland et al., 2009) and validity (Campbell et al., 2013). It is therefore a rich and useful resource and will be used in this study. The five key questions have been identified as phone, open, quick, ahead and preferred. The possibility of combining measures into ‘access’ based on phone, open and quick and those related to continuity of care ahead and preferred should be considered.
Chapter 9

Variable selection

9.1 Introduction

Chapter 1 identified possible risk factors associated with CHD mortality and Chapters 6, 7 and 8 have identified and evaluated data which measure the prevalence of these risk factors in practice population. This chapter considers the process of variable selection adopted in this thesis and the resulting set of explanatory variables included in the initial analysis.

9.2 Variable selection in this data set

The approach to variable selection in this thesis is based on previous research, careful consideration of available data, examination of QOF indicators and the underlying theory linking primary care to mortality presented in Chapter 1. Characteristics of the practice population are considered the main predictors of CHD mortality (all-age and premature). This research considers the role primary care can have on modifying this predictive effect. Therefore, it is important to include population based variables in the model prior to the inclusion of primary care variables. This approach is sometimes termed hierarchical regression, although this term is confusing as it is more commonly applied to multilevel models, or ‘forced’ variables. The data describing population characteristics are limited and therefore the quality and availability of data are the key factors driving variable selection. Within primary care there are many potential measures of quality of primary care, hence the process of variable selection raises more
challenges. This process is described in Sections 9.2.2 and 9.2.3.

9.2.1 Population characteristics

Population factors which have been previously identified and available data are summarised in Table 9.1. Data availability, discussed in Chapter 6, results in a reduction in potential variables for inclusion from 13+ to five.

<table>
<thead>
<tr>
<th>Factor type</th>
<th>Possible factors</th>
<th>Appropriate practice level data available identified in Chapter 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unmodifiable biological risk factors</td>
<td>age, sex, prevalence of high risk genes, and ethnicity</td>
<td>age, sex and ethnicity</td>
</tr>
<tr>
<td>Modifiable biological risk factors</td>
<td>diabetes, obesity, serum cholesterol and high blood pressure</td>
<td>diabetes</td>
</tr>
<tr>
<td>Health behaviours identified with high CHD</td>
<td>smoking, alcohol, physical (in)activity, and poor nutrition</td>
<td>None readily available(^1)</td>
</tr>
<tr>
<td>Socioeconomic deprivation</td>
<td>potentially many variables</td>
<td>Index of Multiple Deprivation (2007)</td>
</tr>
</tbody>
</table>

\(^1\)Smoking data is discussed further in Chapter 15. Obesity, alcohol, physical (in)activity and poor nutrition are available at local area level.

Table 9.1: Summary of population characteristics which have been considered and availability of data.

9.2.2 Primary care - clinical management

The availability of QOF data means there is a multitude of data available to describe clinical management (discussed in Chapter 7). Fig. 9.1 describes how the 80 indicators included in the clinical domain can be reduced to 17 candidate variables.
Consider all indicators within the clinical domain

80 indicators

Only consider indicators from relevant disease areas: CHD (10), stroke (8), diabetes (16), hypertension (3), obesity (1) and depression (2).

40 indicators

Indicators within depression\textsuperscript{1} and obesity\textsuperscript{1} areas not considered useful.

37 indicators

Indicators which relate to ‘record keeping’ or referrals considered less relevant than indicators relating to achievement of targets\textsuperscript{2}

17 indicators

Figure 9.1: QOF clinical indicators: process of variable selection. \textsuperscript{1} and \textsuperscript{2} discussed in previous chapters
Two key studies, Levene et al (2012) and Kiran et al (2010) suggest possible approaches to selecting potential variables. Levene et al use two key clinical indicators which they describe as being applicable to the majority of people on the CHD register and therefore of particular relevance. Based on an earlier paper (Levene et al., 2010) they include QOF hypertension registers, which can be seen as a measure of detection of hypertension (as discussed in Section 6.4.3) and therefore CHD prevention. Kiran et al (2010) use a CHD achievement score, a mean measure of quality of CHD care (Table 7.8). In addition, the division of primary care into prevention, detection and clinical management used throughout this thesis means it would be preferable to have at least one measure of each.

9.2.3 Primary care - access and sustained relationships

The potential to use the GPPS (GP patient survey) to measure access and continuity of care has been described in Chapter 8. A commonly used measure of access is GPs per capita. The percentage of patients who recall being able to make an appointment with a preferred GP, a question on the GPPS, is a useful measure of sustained relationships.

9.3 Approach to be used in this model

Having identified candidate variables based risk factors associated with CHD mortality and available data; this study uses a variable selection method in part based previous research. Variables which other researchers have considered to be potentially important when trying to understand the role primary care may have in modifying population based risk of CHD mortality have been selected. Variables have not been selected based on whether a study has found a variable to be significant as this involved accepting their method of variable selection and the arbitrary nature of significance testing; the focus will be on whether or not the variables were thought to be potentially important with respect to the underlying theory and/or clinically relevant. There have been no previous studies on premature CHD mortality and there is insufficient evidence to suggest whether particular aspects of primary care may be more important in explaining premature mortality. Hence, initially two models will be considered; both include four population characteristics (age, sex, ethnicity and diabetes prevalence) but vary in the variable describing primary care. These are summarised in Table 9.2.
Table 9.2: Summary of variables included in two models considered in this thesis.

<table>
<thead>
<tr>
<th>Model A</th>
<th>Model B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Four population characteristics (age, sex, ethnicity and diabetes prevalence).</td>
<td>Hypertension register as a measure of hypertension detection.</td>
</tr>
<tr>
<td>Smoking cessation indicator as a measure of prevention.</td>
<td></td>
</tr>
<tr>
<td>Percentage of CHD patients who have serum cholesterol 5mmol/l or less (CHD08).</td>
<td></td>
</tr>
<tr>
<td>Percentage of CHD patients who have been treated with aspirin (or alternative) (CHD09).</td>
<td>Kiran CHD achievement score.</td>
</tr>
<tr>
<td>GPs per capita.</td>
<td></td>
</tr>
<tr>
<td>Percentage of patients who recall being able to make an appointment with a preferred GP.</td>
<td></td>
</tr>
</tbody>
</table>

9.4 Conclusions

There is no one method of variable selection which will result in the ‘best’ model, mainly because there is no universal definition of a ‘best’ model. All methods have disadvantages and these should be taken into account when models are interpreted. Variables have therefore been chosen because there is a clear expectation that they may influence the response; the strength of influence of each predictor is the aim of study. The strategy adopted will be to fit the full model, and then interpret the sizes of the estimates of the parameters in terms of their importance (O’Hara and Sillanpää, 2009).

Chapter 10 describes the initial analysis, based on the variables described as Model A above. Results of both models are presented in Chapter 11. In Part V alternative approaches to modelling strategies, including sensitivity analyses, and additional or alternative explanatory variables are discussed.
Part IV

Initial Model - Approach and Results
Chapter 10

Modelling CHD mortality using count models

10.1 Introduction

In this chapter features of primary care associated with premature CHD mortality are modelled. Approaches to modelling deaths were summarised in Chapter 3. Count models have been identified as the most appropriate approach for modelling deaths. In this chapter count models are applied to model the associations between premature CHD mortality and population characteristics and aspects of primary care. Initially a Poisson model is considered. As there is evidence of over-dispersion, an over-dispersed Poisson is also considered. The majority of the chapter considers the use of the negative binomial model, model fit and consideration of outliers. The explanatory variables included in the model are those described in Section 9.3 described as ‘Model A’.

In the following Chapter (11) the results of modelling are presented and briefly discussed. This includes comparisons between premature and all-age mortality as well as the use of the combined CHD achievement score in comparison to individual QOF indicators. In Part V possible extensions of and alternatives to the model are discussed.

10.2 Summary of explanatory variables

The explanatory variables used in the following model are summarised in Table 10.1; there is evidence of skewness in all variables, with the exception of CHD08 (serum
cholesterol). The percentage of patients who are aged 65 and over and who are male are included in the model as a measure of the general risk of the practice population to CHD. These would not be expected to be included in the model if the outcome variable was standardised by age and sex. There is evidence of correlation between

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Practice characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Deprivation indices</td>
<td>16.2 (10.0, 27.0)</td>
</tr>
<tr>
<td>% of GP list on diabetes register</td>
<td>3.8 (3.3, 4.4)</td>
</tr>
<tr>
<td>% White ethnicity</td>
<td>89.9 (77.5, 94.1)</td>
</tr>
<tr>
<td>% of population who are 65 and over</td>
<td>14.7 (12.1, 17.0)</td>
</tr>
<tr>
<td>% of population who are male</td>
<td>50.2 (49.5, 51.2)</td>
</tr>
<tr>
<td><strong>Service characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>GP's per 1000</td>
<td>0.55 (0.47, 0.64)</td>
</tr>
<tr>
<td>% patients with recalled perception of being able to see preferred GP</td>
<td>88 (80, 93)</td>
</tr>
<tr>
<td>% of GP registered list on hypertension register</td>
<td>12.3 (11.0, 14.7)</td>
</tr>
<tr>
<td>The percentage of patients with any or any combination of the following conditions: coronary heart disease, stroke or TIA, hypertension, diabetes, COPD or asthma who smoke whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered within the previous 15 months. (SM02)</td>
<td>94 (92, 96)</td>
</tr>
<tr>
<td>The percentage of patients with CHD whose last measured total cholesterol level (measured in the previous 15 months) is $\leq 5$ mmol/l. (CHD08)</td>
<td>82 (78, 87)</td>
</tr>
<tr>
<td>The percentage of patients with CHD with a record in the previous 15 months that aspirin, an alternative antiplatelet therapy, or an anti-coagulant is being taken (unless a contraindication or adverse effects are recorded). (CHD09)</td>
<td>95 (93, 97)</td>
</tr>
</tbody>
</table>

Table 10.1: Summary statistics for explanatory variables used.

some pairs of variables. In particular there are moderate to strong correlations between the percentage of the practice population who are white, have diabetes and are deprived (see Table 10.2). There are also correlations between patient’s recall of being able to see their preferred GP and deprivation ($R_S = -0.43$); hypertension detection and aged 65 or over($R_S = +0.63$) and the two QOF CHD indicators ($R_S = +0.49$).

Although these correlations may cause some concern, it is to be expected in a model of this type. In order to confirm whether collinearity is a problem in this model; variance inflation factors were calculated, following a multiple regression model, the values do not trigger concern of collinearity (Everitt and Rabe-Hesketh, 2003, p63; discussed in Section 3.9).
Table 10.2: Pairwise correlations $R_S$ between key explanatory variables ($p \leq 0.0001$ in all cases).

<table>
<thead>
<tr>
<th>Deprivation indices</th>
<th>% of GP list on diabetes register</th>
<th>% white</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.42</td>
<td>1.00</td>
<td>-0.61</td>
</tr>
</tbody>
</table>

10.3 Poisson model

Initially the simplest count model, the Poisson model, has been used to model the data. Analysis was carried out using Stata 11, with the population of the practice aged under 75 as a measure of exposure. The results are shown in Table 10.3.

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>IRR*</th>
<th>95% CI</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage white patients</td>
<td>1.007</td>
<td>(1.003, 1.012)</td>
<td>0.001</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.017</td>
<td>(1.012, 1.022)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>1.106</td>
<td>(1.029, 1.189)</td>
<td>0.006</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>1.057</td>
<td>(1.038, 1.077)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>1.066</td>
<td>(1.034, 1.100)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>1.198</td>
<td>(0.911, 1.576)</td>
<td>0.196</td>
</tr>
<tr>
<td>Hypertension detection 2006/07</td>
<td>0.980</td>
<td>(0.955, 1.007)</td>
<td>0.142</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>1.001</td>
<td>(0.993, 1.009)</td>
<td>0.737</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>0.991</td>
<td>(0.983, 0.999)</td>
<td>0.026</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>1.001</td>
<td>(0.983, 1.020)</td>
<td>0.880</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>0.994</td>
<td>(0.990, 1.000)</td>
<td>0.019</td>
</tr>
</tbody>
</table>

*IRR: Incident Rate Ratios - for every one unit increase in the predictor, the predicted count changes by (IRR - 1)× 100%

Table 10.3: IRR and 95% confidence intervals for premature CHD mortality - after fitting a Poisson model.

The IRRs shown in Table 10.3 suggest that increasing the percentage of patients who are white, the level of socioeconomic deprivation, the percentage aged 65 or over, the percentage who are male, the number of GPs per 1000 patients, the percentage of patients with particular conditions who have been offered smoking cessation advice and the percentage of the CHD register who are taking aspirin (or equivalent) are all associated with an increase in premature mortality count. However, the confidence intervals for the number of GPs per 1000 patients and percentage taking aspirin include one; the CI for GPs per 1000 patients is particularly wide. The IRRs for the percentage who are on aspirin and those offered smoking cessation advice are close to one and may
reflect no association with premature mortality.

Conversely, increasing hypertension detection, an increase in serum cholesterol control and an increase the percentage of patients who can recall being able to see their preferred GP are associated with a decrease in premature mortality count. The confidence interval for hypertension detection includes one.

Both the deviance and Pearson goodness of fit tests suggest that the model does not adequately fit the data ($\chi^2_D = 303.5; \chi^2_P = 303.5$ $p < 0.001$ in both cases). The lack of fit could be due to over-dispersion, both the Lagrange Multiplier test ($\chi^2 : 11986$ $p < 0.0001$) and the Z-test ($0.168 p = 0.029$); provide evidence that the data are over-dispersed. Figure 10.1 provides some further evidence that the relationship between the mean and the variance is not linear. The variance of groups of practices ($n = 10$) has been plotted against the mean and there is evidence that the variance increases more rapidly than the mean, particular for the group of practices with the highest mean value, suggesting over-dispersion (Stroup, 2012, p241). Similar results were found when the all-age mortality was modelled.

![Figure 10.1: Plot of variance against mean - practices divided into 10 groups (Poisson model).](image-url)
10.4 Over-dispersed Poisson Model

There is evidence that the Poisson model is not a satisfactory model and that the data are over-dispersed. In this section a quasi-Poisson model is fitted to the data. This involves estimating the value of the dispersion parameter ($\phi$) from the Poisson model and then including this in the model.

The estimate of $\phi$ is based on the Pearson Deviance, which in this case is 288.88.

$$\phi = \frac{\text{Pearson deviance}}{\text{degrees of freedom}} = \frac{288.88}{217} = 1.33$$

The scale parameter of 1.33 indicates that the variance is 33% greater than the mean. Fitting a quasi-Poisson model does not impact considerably on the interpretation of the model. Interpretation based on the $p$ values suggest that the serum cholesterol control indicator is no longer a ‘significant’ variable in the model. The upper bound of the confidence interval is still one. Figure 10.2 suggests that the variance function used in the over-dispersed Poisson is a reasonably good fit to the data, but still does not capture the high variance of practices with highest numbers of deaths.

Figure 10.2: Plot of variance against mean - practices divided into 10 groups (quasi-Poisson model).
10.5 Negative Binomial Model

When the data are modelled using negative binomial regression, interpretation is similar to that with the Poisson model and the quasi-Poisson model. Analysis was carried out using Stata 11, with the population of the practice aged under 75 as a measure of exposure. The dispersion parameter is a function of the expected mean. Results of this model are shown in Table 10.4. As with the Poisson models, the results suggest that all population characteristics are positively associated with premature CHD mortality. There are also no changes to the direction of association, or interpretation, of primary care characteristics. Results are discussed further in Chapter 11. The following section evaluates the fit of the model.

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>IRR</th>
<th>95% CI</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage white patients</td>
<td>1.008</td>
<td>(1.003, 1.012)</td>
<td>0.002</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.017</td>
<td>(1.011, 1.024)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>1.114</td>
<td>(1.028, 1.208)</td>
<td>0.009</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>1.059</td>
<td>(1.038, 1.081)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>1.067</td>
<td>(1.038, 1.103)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>1.197</td>
<td>(0.885, 1.619)</td>
<td>0.244</td>
</tr>
<tr>
<td>Hypertension detection 2006/07</td>
<td>0.978</td>
<td>(0.950, 1.007)</td>
<td>0.133</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>1.002</td>
<td>(0.993, 1.011)</td>
<td>0.712</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>0.991</td>
<td>(0.981, 1.000)</td>
<td>0.044</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>1.002</td>
<td>(0.982, 1.022)</td>
<td>0.884</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>0.994</td>
<td>(0.989, 1.00)</td>
<td>0.036</td>
</tr>
</tbody>
</table>

Table 10.4: IRR and 95% confidence intervals for premature mortality - after fitting a Negative Binomial Model.

10.5.1 Discussion of model fit

There is good agreement between observed numbers of deaths and numbers estimated from the model. Lin’s concordance coefficient is 0.85 (Lin, 1989) and the mean different is 0.05 and 95% limits of agreement are -6.57 to 6.77 (Bland and Altman, 1986). The pseudo-$R^2$ (McFadden’s $R^2$) is 0.075, which is very low, and does not suggest a good fit. However, the interpretation is not straightforward and does not provide evidence of a very poor fit; this was discussed in Section 3.5.
Over-dispersion

Fitting a negative binomial model is only useful if the dispersion parameter $\alpha$ is greater than zero, as when the dispersion parameter is zero the model reduces to a Poisson. In this case there was good evidence that the data are over-dispersed. The estimated value of $\alpha$ is 0.0233. Figure 10.3 shows that the variance function in the negative binomial model captures the higher variance of the larger practices better than the quasi-Poisson, although the improvement is not substantial.

![Plot of variance against mean - observations divided into 10 groups.](image)

Figure 10.3: Plot of variance against mean - observations divided into 10 groups.

Systematic departures from the model.

Both the link and variance functions are assessed initially. Residual plots as suggested by McCullagh and Nelder (1989) were considered. Fig. 10.4 shows residuals against transformed fitted values and identifies no particular pattern in the residuals, although it does clearly identify outliers. The lack of pattern suggests that the link function is appropriate. The plots of standardised deviance and Anscombe residuals are similar, so only standardised deviance residuals are used for the remaining plots. Plots of absolute values of standardised deviance residuals, shown in Figure 10.5, suggest a random pattern; suggesting the variance function is appropriate. Plots of standardised residuals against each explanatory variable are shown in Figures 10.6 and 10.7. Again the plots
highlight clear outliers, and the possibility of some concern for certain explanatory
variables. In particular there is some evidence that the range is greater at low IMD
scores. Added variable plots (Figures 10.8 & 10.9) suggest that all variables make a
contribution to the model.

Figure 10.4: Standardised deviance and Anscombe residuals plotted against trans-
formed fitted values.

Figure 10.5: Absolute residuals plotted against transformed fitted values.
Figure 10.6: Standardised deviance residuals plotted against explanatory values. A: Population characteristics.

Figure 10.7: Standardised deviance residuals plotted against explanatory values. B: Service characteristics.
Figure 10.8: Added variable plot. A: Population characteristics. 1: Percentage white 2: Deprivation (IMD) 3: Diabetes prevalence (%) 4: Percentage 65 and over 5: Percentage male.

Figure 10.9: Added variable plot. B: Service characteristics. 6: Hypertension detection 7: GPs per 1000 patients 8: Preferred GP - patients recall 9: SM02 10. CHD08 11. CHD09
Individual departures from the model.

Initially a plot of standardised residuals against transformed fitted values was examined to identify potential outliers (as Fig 10.4), clear outliers can be identified and these are shown in Figure 10.10. In addition to identifying outliers based on large standardised residuals; Figure 10.11 highlights 26 practices with high leverage (as measured by standardised hat values greater than 2; described in Section 3.5 (McCullagh and Nelder, 1989, p405)); five practices have standardised hat values greater then four.

![Figure 10.10: Residuals against transformed fitted values - identification of outliers.](image)

To determine if practices with large residuals and high leverage have high influence, Cook’s distances were calculated using Stata’s \texttt{glm} post-estimation commands. Cook’s distances greater than 0.0174 ($4/n$) can be considered as worth investigating and greater than 0.0184 ($4/n - p - 1$) as potentially problematic. 17 practices have Cook’s distances greater than 0.0174, all distances are also greater than 0.0184. Index plots of both Cook’s distances (Fig. 10.12) show that several practices clearly have high influence, particular practices stand out as being particularly high in comparison to others. Practices labeled A, C and I can be identified as having particularly high influence in comparison to other practices.
Discussion of outliers

Three practices have been identified as having high influence, as measured through Cook’s distances. Initial observations suggested that two practices serve an atypical
practice list. To investigate this key characteristics of the practices were compared to the range in the sample of 229 practices to identify unusual features.

Practice A serves a university campus and has an unusually high percentage of male patients (61.23%, compared to the IQR for whole sample (49.5% - 51.18%)) and young patients (80.64% are aged under 25 - IQR: (28.10% - 32.96%)) and a relatively low percentage of white patients (44.17% - IQR: (77.51% - 94.07%)).

Practice I serves a local population but aims to treat ‘hard to reach’ patients, including travellers and patients known to be drug dependent. Again, there is a relatively high percentage of young patients (36.66% are aged under 25 and 51.33% are aged between 25 and 44 - IQR for the whole sample is 26.82% to 30.92%) and a high percentage of male patients (58.51%). The third practice is a relatively small practice (lowest 10% of practice population aged under 75) and has a high percentage of older patients (13.13% are aged between 65 and 74 - IQR: (6.39% - 8.88%)).

Data were modelled with and without these practices in order to evaluate the impact of these practices on the level of over-dispersion. Table 10.5 shows the impact of the exclusion of each practice identified as potentially problematic on the estimation of the dispersion parameter-\( \alpha \). The exclusion of practices C and I has a minimal impact on \( \alpha \) and does not suggest that the over-dispersion may be due to outliers. However, the exclusion of practice A reduces \( \alpha \) and there is insufficient evidence to reject the null hypothesis that \( \alpha = 0 \). A similar result is found when all three outlying practices are deleted, suggesting that the over-dispersion may be apparent rather than actual (Hilbe, 2011, p. 141).

As practices A and I serve an atypical patient list it may be justifiable to exclude the practices from the analysis. Deleting the two ‘unusual’ outlying practices means there is no longer sufficient evidence to reject the null hypothesis that \( \alpha = 0 \). This suggests that negative binomial regression may not be the most appropriate model for the data (Table 10.6).

Table 10.7 shows the IRRs and corresponding confidence intervals when over-dispersion
Table 10.6: Impact of deleting two ‘unusual’ outlying practices on overdispersion.

<table>
<thead>
<tr>
<th>Practice excluded</th>
<th>$\alpha$</th>
<th>$p$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0.023</td>
<td>0.019</td>
</tr>
<tr>
<td>A &amp; I</td>
<td>0.016</td>
<td>0.068</td>
</tr>
</tbody>
</table>

is allowed for by using a negative binomial model, and when outliers are deleted and the data are modelled using Poisson regression. The size and interpretation of the IRRs for the majority of explanatory variables is not affected by the deletion of the two ‘unusual’ outlying practices. However, the effect of diabetes prevalence is reduced and the confidence interval for the IRR relating to serum cholesterol now includes one meaning the evidence that this is associated with premature mortality is less clear cut. Results are similar when a Poisson model with the two ‘unusual’ outlying practices are deleted.

One practice has been excluded from the sample because is serves a highly restricted patient list; this was described in Section 5.2. The excluded practice is a very small practice and served an entirely restricted patient list; offering no services to the general population. The target population is also known to have an increased risk of adverse health outcomes (Hewett et al., 2012). Practices A and I have an atypical patient list, but members of the general population who were living in the catchment were eligible to register with the practice during the period of study. Although these two practices have been discussed in detail, there are other practices which have high influence and/or atypical practice lists, and removal of two practices may result in further practices being investigated.

Whilst acknowledging that it is a matter of judgement, the outliers will be included in the model for future analysis. However, the interpretation of certain explanatory variables are sensitive to model choice and outlier deletion.
Table 10.7: Impact of deleting two ‘unusual’ outlying practices on incident rate ratios.

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Negative binomial model</th>
<th>Negative binomial model</th>
<th>Poisson</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 229</td>
<td>A &amp; I excluded</td>
<td>A &amp; I excluded</td>
</tr>
<tr>
<td></td>
<td>IRR (95% CI)</td>
<td>IRR (95% CI)</td>
<td>IRR (95% CI)</td>
</tr>
<tr>
<td>Percentage white patients</td>
<td>1.008 (1.003, 1.012)</td>
<td>1.007 (1.002, 1.011)</td>
<td>1.006 (1.002, 1.011)</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.017 (1.011, 1.024)</td>
<td>1.017 (1.011, 1.023)</td>
<td>1.017 (1.011, 1.022)</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>1.114 (1.028, 1.208)</td>
<td>1.097 (1.013, 1.187)</td>
<td>1.091 (1.014, 1.173)</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>1.059 (1.038, 1.081)</td>
<td>1.060 (1.039, 1.081)</td>
<td>1.059 (1.039, 1.078)</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>1.067 (1.038, 1.103)</td>
<td>1.072 (1.036, 1.110)</td>
<td>1.073 (1.038, 1.109)</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>1.197 (0.885, 1.619)</td>
<td>1.107 (0.821, 1.491)</td>
<td>1.113 (0.843, 1.470)</td>
</tr>
<tr>
<td>Hypertension detection 2006/07</td>
<td>0.978 (0.950, 1.007)</td>
<td>0.978 (0.950, 1.006)</td>
<td>0.979 (0.954, 1.005)</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>1.002 (0.993, 1.011)</td>
<td>1.005 (0.996, 1.014)</td>
<td>1.004 (0.996, 1.013)</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>0.991 (0.981, 1.000)</td>
<td>0.994 (0.985, 1.004)</td>
<td>0.994 (0.986, 1.003)</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>1.002 (0.982, 1.022)</td>
<td>0.996 (0.976, 1.016)</td>
<td>0.996 (0.978, 1.015)</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>0.994 (0.989, 1.000)</td>
<td>0.994 (0.989, 1.000)</td>
<td>0.994 (0.990, 0.999)</td>
</tr>
</tbody>
</table>
10.6 Conclusions

In this chapter it has been shown that the data are over-dispersed and there is evidence that the model fit is improved when a negative binomial model is used in contrast to a standard Poisson model. The difference in the results when different modelling approaches are used to model the count of deaths are minor and do not affect interpretation. Three key outliers have been identified, two of these are practices which serve an atypical patient list, and will be retained in the sample.

In the following chapter the results of the model described in this chapter are reported and compared with models of all-CHD mortality. Comparisons between two measures of primary care are presented.
Chapter 11

Results of initial model

This chapter summarises the results of using negative binomial regression with 229 practices using the variables described in Section 9.3, as described in the previous chapter. Similarities and differences between the use of individual QOF indicators and the CHD achievement score as measures of primary care are described (Section 11.1). A similar approach is adopted for premature and all-age mortality (Section 11.2). An approach to presenting results to aid interpretation and meaning for practitioners and policy makers is used in addition to more standard reporting approaches (Section 11.3). A discussion of the implications of the results are not included here. In Part V the model is further developed, through sensitivity analyses and consideration of alternative and additional variables, and this will have an impact on the results and subsequent implications. The implications of the developed model are discussed in Chapter 18.

11.1 Associations between quality of primary care and premature CHD mortality

Table 11.1 shows the incident rate ratios (IRRs), 95% confidence intervals and \( p \)-values when population characteristics and aspects of the quality of care are included in a model with premature CHD mortality as the dependent variable. Two models are presented; the first model includes two individual QOF indicators and the second includes a combined CHD achievement score.

All characteristics of the practice population are positively associated with premature
mortality. This is as expected as they are known risk factors for CHD mortality, apart from percentage white. Access, as measured through GPs per 1000 patients; prevention, as measured by smoking cessation advice; and one aspect of clinical management, treatment with aspirin, are all positively associated with mortality, although the confidence intervals are wide and include one. Hypertension detection, control of serum cholesterol and continuity of care, as measured by patients recall of being able to see their preferred GP, are negatively associated with premature mortality.

The CHD achievement score is also associated with lower mortality counts. Other IRRs are only affected at the third decimal point and interpretation would be the same as when two indicators are included in the model, unless there is a reliance on \( p \)-values, in which case the measure of continuity of care is no longer ‘significant’.
<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Separate QOF indicators</th>
<th>CHD achievement score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Percentage white patients</td>
<td>1.008</td>
<td>(1.003, 1.012)</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.017</td>
<td>(1.011, 1.024)</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>1.114</td>
<td>(1.028, 1.208)</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>1.059</td>
<td>(1.038, 1.081)</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>1.067</td>
<td>(1.038, 1.103)</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>1.197</td>
<td>(0.885, 1.619)</td>
</tr>
<tr>
<td>Hypertension detection 2006/07</td>
<td>0.978</td>
<td>(0.950, 1.007)</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>1.002</td>
<td>(0.993, 1.011)</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>0.991</td>
<td>(0.981, 1.000)</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>1.002</td>
<td>(0.982, 1.022)</td>
</tr>
<tr>
<td>CHD achievement score</td>
<td>0.988</td>
<td>(0.978, 0.999)</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>0.994</td>
<td>(0.989, 1.000)</td>
</tr>
</tbody>
</table>

**IRRs are interpreted as follows: for every one unit increase in the predictor, the predicted count changes by \((IRR - 1)\times 100\%\).**

Table 11.1: IRRs, 95% confidence intervals and \(p−\)values for premature mortality - inclusion of CHD achievement score (after fitting a Negative Binomial Model).
11.2 Comparison between premature and all-age mortality

Table 11.2 presents the results when all-age CHD mortality is the dependent variable and, for ease of comparison, repeats the results for premature mortality. The associations between population characteristics and mortality are similar, whether all-age or premature mortality is the dependent variable. The percentage of patients who are male is not significantly associated with all-age mortality; the confidence interval is wide and includes one. The size of the association between diabetes prevalence and all-age mortality is smaller than with premature mortality. The associations between service characteristics and mortality are also similar for both all-age and premature mortality; although treatment with aspirin is negatively associated with all-age mortality and the measure of continuity of care, being able to see the preferred GP, is not significantly associated with all-age mortality and the IRR is close to one.
<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Premature mortality</th>
<th>All-age mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Percentage white patients</td>
<td>1.008</td>
<td>(1.003, 1.012)</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.017</td>
<td>(1.011, 1.024)</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>1.114</td>
<td>(1.028, 1.208)</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>1.059</td>
<td>(1.038, 1.081)</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>1.067</td>
<td>(1.038, 1.103)</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>1.197</td>
<td>(0.885, 1.619)</td>
</tr>
<tr>
<td>Hypertension detection 2006/07</td>
<td>0.978</td>
<td>(0.950, 1.007)</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>1.002</td>
<td>(0.993, 1.011)</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>0.991</td>
<td>(0.981, 1.000)</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>1.002</td>
<td>(0.982, 1.022)</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>0.994</td>
<td>(0.989, 1.00)</td>
</tr>
</tbody>
</table>

**IRR**s are interpreted as follows: for every one unit increase in the predictor, the predicted count changes by \((\text{IRR} - 1)\times 100\%\).  

Table 11.2: IRRs, 95% confidence intervals and *p*–values for **premature mortality** and **all-age** mortality (after fitting a Negative Binomial Model).
11.3 Effect of changes in explanatory variables

For every one unit increase in the value of an explanatory variable there is a change in the predicted count of mortality of (IRR-1)x100%. A one unit increase may not be meaningful to practitioners or policy makers. To make the results more meaningful, the impact of a change in the value of the explanatory variable equivalent to a change from ‘above’ or ‘below’ average to ‘average’ has been considered. For population characteristics the impact of reducing the ‘population burden’ from that experienced by a practice at the 75th percentile to one with a median level of ‘burden’ is determined. For service characteristics an improvement from a practice with below average performance, at the 25th percentile, to a practice with a median level of performance is more likely to be the change of interest. These changes are also shown graphically in Figures 11.1 and 11.2. These graphs have been developed to aid interpretation by practitioners and policy makers; further development is needed to ensure that they are easy to understand.

The effect of a one unit changes in explanatory variables and the effect of decrease in population burden or improvement in primary care on premature mortality count are shown in Table 11.3. The IRR suggests that a one unit change in diabetes prevalence potentially has the largest impact on premature mortality count, when only explanatory variables with strong evidence of association are considered. However, when a decrease in ‘population burden’ is considered, both deprivation and percentage aged 65 and over have much larger potential impacts. Comparing practices with different levels of deprivation, even within the 25% of practices within the middle of the range, has the potential to affect mortality count by over 18%. This has important implications for comparing practices, as has been discussed by Sullivan et al 2007. Improving the quality of care in terms of both achieving serum cholesterol target and patients’ recall of seeing their preferred GP have the potential to decrease premature mortality count by over 4%, although this depends on the model being correct, the other explanatory variables in the model and that the link is causal.
<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Effect of a 1% increase in explanatory variable</th>
<th>Effect of improvement in primary care or decrease in population burden - potential reduction in mortality count (%) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage white patients</td>
<td>0.8 (0.3, 1.2)</td>
<td>Decrease of 4.2 in percentage white patients.</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.7 (1.1, 2.4)</td>
<td>Decrease of 10.9 units on scale.</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>11.4 (2.8, 20.8)</td>
<td>Decrease of 0.5% in diabetes prevalence.</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>5.9 (3.8, 8.1)</td>
<td>Decrease of 2.3% in percentage over 65.</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>6.7 (3.8, 8.1)</td>
<td>Decrease of 1.0% in percentage male patients.</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>19.7 (-11.5, 61.9)</td>
<td>Increase of 0.08 GPs per 1000 patients.</td>
</tr>
<tr>
<td>Hypertension detection 2006/07</td>
<td>-2.2 (-5.0, 0.7)</td>
<td>Increase of 1.3% in hypertension detection.</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>0.2 (-0.7, 1.1)</td>
<td>Increase of 2.1% offered advice.</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>-0.9 (-1.9, 0.0)</td>
<td>Increase of 4.9% in achieving serum cholesterol target.</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>0.2 (1.8, 2.2)</td>
<td>Increase of 2.1% in aspirin treatment.</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>-0.6 (-1.1, 0.0)</td>
<td>Increase of 8.0% in patients recalling being able to see preferred GP.</td>
</tr>
</tbody>
</table>

Table 11.3: Effects on premature CHD mortality count of a one unit increase in the value of the explanatory variables and the impact of an improvement in primary care or decrease in population burden. Percentage change in premature count (95% confidence interval) (n = 228) (after fitting a Negative Binomial Model).
Figure 11.1: Effects on premature CHD mortality count of a decrease in population burden. Percentage change in premature count (95% confidence interval).

Figure 11.2: Effects on premature CHD mortality count of an improvement in primary care. Percentage change in premature count (95% confidence interval).
11.4 Conclusions

Results of the initial model suggest that the aspects of practice populations known to be associated with CHD mortality are positively associated with premature and all-age mortality, although the percentage of the practice population who are male is not significantly associated with all-age mortality. The percentage of patients who are white is positively associated with both premature and all-age mortality, which is not the expected direction of association. Certain aspects of quality of primary care are negatively associated with premature mortality; in particular continuity of care and control of serum cholesterol. The evidence that hypertension detection is associated with mortality is less clear cut. There is less evidence that continuity of care is associated with all-age mortality.

When changes between above or below average to average levels of ‘population burden’ or quality of primary care are considered, deprivation and percentage male have the largest potential impact on premature mortality count.

In the following chapters, various sensitivity analyses are carried out, the possibility of other measures of smoking prevalence and hypertension detection and including area level estimates and their impact on results are considered. The results of a final model are presented in Chapter 17, with a detailed discussion in Chapter 18.
Part V

Developing and extending the initial model
Chapter 12

Sensitivity Analyses

12.1 Introduction

This chapter considers the key variables included in the model described in Chapter 10. The percentage aged over 65 and the percentage who are male are both included in that model to account for an increased risk in these groups. Sensitivity analysis is carried out to determine the impact of the practice population estimate used to determine these percentages. In addition the impact of different age variables is considered. Approaches to missing data are also discussed.

12.2 Age categories in count models

The current model has one covariate directly relating to age, the percentage of the practice population aged over 65. This is included to take into account the increased risk of death in those over 65 and those under 65, as illustrated in Fig. 12.1.

12.2.1 Denominator for the percentage aged 65 and over

The percentage aged 65 and over used in the initial model is based on the total practice population. It may be more appropriate to use the ‘at risk’ population, which, in a model of premature deaths, would be the practice population aged under 75.
There is extremely high correlation between the percentage aged 65 and over of the total practice population (%65plusALL) and the percentage aged 65 and over of the practice population aged under 75 (%65plusU75) (R² : 0.97, p < 0.0001). Including %65plusU75 in the model instead of %65plusALL does not improve the overall fit of the model, for example deviance increases from 255.3 to 257.4; there are minor changes in the information criteria. There are differences in the IRRs for the two versions of the covariates, shown in Table 12.1. The lower IRR for %65plusALL suggests using the total practice population as the denominator may lead us to under-estimate the impact of increased age on premature mortality count. In Section 11.3 the interpretation of results based on considering the impact of a decrease from an ‘above average’ to ‘average’ practice was described. In this case the IQR is narrower for %65plusU75 than for %65plusALL and therefore the potential impact of a change from the 75th percentile to the median is greater for %65plusALL (Table 12.2).

<table>
<thead>
<tr>
<th>Age covariate</th>
<th>IRR (95%CI)</th>
<th>Effect of a 1% increase in explanatory variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>%65plusU75</td>
<td>1.111 (1.071, 1.153)</td>
<td>11.1% (7.1%, 15.3%)</td>
</tr>
<tr>
<td>%65plusALL</td>
<td>1.059 (1.038, 1.081)</td>
<td>5.9% (3.8%, 8.1%)</td>
</tr>
</tbody>
</table>

Table 12.1: IRR and 95% confidence intervals for age covariates based on total population and population under 75 (model includes 11 other explanatory variables).

It is recommended that if premature mortality is the only outcome %65plusU75 should be the covariate in the model describing the age profile of the practice population. If
Age covariate | Decrease from 75th percentile to median | Potential affect on mortality count of decrease in population burden (%) (95% CI)
--- | --- | ---
%65plus\textsubscript{U75} | Decrease of 1.41% in percentage over 65 | 15.65 (10.01, 21.57)
%65plus\textsubscript{ALL} | Decrease of 3.24% in percentage over 65 | 19.12 (12.31, 26.24)

Table 12.2: Impact of decrease in percentage aged 65 and over on premature mortality account, based on total population and population under 75 (model includes 11 other explanatory variables).

comparisons with all age mortality are being considered it may be appropriate to use \%65plus\textsubscript{all} to allow more direct comparisons.

12.2.2 Which age groups should be included in the model?

The percentage aged 65 and over was used by Levene et al (2010; 2012). The use of 65 and over as a measure of risk is supported by the increase in risk of CHD at this age (Fig. 12.1) and other studies which have used 65 and over as key age category (Ashworth et al., 2013). Here the use of other age groups is considered as a sensitivity analysis.

A sensitivity analysis was carried out to determine the impact on the results and their interpretation if different age groups were included in the model. The following age groups were considered: % 45 plus; % 50 plus; % 55 plus; % 60 plus; % 65 plus.

A comparison of information criteria and deviance does not suggest any improvement in model fit if different age groups are used. There is a gradual decrease in IRR as the boundary for the age covariate decreases, shown in Table 12.3.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>IRR (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>65 plus</td>
<td>1.059 (1.038, 1.081)</td>
</tr>
<tr>
<td>60 plus</td>
<td>1.056 (1.038, 1.075)</td>
</tr>
<tr>
<td>55 plus</td>
<td>1.054 (1.038, 1.070)</td>
</tr>
<tr>
<td>50 plus</td>
<td>1.048 (1.034, 1.063)</td>
</tr>
<tr>
<td>45 plus</td>
<td>1.045 (1.032, 1.059)</td>
</tr>
</tbody>
</table>

Table 12.3: IRR and 95% confidence intervals for different age groups in a model which includes 11 other explanatory variables.

There is an impact on the interpretation of two key other covariates in the model
when the upper age boundary is changed. Table 12.4 shows how the IRRs for diabetes prevalence and hypertension detection change as the upper age boundary changes. A change in the age covariate has no effect on the direction of association for either variable, but the interpretation of the importance of the association is affected by the age covariate. Hence, the results of the model and their interpretation are sensitive to the age covariate which is included in the model.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Diabetes Prevalence</th>
<th>Hypertension Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>65 plus</td>
<td>1.114 (1.028, 1.208)</td>
<td>0.978 (0.950, 1.007)</td>
</tr>
<tr>
<td>60 plus</td>
<td>1.104 (1.102, 1.195)</td>
<td>0.970 (0.943, 0.999)</td>
</tr>
<tr>
<td>55 plus</td>
<td>1.082 (1.000, 1.172)</td>
<td>0.968 (0.941, 0.996)</td>
</tr>
<tr>
<td>50 plus</td>
<td>1.062 (0.980, 1.152)</td>
<td>0.972 (0.944, 1.000)</td>
</tr>
<tr>
<td>45 plus</td>
<td>1.047 (0.964, 1.136)</td>
<td>0.975 (0.948, 1.003)</td>
</tr>
</tbody>
</table>

Table 12.4: IRR and 95% confidence intervals for diabetes prevalence and hypertension detection when different upper age boundaries are used, (in a model which includes other specified explanatory variables).

12.2.3 Use of more than one age group

When mortality rates are standardised the different expected rates of many age groups are used. It is possible that including more than one age group variable in the model will improve the overall model fit.

Initial exploration of the percentage in key age groups shows that there is a strong correlation between different age groups apart from between those in the 0 to 24 age group and the 25 to 44 age group.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>0 - 24</th>
<th>25 - 44</th>
<th>45 - 54</th>
<th>55 - 64</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 - 44</td>
<td>-0.168</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45 - 54</td>
<td>-0.73</td>
<td>-0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>55 - 64</td>
<td>-0.76</td>
<td>-0.58</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>65 - 74</td>
<td>-0.68</td>
<td>-0.59</td>
<td>0.61</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Table 12.5: Correlations between different age groups in practice populations ($p < 0.0001$ in all cases except where $p = 0.80$).

Table 12.6 shows how the deviance of the model is reduced when additional certain age groups are added to the model. The decrease is not evident when the age group 45-54 is added to the model. However, there is some evidence that the inclusion of three
Table 12.6: Deviance of models with additional age groups in the model.

<table>
<thead>
<tr>
<th>Age groups included in model</th>
<th>Deviance</th>
<th>Pearson Deviance</th>
</tr>
</thead>
<tbody>
<tr>
<td>65−74&lt;sub&gt;75&lt;/sub&gt;</td>
<td>257.45</td>
<td>244.04</td>
</tr>
<tr>
<td>65−74&lt;sub&gt;75&lt;/sub&gt; &amp; 55−64&lt;sub&gt;75&lt;/sub&gt;</td>
<td>255.50</td>
<td>241.81</td>
</tr>
<tr>
<td>65−74&lt;sub&gt;75&lt;/sub&gt; &amp; 55−64&lt;sub&gt;75&lt;/sub&gt; &amp; 45−54&lt;sub&gt;75&lt;/sub&gt;</td>
<td>255.45</td>
<td>242.22</td>
</tr>
<tr>
<td>65−74&lt;sub&gt;75&lt;/sub&gt; &amp; 55−64&lt;sub&gt;75&lt;/sub&gt; &amp; 45−54&lt;sub&gt;75&lt;/sub&gt; &amp; 25−44&lt;sub&gt;75&lt;/sub&gt;</td>
<td>249.14</td>
<td>240.04</td>
</tr>
</tbody>
</table>

additional age groups does lead to a marked reduction in deviance. The reduction in deviance when age group 55-64 is included in the model is small and does not suggest an improvement in model fit. There is some limited evidence that including the age group 25-44 to the model may improve the model fit; reductions in deviance are small. The concordance between the predicted count of deaths and the actual count of deaths is increased slightly when additional age groups are included in the model and average difference and 95% limits of agreement are reduced. However, these changes are minor.

Table 12.7: Concordance and 95% limits of agreement between actual count of deaths and predicted count of models with additional age groups in the model.

<table>
<thead>
<tr>
<th>Age groups included in model</th>
<th>Concordance</th>
<th>Average difference</th>
<th>95% limits of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>65−74&lt;sub&gt;75&lt;/sub&gt;</td>
<td>0.853</td>
<td>-0.05</td>
<td>(-6.63, 6.53)</td>
</tr>
<tr>
<td>65−74&lt;sub&gt;75&lt;/sub&gt; &amp; 55−64&lt;sub&gt;75&lt;/sub&gt;</td>
<td>0.857</td>
<td>-0.036</td>
<td>(-6.52, 6.45)</td>
</tr>
<tr>
<td>65−74&lt;sub&gt;75&lt;/sub&gt; &amp; 55−64&lt;sub&gt;75&lt;/sub&gt; &amp; 25−44&lt;sub&gt;75&lt;/sub&gt;</td>
<td>0.871</td>
<td>-0.005</td>
<td>(-6.11, 6.10)</td>
</tr>
</tbody>
</table>

In conclusion, the evidence suggests that including more age groups in the model does not improve the fit of the model sufficiently to warrant the addition of additional parameters in the model.

12.3 Sex covariate in model

The model includes the percentage of the practice population who are male as a ‘standardisation’ covariate since mortality rates are higher in males. In the initial model the percentage is based on the total practice population (%male<sub>all</sub>) and not the population aged under 75 (%male<sub>U75</sub>). A sensitivity analysis suggests no evidence of improved model fit when %male<sub>U75</sub> was included in the model. The impact on other variables were minor.
12.4 Missing data

The majority of practices have complete data sets describing both their population characteristics and primary care. However, there are data missing in relation to two of the 11 covariates: GPs per 1000 patients and percentage of patients with a preferred GP. In total 12 practices, 5.24%, have missing data and no practice has data missing for more than one explanatory variable.

12.4.1 GPs per 1000 patients

Six practices have no data relating to the number of GPs per 1000 patients; in addition two practices have a data entry of zero for this covariate. Values of zero are considered ‘missing’ in this study as this reflects an administrative approach to the recording of GPs in practices, not the actual number of GPs working in a practice. In the model described in Chapter 10 median imputation was used; that is the missing values were replaced with the median value of the remaining practices. The use of alternative approaches is considered as a sensitivity analysis. Results of the approaches are shown in Table 12.8.

When a complete case analysis (He, 2010) was used, that is only practices with no missing data are included, there are minor changes to the results and their interpretation. This is the case for the confidence intervals of the percentage of CHD register with serum cholesterol below target (CHD08) and patients recall of being able to see their preferred GP, with the upper limit of both CIs increasing to 1.001.

Two approaches to imputation were compared with median imputation: replacement with extreme values; and regression mean imputation (LSHTM, 2014). Conclusions did not depend on the method of imputation and no one method was associated with improved model fit. In this study an approach of simple median imputation was adopted, although this method has known disadvantages. For example, variability may be underestimated, other information is not used to impute the missing data and relationships between variables are not conserved (Zhou et al., 2001). However, in this case, sensitivity analysis suggests that conclusions do not depend on the method on imputation.
<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Median imputation</th>
<th>Complete case analysis</th>
<th>Imputed with minimum value</th>
<th>Imputed with maximum value</th>
<th>Regression with mean imputation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR (95% CI)</td>
<td>IRR (95% CI)</td>
<td>IRR (95% CI)</td>
<td>IRR (95% CI)</td>
<td>IRR (95% CI)</td>
</tr>
<tr>
<td>Percentage white patients</td>
<td>1.008 (1.003, 1.012)</td>
<td>1.008 (1.003, 1.012)</td>
<td>1.008 (1.003, 1.012)</td>
<td>1.008 (1.003, 1.012)</td>
<td>1.008 (1.003, 1.012)</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.017 (1.011, 1.024)</td>
<td>1.018 (1.012, 1.024)</td>
<td>1.017 (1.011, 1.024)</td>
<td>1.017 (1.011, 1.024)</td>
<td>1.017 (1.011, 1.023)</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>1.114 (1.028, 1.208)</td>
<td>1.122 (1.033, 1.219)</td>
<td>1.112 (1.026, 1.206)</td>
<td>1.115 (1.028, 1.209)</td>
<td>1.115 (1.029, 1.209)</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>1.059 (1.038, 1.081)</td>
<td>1.059 (1.038, 1.082)</td>
<td>1.059 (1.037, 1.081)</td>
<td>1.060 (1.038, 1.082)</td>
<td>1.065 (1.038, 1.081)</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>1.067 (1.038, 1.103)</td>
<td>1.077 (1.040, 1.116)</td>
<td>1.068 (1.032, 1.105)</td>
<td>1.063 (1.029, 1.099)</td>
<td>1.068 (1.032, 1.104)</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>1.197 (0.885, 1.619)</td>
<td>1.213 (0.898, 1.638)</td>
<td>1.173 (0.878, 1.568)</td>
<td>1.118 (0.890, 1.405)</td>
<td>1.218 (0.901, 1.647)</td>
</tr>
<tr>
<td>Hypertension detection 2006/07</td>
<td>0.978 (0.950, 1.007)</td>
<td>0.977 (0.948, 1.006)</td>
<td>0.979 (0.951, 1.008)</td>
<td>0.977 (0.949, 1.006)</td>
<td>0.978 (0.950, 1.007)</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>1.002 (0.993, 1.011)</td>
<td>1.002 (0.993, 1.010)</td>
<td>1.002 (0.993, 1.011)</td>
<td>1.001 (0.993, 1.010)</td>
<td>1.002 (0.993, 1.011)</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>0.991 (0.981, 1.000)</td>
<td>0.991 (0.982, 1.001)</td>
<td>0.990 (0.981, 1.000)</td>
<td>0.991 (0.982, 1.000)</td>
<td>0.991 (0.981, 1.000)</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>1.002 (0.982, 1.022)</td>
<td>1.000 (0.980, 1.021)</td>
<td>1.002 (0.982, 1.022)</td>
<td>1.001 (0.981, 1.021)</td>
<td>1.002 (0.982, 1.022)</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>0.994 (0.989, 1.000)</td>
<td>0.996 (0.990, 1.001)</td>
<td>0.994 (0.989, 1.000)</td>
<td>0.994 (0.989, 1.000)</td>
<td>0.994 (0.989, 1.000)</td>
</tr>
<tr>
<td>Deviance</td>
<td>255.3</td>
<td>247.7</td>
<td>255.1</td>
<td>255.8</td>
<td>255.2</td>
</tr>
<tr>
<td>Pearson</td>
<td>240.0</td>
<td>230.1</td>
<td>239.8</td>
<td>240.4</td>
<td>239.8</td>
</tr>
</tbody>
</table>

Table 12.8: Impact of different imputation approaches for missing data for GPs per 1000 patients.
12.4.2 Percentage of patients being able to see preferred GP

Four practices have missing data relating to the GP patient survey (GPPS) in 2006/07, although they all have complete data for the GPPS in 2007/08 and 2008/09.

Initially information from 2007/08 and 2008/09 was used to impute missing data for 2006/07. Practices were ranked on the basis of the percentage of patients who recalled being able to see their preferred GP, the ranks of the practices with missing data were noted and the data for percentage of equivalently ranked practices in 2006/07 were used to replace the missing data. This method is simple but depends on no change in practice performance over time.

Sensitivity analysis was used to determine the importance of the imputation method; all results are compared to the method of ranking practices described above. A complete case analysis had no impact on conclusions. Use of extreme values showed that the IRR for the percentage of patients who recalled seeing their preferred GP is sensitive to the method of imputation used. When the lowest values are used the IRR is increased from 0.994 to 0.997 and the 95% confidence interval includes one. When the maximum values are imputed the IRR is the same to three decimal places, the upper bound of the 95% CI is 0.999 compared to 1.000 in the original analysis.

Two regression based methods were also considered (LSHTM, 2014). One method was based on using the percentage recalling being able to see their preferred GP in 2007/08 and 2008/09 to predict values for 2006/07, a second method was to use all the response for 2007/08, including overall response rate, to predict values for 2006/07. The results based on these methods of imputation are extremely similar, and similar to those found when using the ranking method, shown in Table 12.9. The difference in imputed values is greatest for practice C. The results and their interpretation are not affected when these regression methods of imputation are adopted.

Given the results are not sensitive to the method of imputation based on 2007/08, the ranking approach described above will be used. There are concerns about using the information from 2007/08 to impute the missing the data as the data for 2006/07 are highly unlikely to be missing at random. Practices may have opted out of the GPPS in 2006/07 and this may have been because they expected poor results, however it may also have been because they were already using an effective survey of their patients.
Practice | Method A | Method B | Method C
---|---|---|---
A | 68.78 | 70.87 | 70
B | 82.77 | 82.44 | 80
C | 78.52 | 81.15 | 84
D | 86.83 | 84.73 | 86

A. Regression imputation based on GPPS responses from 2007/08 and 2008/09.
B. Regression imputation based on GPPS responses from 2007/08.
C. Imputation based on rank of practice in 2007/08 (approach adopted in this analysis).

Table 12.9: Imputed response to the GPPS question related to being able to see a preferred GP for practices with missing data from three alternative methods.

### 12.5 Conclusions

Different approaches to missing data imputation and covariate selection have minor effects on model fit and interpretation. However, the changes are minor and do not suggest a change in approach is necessary, although alternative approaches may be justified.
Chapter 13

Multiple linear regression of standardised mortality rates

13.1 Introduction

Chapter 3 describes the concerns associated with multiple linear regression of standardised mortality rates. However, research based on this approach is still relatively common and can be seen in journals (Muller, 2002; Kiran et al., 2010); it is therefore important to consider the impact of different modelling approaches on results and their interpretation. In this chapter standardised premature mortality is modelled using unweighted and weighted multiple linear regression. In addition, the log(SMR) is considered as a dependent variable. The set of explanatory variables are those described in Chapter 10. The results and their interpretation are discussed in comparison to the negative binomial model summarised in Table 11.1 in Chapter 11.

13.2 Unweighted regression

Indirectly standardised premature mortality rates were modelled using multiple linear regression, with ordinary least squares (OLS) as the method of estimating the parameters. The explanatory variables are summarised in Section 10.2. The results of unweighted multiple linear regression are shown in Table 13.1. Overall the direction of associations are similar to those found when count data are modelled using a negative binomial approach. There are two exceptions; in this model diabetes prevalence and
patents offered smoking cessation advice are now negatively associated with premature mortality, although the confidence interval for diabetes prevalence is wide and includes zero. The interpretation of these results would be different from those produced using negative binomial regression. The confidence intervals for white ethnicity now includes zero. In addition the number of GPs per 1000 patients would be considered a significant predictor, as would the percentage of patients offered smoking cessation advice. The measure of continuity of care would no longer be considered significant.

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Unweighted OLS</th>
<th>Negative binomial regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$ (95% CI)</td>
<td>$p$ value</td>
</tr>
<tr>
<td>Percentage white patients</td>
<td>0.22 (-0.43, 0.87)</td>
<td>0.500</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>2.43 (1.52, 3.33)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>-2.51 (-13.51, 8.48)</td>
<td>0.653</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>2.37 (0.094, 4.65)</td>
<td>0.041</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>9.13 (5.06, 13.21)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>42.29 (1.81, 82.77)</td>
<td>0.041</td>
</tr>
<tr>
<td>Hypertension detection 2006/07</td>
<td>-3.61 (-7.68, 0.47)</td>
<td>0.083</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>-1.34 (-2.60, -0.07)</td>
<td>0.039</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>-1.38 (-2.69, -0.07)</td>
<td>0.039</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>1.94 (-0.81, 4.69)</td>
<td>0.166</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>-0.77 (-1.67, 0.14)</td>
<td>0.097</td>
</tr>
</tbody>
</table>

Table 13.1: Estimated regression coefficients for unweighted linear regression showing the association between population and service characteristics and indirectly standardised premature (U75) CHD mortality rate ($R^2 : 0.38$).

Model diagnostics suggest that there are concerns with using unweighted OLS regression. There is evidence of heteroscedasticity in the variance of the residuals, variance increases as fitted values increase, this can be seen in Fig.13.1 (a) and more clearly in Fig.13.1 (b), in which the range of the $y – axis$ is limited, excluding a key outlier. An alternative approach is weighted least squares as a method of estimation.
Figure 13.1: Studentised residuals for unweighted regression.
13.3 Weighted least squares regression

OLS gives each practice equal weighting in the final model, but in this case it may not reasonable to assume that every observation should be treated equally. There is considerable uncertainty in the standardised mortality rates, particularly in smaller practices or in practices where the expected numbers of deaths are low for another reason, Fig.13.2 shows the distribution of the expected number of deaths across practices (for a three year period). The confidence intervals for the mortality rates are generally wide and in some cases they are extremely wide. The width of confidence intervals for 26% of practices is greater than 200 and for 12% it is over 300.

![Figure 13.2: Expected number of deaths in individual practices.](image)

Kiran et al (2010) used the inverse of the variance of the outcome ‘to account for the uncertainty in the standardised CHD admission and mortality rates’. One challenge with using an approach based on the inverse of the variance is that there are two practices which have zero observed deaths in the period of study and therefore have both an SMR and a variance of zero. The two practices in question are practice A identified in Chapter 10 as a practice with high influence and a patient list with unusual characteristics. The second is a relatively small practice (patient list approximately 3000 in each year). There are two key strategies which can be adopted; one is to
exclude these two observations from the model and a second is to substitute a variance for each practice. Substituting a variance for each practice is not straightforward. The practices are not similar and it is difficult to determine whether the practices should be treated in the same way. For example, practice A has no observed deaths and 0.70 expected deaths; the second practice has no observed deaths and 3.43 expected deaths.

Analysis was carried out using no substitution, excluding the two practices with no observed deaths, substituting the variance with a value of 0.1 (which is the median variance) assuming one observed death in each of the two practices. This final method results in a variance of 2.04 for practice A and 0.085 for the second practice. The results of these different weighting approaches are shown in Table 13.2.

The majority of $\beta$-coefficients are not sensitive to the method of substitution. However, the $\beta$ coefficients and confidence intervals for diabetes prevalence and percentage male are affected by the method of substitution. Diabetes prevalence is negatively associated with mortality if there is no substitution or the substitution is based on one observed death, but positively associated with mortality if the substitution is based on the median variance, although the confidence interval includes zero in all cases. The size of the association between percentage male is affected by substitution approach and the confidence interval includes zero when the approach is based on the median variance. One reason for the difficulty in determining the most appropriate substitution process is that one of the practices is a known outlier with high influence. Any weighting which increases the influence of this outlier the may mean the results are inferior to an unweighted least squares analysis. Honeyford et al (2013) adopted the median substitution approach; but it should be noted that the results and their interpretation are sensitive to the substitution approach.

**Comparison of weighted and unweighted least squares estimation**

The key differences between weighted and unweighted regression are the $\beta$ coefficients and confidence intervals for service characteristics. When there is no weighting GPs per 1000 is positively associated with mortality rates, and the confidence interval does not include zero, with weighted regression the confidence interval includes zero and therefore the evidence of any association is less clear cut. The percentage of patients offered smoking cessation advice is negatively associated with mortality rates when the regression is not weighted, but the association is positive and the confidence interval includes zero when the regression is weighted. When weighted regression is used the confidence interval for the percentage of people whose serum cholesterol is at or below
<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>No substitution, $n = 227$</th>
<th>Substitution $\text{var} = 0.1$, $n = 229$</th>
<th>Substitution $O = 1$, $n = 229$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Percentage white patients</strong></td>
<td>$\beta$ (95% CI)</td>
<td>$\beta$ (95% CI)</td>
<td>$\beta$ (95% CI)</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>0.18 (-0.27, 0.64)</td>
<td>0.36 (-0.08, 0.80)</td>
<td>0.18 (-0.27, 0.64)</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>-0.64 (-9.11, 7.82)</td>
<td>2.53 (-5.69, 10.75)</td>
<td>-0.72 (-9.20, 7.76)</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>2.00 (0.10, 3.89)</td>
<td>2.06 (0.13, 3.99)</td>
<td>2.00 (0.10, 3.91)</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>5.60 (1.49, 9.70)</td>
<td>2.95 (-0.78, 6.68)</td>
<td>5.23 (1.14, 9.33)</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>13.78 (-17.72, 45.27)</td>
<td>16.63 (-15.39, 48.64)</td>
<td>16.93 (-14.64, 48.51)</td>
</tr>
<tr>
<td>Hypertension detection 2006/07</td>
<td>-2.87 (-5.70, -0.04)</td>
<td>-2.74 (-5.61, 0.14)</td>
<td>-2.57 (-5.40, 0.27)</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>0.75 (-0.06, 1.56)</td>
<td>0.67 (-0.15, 1.50)</td>
<td>0.69 (-0.12, 1.51)</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>-0.08 (-0.99, 0.83)</td>
<td>-0.24 (-1.47, 1.95)</td>
<td>-0.03 (-0.94, 0.88)</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>0.12 (-1.56, 1.80)</td>
<td>0.24 (-1.47, 1.95)</td>
<td>0.21 (-1.48, 1.89)</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>-0.86 (-1.40, -0.32)</td>
<td>-0.88 (-1.43, -0.33)</td>
<td>-0.87 (-1.41, -0.33)</td>
</tr>
</tbody>
</table>

Table 13.2: Estimated regression coefficients for weighted linear regression showing the association between two CHD indicators and indirectly standardised premature (U75) CHD mortality rate.
target includes zero, suggesting the evidence for the importance of this service characteristic is not clear cut. However, the evidence that the percentage of people with recalled perception of being able to see their preferred GP is more clear cut when the regression is weighted.

Adjusted $R^2$ is 0.38 when unweighted least squares estimation is used; this reduces to 0.33 when weighted least squares estimation is used. This suggests that the variables included in the weighted model account for 33% of the variation in standardised mortality rates.

There is strong evidence that the percentage of patients aged over 65 is positively associated with premature mortality, when both OLS and weighted least squares are used, this is an interesting result as the mortality rates have been standardised by age and sex. This finding may provide evidence that practices with higher percentage of patients who are older are experiencing greater pressure on their practices and are therefore experiencing higher mortality rates. It is known that there is an increase in chronic conditions and multimorbidity in older people (Barnett et al., 2012) and this may create an ‘additional burden’ in some practices, leading to a lower levels of care (Anwar et al., 2012). Alternatively, it may provide evidence that the standardisation process is not satisfactory.

**Assessment of model fit**

One reason for using weighted regression is that it may help when there is evidence of heteroscedasticity in the residuals. Absolute residuals from weighted and unweighted regression are compared, as it is not straightforward to extract standardised residuals after weighted regression. Figure 13.3 shows all residuals and the presence of one major outlier makes it difficult to interpret whether there is a pattern in the residuals, Figure 13.4 allows an easier comparison, and shows that there is evidence that the heteroscedasticity of the residuals has been reduced when weighted regression is used, although there are still some concerns. The outlier labelled ‘I’ is a concern and was identified as having high influence in unweighted regression. Fitting the weighted regression with this practice deleted does not considerably improve the fit of the model and has low impact on the $\beta$ coefficients.

Analysis of residuals suggests that the residuals are not normally distributed. Tests for collinearity do not cause major concerns. No variable has a VIF of greater than four, and the mean value of two suggests supports this conclusion (Everitt and Rabe-
Hesketh, 2003, p63). Plots of residuals against predictor variables do not lead to major concerns that the relationship between predictor variables and SMR is not linear.

Figure 13.3: Absolute residuals of unweighted and weighted regression.

Figure 13.4: Absolute residuals of unweighted and weighted regression, y-axes limited to clarify patterns.
13.3.1 Weighted least squares regression compared to count models

The patterns of associations between weighted least squares and count models are the same for all covariates with the exception of diabetes prevalence. When a count model is used there is strong evidence of a positive association between diabetes prevalence and increased death count, in accordance with previous findings (Levene et al., 2010) and clinical evidence. However, when weighted least squares is used there is limited evidence of a negative association between diabetes prevalence and mortality rates; the relationship is not robust to substitution approaches.

The interpretation of the results are influenced by the modelling approach, although for the majority of variables the interpretations would have been similar. There is strong evidence of a positive association between increased death rates/counts and deprivation and the percentage of the patients aged 65 and over. Both modelling approaches provide clear cut evidence that increases in the percentage of patients who can recall being able to see their preferred GP are associated with decreased death counts/rates. Neither modelling approach provides clear cut evidence that GPs per 1000 patients, hypertension detection, the percentage offered smoking cessation advice or being treated with aspirin (where appropriate) are associated with decreases or increases in mortality counts/rates.

For two variables, the evidence for an association is dependent on the model. When a count model is used there is strong evidence that as the percentage of patients who are white increases the mortality count increases; the association is positive when weighted least squares is used although the confidence interval includes zero. When weighted least squares is used there is limited evidence that the percentage of patients whose cholesterol is at or below target is limited, the confidence interval for the $\beta$ coefficient includes zero. Whilst both Poisson model and negative binomial model suggest that this association is evident, when two key outliers are excluded from the model the evidence is more limited and the confidence interval for the IRR includes one.

The interpretation of the importance of both the percentage aged 65 and over and who are male is affected by the dependent variable. When the number of deaths is the dependent variable the importance of these two variables is easily explained by the higher risk of death from in older people and males. When an age and sex weighted mortality rate is the dependent variable the interpretation of the association is less straightforward. For example, the increased burden of older patients on general
practice may be important, or there may be an increased risk of CHD mortality in this area in comparison to England as a whole.

13.4 Use of log of SMR

An alternative approach to modelling standardised mortality rates is to model log(SMR) rather than SMR. The two practices which have zero observed deaths and therefore have an SMR of zero have been excluded from this analysis. Results are shown in Table 13.3.

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>$\beta$ (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage white patients</td>
<td>0.0040 (-0.0014, 0.0095)</td>
<td>0.143</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>0.024 (0.016, 0.031)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>0.029 (-0.063, 0.12)</td>
<td>0.532</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>0.027 (0.0039, 0.051)</td>
<td>0.023</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>0.068 (0.033, 0.12)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>0.17 (-0.17, 0.50)</td>
<td>0.334</td>
</tr>
<tr>
<td>Hypertension detection 2006/07</td>
<td>-0.048 (-0.082, -0.014)</td>
<td>0.006</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>0.0048 (-0.0058, 0.015)</td>
<td>0.372</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>-0.0062 (-0.017, 0.0048)</td>
<td>0.267</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>0.0030 (-0.020, 0.026)</td>
<td>0.798</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>-0.0085 (-0.016, -0.00097)</td>
<td>0.027</td>
</tr>
</tbody>
</table>

Table 13.3: Estimated regression coefficients for unweighted linear regression of log(SMR), $n = 227$ ($R^2 : 0.35$).

When these results are compared to weighted regression of 227 practices the results are very similar in terms of direction of association and interpretation of confidence intervals. Deprivation, percentage 65 and over, percentage male and the percentage of patients with recalled perception of being able to see preferred GP are robust to the modelling approach. The confidence interval for hypertension detection does not now include zero, suggesting that this variable is sensitive to the modelling approach.

A plot of absolute residuals provides evidence of homoscedasticity. Formal tests are inconclusive; White’s test (UCLA: Statistical Consulting Group, 2014) suggests that the null hypothesis of homoscedasticity should be rejected, whereas the Breusch-Pagan test (Fahrmeir et al., 2013, p.182) suggests the null hypothesis can be accepted. These
tests are sensitive to the assumption of Normality of the residuals, which is in doubt. Plots of residuals against predictor variables suggest that the relationship between predictor variables and log(SMR) is approximately linear. The concern that residuals are lower for larger practices is less evident when log(SMR) is the dependent variable in comparison to weighted least squares, although there is still some evidence of this pattern. There is some evidence that OLS of log(SMR) may be resulting in an improved fit when compared to weighted least squares. A comparison of fitted values to observed SMRs from both approaches suggests that fitted values derived from weighted least squares do not increase as rapidly as observed SMR; this concern does not appear to be as evident when OLS of log(SMR) is used; shown in Figure 13.6.

Figure 13.5: Absolute residuals for unweighted linear regression of log(SMR), n = 227.

13.5 Conclusions

Whilst multiple linear regression initially appears to be a straightforward approach, diagnostics suggest that unweighted least squares is not appropriate. Two practices which have zero deaths over the three year period result in challenges when the variance
Figure 13.6: Observed SMR compared to fitted values; comparing weighted least squares and OLS of log(SMR).

is used to weight the practices and when log(SMR) is the dependent variable. Overall results are broadly similar to those of count modelling approaches; but are sensitive to approaches to weighting and substitution when the SMR is zero. The significant association of the percentage aged over 65 and percentage male after mortality rates have been standardised suggests an area for further work.
Chapter 14

Multilevel Modelling Approach

14.1 Introduction

General practices collect considerable amounts of relevant information about their patients. However, data about some factors are not collected at practice level but may be available for other population groupings. For example, the Health Survey for England (National Centre of Social Research, 2012) collects information about lifestyles such as diet and exercise which are used to model estimates at district level. More recently the Integrated Household Survey (IHS) (ONS, 2014) has started collecting information about a range of factors at local authority district level and smaller areas. The Census collects detailed information for households which is aggregated to different organisational levels.

It may be useful to include these data in models which analyse health outcomes; multilevelling modelling may be a useful approach; allowing the hierarchical nature of the data to be taken into account by including variance at each level in the hierarchy. Multilevel modelling is relatively common in health services research in studies where variability between practices and patients is of interest and researchers have access to both practice- and patient-level data or in studies of neighbourhood effects, when patient-level data are available (for examples see Sundquist et al., 2004; Boggon et al., 2013). In addition, Ashworth et al (2013) utilised multilevel modelling to determine if practice characteristics were clustered at PCT level.

This chapter considers the use of multilevel modelling to include information available at local area level but not practice level. Practices can be seen to be located within PCTs or CCGs or within LAs which exist at different organisational levels. In this
chapter smoking prevalence estimated from the IHS, available at district level, is used as a case study. Premature CHD mortality is the health outcome of interest. Initially multilevel linear regression is considered followed by multilevel Poisson regression. Only population characteristics are considered.

14.2 Linear regression

The simplest one-level model includes an intercept but no explanatory variables:

\[ SMR_{ij} = \beta_0 + e_{ij}, \]

where

- \( SMR_{ij} \): standardised mortality ratio (SMR) of premature CHD mortality of practice;
- \( \beta_0 \): overall mean SMR across practices;
- \( e_{ij} \): error term;
- \( i \) and \( j \) index the \( ith \) practice in the \( jth \) district.

The estimate of \( \beta_0 \) is 107.73 (95% CI:(98.55, 116.91)), the estimate of variance of the error term is \( \hat{\sigma}^2 = 70.88^2 \) and the log-likelihood of the model is \(-1297.73\). Subsequent two-level models will be compared to this model.

14.2.1 Null model of premature CHD mortality rates, adding district effects.

Initially the simplest two level model, the null model, is considered. This allows for district effects on mortality without any explanatory variables.

\[ SMR_{ij} = \beta_0 + u_{0j} + e_{ij}, \]

where

- \( SMR_{ij} \): SMR of premature CHD mortality of practice
- \( \beta_0 \): overall mean SMR across practices
- \( u_{0j} \): the effect of district \( j \) on SMR (district level residual)
- \( e_{ij} \): practice level residual.
The model is fitted using STATA:

```
xtmixed CHD_U75_SMR_all || LT_LA_code:, mle
```

where \( CHD_U75_SMR_all \) is the premature CHD SMR and \( LT_LA \) denotes the district.

The overall mean SMR is estimated as \( \hat{\beta}_0 = 95.35 \), 95% CI: (80.20, 110.50). Estimates of between district variance (level 2) in SMR is estimated as \( \hat{\sigma}^2_{u0} = 64.64^2 \), and within district, between practice (level 1) variance as \( \hat{\sigma}^2_e = 23.33^2 \). The total variance is estimated from:

\[
\hat{\sigma}^2_{u0} + \hat{\sigma}^2_e = 64.64^2 + 23.33^2 = 4722.62.
\]

The variance can be partitioned between level 1, practices, and level 2, districts, by using the Variance Partition Coefficient (VPC) (Goldstein, 2011):

\[
VPC = \frac{\hat{\sigma}^2_{u0}}{(\hat{\sigma}^2_{u0} + \hat{\sigma}^2_e)}.
\]

The calculated VPC is 0.115, suggesting 11.5% of the variance in CHD mortality rate can be attributed to differences between districts, in the null model, with no explanatory variables.

To test the significance of district effects, this model is compared with a null single level model using a likelihood ratio test (LRT).

The log likelihood value for the null model with district (level 2) effects is \(-1287.03\) and for the null single level model is \(-1300.19\). The likelihood ratio statistic is therefore 26.32 on one degree of freedom, as there is one parameter difference between the models, suggesting there are district effects on CHD mortality rates.

### 14.2.2 Random intercept models: inclusion of explanatory variables at practice level.

**Model 1: two level model with explanatory variables**

Level 1 explanatory variables can be modelled. When the five population characteristics used in previous models (% white; % aged 65 and over; % male; diabetes prevalence (DM); and deprivation (IMD)) are included in the model, between district variance of \( u_{0j} \) is reduced to \( 5.97 \times 10^{-8} \), which is negligible. This suggests using a two level
model in this situation is inappropriate, as the amount of variance in CHD mortality
rates attributable to differences between districts approaches zero when population
characteristics at practice level are taken into account.

$$SMR_{ij} = \beta_0 + \beta_1(\%)_{ij} + \beta_2(\%)_{ij} + \beta_3(\%)_{ij} + \beta_4(\%)_{ij} + \beta_5(IMD)_{ij} + u_{0j} + e_{ij}.$$  

STATA command line:

```
xtmixed CHD_U75_SMR_all white 65plus male DM IMD || LT_LA_code:, mle
```

**Model 2: two level model with explanatory variables excluding deprivation**

To allow a thorough examination of the potential of using a multilevel model, a demonstra-
tion model including practice level explanatory variables but without deprivation
(IMD) has been developed (Model 2). In this case, including four level 1 explanatory
variables describing characteristics of the patients (% white; % aged 65 and over;
%male; and DM), reduces between district variance from 11.5% to 8.7%, summarised
in Table 14.1.

$$SMR_{ij} = \beta_0 + \beta_1(\%)_{ij} + \beta_2(\%)_{ij} + \beta_3(\%)_{ij} + \beta_4(\%)_{ij} + u_{0j} + e_{ij}.$$  

STATA command line:

```
xtmixed CHD_U75_SMR_all white 65plus male DM || LT_LA_code:, mle
```

### 14.2.3 Random intercept models: inclusion of explanatory variables at district level.

Including level 1 explanatory variables reduces between district variance to 8.7%. To
explore district variance, a level 2 explanatory variable may be included. This may be
considered a contextual variable and may be an average of level 1 variables. Here, a
measure of smoking prevalence in the local area is included as a level 2 variable due to
Table 14.1: Key results from different two level models.

<table>
<thead>
<tr>
<th></th>
<th>Null model</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>95.35 (80.20, 110.50)</td>
<td>-353.46 (-594.67, -112.26)</td>
<td>-323.29 (-587.97, -58.61)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>n/a</td>
<td>-0.034 (-0.65, 0.58)</td>
<td>-0.14 (-12.78, 7.16)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>n/a</td>
<td>0.56 (-1.77, 2.90)</td>
<td>-1.25 (-3.86, 1.35)</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>n/a</td>
<td>8.36 (4.44, 12.29)</td>
<td>9.13 (4.88, 13.39)</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>n/a</td>
<td>-7.63 (-16.74, 1.48)</td>
<td>-2.81 (-12.78, 7.16)</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>n/a</td>
<td>3.23 (2.39, 4.07)</td>
<td>n/a</td>
</tr>
<tr>
<td>$u_{0j}$</td>
<td>23.332</td>
<td>5.97 $\times$ 10$^{-8}$</td>
<td>18.642</td>
</tr>
<tr>
<td>$\epsilon_{ij}$</td>
<td>64.642</td>
<td>56.772</td>
<td>60.462</td>
</tr>
<tr>
<td>VPC</td>
<td>11.5%</td>
<td>1.8 $\times$ 10$^{-11}$%</td>
<td>8.7%</td>
</tr>
<tr>
<td>log likelihood</td>
<td>-1287.03</td>
<td>-1249.86</td>
<td>-1270.35</td>
</tr>
</tbody>
</table>

Model 3: two level model with explanatory variables excluding deprivation; including a random intercept (smoking)

$$SMR_{ij} = \beta_0 + \beta_1(\text{white})_{ij} + \beta_2(\text{65plus})_{ij} + \beta_3(\text{male})_{ij} + \beta_4(\text{DM})_{ij} + \beta_6(\text{smoking})_{j} + u_{0j} + \epsilon_{ij}$$ (14.1)

STATA command line:

```stata
xtmixed CHD_U75_SMR_all white 65plus male DM IMD level2_sm || LT_LA_code:, mle
```

Including district level smoking in a multilevel model of premature CHD mortality reduces between district variance from 8.7% to 2.7% (Table 14.2); suggesting that 69% of between district variance can be explained by differences in smoking prevalence in this demonstration model, which omits deprivation. Estimates from subsequent models should not be interpreted as meaningful.

14.2.4 Random intercept and random slopes.

Models 1, 2 and 3 assume the relationship between premature CHD mortality and explanatory variables is the same for each district, that is $\beta_1$, $\beta_2$, $\beta_3$ and $\beta_4$ are fixed.
Table 14.2: Key results from different two level models - inclusion of level 2 smoking variable.

across districts. The slopes can be allowed to vary across districts. Initially, the association between gender and premature CHD mortality is allowed to vary across districts. Gender has been selected because the confidence interval for the \( \beta \) coefficient for patients who are male does not include zero, this is not the case for the \( \beta \) coefficients for the remaining level 1 explanatory variables, see Model 4a. Given the priority of diabetes in Leicester City and Leicestershire (University of Hospitals of Leicester, 2011), diabetes has been selected as a second example of a variable varying across districts, see Model 4b.

Model 4a - two level model with explanatory variables excluding deprivation; including a random intercept (smoking) and random slope (gender)

\[
SMR_{ij} = \beta_0 + \beta_1(\%white)_{ij} + \beta_2(\%65plus)_{ij} + \beta_3(\%male)_{ij} + \beta_4(\%DM)_{ij} + \beta_6(\%smoking)_j + u_{1j}(\%male)_{ij} + u_{0j} + e_{ij}
\]

(14.2)

STATA command line:

```
xtmixed CHD_U75_SMR_all white 65plus male DM IMD level2_sm || LT_LA_code: male, mle
```

An LRT is useful in comparing models. In this case the log likelihood statistic for Model 4a is -1248.30, this compares with the log likelihood statistic for Model 3: -1266.35. The likelihood ratio statistic is 36.10 which is above the 5% point of the \( \chi^2 \)
distribution on two degrees of freedom (5.99); this is evidence that the gender effect differs across districts.

**Model 4b - two level model with explanatory variables excluding deprivation; including a random intercept (smoking) and random slope (diabetes)**

\[
\text{SMR}_{ij} = \beta_0 + \beta_1(\%\text{white})_{ij} + \beta_2(\%65plus)_{ij} + \beta_3(\%\text{male})_{ij} + \beta_4(\%DM)_{ij} + \beta_6(\%\text{smoking})_j + u_{1j}(\%DM)_{ij} + u_{0j} + e_{ij}
\]

(14.3)

**STATA command line:**

```
xtmixed CHD_U75_SMR_all white 65plus male DM IMD level2_sm || LT_LA_code: DM, mle
```

The log likelihood for this Model 4b is -1262.56, comparing with Model 3, the likelihood ratio statistic is 7.58, providing evidence that diabetes effects vary between districts.

Table 14.3 shows that for some districts the relationship between diabetes prevalence and premature mortality is positive, that is increased prevalence is associated with an increase in mortality; this is the case for Charnwood, Daventry, Kettering, Hinkley and Bosworth and Harborough. However, in other districts the association is negative, an increase in diabetes prevalence is associated with a decrease in mortality. This is particularly the case for Northampton, Wellingborough, North West Leicestershire, Blaby, Leicester and Melton.

VPC has not been calculated for Model 4a and 4b; when random slopes are included in multilevel models the VPC becomes a function of the predictor variables and its calculation and interpretation are more complex (Goldstein et al., 2002). This would be a useful area for further work if the model was realistic and informative.

### 14.2.5 Summary of multilevel linear regression

Allowing for district level random effects by using a multilevel model does not improve the model when practice level demographic variables including socio-economic deprivation, as measured by IMD, are included in the model. When IMD is excluded for demonstration purposes, a multilevel model may be a useful approach when practice level variables are not available. There is a significant positive relationship between
Table 14.3: Estimate of district level slope for diabetes.

district level smoking prevalence and premature CHD SMRs. The variance in SMR attributable to district level reduces from 8.7% to 2.7% when district level smoking prevalence is included in the model. An advantage of multilevel models is the possibility that the association between explanatory variables and mortality rates may vary across districts. There is evidence that the association between diabetes prevalence and premature CHD mortality rates varies between districts (without IMD in the model). This is a possible area for further research, considering explanations such as primary care or demographic features such as deprivation and ethnicity.
14.3 Poisson regression

14.3.1 Null model of premature CHD mortality COUNTS, with district effects

As previously described (Chapter 10), mortality data are counts and should be modelled appropriately, for example using a multilevel Poisson model. Using a similar approach to that used above, initially the simplest two level model, the null model is considered:

**Null model (Poisson)**

\[
\text{Log(No.of deaths)} = \beta_0 + u_{0j}.
\]

There is no level 1 residual variance \( (e_{ij}) \); the variance depends on the expected count in Poisson models.

Stata command line:

```
xtmepoisson CHD_U75_count || LT_LA_code:, exposure(_n_U75) covariance(unstructured)
```

The LRT statistic comparing this model with Poisson regression without the possibility of district level variance is 27.02, providing strong evidence that the between district variance is non-zero \( (p < 0.0001) \). The variance of \( u_{0j} \) is 0.02. However, as there is no residual variance at practice level it is not possible to estimate the VPC in the same way as with OLS.

14.3.2 Partitioning variance in multilevel count models

Although we do not have level 1 residual variance, it would be useful to be able to quantify the variance in mortality count attributable to differences at district level. Often \( \sigma_u^2 \approx 0 \) is interpreted as no variance at district level and when \( \sigma_u^2 \) is large this is interpreted as lots of variance at district level; it would be useful to quantify \( \sigma_u^2 \) more systematically (Stryhn et al., 2006).

The VPC in linear multilevel models is used to describe the proportion of variance explained at different levels (Goldstein, 2011):
\[
VPC = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_e^2}.
\]

An alternative measure is the intraclass correlation (ICC); the correlation between two observations in the same level 2 unit (district). In linear models the ICC and VPC are equivalent (Merlo et al., 2005).

VPC depends on both level 1 and level 2 variance; in discrete response models, including count models, the variance depends on the expected value, which is a function of the predictors, and a simple VPC is not available (Merlo et al., 2006). Goldstein et al. (2002) describe three methods for partitioning variance in multilevel models.

- model linearisation
- simulation method
- normal distribution model

Stryhn et al (2006) argue an exact method of calculating ICC is preferable for Poisson models. The ICC for two observations with common predictors within the same level 2 unit (district) is given by:

\[
ICC = \frac{\sigma_{(2)}^2}{\sigma_{(1)}^2 + \sigma_{(2)}^2},
\]

where

\[
\sigma_{(2)}^2 = \exp(2X\beta + 2\sigma_{u0}^2) - \exp(2X\beta + \sigma_{u0}^2)
\]

\[
\sigma_{(1)}^2 = \exp(X\beta + \frac{\sigma_{u0}^2}{2}).
\]

ICCs are calculated using modelled linear predictors for a particular set of practices. The estimates of ICCs for the same practices can be compared.

Adopting Stryhn’s exact method, ICCs are calculated for the null model. The estimate of \(\sigma_{u0}^2\) for the null model is 0.02. Estimates of the linear predictor \((X\beta)\) range from 0.66 to 3.73. Here, the ICC is calculated for three practices, in this case practices with linear predictor at the LQ, median and UQ. Table 14.4 summarises this information. The intraclass coefficients are the same order of magnitude to those estimated by OLS, where the null model estimates that 11.5% of the variance is attributed to between district variance. As expected, ICC increases as the linear predictor increases, due to the direct link between the linear predictor and the variance (Merlo et al., 2006).
Table 14.4: Estimate of ICC for three practices.

<table>
<thead>
<tr>
<th>Practice</th>
<th>District</th>
<th>linear predictor</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>84310</td>
<td>Northampton</td>
<td>1.54</td>
<td>0.087</td>
</tr>
<tr>
<td>83256</td>
<td>Harborough</td>
<td>2.056</td>
<td>0.138</td>
</tr>
<tr>
<td>83273</td>
<td>Blaby</td>
<td>2.532</td>
<td>0.204</td>
</tr>
</tbody>
</table>

14.3.3 Random intercept models: inclusion of explanatory variables at practice level

Model $1_{\text{poiss}}$: two level model with explanatory variables

$$\log(\text{No.of deaths}) = \beta_0 + \beta_1(\%\text{white})_{ij} + \beta_2(\%65\text{plus})_{ij} + \beta_3(\%\text{male})_{ij} + \beta_4(\%\text{DM})_{ij} + \beta_5(\text{IMD})_{ij} + u_{0j}$$ (14.4)

When five key demographic practice level (level 1) explanatory variables are included in the model the estimate of the variance of $u_{0j}$ reduces to 0.0068. Table 14.5 shows the estimates of ICCs for the same three practices reported in Table 14.4; the ICCs are considerably lower. In line with the results of linear multilevel regression with explanatory variables in the model the proportion of variance attributable to district level is reduced.

Model $2_{\text{poiss}}$: two level model with explanatory variables excluding deprivation

To mirror the approach taken with OLS, a random intercept model with practice level explanatory variables, but excluding socio-economic deprivation is fitted to the data.

$$\log(\text{No.of deaths}) = \beta_0 + \beta_1(\%\text{white})_{ij} + \beta_2(\%65\text{plus})_{ij} + \beta_3(\%\text{male})_{ij} + \beta_4(\%\text{DM})_{ij} + u_{0j}$$ (14.5)

The variance of $u_{0j}$ is 0.026, greater than when the null model is fitted. Estimates of ICCs increase to values similar to those of the Null model (Table 14.5), suggesting that when deprivation is excluded from the model there is still variance at district level.
14.3.4 Random intercept models: inclusion of explanatory variables at district level - Poisson Model

Smoking at district level (level 2) is included in the model to explore possible reasons for district level variance.

Model $3_{poiss}$: two level model with explanatory variables excluding deprivation; including a random intercept (smoking)

$$
\log(\text{No. of deaths}) = \\
\beta_0 + \beta_1(%\text{white})_{ij} + \beta_2(%\text{65plus})_{ij} + \beta_3(%\text{male})_{ij} + \beta_4(%\text{DM})_{ij} + \\
\beta_6(%\text{smoking})_j + u_{0j} + e_{ij}
$$

When district level smoking prevalence is included in the model there is a significant positive association between smoking level and CHD mortality count. The estimate of the variance of $u_{0j}$ reduces to 0.015, suggesting that including smoking prevalence at district level accounts for some of the between district variance. A comparison of log likelihood statistics gives an LRT statistic of 6.9 suggesting that the inclusion of the district level smoking data improves the model. The estimates of ICCs reduce for each of the three practices: 84310 - 0.065; 83256 - 0.106 and 83273 - 0.156. These reductions provide further evidence that district level smoking prevalence accounts for some of the between district variance.

14.3.5 Random intercept and random slopes - Poisson Model

In parallel with multilevel linear regression, two additional models are considered which allow the effects of gender and diabetes prevalence to vary across districts:

<table>
<thead>
<tr>
<th>Practice</th>
<th>District</th>
<th>Null Model</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_{u0j}$ variance</td>
<td>0.020</td>
<td>0.0068</td>
<td>0.026</td>
</tr>
<tr>
<td>84310</td>
<td>Northampton</td>
<td>0.087</td>
<td>0.028</td>
<td>0.089</td>
</tr>
<tr>
<td>83256</td>
<td>Harborough</td>
<td>0.138</td>
<td>0.051</td>
<td>0.172</td>
</tr>
<tr>
<td>83273</td>
<td>Blaby</td>
<td>0.204</td>
<td>0.067</td>
<td>0.227</td>
</tr>
</tbody>
</table>

Table 14.5: Estimate of ICC for three practices for three Poisson models.
Model $4_{a_{poiss}}$

\[
\log(\text{No.of deaths}) = \beta_0 + \beta_1(\%\text{white})_{ij} + \beta_2(\%\text{65plus})_{ij} + \beta_3(\%\text{male})_{ij} + \beta_4(\%\text{DM})_{ij} + \\
\beta_6(\%\text{smoking})_j + u_{1j}(\%\text{male})_{ij} + u_{0j}
\]  

(14.7)

log likelihood = -602.05

Model $4_{b_{poiss}}$

\[
\log(\text{No.of deaths}) = \beta_0 + \beta_1(\%\text{white})_{ij} + \beta_2(\%\text{65plus})_{ij} + \beta_3(\%\text{male})_{ij} + \beta_4(\%\text{DM})_{ij} + \\
\beta_6(\%\text{smoking})_j + u_{1j}(\%\text{DM})_{ij} + u_{0j}
\]  

(14.8)

log likelihood = -603.60

An LRT tests whether the association between gender and premature mortality count varies across districts; comparing the log likelihood with Model $3_{poiss}$ (-606.83) suggests there is evidence to support the more complex model. This is also the case for the inclusion of a random slope for district level for diabetes prevalence.

When diabetes prevalence is included in the model the confidence interval for $\beta_4$ - the coefficient for level 1 diabetes prevalence includes 1. When compared to the estimate $u_1$ values for multilevel linear regression there are similarities. Kettering, Daventry and Charnwood have a positive association between diabetes prevalence and premature CHD mortality count, compared to Wellingborough, and Northampton, which have a negative association between diabetes prevalence and premature CHD mortality count. However, Northampton is a less extreme result and district level diabetes effect for Leicester is positive rather than negative.

14.3.6 Summary of multilevel Poisson regression

Results are similar to multilevel linear regression although the coefficients for explanatory variables are different. Interpretation of the district level random effects is more difficult as VPC cannot be calculated. The calculation of ICC as an alternative to VPC is relatively simple, but the interpretation is challenging.
14.4 Conclusions

Multilevel modelling is an important methodology both when data are unavailable at practice level and to consider the impact of explanatory variables at higher level geographical groupings, providing possible areas for future research. Results of linear multilevel regression including VPCs are simple to interpret. ICCs can be used as an alternative for non-linear models and are relatively simple to calculated using an exact method; their interpretation can be difficult.

It has previously been shown negative binomial models are more appropriate, owing to overdispersion. Multilevel negative binomial models have not been considered here; this is an area for further work.

When socio-economic deprivation is included in the model the proportion of variance attributable to district level is negligible; when excluded from the model, potentially interesting district level associations can be considered.
Chapter 15

Estimating smoking prevalence in general practices: an evaluation of data from the Quality and Outcomes Framework (QOF).

15.1 Introduction

The lack of information about smoking prevalence in practice populations, described in Chapter 6, has led to a variety of measures being used in research (Soljak et al., 2011; Purdy et al., 2011) and some research omitting smoking prevalence (Bankart et al., 2011; Kiran et al., 2010; Honeyford et al., 2013), despite its recognised importance (Twigg et al., 2004). The addition of a new indicator in the 2012/13 QOF has led to a novel approach to estimating smoking prevalence being developed. Data were published in October 2013.

The national Quality and Outcomes Framework (QOF) has been described in Chapter 7. Since its inception in 2004, QOF has included indicators relating to smoking. The underlying aim of these indicators has not changed over the years: a) practices should record smoking status in patient notes and, b) for those who smoke, smoking cessation advice/support/treatment should have been offered. In 2012/13 indicators covered the total practice population for the first time; until this point the focus was on targeting smoking cessation advice to those with chronic conditions.
The QOF indicators have not been designed to determine smoking prevalence within the practice population; indeed it is clearly stated that *QOF provides no information on numbers of smokers and non-smokers* (HSCIC, 2014d). This limitation is mainly attributed to the condition-specific nature of the indicator. The wording has not changed since the inclusion of the two new indicators which apply to the general population and are not condition-specific.

In this chapter the extent to which underlying data published as part of QOF can be used to estimate smoking prevalence within practice populations is evaluated. The impact of including these estimates in a model considering associations between premature CHD mortality and population service characteristics is discussed.

### 15.2 Summary of QOF smoking indicators

In 2003/04, 2004/05 and 2005/06 there were smoking indicators within the following disease areas: CHD; stroke; hypertension; diabetes; COPD and asthma. The indicators followed the format shown below for CHD:

- **CHD03** The percentage of patients with coronary heart disease whose notes recorded smoking status in the past 15 months, except those who have never smoked where smoking status needs to be recorded only once.
- **CHD04** The percentage of patients with coronary heart disease who smoke, whose notes contain a record that smoking cessation advice has been offered within the last 15 months.

In 2006/07 a smoking domain was added to QOF, replacing the smoking indicators within the separate disease areas. Table 15.1 summarises smoking indicators which have been included in QOF since 2006/07.
### General form of indicator

| % of patients whose notes record smoking status<sup>2</sup> | SM01: 2006/07 & 2007/08  
SM03: 2008/09 to 2011/12  
SM05: 2012/13  
Records 23: 2008/09 to 2011/12  
SM07: 2012/13  
SM001: 2013/14 (retired in 2014/15) |
| % of patients who are recorded as current smokers whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered within the previous 15 months<sup>3</sup> | SM02: 2006/07 & 2007/08  
SM04: 2008/09 to 2011/12  
SM06: 2012/13  
SM005: 2013/14 to 2014/15 | SM08: 2012/13  
SM004: 2013/14 to 2014/15 |
| The practice supports smokers in stopping smoking by a strategy which includes providing literature and offering appropriate therapy. | Information 5: 2006/07 to 2011/12  
SM003: 2012/13 to 2014/15 | |

<sup>1</sup>In 2008/09 CKD, asthma, schizophrenia, bipolar affective disorder or other psychoses were added to the list of chronic conditions and in 2012/13 PAD was added.

<sup>2</sup>For those with chronic conditions, the record must have been made in the past 15 months, reduced to 12 months in 2013/14, for all patients the period is 27 months, reduced to 24 months in 2013/14.

<sup>3</sup>In 2012/13 this changed to who have a record of an offer of support and treatment within the preceding 15 months, the period is 27 months for all patients, reduced to 12 months and 24 months respectively in 2013/14.

Table 15.1: Summary of smoking indicators for which underlying achievement is published; QOF guidance is available from www.nhsemployers.org (2014).
15.3 Estimating smoking prevalence from QOF data

In order to estimate smoking prevalence, both the number of people who smoke and the population from which this number is drawn must be known. Two key QOF indicators are used in the calculations of smoking prevalence in the total practice population:

- SM07 The percentage of patients aged 15 years and over whose notes record smoking status in the preceding 27 months.
- SM08 The percentage of patients aged 15 years and over who are recorded as current smokers who have a record of an offer of support and treatment within the preceding 27 months.

These can be summarised as follows:

\[
\text{Smoking status indicator (SM07)} = \frac{\text{No. of patients who have their smoking status recorded}}{\text{No. of eligible patients in the practice}}
\]

(15.1)

\[
\text{Smoking cessation indicator (SM08)} = \frac{\text{No. of patients who have a record of cessation support}}{\text{No. of patients recorded as current smokers}}
\]

(15.2)

The denominator of the smoking status indicator (SM07) provides an estimate of the sample of the practice population whose smoking status should be recorded. This includes the whole practice population aged over 15, with the exception of people who have joined the practice in the three months prior to the data extraction point and patients who refuse to provide their smoking status. The denominator of the smoking cessation indicator (SM08) provides an estimate of those who are recorded as current smokers.

Using the data given for these indicators, available from the Health and Social Care Information Centre (HSCIC, 2014e), it is possible to estimate the smoking prevalence in a practice population. Table 15.2 gives worked examples of how the calculations. For example, for practice A the denominator for SM07 is 3721 the number of people for whom smoking status should be recorded. The denominator for SM08 is 1129 - indicating that there are 1129 registered patients recorded as smokers. Hence smoking prevalence can be estimated as 1129/3721 or 30.3%, see Table 15.2. This method was
used to estimate smoking prevalence for the total practice population in 2013/14.

The percentage of those with chronic conditions who are smokers can be determined in a similar way using indicators SM05 and SM06. Using these indicators the smoking prevalence in those with chronic conditions was calculated from 2006/07 to 2013/14.
### Example practices

<table>
<thead>
<tr>
<th>QOF description</th>
<th>Interpretation for purposes of calculating smoking prevalence</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM07 Points</td>
<td></td>
<td>11</td>
<td>10.5</td>
<td>10.8</td>
<td>9.6</td>
<td>11</td>
</tr>
<tr>
<td>SM07 Numerator</td>
<td>Patients(^1) whose notes contain a record of smoking status</td>
<td>3450</td>
<td>1319</td>
<td>6276</td>
<td>31948</td>
<td>6504</td>
</tr>
<tr>
<td>SM07 Denominator</td>
<td>Patients who are eligible to be included in this indicator(^2)</td>
<td>3721</td>
<td>1497</td>
<td>7033</td>
<td>37654</td>
<td>7212</td>
</tr>
<tr>
<td>SM07 UA</td>
<td></td>
<td>92.70%</td>
<td>88.10%</td>
<td>89.20%</td>
<td>84.80%</td>
<td>90.20%</td>
</tr>
<tr>
<td>SM08 Points</td>
<td></td>
<td>12</td>
<td>9.9</td>
<td>12</td>
<td>8.9</td>
<td>12</td>
</tr>
<tr>
<td>SM08 Numerator</td>
<td>Patients who are recorded as current smokers and have a record of an offer of support etc</td>
<td>1024</td>
<td>325</td>
<td>1578</td>
<td>8439</td>
<td>2165</td>
</tr>
<tr>
<td>SM08 Denominator</td>
<td>Patients who are recorded as current smokers</td>
<td>1129</td>
<td>401</td>
<td>1586</td>
<td>10931</td>
<td>2373</td>
</tr>
<tr>
<td>SM08 UA</td>
<td></td>
<td>90.70%</td>
<td>81.00%</td>
<td>99.50%</td>
<td>77.20%</td>
<td>91.20%</td>
</tr>
<tr>
<td>Calculation to determine percentage who are smokers</td>
<td>SM08 den/SM07 den</td>
<td>1129/3721</td>
<td>401/1497</td>
<td>1586/7033</td>
<td>10931/37654</td>
<td>2373/7212</td>
</tr>
<tr>
<td>Estimate of smoking prevalence</td>
<td></td>
<td>30.30%</td>
<td>26.80%</td>
<td>22.60%</td>
<td>29.00%</td>
<td>32.90%</td>
</tr>
</tbody>
</table>

\(^1\) Patients aged over 15.
\(^2\) For example patients who are newly registered with a practice (< three months) are excluded from the indicator.

Table 15.2: Example of QOF data from 2012/13, showing how it can be used to calculate smoking prevalence for individual practices. All data available from HSCIC (HSCIC, 2014e)
15.4 Estimates of smoking prevalence in practice populations.

The nature of the QOF indicators means it is only possible to estimate smoking prevalence in the general practice population for 2012/13. Smoking prevalence in those with chronic conditions can be calculated for the years 2006/07 to 2012/13. Smoking prevalence was calculated for 215 practices; these practices had QOF data available for seven financial years (2006/07-2012/13). The median estimate of smoking prevalence in practice populations was 19.2%, ranging from 5.8% to 43.0% (IQR: (15.1%, 22.9%)). In the same period median estimated smoking prevalence in those with chronic conditions was 15.4% (IQR: 12.6% to 19.4%), ranging from 7.1% to 51.5%. Smoking prevalence in those with chronic conditions was lower than the practice population for 195 practices (91%). Patterns in smoking prevalence in those with chronic conditions was has been similar over seven year period, shown in Table 15.3.

<table>
<thead>
<tr>
<th>Year</th>
<th>Median (IQR)</th>
<th>(min, max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006/07</td>
<td>15.05 (12.53, 18.56)</td>
<td>(4.28, 44.58)</td>
</tr>
<tr>
<td>2007/08</td>
<td>14.81 (12.23, 18.72)</td>
<td>(4.94, 48.52)</td>
</tr>
<tr>
<td>2008/09</td>
<td>15.40 (12.71, 19.33)</td>
<td>(5.62, 53.82)</td>
</tr>
<tr>
<td>2009/10</td>
<td>15.88 (12.77, 19.29)</td>
<td>(7.08, 53.09)</td>
</tr>
<tr>
<td>2010/11</td>
<td>15.62 (12.72, 19.61)</td>
<td>(6.15, 53.10)</td>
</tr>
<tr>
<td>2011/12</td>
<td>15.38 (12.69, 19.93)</td>
<td>(6.08, 52.10)</td>
</tr>
<tr>
<td>2012/13</td>
<td>15.45 (12.56, 19.37)</td>
<td>(7.12, 51.51)</td>
</tr>
</tbody>
</table>

Table 15.3: Smoking prevalence (%) in those with chronic conditions (n = 215).

15.5 Usefulness of estimates

15.5.1 Comparison with other sources of data

The Integrated Household Survey (IHS) is the most useful source of data for comparisons and is published at local authority district level for 2011/12 (Department of Health, 2012). Data at practice level have been combined to create district level QOF-based estimates based on the post code of the practice to allow comparisons with IHS estimates.

Estimates of smoking prevalence were in line with estimates derived from the IHS. Aggregating over the total area, smoking prevalence was 19.5%, compared to 19.3%
when IHS district level data were aggregated over the same area. When practice data were combined to give estimates of smoking for local authority districts, there was good agreement (mean difference: 0.39%; 95% limits of agreement (-3.77, 4.55)) between estimates based on QOF registers and IHS estimates (Figure 15.1). The good level of agreement suggests that estimates based on QOF registers are useful estimates of smoking prevalence within practice populations. The potential utility of QOF registers to estimate smoking prevalence in local areas is briefly described by Honeyford et al (2014).

Figure 15.1: Relationship between aggregated QOF estimates and IHS estimates for local authority districts (QOF estimates based on 2012/13 data; IHS estimates based on 2011/12 survey).

15.5.2 Associations between recording of smoking status and prevalence

There was a strong positive correlation between recording of smoking status in the general population and in those with chronic conditions (underlying achievement for SM07 and SM05 respectively) ($R_p = 0.74$, $p < 0.0001$). There was no evidence of an association between smoking prevalence in the general population and recording of smoking status ($R_p = -0.07$, $p = 0.28$) or the percentage with a chronic condition ($R_p = 0.03$, $p = 0.67$).
15.5.3 Associations between smoking prevalence in the general practice populations and those with chronic conditions

There is a strong correlation between estimates of smoking prevalence in the general population and in those with chronic conditions ($R_p = 0.91, p < 0.0001$). The relationship suggested by linear regression is shown below:

$$SMOKING_{general} = 3.30 + 0.99(SMOKING_{chronic})$$

This suggests that smoking prevalence in those with chronic conditions is a useful proxy for smoking prevalence in the general population in 2012/13. There is high concordance between estimates of smoking prevalence in those with chronic conditions for all years; Lin’s concordance coefficient was greater than 0.92 and mean difference was less than one in all cases. This suggests that smoking prevalence in those with chronic conditions is likely to be a useful proxy for smoking prevalence in the general population in previous years.

15.6 Including smoking prevalence estimates in models of mortality

The estimate of smoking prevalence in those with chronic conditions was included in the ‘initial model’ described in Chapter 11. Table 15.4 shows incident rate ratios (IRRs), 95% CIs and associated $p$-values for the original and modified models. Inclusion of the smoking prevalence variable in the model reduced the strength of the associations between deprivation and premature mortality, and percentage white and premature mortality. A one unit increase in smoking prevalence was associated with an increase of 2.9% in expected premature CHD mortality count. If a practice with a moderately high smoking prevalence (75th percentile: 18.86%) is compared to one with a median level of smoking prevalence (15.09%), a difference of 10.93% in premature CHD mortality count can be expected, after adjusting for the other variables in the model.
15.7 Discussion

These results show how QOF registers can be used to estimate smoking prevalence in practice populations and that these estimates are useful when analysing patterns of mortality.

When smoking prevalence is estimated in the general population using QOF indicators there is good agreement with estimates of IHS smoking prevalence for similar geographical areas. QOF data can also be used to estimate smoking prevalence in those with chronic conditions, which is generally lower than smoking prevalence in the general population. There is good agreement between the estimates in successive years. The correlation between estimates of smoking prevalence in the general population in 2012/13 and those with chronic conditions is strong. These strong correlations suggest that the estimates based on previous years can be used in place of smoking prevalence in the general population in models where the aim is to correct for the effect of smoking prevalence. Regression analysis suggests that smoking prevalence in those with chronic conditions can be used to predict smoking prevalence in the general practice population, for practices with a typical patient list; these predictions could be used if estimates of the effects of smoking were sought.

When an estimate of smoking prevalence in those with chronic conditions was used in a study of the association between premature CHD mortality and various population and service characteristics, an important positive association between CHD mortality and smoking prevalence was shown, this is discussed in more detail in Chapter 18.

Strengths and weaknesses

The agreement between IHS based area estimates of smoking prevalence and estimates based on combining QOF data suggests that manipulating QOF data results in a useful measure of smoking prevalence within practice populations, in parallel with Szatkowski et al.’s (2012) finding of good agreement between national smoking prevalence predicted by patient records and the General Household Survey.

Since the percentage of patients who do not have their smoking status recorded varies from 40% to less than 1% and the characteristics of these patients are not known the estimates obtained may vary for artifactual reasons since men, young people and those without chronic conditions are less likely to have their smoking status recorded (HSCIC, 2004; Simpson et al., 2010; Wong et al., 2012). However, the analysis presented here did
not find an association between, for example, the percentage with a chronic condition and the recording of smoking status in the total population or the estimate of smoking prevalence.

Since practices are only asked to record smoking status in the preceding 27 months, estimates may have limited for assessing the effectiveness of interventions, unless practices commit to more regular recording.

Although there is a strong association between smoking prevalence in the general practice population and those with chronic conditions in 2012/13, it is not known whether this was the case in earlier years, so some caution in using earlier estimates is appropriate.

Practice level data have been aggregated to local authority districts based on practice postcode. General practice catchments are not constrained by local authority boundaries. It ha been shown that 80% of patients live within a 10 minute car journey of their general practice (Haynes et al., 2003), suggesting that patients choose practices close to where they live. It is relatively common for practice postcodes to be used as a proxy for patient postcodes; when used to estimate deprivation they have been found to underestimate relationships between deprivation and health outcomes (Strong et al., 2006; McLean et al., 2008). Further work using individual patient records is necessary to analyse the frequency of recording of smoking status and the characteristics of patients for whom no smoking status is recorded. Sensitivity analysis considering the impact of exception reporting indicates no impact on interpretation.

15.8 Conclusions

Data published through QOF allows useful estimations of smoking prevalence within practice populations and in those with chronic conditions to be made. There is strong evidence to support the use of estimates of smoking prevalence in those with chronic conditions as a proxy for the general population when modelling health outcomes. Given the known association between smoking and CHD, the model considered in Chapter 10 is improved by the inclusion of an estimate of smoking prevalence.
<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Without smoking prevalence</th>
<th>With smoking prevalence</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR</td>
<td>95% CI</td>
<td>p value</td>
</tr>
<tr>
<td>Percentage white patients</td>
<td>1.008</td>
<td>(1.003, 1.012)</td>
<td>0.002</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.017</td>
<td>(1.011, 1.024)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>1.114</td>
<td>(1.028, 1.208)</td>
<td>0.009</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>1.059</td>
<td>(1.038, 1.081)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>1.067</td>
<td>(1.038, 1.103)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>1.197</td>
<td>(0.885, 1.619)</td>
<td>0.244</td>
</tr>
<tr>
<td>Hypertension detection 2006/07</td>
<td>0.978</td>
<td>(0.950, 1.007)</td>
<td>0.133</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>1.002</td>
<td>(0.993, 1.011)</td>
<td>0.712</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>0.991</td>
<td>(0.981, 1.000)</td>
<td>0.044</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>1.002</td>
<td>(0.982, 1.022)</td>
<td>0.884</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>0.994</td>
<td>(0.989, 1.000)</td>
<td>0.036</td>
</tr>
<tr>
<td>Smoking prevalence estimate: QOF 2006/07</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Table 15.4: IRRs, 95% confidence intervals and p-values for premature mortality - inclusion of smoking prevalence estimate (after fitting a Negative Binomial Model model).
Chapter 16

Consideration of alternative explanatory variables

In this chapter various additional and alternative explanatory variables are considered. The model described in Chapter 10 includes explanatory variables based on previous work, predominantly the work of Levene et al (2010; 2012) and Kiran et al (2010). Three explanatory variables explore measures of obesity (Section 16.1), hypertension detection (Section 16.2) and measures of access and continuity of care (Section 16.3). Estimated smoking prevalence has been included in all models as evidence from Chapter 15 suggests inclusion results in an improved model.

16.1 Obesity

QOF obesity registers are the only source of information about obesity prevalence in practice populations. Section 6.4.2 discusses the QOF obesity register and concludes that the registers are not a useful measure of obesity as there is evidence that the registers under-estimate obesity prevalence. Modelled estimates of obesity prevalence are available at MSOA level (APHO, 2010). These measures are included in the model presented in Chapter 11, with the addition of smoking prevalence.

Results

When QOF obesity prevalence is included in the model there is no evidence of improved model fit. There is evidence of a negative association between obesity prevalence as
measured through QOF registers; the incident rate ratio is 0.986. This negative association suggests that QOF obesity registers may be a measure of obesity detection than prevalence. However, the confidence interval is wide and includes one (95%CI:(0.958, 1.014)) and the evidence is therefore not clear cut.

Practice postcodes were linked to MSOAs and modelled estimates of obesity prevalence were included in the model. Again, there is no evidence of improved model fit. In this case the association is positive, although the confidence interval is wide and includes one (IRR:1.005; 95%CI:(0.989, 1.022)). The positive association between the modelled estimates and premature mortality further supports the suggestion that QOF registers are a measure of obesity detection rather than prevalence.

Inclusion of either measure has only minor effects on the IRRs of other explanatory variables.

There is insufficient evidence to suggest that including either measure of obesity in the model improves the model or aids its interpretation.

### 16.2 Hypertension

The measure of hypertension detection used in Chapter 10 is the percentage of the practice population who have been diagnosed with hypertension and have been included in the QOF hypertension register. This approach was taken by Levene et al (2010; 2012) who found negative associations between CHD mortality and QOF hypertension registers. This association supports the premise that QOF hypertension registers are a useful measure of detection in studies of this type. QOF hypertension registers are affected by both the detection rate and the underlying prevalence of hypertension in the practice population, this is discussed in more detail in Section 18.5.2. An alternative measure is to use undiagnosed hypertension prevalence, an approach similar to that taken in studies of emergency admissions for CHD (Purdy et al., 2011), stroke (Soljak et al., 2011) and COPD (Calderon-Larranaga et al., 2011).

Modelled estimates of hypertension prevalence have been developed by the Association of Public Health Authorities (APHO), described in Chapter 6 (APHO, 2009; APHO, 2013). These have been used to estimate undiagnosed hypertension prevalence for each practice. APHO estimates are only available for 228 practices and therefore only these are included in the analysis. Undiagnosed hypertension prevalence ranges from 4.13% to 20.84%. The median level of undiagnosed hypertension is 10.58% (IQR:(9.19%, 20.4%))
Three models are considered:

**Model A** No measure of hypertension detection.

**Model B** Hypertension detection - QOF register 2006/07.

**Model C** Undiagnosed hypertension prevalence (APHO estimate-QOF register)

The results of Models B and C are shown in Table 16.1. In comparison to model A, there is no evidence of improved model fit when hypertension detection is included in the model. A likelihood ratio test suggests that there is improved model fit when undiagnosed hypertension prevalence is included in the model. Inclusion of undiagnosed hypertension has minor effects on the IRRs for other explanatory variables. The size of the association between the percentage aged 65 and over is reduced when undiagnosed hypertension prevalence is included in the model, the IRR decreases from 1.067 to 1.047.

Hypertension detection is negatively associated with premature mortality, however the confidence interval for IRR includes one and the evidence is not clear cut. Undiagnosed hypertension prevalence is positively associated with premature mortality; a 1% increase in undiagnosed hypertension is associated with a 3.5% increase in premature mortality count. Increasing levels of detection are therefore associated with reduced premature mortality. If a practice with above average levels of undiagnosed hypertension (75th percentile: 12.06%), improves to ‘average’ levels of diagnosis (median: 10.58%) this could be associated with a decrease in premature mortality count of 5.18%, given the current model and assuming that the relationship is causal and modelled estimates are a reasonable measure of expected hypertension prevalence.

When all-age mortality is considered there is no evidence that including a measure of hypertension detection (either QOF registers or undiagnosed prevalence) there is no evidence of improved model fit. The direction of associations are the same for both measures, however the confidence intervals for both IRRs include one.
<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>QOF register 2006/07</th>
<th>Undiagnosed hypertension</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Percentage white patients</td>
<td>1.002</td>
<td>(0.99, 1.008)</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.006</td>
<td>(0.996, 1.016)</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>1.102</td>
<td>(1.016, 1.195)</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>1.067</td>
<td>(1.045, 1.089)</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>1.053</td>
<td>(1.018, 1.090)</td>
</tr>
<tr>
<td>Smoking prevalence estimate: QOF 2006/07</td>
<td>1.029</td>
<td>(1.009, 1.048)</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>1.110</td>
<td>(0.820, 1.503)</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>1.005</td>
<td>(0.996, 1.014)</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>0.993</td>
<td>(0.984, 1.002)</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>0.998</td>
<td>(0.978, 1.018)</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>0.994</td>
<td>(0.989, 0.999)</td>
</tr>
<tr>
<td>Hypertension indicator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QOF register 2006/07</td>
<td>0.982</td>
<td>(0.954, 1.010)</td>
</tr>
<tr>
<td>Estimate of undiagnosed prevalence</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 16.1: IRRs, 95% confidence intervals and p-values for premature mortality - inclusion of hypertension detection and undiagnosed hypertension (n = 228) (after fitting a Negative Binomial Model model).
16.3 Access and continuity of care variables

In Chapter 8, five key questions on the GP patients survey (GPPS) are discussed in detail. Analysis of correlations and principal component analysis suggests that the five questions could be grouped into ‘access’ based on phone, open and quick and those related to continuity of care - ahead and preferred.

Using this suggestion a mean ‘access’ and ‘continuity of care’ measure were calculated. The response to phone, open and quick were summed and divided by three to give a measure of mean ‘access’; a similar approach was used with ahead and preferred to create a ‘continuity of care measure’. These were included in the model instead of one simple explanatory variable. 225 practices have complete GPPS data. Summary of the results this approach are shown in Table 16.2.

<table>
<thead>
<tr>
<th>Key variable</th>
<th>IRR (95%CI)</th>
<th>Deviance</th>
<th>Pearson deviance</th>
</tr>
</thead>
<tbody>
<tr>
<td>No GPPS based variable</td>
<td></td>
<td>248.95</td>
<td>231.85</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>0.994 (0.988, 0.999)</td>
<td>250.18</td>
<td>235.41</td>
</tr>
<tr>
<td>Combined ‘access’ measure</td>
<td>0.997 (0.992, 1.001)</td>
<td>250.78</td>
<td>234.24</td>
</tr>
<tr>
<td>Combined ‘continuity of care’ measure</td>
<td>0.996 (0.993, 1.000)</td>
<td>253.58</td>
<td>239.41</td>
</tr>
</tbody>
</table>

Table 16.2: Comparison of access and continuity of care variables (n = 225).

There are negative associations between ‘access’ and ‘continuity of care’ variables and premature mortality. Inclusion of the ‘access’ variable shows no evidence of improved model fit and the confidence interval for the IRR includes one. Both the single measure preferred and the combined ‘continuity of care’ measure show evidence of improved model fit. There is no evidence that using the combined measure is preferable over the single measure preferred, particulary as interpretation of combined measures is less straight forward.
16.4 Conclusions

In this chapter alternative approaches for three explanatory variables were considered. There is no evidence that the available obesity prevalence measures improve model fit or, in the case of QOF obesity registers, are a useful measure of prevalence. The lack of reliable information about obesity prevalence in practice populations is a weakness in work of this type. The use of undiagnosed hypertension prevalence as a measure of quality of care provides useful evidence that hypertension detection is an important aspect of primary care; this is discussed in more detail in Section 18.5.2. The potential to use combined measures of access and continuity of care, based on responses to the GPPS, have been explored and this may be a useful approach in other studies.
Chapter 17

Developing and extending the initial model - results

17.1 Introduction

Chapter 11 describes the results of the model which was originally proposed. In this chapter the results of an improved model are described. This model has been developed through the consideration of alternative measures of hypertension detection and the novel method of estimating smoking prevalence. Table 17.1 summarises the results of the negative binomial model when individual QOF indicators are included in the model and compares premature and all-age mortality. Table 17.2 includes the alternative measure of quality of care; the CHD achievement score. Finally, Table 17.3 demonstrates the effect size of each explanatory variable on premature mortality. These changes are shown graphically in Figures 17.1 and 17.2. This is shown as a one unit change in the explanatory variable and the impact of changes within the range of the variable. In particular, the impact of an improvement in service from below average (25\textsuperscript{th} percentile) to ‘average’ (median). For population characteristics the impact of a decrease in population burden from above average 75\textsuperscript{th} percentile to the median is also shown. Results of multiple linear regression are not included as they have been presented and discussed in Chapter 13. As the model is not deemed appropriate for the data, this model is not considered again here. However, a table summarising all the models which have been considered is included at the end of this chapter for reference (Table 17.4).
17.2 Results

Table 17.1 shows the estimates of incident ratios (IRRs) in the negative binomial model, based on 228 practices. 95% confidence intervals and $p$ values are also given for all explanatory variables. Results are for both premature and all-age mortality. The pseudo-$R^2$ is low, (0.09 for premature mortality; 0.11 for all-age mortality). Low values are common in models of this kind (McCullagh and Nelder, 1989); a value of 0.2 is considered to be indicative of very good fit (Louviere et al., 2000) so there is still concern that model fit is not very good. The concordance between predicted counts of death, based on the model, and actual counts is high (Lin’s concordance coefficient: 0.85 - premature mortality; 0.93 - all-age mortality) (Lin, 1989). Table 17.2 gives results when an overall CHD achievement score is included in the model instead of individual indicators. Estimates of pseudo-$R^2$ and concordance are similar for these models.

17.2.1 Population characteristics

The main predictors of variation in premature CHD mortality of the population characteristics included in the model were prevalence of diabetes, percentage aged over 65, percentage male and the prevalence of smoking. Table 17.3 suggests that a reduction in smoking prevalence from ‘above average’ to ‘average’ is potentially associated with 10.73% decrease in premature mortality count. A 1% change in diabetes prevalence is associated with the biggest potential change in premature CHD mortality, however the difference between the 75th percentile and the median is only 0.5%. This change is still potentially associated with a decrease in 5.05% in mortality count.

Percentage aged over 65, diabetes and smoking prevalence are the main predictors of all-age mortality. The IRRs for both smoking and diabetes prevalence are nearer to one for all-age mortality, suggesting a smaller effect. Percentage male is no longer a significant predictor. Percentage white is positively associated with all-age mortality.

The results are discussed in detail in Chapter 18

17.2.2 Service characteristics

Increase in serum cholesterol control in CHD patients (CHD08) and the percentage of patients who can recall being able to see their preferred GP are both associated
with a decrease in premature mortality count. GPs per capita, percentage of patients with chronic conditions given smoking cessation advice (SM02) and percentage of CHD patients treated with aspirin are all positively associated with premature mortality, although the evidence for an association is limited, as confidence intervals include one. Undiagnosed hypertension is positively associated with premature mortality, demonstrating that improved hypertension detection is associated with decreased premature mortality. A reduction in undiagnosed hypertension prevalence from above average to ‘average’ (median) is potentially associated with a 5.25% reduction in premature CHD mortality. This is similar to the potential affect of improving the percentage of patients who recall being able to see their preferred GP from below average to ‘average’ (median); a potential reduction of 4.8%.

Service characteristics are less clearly linked to all-age mortality. Patterns of association are the same for all indicators; apart from the percentage of CHD patients treated with aspirin. This indicator is negatively associated with all-age mortality although the evidence for a relationship is not clear cut. The confidence interval for the percentage of patients who can recall seeing their preferred GP is wide and includes one. The IRR for undiagnosed hypertension prevalence is lower for all-age mortality, and the confidence interval now includes one.

Increased levels of the CHD achievement score is associated with decreased premature and all-age mortality, although the IRR for both includes one.
<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Premature mortality</th>
<th>All-age mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage white patients</td>
<td>1.001 (0.996, 1.007)</td>
<td>1.004 (1.000, 1.008)</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.003 (0.993, 1.014)</td>
<td>1.005 (0.998, 1.013)</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>1.101 (1.025, 1.182)</td>
<td>1.085 (1.030, 1.143)</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>1.047 (1.025, 1.068)</td>
<td>1.067 (1.052, 1.083)</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>1.049 (1.014, 1.085)</td>
<td>1.002 (0.977, 1.027)</td>
</tr>
<tr>
<td>Smoking prevalence estimate: QOF 2006/07</td>
<td>1.029 (1.010, 1.049)</td>
<td>1.014 (1.000, 1.029)</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>1.138 (0.839, 1.542)</td>
<td>1.015 (0.817, 1.262)</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>1.005 (0.996, 1.014)</td>
<td>1.007 (1.000, 1.013)</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>0.993 (0.983, 1.001)</td>
<td>0.993 (0.987, 1.000)</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>1.002 (0.982, 1.022)</td>
<td>1.002 (0.981, 1.009)</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>0.994 (0.989, 0.999)</td>
<td>0.998 (0.994, 1.002)</td>
</tr>
<tr>
<td>Estimate of undiagnosed hypertension prevalence</td>
<td>1.035 (1.008, 1.063)</td>
<td>1.016 (0.998, 1.035)</td>
</tr>
</tbody>
</table>

Table 17.1: IRRs, 95% confidence intervals and p-values for premature mortality - inclusion of estimated smoking prevalence and undiagnosed hypertension and individual CHD indicators (n = 228) (after fitting a Negative Binomial Model).
<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Premature mortality</th>
<th>All-age mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Percentage white patients</td>
<td>1.002</td>
<td>(0.996, 1.007)</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.005</td>
<td>(0.995, 1.015)</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>1.098</td>
<td>(1.023, 1.178)</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>1.046</td>
<td>(1.025, 1.068)</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>1.046</td>
<td>(1.011, 1.081)</td>
</tr>
<tr>
<td>Smoking prevalence estimate: QOF 2006/07</td>
<td>1.026</td>
<td>(1.007, 1.045)</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>1.121</td>
<td>(0.826, 1.521)</td>
</tr>
<tr>
<td>CHD achievement score</td>
<td>0.993</td>
<td>(0.982, 1.004)</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>0.994</td>
<td>(0.989, 1.000)</td>
</tr>
<tr>
<td>Estimate of undiagnosed hypertension prevalence</td>
<td>1.032</td>
<td>(1.006, 1.059)</td>
</tr>
</tbody>
</table>

Table 17.2: IRRs, 95% confidence intervals and p-values for premature mortality - inclusion of estimated smoking prevalence and undiagnosed hypertension (n = 228) (after fitting a Negative Binomial Model).
<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Effect of a 1% increase in explanatory variable</th>
<th>Effect of improvement in primary care or decrease in population burden - potential reduction in mortality count (%) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage white patients</td>
<td>0.1 (-0.4, 0.7)</td>
<td>Decrease of 4.2% in percentage white patients.</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>0.3 (-0.7, 1.4)</td>
<td>Decrease of 10.9 units on scale.</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>10.1 (2.5, 18.2)</td>
<td>Decrease of 0.5% in diabetes prevalence.</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>4.7 (2.5, 6.8)</td>
<td>Decrease of 2.3% in percentage over 65.</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>4.9 (1.4, 8.5)</td>
<td>Decrease of 1.0% in percentage male patients.</td>
</tr>
<tr>
<td>Smoking prevalence estimate: QOF 2006/07</td>
<td>2.9 (1.0, 4.9)</td>
<td>Decrease of 3.7% in smoking prevalence.</td>
</tr>
<tr>
<td>Number of GPs per 1000 patients</td>
<td>13.8 (-16.1, 54.2)</td>
<td>Increase of 0.08 GPs per 1000 patients.</td>
</tr>
<tr>
<td>% patients offered smoking cessation advice (SM02)</td>
<td>0.5 (-0.4, 1.4)</td>
<td>Increase of 2.1% offered advice.</td>
</tr>
<tr>
<td>% serum cholesterol (CHD08)</td>
<td>-0.7 (-1.7, 0.1)</td>
<td>Increase of 4.9% in achieving serum cholesterol target.</td>
</tr>
<tr>
<td>% aspirin (CHD09)</td>
<td>0.2 (-1.8, 2.2)</td>
<td>Increase of 2.1% in aspirin treatment.</td>
</tr>
<tr>
<td>% of patients with recalled perception of being able to see preferred GP</td>
<td>-0.6 (-1.1, -0.1)</td>
<td>Increase of 8.0% in patients recalling being able to see preferred GP.</td>
</tr>
<tr>
<td>Estimate of undiagnosed hypertension prevalence</td>
<td>3.5 (0.8, 6.3)</td>
<td>Decrease of 1.5% in undiagnosed hypertension prevalence.</td>
</tr>
</tbody>
</table>

Table 17.3: Effects on premature CHD mortality count of a one unit increase in the value of the explanatory variables and the impact of an improvement in primary care or decrease in population burden. Percentage change in premature count (95% confidence interval) (n = 228) (after fitting a Negative Binomial Model).
Figure 17.1: Effects on premature CHD mortality count of a decrease in population burden. Percentage change in premature count (95% confidence interval).

Figure 17.2: Effects on premature CHD mortality count of an improvement in primary care. Percentage change in premature count (95% confidence interval).
<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Poisson model</th>
<th>NBM</th>
<th>Unweighted</th>
<th>Weighted</th>
<th>log(SMR)</th>
<th>NBM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage white patients</td>
<td>1.007 (1.003, 1.012)</td>
<td>1.008 (1.003, 1.012)</td>
<td>0.22 (-0.43, 0.87)</td>
<td>0.36 (-0.08, 0.80)</td>
<td>0.0040 (-0.0014, 0.0095)</td>
<td>1.001 (0.996, 1.007)</td>
</tr>
<tr>
<td>Deprivation score</td>
<td>1.017 (1.012, 1.022)</td>
<td>1.017 (1.011, 1.024)</td>
<td>2.43 (1.52, 3.33)</td>
<td>2.30 (1.59, 3.01)</td>
<td>0.024 (0.016, 0.031)</td>
<td>1.003 (0.993, 1.014)</td>
</tr>
<tr>
<td>Prevalence of diabetes 2006/07</td>
<td>1.106 (1.029, 1.208)</td>
<td>1.114 (1.028, 1.208)</td>
<td>-2.51 (-13.31, 8.48)</td>
<td>2.53 (-5.69, 10.75)</td>
<td>0.029 (-0.063, 0.12)</td>
<td>1.101 (1.025, 1.182)</td>
</tr>
<tr>
<td>Percentage over 65</td>
<td>1.057 (1.038, 1.077)</td>
<td>1.059 (1.038, 1.081)</td>
<td>2.37 (0.99, 4.65)</td>
<td>2.06 (0.13, 3.99)</td>
<td>0.027 (0.0039, 0.051)</td>
<td>1.047 (1.025, 1.068)</td>
</tr>
<tr>
<td>Percentage male patients</td>
<td>1.006 (1.034, 1.100)</td>
<td>1.067 (1.038, 1.103)</td>
<td>9.13 (5.06, 13.21)</td>
<td>2.95 (-0.78, 6.68)</td>
<td>0.068 (0.038, 0.10)</td>
<td>1.049 (1.014, 1.085)</td>
</tr>
</tbody>
</table>

**Smoking prevalence estimate: QOF 2006/07**
- Not included

**Number of GPs per 1000 patients**
- 1.198 (0.911, 1.576) | 1.197 (0.885, 1.619) | 42.29 (1.81, 82.77) | 16.63 (-15.39, 48.64) | 0.17 (-0.17, 0.50) | 1.138 (0.839, 1.542) |

**% patients offered smoking cessation advice (SM02)**
- 1.001 (0.993, 1.009) | 1.002 (0.993, 1.011) | -1.34 (-2.60, -0.07) | 0.67 (-0.15, 1.50) | 0.0048 (-0.0058, 0.15) | 1.005 (0.996, 1.014) |

**% serum cholesterol (CHD08)**
- 0.991 (0.983, 0.999) | 0.991 (0.981, 1.000) | -1.38 (-2.69, -0.07) | -0.24 (-1.47, 1.95) | -0.0062 (-0.017, 0.0048) | 0.993 (0.983, 1.001) |

**% aspirin (CHD09)**
- 1.001 (0.983, 1.020) | 1.002 (0.982, 1.022) | 1.94 (-0.81, 4.69) | -0.24 (-1.47, 1.95) | 0.0030 (-0.020, 0.026) | 1.002 (0.982, 1.022) |

**% of patients with recalled perception of being able to see preferred GP**
- 0.994 (0.990, 1.000) | 0.984 (0.989, 1.000) | -0.77 (-1.67, 0.14) | -0.88 (-1.43, -0.33) | -0.0085 (-0.016, -0.0010) | 0.994 (0.989, 0.999) |

**Hypertension detection - QOF hypertension register**
- 0.980 (0.955, 1.007) | 0.978 (0.950, 1.007) | -3.61 (-7.68, 0.47) | -2.74 (-5.61, 0.14) | -0.048 (-0.082, -0.014) | not included |

**Estimate of undiagnosed hypertension prevalence**
- Not included

**Other relevant information**
- Does not account for over-dispersion. When an additional parameter is included to account for over-dispersion (over-dispersed model) the only change the p-value for CHD08 is no longer than 0.05. See Sections 10.3 and 10.4.

Good agreement between observed and expected counts of death, although $R^2$ is low. When outliers are removed there is less evidence that the data is over-dispersed, these are discussed in Section 10.5.1.

Patterns in the residuals provide evidence that this model is not appropriate.

Improved model fit. Results are similar to count models, although there are differences (Sect 13.3).

Good model fit. Results are similar to weighted OLS and count models (Sect. 13.4).

Final model, includes improved measure of hypertension detection and smoking prevalence, described in detail in Chapter 17 and discussed in Chapter 18.

Table 17.4: IRRs or $\beta$ coefficients (and 95% confidence intervals) for all models considered in this thesis. Premature CHD mortality is the dependent variable throughout. Numbers in **bold** indicate where confidence intervals do not include one (IRR) or zero ($\beta$ coefficients). (*NBM: Negative Binomial Model*)
Part VI

Discussion, recommendations and conclusions
Chapter 18

Discussion, recommendations and conclusions

18.1 Introduction

This thesis aims to describe and quantify associations between aspects of primary care and health outcomes of practice populations. Premature CHD mortality has been used as an example health outcome throughout the thesis, although all-age CHD mortality has also been considered. This chapter discusses key results and observations which have been identified in this thesis and discusses both methodological and substantive issues.

The conceptual model underlying this study was proposed by Levene et al (2012) and suggests that primary care can play a key role in altering the effect of population characteristics on the proportion of any given population transitioning from healthy or morbid to deceased. The complex interactions between population characteristics and primary care were illustrated in Figure 1.2, which is reproduced below for reference. This figure also illustrates the many aspects of both populations and primary care which have been proposed to have an impact on premature CHD mortality.

Key challenges in quantifying associations between primary health care and health outcomes include the selection of an appropriate modelling strategy (Chapters 3, 10, 13 and 14), the identification of relevant and useful data and to subsequently select the key explanatory variables to include in statistical models (Chapters 6, 7 and 8).

The results of the extended model, which includes an estimate of smoking prevalence
and a revised measure of hypertension detection, are discussed here having been presented in Chapter 17. Findings which are of particular interest, either because they confirm or contradict previous research or have important implications for practices, have been highlighted. Relationships to other studies and implications for further research, practitioners and policy makers are discussed. Strengths and weaknesses of the approach are highlighted throughout the discussion, with more general points covered in Section 18.8.
Figure 18.1: Model to describe role of primary care in improving health outcomes in populations.
18.2 Principal findings

Principal findings, both substantive and methodological, are described here, with more detailed discussion of key findings in following sections. In the final model, including estimates of smoking prevalence and undiagnosed hypertension (Table 17.1), indicators of both population characteristics and the quality of primary care were found to be associated with variations in premature CHD mortality. Increasing percentages of the practice population who were on the practice diabetes register, who were over 65 and who were male as well as the estimated prevalence of smoking in those with chronic conditions were all associated with increasing levels of premature CHD mortality. Control of serum cholesterol levels in those with CHD and the percentage of patients who could recall being able to see their preferred GP were both associated with decreased levels of premature CHD mortality. Increasing levels of undiagnosed hypertension were associated with increased levels of mortality.

The inclusion of smoking prevalence in the model reduced the predictive effect of both the percentage of the population who are white and levels of socioeconomic deprivation.

The combined measure of the quality of primary care was associated with a decrease in both all age and premature CHD mortality, although the evidence of an association is not clear cut. It is difficult to determine which individual indicators within this measure are key to reducing CHD mortality.

Similar results were found when all-age mortality was considered, although there is less evidence that service characteristics are important predictors of all-age mortality.

The availability of data to describe lifestyle factors and the prevalence of certain risk factors, such as obesity and serum cholesterol, limits our understanding of the potential modifying effect of primary care.

Different modelling approaches yielded qualitatively similar results. However, detailed interpretation of the results would be model dependent, particularly if statistical significance were rigidly applied as a criterion of importance.

18.3 Approaches to modelling

This thesis considers the impact of modeling approach on the results and their interpretation. Whilst multiple linear regression of standardised mortality (or hospitalisation)
rates can be found in the literature, as described in Chapter 5, Poisson models of count data are a more natural and appropriate approach, particularly when age standardised explanatory variables are not available (Rosenbaum and Rubin, 1984). There was evidence of overdispersion meaning that Poisson models were not appropriate and negative binomial models or quasi-Poisson models were considered as useful alternatives, the negative binomial model preferred as it has a distributional form. The exclusion of outliers, discussed in Section 10.5.1 indicated that the overdispersion may have been due to the presence of outliers, which were retained.

Negative binomial models of count data were compared with multiple linear regression of standardised rates and, whilst the directions of associations and broad magnitude of association were generally not affected by the choice of model, the statistical significance of the results varied between models for several key variables. The direction of association for prevalence of diabetes was particularly sensitive to the modelling approach used, being affected by the approach to weighting, substitution and whether SMR or log(SMR) was modelled. QOF related quality of care indicators (control of serum cholesterol and CHD achievement score) would not be considered significantly associated with premature mortality using a weighted linear regression model for SMRs, in contrast to the significant association found using a negative binomial model of death counts. The percentage of patients with recalled perception of being able to see their preferred GP was consistently, negatively associated with premature CHD mortality, regardless of modelling approach. When estimates of smoking prevalence are not included in multiple linear regression or count models, socio-economic deprivation was consistently positively associated with premature CHD mortality.

Patterns in the residuals suggest that the model linear regression model fit for standardised rates improved as practice size increased. There was evidence that the model fit was better when a negative binomial model of count data was used.

The potential to use multilevel models was described in Chapter 14. Multilevel models are an important modelling approach for practice based studies, particularly when data are available at patient level or at local area levels. It is relatively easy to use multi-level Poisson models, although interpretation of the partition of variance can be difficult. An area for further work is multilevel negative binomial models.
Relation to other studies

Papers which compare and discuss the impact of modelling choice on results and their interpretation are rare. Milyo and Mellor (2003) compared different approaches to age adjustment in multiple linear regression of mortality rates and various social determinants and found that the results were sensitive to the method of age adjustment and weighting approach, as confirmed here.

Levene et al analysed associations between CHD mortality and indicators of population characteristics and primary care at PCT level, using both multiple linear regression of SMRs (2010) and negative binomial regression of mortality counts (2012). Although the variable selection method differed in the two papers, the results and their interpretation were similar for key variables. The percentages of the practice population who were white, had diabetes and were smokers as well as deprivation and detected hypertension levels were all significantly associated with CHD mortality, regardless of modelling approach. Indicators of quality of primary care, other than hypertension detection, were not found to be significantly associated with CHD mortality, again, regardless of modelling approach.

Kiran et al (2010) used multiple linear regression of SMRs in a practice level study and found that quality of primary care was negatively associated with CHD mortality. In contrast, here, with multiple linear regression no significant association between indicators of quality of primary care was seen, but there was with a negative binomial model of death counts. This was the case for both all-age and premature CHD mortality (Honeyford et al., 2013). However, the sample of practices, period of time and other explanatory variables included in the models were different, which may be more important in explaining differences than modelling approach.

Critical review

The potential impact of outliers on over-dispersion and subsequent choice of model, and the impact of weighting in multiple linear regression, and challenges of practices with zero deaths have been considered. Different approaches to age adjustment in multiple linear regression of mortality rates were not addressed because of lack of availability of age-adjusted explanatory variables.

Multilevel models were explored in Chapter 14; however, a case-study approach was utilised focusing on area-level measures of smoking prevalence. Area-level measures
of obesity and activity are available and it would be useful to explore these using multilevel modelling. Patient level data would have meant that multilevel models including patient, practice and local area level models could have been evaluated, which would strengthen the possibility of using multilevel models in studies of this type. Multilevel negative binomial models were not considered and this is an area for further research.

**Implications for further research, practitioners and policy makers**

Exploring different approaches to age-adjustment would add to our understanding of the use of multiple linear regression of standardised rates, although count models may be considered a more natural and appropriate approach. Further work to enable direct comparisons of the magnitude of coefficients when negative binomial and multiple linear approaches are used would be an interesting area of research.

The sensitivity of results to modelling approaches, including weighting, means that researchers, practitioners and policy makers should be cautious when interpreting the results of observational studies. When publishing research, it is recommended that authors include sensitivity analyses exploring the impact of modelling choice.

### 18.4 Population characteristics

Throughout this thesis population characteristics have been considered in terms of the categories shown in Table 18.1.

The limited availability of data to quantify many of these characteristics was described in Chapter 6.

In the final model sex is described as the percentage of practice population who are male and age is summarised as the percentage of the practice population aged 65 and over. Both increasing age and being male are known to increase the risk of CHD and are found to be associated with premature CHD mortality.

Ethnicity is summarised as the percentage of the practice population who are white. This is a limited measure of ethnicity; particular ethnic groups are known to be associated with higher risk of CHD mortality, but this is not reflected in the categories of data available. Although there is an association between white ethnicity and premature mortality when there is no measure of smoking prevalence, when they are included in
<table>
<thead>
<tr>
<th>Factor type</th>
<th>Possible factors</th>
<th>Appropriate practice level data available identified in Chapter 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unmodifiable biological risk factors</td>
<td>age, sex, prevalence of high risk genes, and ethnicity</td>
<td>age, sex and ethnicity</td>
</tr>
<tr>
<td>Modifiable biological risk factors</td>
<td>diabetes, obesity¹, serum cholesterol and high blood pressure²</td>
<td>diabetes</td>
</tr>
<tr>
<td>Health behaviours identified with high CHD</td>
<td>smoking, alcohol, physical (in)activity, and poor nutrition</td>
<td>None readily available³</td>
</tr>
<tr>
<td>Socioeconomic deprivation</td>
<td>potentially many variables</td>
<td>Index of Multiple Deprivation (2007)</td>
</tr>
</tbody>
</table>

¹Obesity was discussed in Chapter 16.
²Hypertension was discussed in Chapter 16 and in Section 18.5.2 below.
³Smoking data was discussed in Chapter 15 and in Section 18.4.1 below.

Table 18.1: Summary of population characteristics which have been considered and availability of data.

The data available to describe the prevalence of modifiable risk factors and health behaviours are limited. Smoking prevalence and hypertension are described in Sections 18.4.1 and 18.5.2 respectively. QOF diabetes registers have been used as a measure of diabetes prevalence; QOF prevalence registers are discussed further in Section 18.5.2. Increased prevalence of diabetes in practice populations is associated with increased premature and all-age CHD mortality as may be expected (NIH National Heart, Lung and Blood Institute, 2011). The association between diabetes prevalence in practice populations and CHD mortality may also be related to health behaviours associated with the higher risk of both diseases such as inactivity and poor nutrition.

The lack of detailed information about risk factor prevalence and health behaviours has been described in Chapters 6 and 7. The introduction of health checks in English general practices has the potential to provide detailed information on both BMI and serum cholesterol levels, as well as self reported dietary and alcohol information. Although the percentage of the population receiving invitations to health checks and the subsequent uptake is relatively low, nearly 100,000 people in the East Midlands received a health check between April 2011 and March 2012 and a further 139,000 in 2012/13 (NHS Health Check, 2013). If these data were collated at practice level and readily available for researchers our understanding of the prevalence of CHD risk
factors could be improved.

### 18.4.1 Inclusion of a smoking prevalence measure

Results described in Chapter 15 show how QOF registers can be used to estimate smoking prevalence. There is good agreement between estimates for the general population based on QOF and estimates of smoking prevalence based on the IHS for similar geographical areas. QOF data can also be used to estimate smoking prevalence in those with chronic conditions, which is generally lower than smoking prevalence in the general population. There is good agreement between the estimates in successive years. The correlation between estimates of smoking prevalence in the general population in 2012/13 and those with chronic conditions is strong, suggesting that the estimates based on previous years can be used in place of smoking prevalence in the general population for some purposes.

When an estimate of smoking prevalence in those with chronic conditions was included in the model, increased smoking prevalence was positively associated with premature and all-age CHD mortality. There were consequential reductions in magnitude of the incident rate ratios (IRRs) for both deprivation and percentage white. This suggests that these may be acting as surrogate markers of other lifestyle factors, such as smoking prevalence.

### Relation to other studies

Smoking prevalence is a known risk factor for cardiovascular diseases and premature mortality (Twigg et al., 2004). However, the lack of reliable information about smoking prevalence means this variable is often not included in practice level studies (Kiran et al., 2010; Honeyford et al., 2013; Bankart et al., 2011). Smoking prevalence, estimated from IHS data, was positively associated with stroke admissions (Soljak et al., 2011) and CHD admissions (Purdy et al., 2011).

This analysis suggests that inclusion of smoking prevalence may affect the apparent relationship between socio-economic deprivation and health outcomes. Similarly, it was found that social class was not linked to hospital admissions for stroke and CHD when rates were adjusted for various factors including smoking (McCartney et al., 2013). However, even with smoking prevalence included in the models, associations with socioeconomic factors - particularly deprivation - have been shown to remain
(Purdy et al., 2011; Brettell et al., 2013). The lack of reliable smoking information may be leading to relative over emphasis being placed on socio-economic deprivation. Smoking prevalence in those with chronic conditions is typically lower than in the general population. This may be due to diagnosis increasing motivation to quit smoking (Brettell et al., 2013), the increase in smoking cessation advice and support (Coleman, 2010) or the age and gender profile of those with chronic conditions. Smoking prevalence in those with chronic conditions has not reduced over the seven year period, possibly suggesting that smoking cessation advice has limited effect. However, the turnover of patients with chronic conditions as a result of both premature mortality and new diagnoses makes interpretation difficult.

In the initial model the percentage of the practice population who were white was associated with increased counts of premature CHD mortality; this result is unexpected as other ethnic groups are known to have a higher risk of CHD mortality. For example, Kiran et al. (2010) found that increased percentages of South Asian patients were associated with both increased CHD admissions and mortality. When smoking prevalence is included in the model there is minimal evidence of an association between white ethnicity and premature CHD mortality. This suggests that the proportion of the practice population who are white may be describing other characteristics of the practice population; in the case of premature CHD mortality percentage white may be acting as a proxy for smoking prevalence. However, even when smoking prevalence is included in the model, there is still evidence of an association between white ethnicity and all-age CHD mortality. This suggests that there may be other important variables which have not been included in the model, for example physical (in)activity or interactions between ethnicity and smoking prevalence.

Critical review

This thesis proposes a novel way of estimating smoking prevalence in practice populations. The estimates have been validated by comparisons with estimates for local authority districts from the Integrated Household Survey (ONS, 2014). Although estimates of smoking prevalence in the practice population can only be made since the introduction of QOF indicators covering the whole practice population in 2012/13, the analysis in Chapter 15 shows that estimates of smoking prevalence in those with chronic conditions can be used as a proxy for the total population. However, there are concerns with using this method to estimate smoking prevalence in practice populations as described in detail in Chapter 15. In summary the percentage of patients
who have their smoking status recorded varies, the QOF data rely on self reported smoking status, there are problems with allocating residential postcodes to practices, and relationships in earlier years are unknown.

In this analysis the estimate of smoking prevalence which has been used is a measure of smoking prevalence in those with chronic conditions, rather than the practice population. The effect size associated with smoking prevalence would be different if an estimate of smoking prevalence in the general practice population was used, since, for example, those with chronic conditions tend to be older in practices where the population mix is different in the younger population. There is no way of determining the smoking prevalence in those aged under 75, which would be particularly useful in this study.

**Implications for further research, practitioners and policy makers**

Having estimates of the smoking prevalence in practice populations is important to those tasked with reducing smoking rates and improving the nations health. CCGs and public health departments in local authorities need them to target smoking cessation and other additional resources. Understanding more about the patient populations would enable similar practices to be compared when considering differences in health outcomes and the apparent effectiveness of interventions (Shipton et al., 2009). The inclusion of smoking prevalence in statistical models impacts on the interpretation of the importance of other variables; again emphasising the importance of valid estimates of smoking prevalence in practice populations.

In order to improve the validity of estimates based on QOF, it is necessary to analyse patient level data. Although considerable work has been done on the characteristics of patients for whom no smoking status is recorded (HSCIC, 2004; Simpson et al., 2010; Wong et al., 2012), the conclusions need to be confirmed in the light of the changing requirements of QOF. It would be useful to determine the reliability of smoking status in QOF given it is self reported and only needs to be recorded every 27 months. In addition, analysis of exception reporting should be carried out, including patient level analysis.

The method described in Chapter 15 was developed to estimate smoking prevalence in practice populations. It is also possible to estimate smoking prevalence in local areas by aggregating practice level data, as was done to evaluate the methodology. The ‘sample size’ covered by QOF indicators is much larger than any national or
local survey. Manipulation of QOF data may, therefore, be an important method of estimating smoking prevalence in a range of local areas. Access to patient level residential postcodes would allow the potential of QOF data to estimate local area smoking prevalence to be explored in more detail.

The stability over time of smoking prevalence in those with chronic conditions raises concerns about the effectiveness of smoking cessation advice. A wide range of smoking cessation advice and support has recently been reviewed by Zwar et al (2014); consideration of how these impact on those with chronic conditions is recommended as a result of this finding.

QOF smoking indicators have changed since 2004 and continue to change. The introduction, in 2012/13, of an indicator which allows estimates of the smoking prevalence within the general population is useful for researchers as well as CCGs and public health officials. The removal of the indicator that covers the recording of smoking status in the total population from QOF in 2014/15 will impact on the methodology described in here, although the number of patients who are recorded as current smokers will continue to be available. The population of the practice will need to be used as the denominator in the calculation of smoking prevalence. Recording of smoking status may decline after the removal of the indicator, although this has not been the case with other indicators (Kontopantelis et al., 2014).

18.5 Primary care - prevention, detection and clinical management

In this study, quality of care, in terms of prevention, detection and clinical management, was generally measured using underlying achievement of QOF indicators (the percentage of patients eligible to be included for whom the target described by the indicator was achieved). Two approaches were considered: individual indicators; and a combined measure using indicators from a range of clinical domains. The use of individual indicators demonstrated that there is evidence of associations between quality of care and premature CHD mortality. Higher levels of serum cholesterol control (CHD08) was consistently found to be associated with lower levels of premature and all-age CHD mortality. The significance of the association was sensitive to the inclusion of estimates of smoking prevalence in the model, although the direction of association was unaffected. There was less evidence that treatment with aspirin, or alternative,
(CHD09) was associated with premature or all-age mortality. In the final model, which included estimates of smoking prevalence and undiagnosed hypertension, the confidence intervals for both premature mortality and all-age mortality included one. The high underlying achievement for treatment with aspirin in comparison to control of serum cholesterol may provide an explanation for the lack of association. These associations were also sensitive to modelling approach.

In the initial model, which did not include smoking prevalence and used QOF hypertension registers as a measure of hypertension detection, the combined indicator provided strong evidence that quality of care was associated with premature CHD mortality (11.1). However, when smoking prevalence and undiagnosed hypertension were included in the model the evidence was no longer clear cut and the confidence interval for the IRR included one. Similar patterns were seen with all-age mortality.

An additional measure of quality of primary care considered was hypertension detection. When hypertension detection based on the percentage of the practice population on the QOF hypertension register was included in the model, higher detection was associated with lower levels of premature and all-age CHD mortality, although the confidence intervals were wide and include one (see Tab. 11.2). Chapter 16 described an alternative measure of quality of hypertension detection; a measure of undiagnosed hypertension prevalence. Increased levels of undiagnosed hypertension prevalence were associated with increased numbers of both premature and all-age CHD mortality; the evidence was stronger for premature mortality. The different approaches to measuring hypertension detection are discussed in more detail in Section 18.5.2.

### 18.5.1 Choice of measure of quality of care

Chapter 7 describes the potential to use QOF data to measure the quality of care delivered by primary care in terms of prevention, detection and clinical managements of both risk factors for CHD and CHD itself. Although QOF provides an abundance of measures which quantify the performance of practices there are considerable challenges in the selection of appropriate measures and the interpretation of these measures in terms of ‘quality of care’. Combined measures, such as the CHD achievement score, or individual indicators can be used. If individual indicators are used the method of selection of indicators needs to be carefully considered. Chapter 9 describes the method of indicator selection adopted in this thesis; indicators were selected on the basis of previous research and careful consideration of the nature of the indicators. The choice
of measure of quality of care impacts on the interpretation of the association of primary care and CHD mortality.

**Relation to other studies**

There are few studies which consider the use of QOF indicators in CVD health outcomes and none which are concerned with premature mortality. The rationale for inclusion of particular indicators is not generally included in research articles although Soljak et al. (2011) explains that ‘the four stroke indicators that best represented outcomes rather than processes of care’ were selected as they were most likely to be associated with admission risk. The combined CHD achievement score was based on four modifiable risk factors (high blood pressure, high blood cholesterol, diabetes and smoking) and also includes indicators relating to treatment of CHD (Kiran et al., 2010). This is described in more detail in Chapter 7.

Levene et al. (2012) found little association between CHD mortality with either serum cholesterol control or aspirin treatment. Similarly they found no evidence of association between stroke mortality and cholesterol control and aspirin treatment in stroke patients. Lack of evidence of association between cholesterol and blood pressure control in CHD patients was found in studies of CHD admissions (Purdy et al., 2011). In contrast, practices with higher CHD achievement scores were found to have lower CHD admission and mortality rates. The relationship was strongest in practices with high levels of deprivation within their practice population, and disappeared in practices with lower levels of deprivation (Kiran et al., 2010). Better cholesterol control in stroke patients has been found to be associated with lower hospital admissions (Soljak et al., 2011).

The lack of agreement between studies will be associated with the different confounding variables included in models, but may also reflect the choice of measure of quality of care, the different health outcomes and the year of study, which varied between studies.

**Critical review**

This thesis clearly sets out the challenges of selecting indicators from QOF to use as measures of quality of primary care. Two approaches have been discussed; the use of individual indicators and a combined score. The interpretation of combined scores is difficult for policy makers and practitioners as it is not clear which areas of clinical
management are most important in reducing premature CHD mortality.

This analysis did not have access to patient level data and therefore could not determine the overlap between indicators - the number of patients on more than one register. It has been shown that 30% of the population aged between 45 and 64 are likely to have multimorbidity and the percentage with multimorbidity increases with increasing socio-economic deprivation (Barnett et al., 2012). The same study found that comorbidity between those with diabetes and CHD is approximately 20%. This suggests overlap, the number of patients on more than one register, is likely to be high.

Exception reporting allows practices to exempt patients from individual indicators; this varies between practices and indicators. Practice level exception reporting was not available in 2006/07 and it was therefore not possible to include this information. It has been argued that when considering underlying achievement the number of patients who have been excluded from reporting through exception reporting should be included as this shows the true level of care received by patients. However, reasons for exception reporting vary and may include patients for whom the treatment is not appropriate as well as patients who have opted out of treatment for a variety of reasons.

QOF data are not published by age group or gender which means it is not possible to determine whether achievement in particular age groups or in males or females is key to improving health outcomes.

Although two approaches to indicator selection were considered, there are other approaches which may have been appropriate. Combinations of indicators into subgroups, such as ‘recording’, ‘referrals’, ‘control’ and ‘medication’ may be informative. Likewise, grouping indicators into ‘blood pressure control’ and using blood pressure control indicators from a range of disease areas may be informative. However, these approaches have inherent complications. Within the area of ‘blood pressure control’, practices with low levels of detection may appear to be effective at management, although the percentage of patients who have hypertension which is controlled may be low. Stepwise selection methods may have been useful, after reducing the number of potential indicators as described in Chapter 9. Concerns with this approach were discussed in Section 3.9. The work of Ashworth et al (2013) has identified particular indicators as playing the most important role in reducing mortality. Using this as guidance for indicator selection would be a possible extension for this work.

Using QOF indicators as measure of quality of care has clear limitations. The achievement of targets which have been identified as having clinical relevance, but which have
been negotiated as part of a pay for performance contract does not necessarily reflect quality of care. In this study the use of QOF is limited to prevention, detection and clinical management and other measures have been included to consider access and continuity of care. Prior to QOF, prescribing data were sometimes used to measure clinical management. However, QOF indicators have advantages over this as they include a much wider range of information. Other measures include the number of specialist clinics or the rate of voluntary disenrollment (Nagraj et al., 2013). QOF is a readily available source of information for observational studies of this type and practices are visited by assessors on an annual basis in 2006/07 (Department of Health, 2003) to confirm that practice activity matches QOF reporting.

Implications for further research, practitioners and policy makers

The rationale and method of indicator selection should be carefully considered as part of any research which utilises QOF data as a measure of quality of care, as different approaches are likely to lead to different interpretations of their importance. The rationale adopted should be included in research publications. Access to patient level data would enable researchers to understand more fully the combinations of targets which are being achieved for individual patients and sub-groups of patients. This may allow a deeper understanding of how QOF indicators are associated with improved health outcomes. In addition, sub-groups of patients who have been excluded from QOF indicators could be identified. Sensitivity analysis could be used to determine the impact of exception reporting on results and their interpretation. A best practice guide for research which utilises QOF indicators could then be developed.

This research adds to a body of research which demonstrates that quality of care is associated with improved health outcomes. Although research is not conclusive in identifying which particular QOF indicators are key, it is unlikely that improving achievement will have a negative impact on health outcomes. Analysis of achievement of indicators in subgroups will support practices and CCGs in targeting their resources.

Currently practice level data regarding health checks are not published; this includes both the level of uptake and the results of tests. These data could provide a rich resource of information on obesity and hypertension prevalence, as well as information about serum cholesterol levels. The potential utility of this data in research should be explored along with the cost of collation and publication.
18.5.2 Hypertension detection

QOF diagnosed hypertension prevalence is negatively associated with premature CHD mortality, although the evidence is not clear cut. Undiagnosed hypertension prevalence, based on modelled estimates of hypertension prevalence and QOF diagnosed hypertension prevalence is positively associated with premature CHD mortality and the evidence for this association is strong. Reducing rates of undiagnosed hypertension from 12.06% to 10.58%, the difference between the 75th percentile and the median, is associated with a potential reduction of 5.25% decrease in premature CHD mortality.

Relation to other studies

Levene et al (2012) have shown that QOF diagnosed hypertension is associated with reduced CHD and stroke mortality. The majority of similar studies which have considered hypertension have focussed on control of hypertension rather than detection. Hypertension control was included in an overall CHD achievement score which was negatively associated with CHD mortality (Kiran et al., 2010). Control of hypertension has been found to be negatively associated with stroke admissions in some analyses (Soljak et al., 2011). Various studies have analysed the impact of QOF on blood pressure control suggesting that blood pressure control has improved substantially and reduced to ‘near disappearance’ the gap between the least and most deprived areas (Ashworth et al., 2008). However, other studies suggest that the recording of blood pressure may have been affected by QOF guidelines, which defines controlled systolic blood pressure to be under 150, for example see Carey et al (2009).

Fewer studies have considered QOF detection. QOF registered hypertension was found to be higher in practices with patients and practice characteristics including more older patients, more patients reporting poor health and fewer patients being able to book an appointment in advance, sometimes used as a measure of continuity of care (Anwar et al., 2012). This may be related to the higher demand placed on practices with a higher proportion of patients with chronic disease. Further analysis has shown that practices with more GPs per 1000 patients had higher QOF registered hypertension and that as practices increased in size detection tended to decline (Bankart et al., 2013).
Critical review

A key strength of this study is that it considers which measure of hypertension detection is most useful in studies of this kind. The inclusion of hypertension detection as an explanatory variable in this analysis is as an indicator of quality of care. Specific measurement of ‘detection’ is not straightforward. Figure 18.2 provides a way of understanding the links between actual prevalence, detection, clinical management and health outcomes, but also confirms their complexity.

![Figure 18.2: Relationship between underlying health condition, prevalence, detection and mortality.](image)

Levene et al (2010), used QOF based prevalence and interpreted associations with premature mortality as evidence that QOF based diabetes prevalence is closely linked to underlying prevalence whereas QOF based hypertension prevalence may be a more useful proxy for detection rates. This interpretation is supported by comparisons between estimates of prevalence based on models and QOF. Martin and Wright (2009) found that QOF diagnosed diabetes rates were closer to modelled estimates than QOF diagnosed hypertension. QOF based diabetes accounted for 82.4% of modelled estimates compared to 69.3% for hypertension, but Holman et al (2011) found lower
estimates for diabetes. Other analysis, based on 2005/06 QOF hypertension registers estimated a mean difference in QOF diagnosed hypertension prevalence and modelled prevalence of 12.15% (Dixon et al., 2012).

It is likely that QOF recorded hypertension reflects both underlying prevalence and the quality of diagnosis. This highlights a limitation in using QOF hypertension registers as a measure of detection, when detection is in turn being used as a measure of quality of primary care. Therefore, an alternative measure of quality of detection based on deriving the level of undiagnosed hypertension has been considered, described in Chapter 16. This measure has been shown to be important in explaining premature CHD mortality.

To determine levels of undiagnosed hypertension prevalence, estimates of prevalence based on models have to be used. The models are based on the Health Survey for England (HSE), and take into account age, sex and ethnicity. It is recognised that for practices which are very different from a ‘typical’ population the model assumptions may not hold (Walford et al., 2011). However, updated estimates based on the same model, are used to inform the National General Practice Profiles published online by Public Health England (Public Health England, 2014b) suggesting they are an important source of information. The use of estimated undiagnosed prevalence, based on modelled estimates of prevalence, has also been used in similar studies (Calderon-Larranaga et al., 2011; Purdy et al., 2011; Soljak et al., 2011). There are no data available to estimate undiagnosed hypertension prevalence by age as neither modelled estimate nor QOF diagnosed prevalence are published by age groups. The HSE has shown that untreated hypertension is highest in those aged between 45-64 (HSCIC, 2010), which suggests that undiagnosed hypertension may be higher in those aged under 75. This study did not have access to data from health checks either at practice level or individual level.

Implications for further research, practitioners and policy makers

The measure of ‘disease detection’ selected for studies of associations between primary care and health outcomes needs to be carefully considered, particularly when this is being considered as a measure of quality of primary care. Research presented here suggests the selected measure can have an impact on results and their interpretation. Analysis of patient level data are important to determine whether particular subgroups of patients have higher levels of diagnosed and undiagnosed hypertension prevalence.
Access to the models used to estimate expected hypertension prevalence would allow expected levels of hypertension prevalence for different subgroups to be estimated. This could be used in conjunction with diagnosed prevalence to develop a detailed understanding of where undiagnosed prevalence is highest.

When research into health outcomes at practice level is being undertaken, the measure of disease detection must be carefully considered, particularly when this is being used as a measure of quality of primary care.

There is clear evidence that increased hypertension diagnosis is associated with lower levels of premature CHD mortality. The need to improve case finding has been recognised by the Government, as has been shown by the introduction of ‘health checks’ in 2009. In 2013/14, only 12% of the eligible population in the East Midlands received a health check (NHS Health Check, 2013). Hence relying on health checks to improve hypertension may not be sensible and the development of strategies to improve levels of diagnosis of hypertension should be considered a priority. Practices have ready access to estimated and diagnosed prevalence within their registered populations and are therefore able to quantify undiagnosed prevalence. Further understanding of diagnosis rates and health check uptake within subsets of their practice populations may allow them to improve diagnosis.

18.6 Primary care - access and continuity of care

This work adds to the increasing body of knowledge on the measurement of continuity of care (Chapter 8) and the association between increased continuity of care and improved health outcomes. The effect size in this study is not large; a 1% increase in the percentage of patients recalling being able to make an appointment with their preferred GP is associated with a 0.6% decrease in premature mortality counts. However, the impact of improving from a below average practice (25th percentile) to an above average practice (75th percentile) is an increase in percentage satisfaction of 13% which would equate to a 7.8% decrease in premature mortality, if the relationship is causal and given other variables remaining the same. There is less evidence to suggest an association between continuity of care and all-age mortality.

Although the relationship between GPs per capita and premature mortality is positive, the confidence interval is wide and the $p$-value is high. This may reflect either that access to primary care is not an important explanatory factor in a country of universal
access or that GPs per capita is not a useful measure of access.

**Relation to other studies**

Continuity of care, based on responses to the GP patient survey (GPPS), has been shown to be associated with improved health outcomes. In a study at PCT rather than practice level, the percentage of patients who recalled being able to see their preferred GP was associated with lower mortality caused by cancer and COPD (but not CHD mortality), with a similar effect size to that found here (Levene et al., 2012). The same measure was associated with lower emergency admission rates (Bankart et al., 2011) although not emergency attendance (Baker et al., 2011). Possible reasons for the importance of continuity of care found here, in contrast to other studies may be due to premature CHD mortality being the outcome of interest here. This is discussed in more detail in Section 18.7.

Kiran *et al.* (2010) found a significant negative association between GPs per capita and CHD mortality, but no measure of continuity of care was included in the study. In a study which included a continuity of care measure, a negative association was found between GPs per capita and COPD admissions (Calderon-Larranaga *et al.*, 2011). However, a positive association between GPs per capita and health outcomes was found in an analysis of stroke admissions (Soljak *et al.*, 2011).

**Critical review**

The most appropriate measure of continuity of care and alternative measures have been carefully considered. For example, the possibility of combining responses to the GPPS were discussed in Chapter 16 and associations between practice characteristics, response rate and positive responses were analysed in Chapter 8. However, these measures are based on the GPPS, which has relatively low response rates, such as 44% in 2006/07. Non-response is more common in men and the young. Kontopantelis *et al.* (2010) have found that over-representation of females is unlikely to have introduced much bias as the effect of gender is small. However, since younger people tend to report lower levels of satisfaction, satisfaction may have been overestimated. It has been shown that responses to the GPPS match ‘mystery shopper’ style experiments where researchers have attempted to make appointments at a variety of practices (Campbell *et al.*, 2013). However, there has not been research to determine if patients’ response to the question ‘were you able to make an appointment with your preferred GP?’ is
associated with actually seeing the same GP over a period of time.

GPs per capita was included in the model as a measure of access, although other measures based on the GPPS were considered (Chapter 16). Evidence that GP supply is associated with health outcomes is mixed. This may be due to the robustness of the data which were used to calculate ‘GPs per capita’, which may not include all relevant staff (NHS Connecting for Health, 2011). In addition, the measure may not adequately reflect the amount of time each GP devotes to patient care. There has been no attempt to quantify the number of nurses, specialist clinics or other provision which might reflect ‘access’ to primary care.

**Implications for future research, practitioners and policy makers**

The ongoing importance of continuity of care in improved health outcomes suggests that this is a key area for future research in England. This is of particular importance at a time when politicians are keen to make promises about access to primary care (Wintour, 2014) and measures of access and continuity of care have been shown to be associated with greater use of out-of-hours primary care (Zhou et al., 2014). Determining whether responses to the GPPS are associated with continuity of care as demonstrated by patients seeing the same doctor on more than one occasion is a priority area for future research. In addition, determining which aspects of continuity of care are most important for improved health outcomes is an important area for further research.

Both policy makers and practitioners should be aware of the potential for continuity of care to improve health outcomes and consider whether appointment systems and policies are supporting this.

**18.7 Comparisons with all-age mortality**

The primary health outcome considered in this study is premature CHD mortality; in this study mortality aged under 75. The same models have been applied to all-age mortality. This analysis suggests that patterns of associations between population characteristics and all-age mortality are broadly similar to those with premature mortality for the majority of explanatory characteristics. However, there is no evidence of an association between the percentage male and all-age mortality. This may be explained by the high proportion of premature mortality which is male (75.3%) in comparison
with all-age mortality (59.2% male), see Table 5.1. The IRR for both diabetes prevalence and smoking prevalence are lower for premature mortality. Smoking prevalence is lower in older age groups; high smoking prevalence in a practice may reflect higher proportions of younger people smoking and not higher levels in all age groups. There is evidence that the model fit is better for all-age mortality.

The difference in the proportion of deaths which are male in patients of any age and those under 75 may explain some of the differences in the association between service characteristics and premature versus all-age mortality. The interpretation of the association between the percentage of patients with recalled perception of being able to see their preferred GP, used in this study as a measure of continuity of care, is different for all-age mortality when compared to premature mortality. There is little evidence of an association between all-age mortality and continuity of care.

The percentage of patients with chronic conditions offered smoking cessation advice (SM02) is significantly, positively associated with all-age mortality. The association with premature mortality is also positive; however the confidence interval includes one. Meaningful interpretations of this association are challenging.

Relation to other studies

The vast majority of studies consider health outcomes for the total population, and do not distinguish between different age groups. This is despite the current emphasis on premature deaths by the UK Government, (see the ‘Longer Lives’ website (Public Health England, 2014a)).

Bottle et al (2008) considered unplanned hospital admissions for CHD (2004/05) in two age groups 45-74 and 75 and over. Unplanned admissions were positively associated with socio-economic deprivation; the association was stronger in the younger age-group. They found no association between unplanned admissions in either age group with a QOF based measure of quality of care. Allender et al (2012) demonstrated a stronger association between socioeconomic deprivation and premature CHD mortality, compared to all-age CHD mortality. Evidence for a similar pattern of association was not clear cut in this analysis. The IRR for deprivation is slightly higher for all age mortality (1.003 compared to 1.005) and the confidence interval is narrower. This, again, may be associated with the higher proportion of premature mortality which is male and the complex relationships between poor health in males and females and socio-economic factors (Stafford et al., 2005).
Other studies have not found significant associations between smoking cessation advice and improved health outcomes. It is possible that the smoking cessation indicator is in some way measuring smoking prevalence. However, under the relevant QOF guidelines smoking cessation advice need only be recorded for those who have chronic conditions and are recorded as current smokers. The proportion of those with chronic conditions who are 75 or over will be high. The underlying achievement of recording smoking status in those with chronic conditions is high (mean: 95.4%). Coleman (2010) expresses concern that the recording of smoking cessation advice does not reflect actual advice received or the receipt of effective cessation support, bringing into doubt the use of this indicator as a measure of prevention.

In this study, no evidence of association was found between all-age mortality and continuity of care. This supports the findings of Levene et al (2012) who found no association between continuity of care and all cause, CHD and stroke mortality. This is contrary to the significant, association between continuity of care and premature CHD mortality found in this study. Continuity of care has been shown to be important in improving health outcomes in older people (Wolinsky et al., 2010; Worrall and Knight, 2011). This difference is not expected as 67% of CHD deaths are in those aged 75 and over. In this study continuity of care is measured by patients perception of being able to see their preferred GP, which may not reflect actual experience of continuity of care. Studies have shown that recall of being able to see a preferred GP increases with age, with over 90% of patients aged 75 and over expressing satisfaction (Kontopantelis et al., 2010).

**Critical review**

Reducing premature mortality is a key target for this Government and many CCGs. The majority of studies do not consider health outcomes by age group and this study shows that the interpretation of the importance of associations between aspects of primary care and mortality may depend on whether all-age or premature mortality are considered as the health outcome. However, none of the explanatory variables in this study were available categorised by age. It is known that many of the variables vary with age and this may affect the interpretation of the results. Although it is now relatively common to use 75 as a boundary age for ‘premature mortality’ (Public Health England, 2014a), a different definition of premature may have given different results. QOF data are also not available by gender; given the high proportion of premature deaths which are male this also limits the possible analyses.
Implications for further research, practitioners and policy makers

The importance of reducing premature mortality means that studies of mortality should aim to consider whether results are similar when premature mortality is considered. Policy makers and practitioners should be aware the research on all-age mortality may have masked important associations between continuity of care and mortality. The focus on premature mortality means the collation and publication of key QOF data by age and gender should be considered.

18.8 Critical review - summary

This thesis has identified and evaluated modelling approaches which are commonly used in research of this type; multiple linear regression of standardised mortality rates; poisson and negative binomial regression of counts; and multilevel linear regression. Alternative approaches are available, for example path analysis which might have allowed causal relationships between variables to be considered (Olobatuyi, 2006).

The data available to describe key factors known to be associated with CHD have been carefully considered and factors for which there is no reliable data available have been highlighted. The thesis has focussed on publicly available data and did not request additional data from PCTs as the research progressed, partly due to the changing organisational structure of PCTs. For example, the proportion of patients resident at individual postcodes which would may have improved the estimates of obesity and allowed smoking prevalence data to be more thoroughly interrogated. This information is now publicly available (HSCIC, 2014b).

The sample of practices included in this study includes only practices within three PCTs in the East Midlands. The geographical area includes rural and inner-city areas and has an ethnic composition similar to England as a whole. However, there are many areas in England which have different ethnic composition, different patterns of socio-economic deprivation and are served by different secondary and tertiary care. These results are applicable to the sample of practices and although many results can be considered in a national context, further research is necessary to confirm which results are generalisable.

Patient level data were not available for this study, nor was patient level analysis the main aim. However, access to patient level data would have enabled a more detailed
understanding of various aspects of this research. In particular it would allow a more
detailed understanding of how QOF indicators affect individual patients and groups of
patients of similar age, ethnicity and those with comorbidities.

Detailed information on exception reporting was not available for 2006/07 QOF data.
From 2007/08 information on rates of exception reporting was readily available and
could have been used to consider rates of exception reporting for different indicators and
individual practices. Including patients who have been excluded from specific indicators
provides an actual percentage of patients for whom a target has been achieved and this
may provide a better measure of quality of care.

This research has focussed on the main effects of key variables. Interactions between
variables may also be of interest. For example, Kiran et al (2010) found interactions
between deprivation and quality of primary care to be important in predicting CHD
mortality; similarly there was evidence of an interaction between booking with a pre-
ferred GP and deprivation in a study of emergency admissions (Gunther et al., 2013).
There are, therefore, reasons to consider interactions between socio-economic depriva-
tion and quality of care. However, when estimates of smoking prevalence were included
in this analysis socio-economic deprivation was no longer a significant predictor of pre-
mature CHD mortality, which would make interpretation of interaction terms which
included deprivation difficult. Whilst there may be interest in interactions between
variables, the level of aggregation of the data affects the interpretation of possible
interactions. Interactions are more likely to be operating at individual patient than
practice level. Interactions between patient level factors, such as ethnicity, and prac-
tice level factors, such as access, may also be important. Access to patient level data
would allow further research into this area. Given the lack of interactions included in
previous research of this type there were no strong \textit{a priori} reasons for considera-
tion of specific interactions.

The mortality data were available for three financial years, but were aggregated due to
small numbers of deaths. An alternative approach would have been to model the three
years separately, using a model which takes into account the excess zeroes. This would
confirm if the results are consistent over the three years. However, the interpretation of
the results would not be straightforward; using QOF data from years contemporaneous
with the period of mortality may not aid in our understanding of the association
between mortality and primary care, as the QOF data may not reflect the experience
of the patients who have died. In addition, the large number of practices for whom
there were no deaths in individual years means it would have been difficult to compare
the analysis of count models with multiple linear regression of SMRs.

Ethnicity is a known risk factor for CHD. For example, CHD accounts for a high proportion of deaths for men born in South Asia but dying in England and Wales; 27% of all deaths compared to 18% for people born in the UK (Scarborough et al., 2010, p14). Mortality trends also show difference between ethnic groups; for the majority of groups mortality trends have decreased since 1979 but for males born in Bangladesh mortality has increased (p19). In contrast, studies of CHD mortality by ethnic group rather than country of birth have shown highest CHD mortality in those who are white (p11). The data which were available necessitated using the percentage who were white in the analysis. However, to summarise ethnicity in this way, based on self reported ethnicity in hospital episode statistics extrapolated to practice populations, does not allow the impact of the proportion of patients who were born outside of the UK and are members of different ethnic groups on CHD mortality to be properly explored. The inclusion of smoking prevalence, socio-economic deprivation and diabetes prevalence will act as proxies for some of the differences in mortality between groups. Data on ethnicity based on the GPPS are now readily available for practices and includes a more detailed summary of ethnicity.

18.9 Recommendations - summary

Little research has focussed on premature mortality and this research suggests conclusions about the impact of the quality of care are different when premature mortality is the outcome measure. There should be further research with premature mortality as the outcome measure to validate these important findings to other areas and other causes of death. Recommendations for further research, policy makers and practitioners are summarised in Table 18.2.
<table>
<thead>
<tr>
<th><strong>Policy makers and practitioners</strong></th>
<th><strong>Further research</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>QOF</strong></td>
<td></td>
</tr>
<tr>
<td>Publication of QOF achievement by ‘groups’ of patients.</td>
<td>Patient level analysis of QOF indicators to identify which groups of patients primary care is most effective in ‘treating’ in terms of QOF.</td>
</tr>
<tr>
<td></td>
<td>Patient level analysis of recording of smoking status for QOF to estimate smoking prevalence in both practice populations are local areas.</td>
</tr>
<tr>
<td><strong>Health check data</strong></td>
<td></td>
</tr>
<tr>
<td>Collation and publication of health check data to provide detailed information on both BMI and serum cholesterol levels, as well as self reported dietary physical (in)activity and alcohol information.</td>
<td>Analysis of health check data to determine its usefulness in observational studies of health care.</td>
</tr>
<tr>
<td><strong>Modelling approaches</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quantification of the size of affect of key variables to allow comparisons between multiple linear regression and count models.</td>
</tr>
<tr>
<td></td>
<td>Research should include sensitivity analyses on the impact of modelling approach.</td>
</tr>
<tr>
<td></td>
<td>Interactions between key variables at practice level, at patient level and between patient and practice level may be an informative area of further research.</td>
</tr>
<tr>
<td><strong>Quality of care</strong></td>
<td></td>
</tr>
<tr>
<td>The increasing evidence that continuity of care is an important aspect of primary care means policy makers need to consider how targets, organisation of primary care and funding affects the continuity of care.</td>
<td>When carrying out research using QOF indicators the rationale and method of indicator selection should be included in research publications.</td>
</tr>
<tr>
<td>Given the known link between hypertension and cardiovascular diseases, including stroke and CHD, improving the diagnosis of hypertension may be associated with decreasing mortality from stroke and CHD and may decrease A&amp;E visits and admissions.</td>
<td>Further research into ‘disease detection’, with a focus on the validity of modelled estimates for practice populations and groups of patients, is also important.</td>
</tr>
<tr>
<td><strong>Developing understanding of practice populations</strong></td>
<td>Researchers should work with practices to ensure data are presented in an accessible format and inferences based on practice level data are valid.</td>
</tr>
<tr>
<td>Practice population characteristics, including age, gender and the prevalence of smoking and diabetes, are important predictors of premature CHD mortality. Practices collect considerable quantities of data about individual patients which can be readily collated into information about practice populations. This can be utilised to support practice staff in understanding outcome measures which are published as part of practice profiles. Practices can also analyse their own QOF data by groups of patients which would enable them to inform their own practice.</td>
<td>Development of graphical representation of statistical analysis to aid interpretation by practitioners and policy makers.</td>
</tr>
<tr>
<td>Research into pathways which support our understanding of ethnicity and health.</td>
<td></td>
</tr>
</tbody>
</table>

Table 18.2: Recommendations for policy makers, practitioners and further research.


18.10 Conclusions

This thesis adds to the body of research demonstrating that high quality primary care is associated with improving health outcomes. Aspects of continuity of care, disease management and detection of chronic disease have been identified as having an impact on reducing premature CHD mortality. The association is weaker for all-age CHD mortality. Whilst the most important individual indicators relating to disease management have not been identified, there is clear evidence that improving achievement in some QOF indicators is associated with decreasing CHD mortality. The importance of continuity of care, again shown here, strongly suggests that this is an area for general practices to prioritise. The findings also suggest that data about outcomes such as premature CHD mortality could be used by practices to monitor and, over several years, plan their care to improve population health.

The lack of reliable practice level information on key areas such as obesity, diet, alcohol and physical (in)activity has an important impact on primary care research and is an important health information issue needing effective attention as the NHS undergoes major changes. The data collated as part of the health check programme could provide valuable information about the prevalence of certain risk factors in the population. In addition, data available for patients aged under 75 would allow an improved understanding of factors affecting premature mortality.

This thesis clearly shows the additional burden some practices face as a result of above average risk factor prevalence in the practice population. However, both smoking and diabetes prevalence are modifiable factors. The lack of reliable information about smoking prevalence in local populations, including practice populations has been a major deficit in studies of this kind. Chapter 15 shows how QOF data can be manipulated to provide estimates of smoking prevalence in practice populations and suggests that this data can be aggregated to provide estimates in local populations. The removal of QOF smoking indicators which apply to the whole practice population in 2014/15 means a valuable source of information about one of the most important factors in premature mortality will be lost.

Multiple linear regression of standardised rates is not appropriate for studies of this type, particularly when age-standardised explanatory variables are not available. In this thesis, negative binomial models of count data were considered the most appropriate, as the data showed evidence of over-dispersion. When results of different modelling approaches were compared, in some cases both the direction of association and statis-

246
tical significance were affected.

To further our understanding of the relative importance of different QOF indicators it would be useful to study individual QOF indicators at patient level and how they interact with each other and with characteristics of individuals. The potential benefit of publication and subsequent analysis of QOF achievement in subgroups of practice populations may outweigh the cost and should, therefore, be considered. It is currently possible for practitioners and CCGs to analyse this data by subgroup; this may lead to an improved understanding of patient populations and may result in improved health care.

Improving the quality of primary care will play an important part in decreasing premature mortality, and there is evidence that high underlying achievements in QOF clinical indicators are a useful measure of quality primary care. Continuity of care, in a country with universal access to health care, is important and should not be underestimated by policy makers and clinicians. The importance of diagnosing hypertension, given its links with all cardiovascular diseases, is a key area of priority.

Understanding meaningful associations between primary care and health outcomes is an important area of research. This is assisted by collecting the right data, using the most appropriate statistical methods and generating relevant models based on up-to-date theoretical frameworks.
Bibliography


Anwar, M. S. et al. (2012). “Chronic disease detection and access: does access improve detection, or does detection make access more difficult?” *Br J Gen Pract* 62.598, pp. 337–343. DOI: 10.3399/bjgp12x641456.


256


Rose, I (2012). Email to Kate Honeyford. email communication. 21 March 2012.


