Examining the effectiveness of Quality Outcomes Framework targets using individual level data: An econometric analysis

Neil Michael Hewitt MA(Hons) MSc

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

OCTOBER 2014

ABSTRACT

Background

The Quality Outcomes Framework (QOF) was introduced in April 2004, and is the largest primary care pay for performance (P4P) scheme in the world[1]. To date research on the QOF to date has principally looked at its effects on process and surrogate outcomes rather than condition specific and clinically significant outcomes. Where research has looked at clinical outcomes, measures have not been condition specific, and were measured at a population or local level, not that of the individual patient. Large administrative databases like Clinical Practice Research Datalink (CPRD) and linked Hospital Episode Statistics (HES) allow researchers to track individuals across Primary and Secondary care to determine the impact of policies in either sphere. This study utilises the links between these two large datasets to determine the impact of evidence based targets in the QOF on clinical outcomes at the individual level

Methods

This thesis sets out to comprehensively examine the effect of physician incentives generally and the QOF specifically on individual level outcomes through a series of structured literature reviews and data analysis. The literature reviews, in conjunction with relevant clinical guidelines, triangulate the evidence around the effectiveness of incentives, their impact on CHD clinical outcomes, and the impact of QOF CHD targets. The selection of QOF targets for analysis is directed by clinical evidence. NICE and SIGN commissioned guidelines are consulted to select from the QOF those targets which had high level evidence of clinical benefit, which were extractable from large administrative datasets. This ensured in the data analysis that clinical benefits arising from treatment to QOF targets would be evident in the outcome measure.

Individuals were selected into the dataset if they had a QOF qualifying Read code for CHD and linked into HES data. Since the QOF has been near universally adopted, 4 years of pre-QOF data were used in addition to 7 years of QOF data, to measure compliance with evidence based targets in the QOF at the individual level. Outcomes were selected ICD 10 I20-25 codes, corresponding to the Ischaemic Heart Disease (IHD) category, and all relate to complications arising from poor condition

management. These had to represent the primary diagnosis for a hospitalisation and, depending on the severity, be either an emergency or non-emergency admission.

Panel data econometric analysis was undertaken at the individual patient level. The outcome variable was specified in two ways, either as a hospitalisation count per year or a binary hospitalisation occurrence variable. Adjustments were made to model specifications for over-dispersion and excess zeros in the outcome variable, correlation between the explanatory variables and the error term, and heteroskedasticity in the error term.

Results

Literature and clinical guideline reviews

The high level evidence base on physician incentives in primary care showed no strong evidence for any particular form of incentivisation, let alone P4P. Incentives were followed up for short durations, and there was a focus on process and surrogate outcome measures. None linked incentives to individual level condition specific clinical outcomes.

The majority of QOF targets do not measure or reward performance to a surrogate or clinical outcome measure. Rather they reward processes which often are not evidence based, and have no direct links to clinical outcomes, making it difficult to determine and attribute clinical benefits to them. In terms of data extractability using large administrative datasets, CHD QOF targets were judged to be most suitable.

Evidence on the effects of physician incentive on clinical outcomes, hard and surrogate, in CHD patients was limited. Most examined their impact over a short time period. None of the studies monitored patient contact with primary and secondary care.

The published literature on CHD and the QOF was very limited in that it largely relied on uncontrolled before and after cross sectional study designs, and one dataset from an inner London PCT.

Econometric analysis

Evidence based CHD QOF targets were found to significantly reduce outcomes in all whole population models and most sub group models ($p \leq 0.05$) within a one year time period. The cholesterol QOF target was additionally found to reduce outcomes in the same period, in most analyses. Having a co-morbidity tended to increase

outcomes, however the results were not always statistically significant. Only Heart Failure (HF) was consistently found to significantly increase outcomes in all analyses (OR & IRR≈2, $p<0.001$)). Worsening levels of deprivation consistently led to a significantly greater number of outcomes with few exceptions. Being treated in a higher attaining practice on QOF CHD targets significantly reduced outcomes in all analyses ($p<0.001$). Having an outcome event prior to CHD coding significant increased outcomes throughout the analyses (OR & IRR≈2.5, $p<0.001$)). Increasing age and being male were generally causes of increased outcomes. Rheumatoid Arthritis, Strategic Health Authority (SHA) region, practice size and GP workload were found to have a minimal or no impact on outcomes.

Conclusion

This thesis has demonstrated at the individual patient level that evidence based CHD targets in the QOF are effective in reducing evidence linked hospital admissions allowing for a one year delay. It has done so using a large sample of patients from across England, over an 11 year period. These results have been tested in various model specifications to correct for heteroskedasticity in, and covariance with, the error term and an excess of zeros and over-dispersion in the outcome variable. In all model specification this finding was found to hold true. This research has demonstrated the benefits of using large administrative datasets and the importance of linking outcomes to high level evidence.

ACKNOWLEDGEMENTS

This work has been conducted under the supervision of Professor Nadeem Qureshi (Division of Primary Care, Medical School, University of Nottingham) and Dr Murray Smith (School of Pharmacy, University of Nottingham). I am extremely grateful to them both for their guidance, encouragement, assistance and insight throughout this work. This work has been supported by a studentship from the National Institute for Health Research (NIHR) School for Primary Care Research (SPCR). I am eternally grateful for their support and the opportunities it has opened up for me.

I would like to thank the Division of Primary Care for facilitating and supporting this work. Acknowledge my debt to Dr Casey Quinn, latterly of the Division of Primary Care, for persuading me to undertake this thesis, and getting it started. Finally big thanks go to family and friends, and the numerous other individuals who, one way or another, have helped me reach this position.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACC	American College of Cardiology
ACE	Angiotensin-Converting-Enzyme
AF	Atrial Fibrillation
AHA	American Heart Association
AMI	Acute Myocardial Infarction
ARB	Angiotensin Receptor Blocker
AII/A2	Angiotensin II
BHS	British Hypertension Society
BMA	British Medical Association
BP	Blood Pressure
CABG	Coronary Artery Bypass Graft
CBA	Controlled Before and After Study
CHD	Coronary Heart Disease
CKD	Chronic Kidney Disease
COPD	Chronic Obstructive Pulmonary Disease
CPRD	Clinical Practice Research Datalink
CVD	Cardiovascular disease
CVD PP	Cardiovascular disease Primary Prevention
DM	Diabetes Mellitus
FFS	Fee for Service
GRADE	Grading of Recommendations, Assessment, Development and Evaluation
GP	General Practitioner
HES	Hospital Episode Statistics
HF	Heart Failure
HMIC	Health Management Information Consortium
HMO	Health Maintenance Organisation
HSCIC	Health and Social Care Information Centre
ICD	International Classification of Diseases
IHD	Ischaemic Heart Disease
IRR	Incident Rate Ratio
JBS	Joint British Societies'

LVD	Left Ventricular Dysfunction
MH	Mental Health
MI	Myocardial Infarction
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
OR	Odds Ratio
PCP	Primary Care Physician
PCT	Primary Care Trust
PMS	Personal Medical Services
PP	Primary prevention
P4P	Pay for performance
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QIP	Quality Incentive Programme
QMAS	Quality Management Analysis System
QOF	Quality Outcomes Framework
RA	Rheumatoid Arthritis
RCT	Randomised Control Trial
SHA	Strategic Health Authority
SIGN	Scottish Intercollegiate Guidelines Network
SMR	Standardised Mortality Ratio
STIA	Stroke and Transient Ischaemic Attack
THIN	The Health Improvement Network
UK	United Kingdom
VA	Veterans Affairs

Chapter 1 Introduction

1.1 Introduction

This thesis sets out to examine the effect of physician incentives generally, pay for performance (P4P) specifically, on health outcomes. It does so in the context of changes made to the UK primary healthcare sector in 2004 which saw the introduction of a P4P scheme referred to as the Quality Outcomes Framework (QOF). This is examined at an individual patient level using large administrative datasets, with evidence based CHD QOF indicators used in econometric analyses.

1.2 Objective

This chapter sets out the aims and objectives for this thesis and describes and appraises the key components that are covered within it. These are:

1. The QOF, the physician incentive scheme which is the focus of this thesis.
2. CHD, the condition that is used to examine the effectiveness of evidence based targets in the QOF.
3. Large proprietary datasets, which were used to source data to model and analyse the QOF.

1.3 Overview of the QOF

1.3.1 Background

The Quality Outcomes Framework (QOF) was introduced in April 2004 as part of the new General Medical Services (GMS) contract, 'Delivering Investment in General Practice', published in December 2003[2]. It is considered to be one of the most ambitious pay for performance schemes ever undertaken anywhere in the world in primary care[3]. The GMS contract introduced new funding arrangements for primary care, and saw the abolition of the Red Book (Statement of Fees and Allowances), which had prior to this set out the payment tariffs for individual treatments. It contained six main funding streams, one of which was the QOF, which was expected to account for around 20% of practice income[4]. Although the scheme was voluntary, from the outset nearly 100% of GP practices chose to participate.

1.3.2 QOF domains

The QOF was designed to incentivise improvements in practice organisation, the delivery of care and provision of additional services in certain designated clinical areas. It consists of four domains: clinical, organisational, patient experience and additional services. The organisational domain rewards practices for good organisational practice and Human Resource management and is based around the Royal College of General Practitioners' Quality Practice Award. Patient experience uses patients' elicited feedback concerning the quality of access and care gathered through a quarterly national survey to determine practice payments. The additional services domain rewards those practices that provide extra functions such as cervical screening and maternity services.

The biggest domain is the clinical domain, accounting for 550 points, just over 50% of the total points on offer in 2004/05. This domain is the focus of this chapter and thesis. Within the clinical domain there were 76 indicators in 11 clinical areas when the QOF was introduced: Coronary Heart Disease (CHD), Left Ventricular Disease (LVD), Stroke or Transient Ischaemic Attack (STIA), Hypertension, Diabetes, Chronic Obstructive Pulmonary Disease (COPD), Epilepsy, Hypothyroidism, Cancer, Mental Health (MH), and Asthma. Each target has 'payment stages' which set lower and upper thresholds for payments. Failure to meet the minimum threshold means payments are not triggered while the upper threshold sets a ceiling on the level of achievement for which the practice will receive QOF payments. For example, QOF target, BP4, payments thresholds were 25% and 90% at its inception. This means that numbers of patients treated below the 25% threshold and above the 90% threshold would not generate QOF points and payments for their practices; and hence there was no financial incentive to treat patients if the lower threshold could not be reached, or to treat patients who fell outside the upper threshold. Payments are made proportionately within those thresholds. Practices are allowed to exempt patients provided certain criteria are met, which reduces the denominator and increases attainment; this is called "exception reporting." Exception reporting is designed to prevent general practices being financially disadvantaged for not treating certain patients: Namely those who were non-compliant or had contraindications.

1.3.3 How does the QOF work?

What practices need to do to comply with QOF targets is detailed in the QOF business disease specific ruleset. These are updated as and when the QOF targets are updated. This thesis used version 17, produced 07/05/2010, as this was the most recent to the study

period. Practices receive payments based on the amount of their patients who meet the QOF target, which is determined by the size of each disease specific register, which, in turn, determines the numerator for the disease targets, and whose recording itself is a QOF target. Calculation of performance works as follows: the first target in each clinical indicator is termed 'Records,' and is a payment made to the practice for initially establishing and then maintaining and updating a register consisting of patients, within the practice, who have the condition for that clinical indicator. For example, the first target in Hypertension, BP1, is that 'The practice can produce a register of patients with established hypertension'. This is the only rule to this target, so attainment should be 100%. This disease register then forms the disease register for the majority, but not all of the subsequent targets. Following on from the disease register there is often an 'Initial Diagnosis' target. This pays the practice based on whether a specified test has been used to verify the existence of the disease or to determine its specific form. This may be merged with, or there may be a separate 'Initial management' section; which does essentially the same thing. After achieving targets for initial diagnosis and investigation, the next section deals with 'Ongoing management targets. Prior to this point practices will have been paid for recording a patient's condition and possibly performing a confirmatory test or assessment. It is in this section where practices start to have to hit specified clinical targets, in the form of surrogate and process outcome measures, to trigger payments. These tend to be for specific blood pressure readings (e.g.≤150/90mmHg); serum cholesterol levels (e.g. ≤5mmol/l); and the prescription of certain pharmaceutical products such as ACE inhibitors, beta blockers, anti-platelets or anti-coagulants. Most of these apply to the whole disease register established in the first disease target minus exceptions; however some apply to sub groups of that population. For instance it could be CHD patients who have an MI coding, or Heart Failure patients with Left Ventricular Dysfunction (LVD). There still remains within this section however a significant number of targets that provide payment regardless of outcome. These are related to and precede the surrogate outcome targets, providing payment if the test or reading took place.

1.3.4 Changes to the QOF

The QOF, when first introduced in April 2004, contained eleven clinical conditions in the clinical domain, including LVD which formed a subgroup of the CHD target group. Since then the QOF has undergone a number of changes and the frequency of those changes has increased over its lifetime. The first and most substantive changes to the QOF occurred in

2006. This saw wholesale changes to lower QOF thresholds, the introduction of nine new clinical conditions and the retirement of LVD, with its targets being applied to the newly introduced Heart Failure (HF) condition group. Although there were minor changes to COPD in 2008, the next major change did not occur until 2009. This saw the introduction of a new clinical area, Cardiovascular Disease Primary Prevention (CVD-PP), and significant changes to the points offered to other clinical areas as well as greater weighting for the clinical domain within the QOF. As a result the clinical domain had come to dominate the QOF accounting for 70% of the points on offer, up from 52.4% at its inception. Changes made to the QOF from its introduction to the year 2010/11, the study period, are shown in Table 1-1 on the following page. LVD prior to the introduction of HF formed a subset of the CHD targets group and therefore its points, when present, are included in CHD's total for that period.

Table 1-1 Changes to the QOF over the study period, 2004-2011

	2004			2006			2008			2009		
	Number of Targets	Total points	%	Number of Targets	Total points	%	Number of Targets	Total points	%	Number of Targets	Total points	%
Clinical	76	550	52.4	80	655	65.5	80	650	65.0	86	697	69.7
CHD	15	121	11.5	10	89	8.9	10	89	8.9	10	87	8.7
STIA	10	31	3.0	8	24	2.4	8	24	2.4	8	24	2.4
Hypertension	5	105	10.0	3	83	8.3	3	83	8.3	3	81	8.1
Diabetes	18	99	9.4	16	93	9.3	16	93	9.3	17	100	10.0
COPD	8	45	4.3	5	33	3.3	5	28	2.8	5	30	3.0
Epilepsy	4	16	1.5	4	15	1.5	4	15	1.5	4	15	1.5
Hypothyroid	2	8	0.8	2	7	0.7	2	7	0.7	2	7	0.7
Cancer	2	12	1.1	2	11	1.1	2	11	1.1	2	11	1.1
MH	5	41	3.9	6	39	3.9	6	39	3.9	6	39	3.9
Asthma	7	72	6.9	4	45	4.5	4	45	4.5	4	45	4.5
Heart Failure				3	20	2.0	3	20	2.0	4	29	2.9
Dementia				2	20	2.0	2	20	2.0	2	20	2.0
Depression				2	33	3.3	2	33	3.3	3	53	5.3
CKD				4	27	2.7	4	27	2.7	5	38	3.8
AF				3	30	3.0	3	30	3.0	3	27	2.7
Obesity				1	8	0.8	1	8	0.8	1	8	0.8
Palliative Care				2	6	0.6	2	6	0.6	2	6	0.6
Learning Disability				1	4	0.4	1	4	0.4	1	4	0.4
Smoking				2	68	6.8	2	68	6.8	2	60	6.0
CVD-PP										2	13	1.3

The net results of changes to the QOF during the study period have been to widen its scope and make initial steps in making the QOF more demanding. More recently as the rate of increase in NHS funding has slowed, monetary concerns have also played a significant role in changes to the QOF and in efforts to make it more exacting[5]. There have been two main factors driving and more recently accelerating these changes. Firstly successive governments have sought to claw-back some or all of the substantial gains witnessed in the early years of the QOF in GP pay and conditions[6]. Secondly the introduction of the NICE QOF committee in 2009 and initiatives emerging from that have led to a more active and evidence based approach in the formulation of new QOF targets, and in the retirement of existing targets[7]. The effect of this greater use of NICE evidence to formulate targets was seen in the decision in 2011 to replace 13 targets with 17 NICE recommended ones in eight clinical areas.

1.3.5 QOF attainment

QOF attainment has been high from the outset and followed an upward trajectory for most of the study period. This is shown for CHD and Overall performance at the English national level on all QOF clinical conditions in Figure 1-1 below. Details on the remaining QOF targets are shown in Table 1-2 on the following page.

Figure 1-1 Performance on QOF targets, CHD and all targets

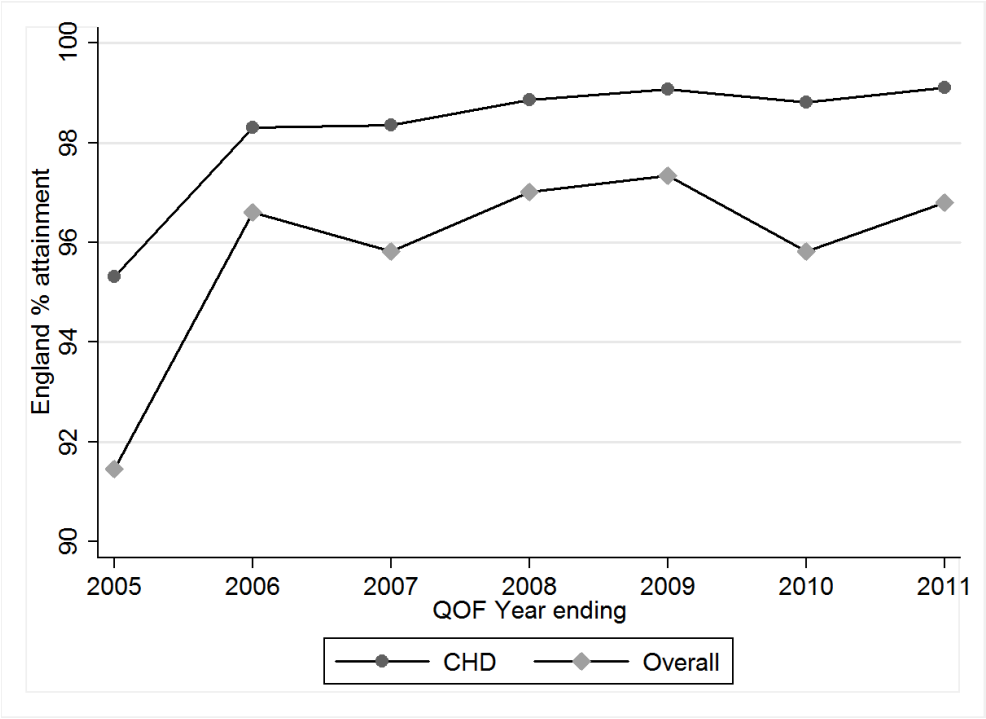


Table 1-2 QOF conditions performance over the study period, attainment at the English national level

Year	Stroke	Hyper	Diabetes	COPD	Epilepsy	Hypothy	Cancer	Mental Health	Asthma	HF	Paliative Care	Dement	Depress	CKD	AF	Obesity	Learn Disab	CVD PP
04/05	92.0%	94.4%	93.2%	86.2%	86.8%	98.2%	89.6%	88.5%	91.6%									
05/06	97.2%	98.1%	97.4%	95.6%	93.7%	99.5%	95.3%	94.6%	97.3%									
06/07	97.3%	98.3%	97.5%	96.0%	93.8%	99.6%	96.2%	91.6%	96.4%	96.1%	90.1%	97.2%	80.8%	97.5%	97.2%	100%	98.3%	
07/08	98.2%	98.8%	98.0%	97.4%	94.3%	99.6%	96.0%	93.9%	97.6%	97.7%	92.3%	97.3%	90.6%	96.8%	98.3%	100%	98.7%	
08/09	98.2%	98.9%	98.4%	96.2%	94.7%	99.6%	96.6%	94.7%	98.2%	98.4%	93.0%	98.0%	92.8%	96.9%	97.8%	100%	99.0%	
09/10	98.0%	98.9%	95.2%	95.8%	94.7%	99.4%	95.8%	94.5%	98.1%	97.3%	89.3%	97.5%	81.7%	94.7%	97.9%	100.%	98.6%	91.2%
10/11	98.6%	99.1%	96.1%	97.4%	95.0%	99.7%	96.9%	95.2%	98.7%	98.0%	91.4%	98.1%	85.1%	96.4%	98.6%	100%	99.2%	93.7%

Hypothy = Hypothyroidism Hyper=Hypertension Dement= Dementia Depress= Depression Lean Disab= Learning Disability

Both Figure 1-1 and Table 1-2 are based on England national level attainment. LVD, which was present in the first two years of the QOF is not reported. This condition applied to a subset of CHD patients and was included in the CHD domain, though its results were reported separately. However it has been used in the calculation for average overall performance for all QOF targets.

Figure 1-1 shows that average attainment over the study period on CHD QOF targets has been higher than the overall average for all QOF targets. Performance on both measures peaked in 2009, with 99.1% attainment on CHD and 97.3% overall. While CHD repeated that level of attainment in 2011, overall performance has slipped since, though it still remains high.

Table 1-2 breaks down the overall figure by the composite conditions making up the QOF, CHD and LVD aside. This shows that for all conditions with the exception of depression and palliative care on average; since 2005/06, QOF practice attainment has been in the mid to high 90% region. In the case of obesity the average practice, allowing for rounding up, has achieved 100% on this target over its lifetime. All the figures point to universally high attainment with little exception, on all targets.

1.3.6 Critique

Too much of the QOF is devoted to rewarding activities such as recording disease status, the presence of disease specific complications, or whether surrogate outcome measures such as cholesterol levels are recorded. Beyond these examples the prescription targets are met simply by dispensing the specified drugs rather than for getting the patient to a desired outcome. Likewise a number of targets relate to reviews of conditions which are also activity based with no specified, discernible and consequently measurable clinical outcome. Where outcomes are specified they are surrogate or process measures rather than clinical outcomes. This is a general issue with any performance related pay scheme, particularly in an area such as primary care. Namely payments need to occur a reasonable time after the activity, in line with salary cycles or annual accounting periods. Consequently items that can be measured with reasonable immediacy, such as prescriptions, surrogate outcomes or reviews of a patient's condition, are favoured as opposed to clinical outcomes which can occur in the years and decades following primary prevention and treatment.

Understandably a pay for performance scheme under which practices average over 91% overall in the first year and then go on to improve on that performance raises questions as to how demanding the targets are. For those funding the scheme it also made it more costly than initially expected. At its introduction, expectations were that the QOF would account for 20% of a GP's income, based upon GPs achieving around 75% of the QOF points on offer on average. However average GP performance on the QOF turned out to be over 90%, accounting for 25% of practice income. This led to a total over spend on the QOF by the government of £1.76 billion in its first three years alone[8]. Despite committing substantial resources to the scheme - in excess of £1 billion a year - no pilot studies were undertaken prior to its implementation; no cost-effectiveness studies were done on any of the targets and, in some instances, neither were the targets based on strong clinical evidence. No baseline measures were taken prior to its implementation in order to determine subsequent performance, which turned out to be much better than predicted[8]. These costs could be absorbed by generous financial settlements for the NHS over the study period, but looking back critically from more 'austere' times it may not have been a good use of scarce resources.

Various reasons have been given for this high level of attainment. Some practitioners have claimed it shows that that they had been providing high quality care and investing in better care, prior to the QOF, but that went unrecognised until it was captured by the QOF[9]. In this sense the QOF was a period of catch up where pay and recognition caught up with previous effort and investment; as well as a continued commitment to quality. Others point out that the government of the day introduced the new contract to address recruitment and retention issues within general practice, and the QOF was part of a wider policy of intensive investment in the NHS and primary care to make general practice a more financially attractive proposition, rather than solely a quality-improvement mechanism[9 10]. In this regard the concern was not to challenge GP's to do more and improve patient outcomes but to pay GP's more for work they were already undertaking. Indeed the QOF was sold to its members by the BMA as 'less work, more pay, a better pension.'[11] In terms of the evidence Campbell et al have found that care for CHD patients did not improve significantly beyond existing trend with the introduction of the QOF based on mean attainment on a range of QOF related measures; and while diabetes and asthma care did improve significantly beyond trend in the QOF's first year,

they subsequently reverted back to trend[12]. Hence the QOF at best resulted in a small incremental improvement on existing trends and practitioners getting to a point on QOF measures a year or two earlier than they would otherwise have done. There is evidence that exception reporting has masked inequalities enabling practices to report uniformly high achievement[13]. There is also evidence to show that it has also been used by some practices to 'game' the system by altering their exception reporting rates in order to improve their QOF score[14]. There has been some suggestion that GP's 'cherry picked' patients with the condition who are the easiest to manage and treat, though there is not strong evidence for this[14]. There is mixed evidence on resource substitution, a process where physicians divert their efforts and attention to incentivised conditions at the expense of those not covered under the incentive scheme. Sutton et al found positive spill over effects into non incentivised areas of 10.9 percentage points for those patients targeted under the QOF[15]. Campbell et al on the other hand found that mean quality scores fell in most conditions on non-incentivised targets following the QOF, widening existing gulfs in attainment[12]. While Steel et al found no significant improvement in non-incentivised compared to a significant improvement in those covered by incentives and linked to incentivised conditions [16]

Targets were set below 100% even after allowing for exception reporting, leading to what Fleetcroft et al termed a pay: performance gap[17]. This gap was the difference between the number of patients eligible for the target and those actually treated. Upper thresholds set below 100% counted for 52% of this gap, exception reporting the rest. Finally, critics have argued that the QOF has simply led to a natural incentive response by GPs to 'follow the money' and better record those areas of care that are incentivised and treat those conditions where points are available; which does not in itself necessarily equate to better quality of overall care[18].

1.4 Coronary Heart Disease

This section discusses Coronary Heart Disease (CHD), the condition that was chosen to analyse the effects of evidence based targets in the QOF in this thesis. The process by which this condition was chosen will be covered in Chapter 3.

1.4.1 What causes Coronary Heart Disease?

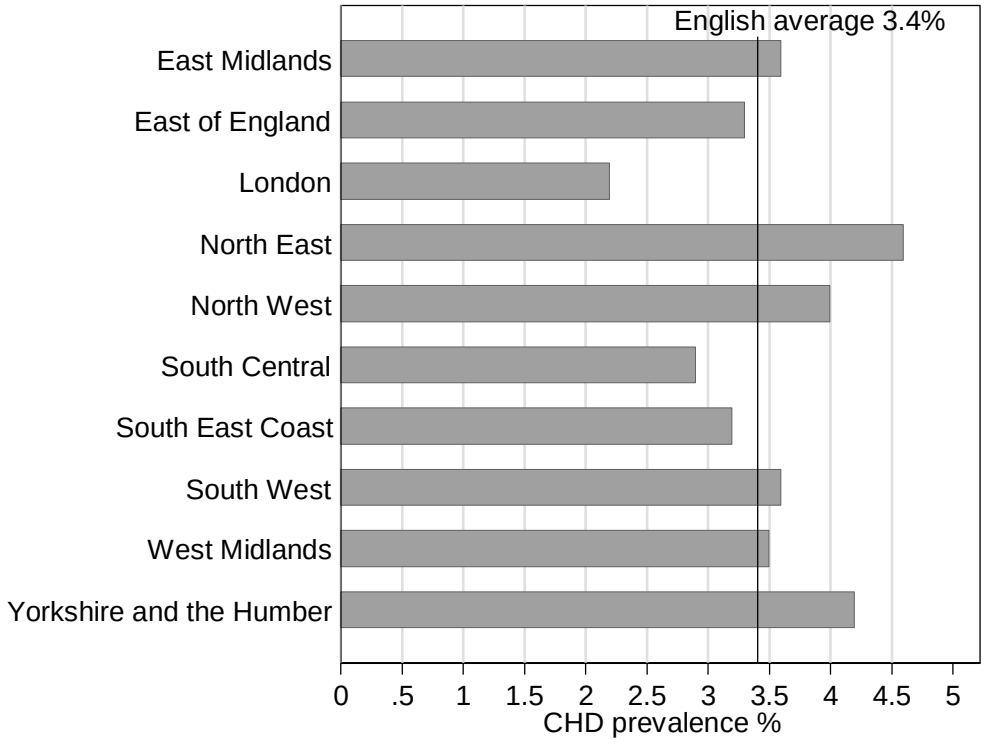
Coronary Heart Disease (CHD) is caused by a build-up of fatty deposits in the walls of the coronary arteries which supply blood to the heart. As these deposits build up this reduces the size of those arteries and limits their ability to pass blood and oxygen to the heart. This causes the heart to work harder particularly during periods of physical exertion, leading to chest pains; an indication of the most common form and usually the first symptom of CHD, Angina. Pain and tightness can also occur in the shoulders, arms, throat or jaw. If this pain comes on gradually, improves with rest and lasts for a couple of minutes it is classed as stable angina. Unstable Angina is more serious where the pains are unpredictable and last longer, and can be a sign that a heart attack is imminent. A heart attack or myocardial infarction refers to the situation where one of the coronary arteries is completely blocked starving the heart of oxygen. Without treatment the patient will die or the heart will be permanently damaged. Due to its severe implications preventing patients from having an MI will be a primary objective of treatment for patients with existent CHD and secondary prevention measures. While CHD is a progressive disease, patients need not have suffered from Angina nor shown any symptoms of CHD or cardiovascular disease prior to having an MI. These patients would not come under the remit of the QOF or this thesis as patients cannot be treated under any of the secondary prevention of CHD QOF targets until CHD is diagnosed.

1.4.2 Coronary Heart Disease: facts and figures

CHD is the most prevalent of the cardiovascular diseases, the term given to the group of diseases that affect the heart and circulatory system. These are the most common cause of death in the UK accounting for nearly a third of all deaths, over 179,000, in 2010. CHD alone accounted for over 80,000 of these making it the largest cause of death in the UK, responsible for 17% of all deaths in males and 12% in females in 2010[19 20]. It is also one of the most common causes of premature death in both sexes, in men the largest cause, accounting for 17% of all premature deaths; in females behind lung cancer and respiratory disease accounting for 8% of all premature deaths[20]. The estimated cost of CHD to the UK economy was over £6.7 billion in 2009[20]. A little over half of this cost is due direct health care and informal care costs; with the remainder down to productivity losses[20].

Based on QOF data for the year 2010/11 it is estimated there are around 2.3 million people suffering with CHD in the UK[20]. Other estimates put the number of people living with CHD at approximately 2.7 million of which around 2 million have Angina, the most common symptom for CHD[19]. It is least prevalent in England with 3.4% of the population on the QOF CHD register and most prevalent in Scotland with 4.4% registered as having CHD[20]. Rates of death from CHD are also highest in Scotland, though as with prevalence these figures hide significant regional variation. A breakdown of CHD prevalence for English SHA regions is shown in Figure 1-2 using 2010/11 data[20]. In relation to the English national prevalence rate of 3.4% clear regional divides are evident. CHD is most prevalent in the Northern SHAs, the North East, North West and Yorkshire and the Humber. The Midlands SHAs and East of England have near English average levels of prevalence, with the East Midlands SHA deviating the most from that average with a CHD prevalence rate of 3.6%. With the exception of the South West, all the southern SHAs have CHD prevalence figures below the English average, London SHA having the lowest CHD prevalence rate of 2.2%.

Figure 1-2 CHD prevalence by English Strategic Health Authority, 2011



Rates of deaths from CHD vary with deprivation, with rates increasing with each quintile increase in deprivation. Research by McCartney et al.[21], reported by the

British Heart Foundation (BHF) shows a clear link between deprivation and death rates per 100,000 population from CHD, between the years of 1994 and 2008[20]. Deprivation was measured using the Carstairs index at a local authority level, grouped by quintile, examined by gender and additionally for all ages and those under 75 years of age. The results throughout the study period showed a clear progressive increase in rates of CHD deaths with each quintile increase in deprivation, for both genders, all ages and those under 75. These rates fell throughout the study period so that by 2008 rates were in all instances over half what they had been in 1994. Women in both age groups had the lowest rates throughout and enjoyed the biggest reductions in rates over the study period. However the differences in rates of deaths between the most and least deprived quintile, expressed as a ratio, was larger for females throughout, suggesting that the most deprived females performed less well compared to their least deprived peers than their male counterparts. Differences in rates between the least and most deprived were also greater for the under 75s in both genders. Furthermore over the study period, the differences in rates of death, expressed as a ratio, between the most and least deprived quintile increased from 1.3 to 1.5 for men of all ages; 1.4 to 1.6 for women of all ages; 1.5 to 1.8 for men under 75; and 1.8 to 2.3 for women under 75.

An alternative and arguably more comprehensive measure of deprivation is the Index of Multiple Deprivation (IMD), which considers seven aspects of deprivation as opposed to the Carstairs index four. This has been measured at the individual level using the 2007 algorithm in this study.

1.3.3 Secondary prevention of Coronary Heart Disease

This work is interested in patients who already have CHD and therefore in the prevention of secondary complications. This will involve the same lifestyle advice around increasing exercise, having a healthy diet and giving up smoking and other modifiable lifestyle factors that increase the risk of having CHD initially; and increase the risk of complications once diagnosed. However compared to primary prevention, following diagnosis individuals will require more intensive control and medical management through the use of drugs such as statins and ACE inhibitors, and these measures form the basis for CHD condition management in the QOF. At the same time reoccurrence and exacerbations in early and mid-stage

complications remain important negative outcomes that secondary prevention will seek to avoid.

Office for national statistics figures reported by the BHF show a rapid increase in the use of anti-hypertensive and heart failure, anti-platelet, and lipid lowering drug classes for the primary and secondary prevention of CVD from 1981 to 2011, which covers the whole of the study period, January 2000 to April 2010[20]. Starting from a low base, with prescriptions volumes below 1000 a year in 1981, 5000 for antihypertensives, volumes increased rapidly for all of these drug classes, most markedly after 2000, beginning in the early 1990s for anti-hypertensives and anti-platelets, and the mid-1990s for lipid lowering drugs. By the end of the period it appears that prescriptions of anti-hypertensives and lipid lowering drugs were nearing a plateau at volumes of approximately 61,000 and 65,000 a year respectively; while those for anti-platelets appear to have peaked in 2009, at volumes of near 40,000 prescriptions per year. The impact of all of these drug classes will be assessed directly or indirectly on the study outcomes in the course of any analysis.

The outcomes measured in this study relate to complications arising out of existent CHD, key among these are myocardial infarctions (MI). Work by Smolina et al.[22], reported by the BHF show the incidence of hospitalisation per 100,000 with a primary diagnoses for Acute Myocardial Infarction (AMI) in England, between 2002 and 2010, covering the majority of the study period[20]. These have been age standardised to the European standard population and are again reported by gender, for all ages, and those under 75. Incidence of AMI hospitalisation has fallen in both genders and age groups between 2002 and 2010, largely on a year by year basis. As expected the rates are highest in males though differences between the genders has narrowed over this period from a difference of 101 per 100,000 (169 male compared to 68 per 100,000 female, 2002) to 71 per 100,000 for the all age group (122 male compared to 51 female, 2010); representing a fall of about a quarter in males and a fifth in females. For those under 75 there has also been a narrowing in the size of the absolute difference between males and females; although the rate of decline was similar at around one third in both genders (123 to 86 per 100,000 male; 41to 29 per 100,000 female).

1.5 Large administrative datasets

1.5.1 Introduction

The QOF and other pay for performance schemes have been developed on the back of advances in information technology which have allowed the collation and analysis of large quantities of patient data. In the case of the QOF a national IT system referred to as the Quality Management Analysis System (QMAS), uses data provided by practices to calculate their QOF attainment. Within practices too there have been considerable investments in IT systems. The 2004 GMS contract formalised and standardised IT procurement arrangements by only agreeing to fund systems from an approved list of providers. Those same advances in IT, and greater demands for information and evidence based decision making, have led to the 'commercialisation' of patient data through the development and growth of a number of administrative datasets that give researchers access to large quantities of patient data collected through GP computer systems. The biggest data warehouses for patient data within primary care are CPRD, THIN, and QResearch. CPRD and THIN use data provided by practices using the Vision system provided by InPractice, and QResearch from practices using systems provided by EMIS. In secondary care data is collected by HES. Provided below is a brief description of each of these data sources.

1.5.2 Clinical Practice Research Datalink (CPRD)

CPRD contains over 52 million patient records for Primary Care Practices in the UK with records going back to 1987. It contains over 650 practices that are considered to be up to research standard, meaning they have met data assurance standards such that the data they provide is considered to be research quality. Within these practices there are over 12 million patients, of whom near 6 million are active [23 24].

1.5.3 The Health Improvement Network (THIN)

THIN was developed as an alternative to CPRD in 2002 by EPIC (UK). Data collection commenced in 2003 however THIN does contain patient records going back as far as 1986, depending on the participating practices adoption of the relevant IT system. THIN, as of July 2013, has electronic medical records on 11.1 million patients, 3.7 million of whom are active, from 562 General Practices in the UK[25].

1.5.4 QResearch

QResearch is a not for profit organisation jointly funded by the University of Nottingham and EMIS. It dates back to the early 1990s and has records on over 13 million patients from 660 General Practices[26]

1.5.5 HES

HES collects data on all admissions to NHS hospitals in England, including private patient, and non-residents; as well as admissions to private providers where the NHS is the purchaser of those services. Data is available from 1989 for admitted care and 2003 for outpatient attendance, with 12 million new episodes of admitted care and 40 million outpatient attendance records added each year[27].

1.5.6 Use of large administrative datasets in research

Improvements in data storage capacities and GP IT systems have enabled the data warehouses to store greater volumes and variety of data increasing the scope of possible research. At the same time improvements in data processing and statistical packages have made it possible for the individual researcher to analyse complex data using relatively inexpensive personal computers. These developments have undoubtedly played a large part in driving the growth in research using data provided by these data warehouses. Now that these datasets are over 20 years old and in some instances nearer 30, they also offer the potential to examine the long term effects of policy and treatment. Examples of studies which have utilised the datasets in this manner include a study by Hayes, looking at 15 years prescribing trends in Bipolar disorder using THIN data[28]; and a study by Scowcrofe looking at thrombobrophylaxis rates in newly diagnosed AF patients over a 10 year period using CPRD data[29]. The advantages and disadvantages of using large pre collected datasets in research will be considered further in Chapter 8.

1.5.7 Use of large administrative datasets in this research

This study uses both CPRD and HES data. The recent introduction of linked patient data across these two datasets has enabled the researcher to track individuals from CPRD into certain HES linked files. This is the first study to our knowledge which has utilised these data linkages to look at individual patient level effects of QOF targets on clinical outcomes. Further details on how these two

datasets work in conjunction with each other and how the study used these linkages are provided in Chapter 6.

1.6 Structure of the thesis

The overall aim of this thesis is:

To determine whether individual patient level performance against QOF target measures had an impact on evidence based clinical outcomes using CPRD and HES data.

There are a number of pertinent points arising from this aim that will be addressed within the thesis. These are:

1. What is the current evidence for the effectiveness of different forms of physician incentive on health outcomes?
2. Which QOF targets had high level evidence and accessible data points to inform clinical outcome measures?
3. What evidence exists for the effectiveness of physician incentives on hard and surrogate clinical outcomes in CHD patients?
4. What research has been conducted on the effectiveness of QOF CHD targets?
5. How can CPRD and HES be used to evaluate CHD QOF targets at an individual patient level?
6. What do the analyses demonstrate about the association between QOF CHD targets and clinical outcomes?

These questions and the overall aim will be addressed within the chapters of this thesis. The manner in which it will be done is outlined below where the aims and descriptions of each chapter are set out as well as their context within the thesis:

Chapter 2 Evidence for the effectiveness of financial incentives in primary care

Objective: To examine the evidence for the effectiveness of financial incentives in primary care, through a review of systematic reviews that report studies using high level study designs

Description: This chapter introduces key economic concepts, and considers the effectiveness of physician incentives on the physician, patient, service implications and policy responses. The main focus is a review of reviews examining the high

level study design evidence, defined as well conducted randomised control trials (RCT), controlled before and after studies (CBA), meta-analyses or systematic reviews of those studies, level 1+ or 1++ studies or equivalent; looking at the impact of financial incentives in primary care[30]. The PICOD used in this review is as follows:

- Population: Primary care physicians
- Intervention: Any form of payment, incentive or reimbursement mechanism that affected the individual physician
- Comparator: A suitable control who did not receive the intervention or a comparator group that received an alternative form of physician incentive.
- Outcome: Any process, surrogate or clinical outcome measure
- Design: RCTs, CBA studies

Context: In examining the QOF it is important to be aware of the evidence for financial incentives in primary care, where P4P fits within it, and how effectiveness has been reported.

Chapter 3 The evidence base for QOF clinical targets and feasibility of extracting linked hospital admissions

Objective: To select from the QOF clinical domain, targets, which have high level evidence of improved clinical outcomes which can be readily extracted and modelled using CPRD and HES data.

Description: This chapter looks at the clinical guideline evidence for QOF targets to determine which of them have high level evidence of clinical effectiveness. It then selects from the targets supported by high level evidence those which can be best measured and modelled using data extracted from CPRD and HES.

Context: In order to demonstrate the effectiveness of the QOF it is important to select targets that have clinical benefits that are established, and can be measured using CPRD and HES data.

Chapter 4 The effectiveness of physician incentives on CHD clinical outcomes

Objective: To examine the effectiveness of financial incentives on hard and surrogate CHD clinical outcomes.

Description: This chapter looks at the evidence for the effects of physician financial incentives on hard and surrogate clinical outcomes in CHD patients. No requirements on study design were specified. The PICO used is as follows:

- Population: Patients who had a known CHD diagnosis or were admitted with complications arising from CHD
- Intervention: Any form of physician financial incentive
- Comparator: Within comparison (before and after) or between comparison against comparable populations in different settings (e.g. Countries, States) where different physician incentives operated.
- Outcome: Condition specific hard and surrogate clinical outcomes reported at the patient level.

Context: The previous chapter identified CHD as the condition by which to assess the effectiveness of QOF targets. That effectiveness will be measured by clinical outcomes using large linked administrative datasets. This chapter therefore sets out to explore the existing study evidence for the effect of incentives on clinical outcomes in CHD patients. Since the review in Chapter 2 revealed a lack of high level research evidence on the effects of physician incentives and limited reporting of clinical outcomes: It was decided to relax the study design inclusion criteria to capture other relevant studies, and undertake a broader review of the literature.

Chapter 5 CHD in the QOF

Objective: To explore what research exists on CHD in the QOF, and what this research can contribute to the existing body of evidence.

Description: This chapter reviews all the literature on CHD and the QOF without any restrictions on study design to determine what exists, how it is reported and what it shows. Studies are included if they met the following PICOD criteria:

- Population: Patients diagnosed with CHD
- Intervention: QOF P4P scheme
- Outcomes: QOF CHD target measures, surrogate measures that mapped to those targets, or clinical outcomes, reported at the patient level
- Design: Had to include a minimum of two time points at least one of which had to be post the QOF's introduction.

Context: Due to the absence of a comparator or control group, studies on the QOF have not been included in previous literature reviews. Given the focus of the thesis is on the QOF it was important to determine what research has already been undertaken on the QOF and CHD specifically. The absence of QOF studies in previous reviews reflected weaknesses in study design, largely a consequence of the way in which the scheme was implemented. This review therefore placed minimal restrictions on study design to ensure that as much of the existing literature was included as practically possible.

Chapter 6 Methodology: Using CPRD and HES data to model the QOF

Objective: To explain how data from CPRD and HES can be used to construct variables relevant to the analysis of the impact of the QOF

Description: This chapter explains how the two databases are organised; data was extracted from their file structures, and used to create QOF and other variables relevant to this study.

Context: As this study uses routinely available administrative data it is important to describe these sources and set out the methodological approaches that were used to create dependent and explanatory variables of interest. This is the first study on the QOF to link individuals across CPRD and HES, and therefore clarity on how data was sourced and variables constructed was important.

Chapter 7 Data description and analysis

Objective: To explore the relationship between evidence based targets in the QOF and other relevant variables with the CHD clinical outcomes of interest.

Description: This chapter introduces econometrics, panel data, and the alternative model specifications available. A descriptive analysis is undertaken on key study variables and their panel properties. Panel data econometric analysis is then conducted using those variables and its results interpreted

Context: Previous chapters identified gaps in the literature with regards to an absence of clinical outcome measures and short follow up of patients and incentives. This chapter examines at the individual level whether there is a relationship between compliance with QOF CHD targets and condition specific clinical outcomes, over an 11 year period. It is the first study to do so using individual patient level linkages between CPRD and HES.

Chapter 8 Discussion and synthesis

Objective: To summarise and synthesize the information gathered in the thesis

Description: This chapter draws together the findings from all the chapters in the thesis to draw conclusions, policy implications, and recommendations for future research.

Context: This is the culmination of all the previous chapters where the main points are brought together to comprehensively review the evidence gathered in the context of the overall research aim, and address its implications for clinical practice

1.7 Summary

The QOF represents a substantial and sustained investment in P4P in primary care in the UK NHS. It is one of the biggest investments in P4P in health care globally. As such it offers the researcher the opportunity to evaluate the effectiveness of this area of financial incentivisation. CHD is a chronic and life threatening condition affecting over 2.5 million in the UK; represents the biggest cause of premature death; and places a significant cost burden on the UK NHS and wider economy. Therefore reducing the complications arising from CHD through better management in primary care will not only directly benefit the individual patient but has the potential to be cost saving by reducing demands on other areas of the health service. In addition to NHS cost savings, improved disease management in primary care has potentially wider societal and economic benefits.

This thesis, over the following chapters, will attempt to determine whether meeting evidence based targets in the QOF for the secondary prevention of CHD has had a noticeable impact on clinical outcomes at the individual patient level. In doing so it should help to answer questions concerning the effectiveness of incentives in improving patient outcomes. This not only has policy implications for the QOF, and financial incentivisation in primary care generally, but ties in with the future direction of the NHS as set out in the White paper of 2010 and enacted by the Health and Social Care Act, 2012[31]. The vision set out in the white paper puts patients at the heart of all decisions and places a relentless focus on clinical outcomes. This is the direction taken in the thesis which focuses research at the individual patient level and is concerned with clinically significant outcomes. It

will also demonstrate the challenges and benefits of modelling the QOF using large medical administrative datasets, and their potential in future research evaluating the effects of policy.

Chapter 2 Evidence for the effectiveness of financial incentives in primary care

2.1 Objective

To explore the economics of financial incentives and examine their effectiveness in primary care, through a review of systematic reviews that report primary studies using high level study designs (RCT or CBA studies)

2.2 Introduction

Pay for performance schemes, like the QOF, are a recent innovation within the UK and other health care systems. To place the QOF in a context this chapter will look at other forms of incentivisation that are also used, as well as different methods of funding and delivering health care. These will be considered in an economics context examining market imperfections in health care, their incentives effects on physicians and service delivery implications. Finally a review of reviews will look at the high level study design evidence on the effects of physician incentives. The research themes this chapter seeks to answer are: What high level study evidence exists for the performance of different forms of physician incentives in primary care? What does it show? How does it inform research on the QOF?

2.3 Health care delivery and physician incentives

2.3.1 Economic context

Policy makers, providers and consumers of health care have faced a constant struggle to contain costs, meet increasing demands and keep up with increased expectations. Different forms of incentivisation have been and still are used, Pay for Performance (P4P) being the most recently favoured; to maximise patient benefit at the lowest cost. These also reflect attempts to deal with market imperfections inherent in the health care system. Key among them is an information asymmetry between consumers and providers. Physicians, the providers of health care, are better informed than consumers of care, patients, meaning there is an incentive for them to create unnecessary demand for their services to boost their income. We refer to this excess demand created by the physician for their services as Supplier Induced Demand[32]. Any health

management system or payment mechanism will attempt to curb this incentive while maximising some notion of patient health and controlling costs.

How health care is delivered reflects national priorities, politics and sensibilities. A free market in health care is not tolerated at a national level as the costs can be high and unpredictable, while some people, usually the most sick and infirm are the least likely to have the means to pay for it. Due to this uncertainty over costs and timing of demand, healthcare as a commodity is something which lends itself to insurance[33]. As those risks and costs are greatest in the sick and elderly and lowest in the younger healthier working population, the latter given free choice would be more willing to forsake insurance or only take it up at a price commensurate to their risk. This would make costs unaffordable for the former, when their means are at their most limited, and therefore some form of compulsion is introduced into all modern health care systems, as it is in a number of other insurance markets. In addition there are also positive externalities associated with healthcare, which are not reflected in its market cost. As a society we all benefit from aspects of public health provision as a healthier population means everyone is less likely to get infectious diseases[34]. Likewise we all benefit from mass immunisation programmes by the protection it confers on the whole of society. Left to the market there would be an under provision of these positive externalities, as individuals thinking only of their own benefits and not those conferred on the rest of society, would be willing to forego treatment imposing negative externalities on others[35].

2.3.2 Paying for and delivering health care

For all the reasons discussed the UK and most other health systems use some form of pooled insurance where people actually or notionally pay into a centrally funded scheme. This fund can then be used to provide comprehensive universal health coverage such as the UK National Health Service (NHS). Alternatively it may be used to provide a universal minimum level which all citizens draw upon, or only those who risk being excluded entirely from health care for the financial reasons previously identified, such as the elderly and unemployed. In systems with a limited state provision additional individual health insurance will usually be purchased by the individual or provided by their employers, with statutory legislation compelling them to do so.

2.3.3 Paying physicians

There are numerous forms of physician incentives common, but used in differing proportions, across the different healthcare delivery systems. There is the salaried system where the physician is paid an annual salary for their services. This may form the whole or part of their annual income. While this controls supplier induced demand, it neither contains costs, incentivises productivity, or rewards higher workload. Indeed it incentives the physician to minimise their workload either by denying care or passing on patients to other areas of the health care system where possible. To supplement payment in a salaried system and to incentivise the physician to contain costs we often find capitation payments. In its simplest form this is a fixed payment for each patient on a practice list, however in reality these will be adjusted for the age, deprivation and the disease prevalence profile of the population to reflect costs and workload. This forces the physician to contain costs within the capitation budget, while theoretically providing an incentive for physicians to recruit and retain patients so they receive that capitation element. Fee for service (FFS) is also used widely but most commonly in private health insurance schemes. This pays the physician a fee for each unit undertaken of the activity that attracts the fee. Under this system the physician is rewarded for their activity so it rewards efficiency and productivity in the fee attracting areas. However as the physician only gets paid if they undertake the activity, and for each unit they undertake, it is also the system where the incentive to induce demand is at its greatest, and where it is most difficult to contain costs. It can also lead to resource substitution away from areas not attracting the fee, or less generous fees, to those which do, to the detriment of care and outcomes in those areas.

More recently we have seen a move towards pay for performance or target payments. These forms of payments are very similar, both rewarding physicians for reaching set targets. In the Cochrane review on target payments they are described as a system where; "a lump sum payment is made, if, and only if, the primary care physician (PCP) reaches a predetermined quantity or target level of care.[36]" In contrast pay for performance (P4P) rewards physicians for meeting certain efficiency or quality measures. The differences are thus subtle as a target can be quality related and often activity targets are included with quality measures such as the event record targets in the QOF. Hence the divisions between the two

are often blurred and target payments and P4P are often seen as a single form of physician payments coming under the P4P banner. As these forms of physician payment are not activity based, unlike FFS, and triggered only, and once only, when the standard or target is met, there is no financial incentive for the physician to over treat or induce demand for their services.

2.4 Review of reviews of primary care incentives

A review of systematic reviews was undertaken to identify the current evidence in well-designed structured reviews, on the effectiveness of different forms of financial incentives in primary care. The Cochrane Library[37]; Embase and Medline databases on the Ovid SP search platform, and PubMed were searched for relevant systematic reviews. The reason behind the review was to gain an appreciation of the evidence for the effectiveness of P4P specifically, but also on physician incentives more generally to see if there were any lessons for primary care as a whole, and research on the QOF specifically. Only clinical or surrogate outcomes, or process measures, were of interest to this review. These relate to quantitative clinical interventions directed at, or measured at, a patient level. Process measures refer to clinical processes such as the prescription of drugs, surrogate outcomes to any measure that maps (correlates) to a clinical outcomes; such as blood pressure control, and clinical outcomes, to hospital based clinical activity. Qualitative performance measures such as patient satisfaction and compliance with guidelines were not of interest and not included. This review compares P4P schemes like the QOF to other forms of physician incentives. To best facilitate this only those reviews which focused on well-designed randomised controlled trials (RCT), and controlled before and after studies (CBA), or high quality meta-analyses of those study designs were included. These are referred to in this chapter, and throughout the thesis, as 'high level study designs.' In terms of national guideline development bodies such as the National Institute of Health and Care Excellence (NICE) and Scottish Intercollegiate Guidelines Network (SIGN), these would closely match level 1+ and 1++ levels of evidence [30]. Since the focus was only on these, what are considered stronger, study designs, this necessarily meant that the large amount of research in this area, which used less 'robust' study designs, was neglected. The rationale for this approach is that the study designs chosen minimise bias and

confounding, and hence where there is evidence of an effect, there would be a significant degree of confidence that this was the result of the intervention.

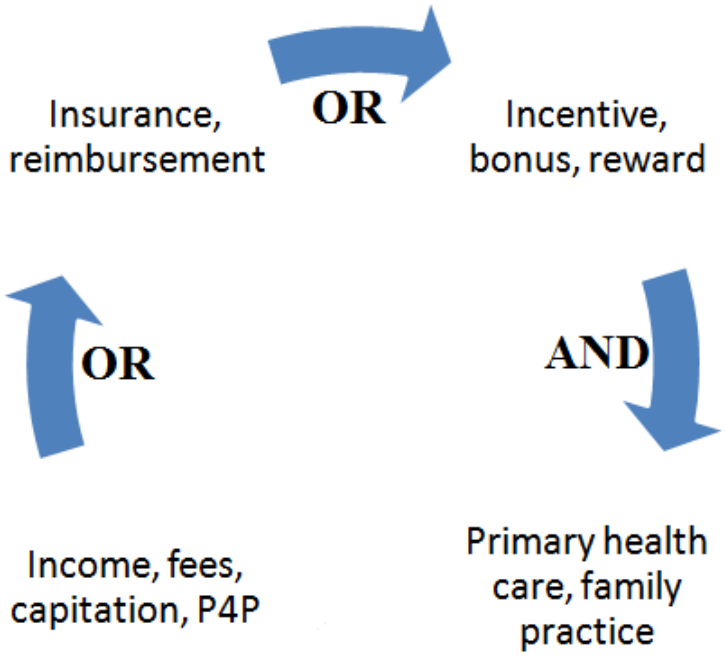
2.4.1 Inclusion criteria

Systematic reviews were included if they met the following criteria:

1. Population: Primary care physicians
2. Intervention: Any form of payment, incentive or reimbursement mechanism that affected the individual physician
3. Comparator: A suitable control who did not receive the intervention or a comparator group that received an alternative form of physician payment.
4. Outcome: Any process, surrogate or clinical outcome measure
5. Design: RCTs, CBA studies

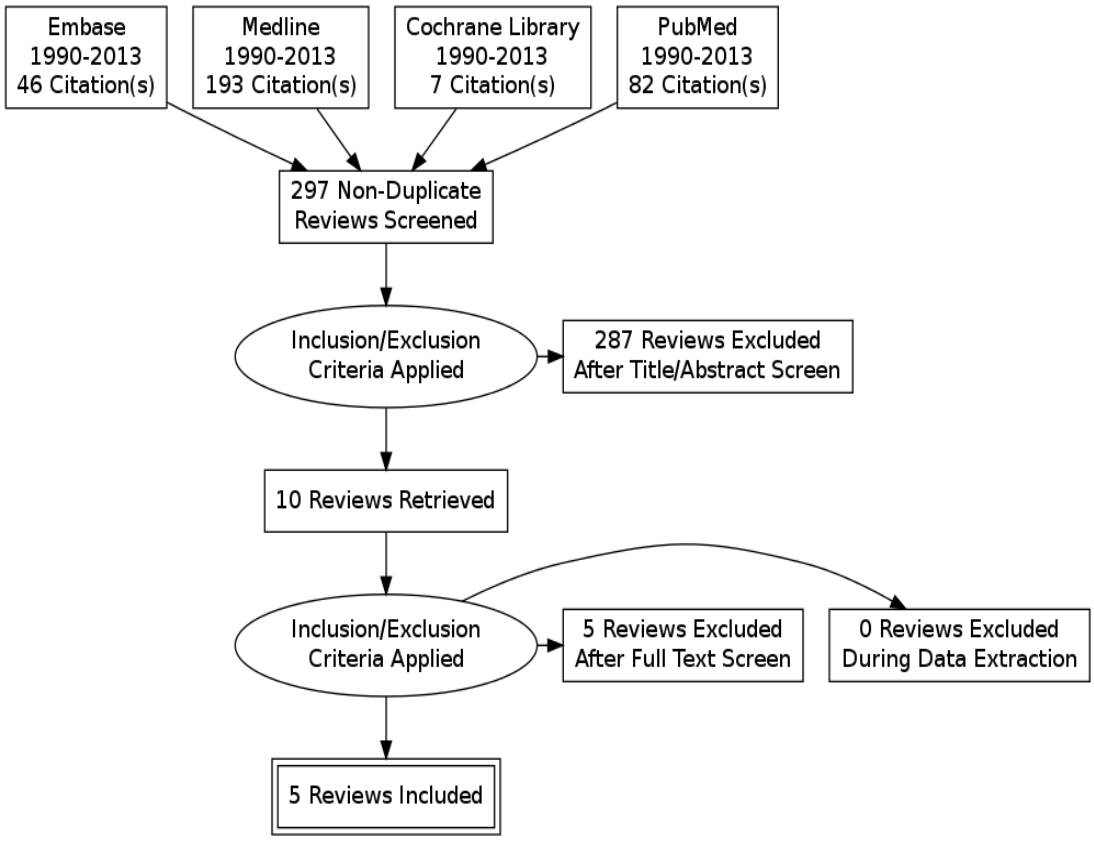
The main search terms are summarised in Figure 2-1 for the OvidSP search platform with full details on all database search terms provided in Appendix 1.

Figure 2-1 Review of reviews literature search terms



The PRISMA diagram in Figure 2-2 details the search for reviews of systematic reviews

Figure 2-2: Review of reviews PRISMA diagram



Four separate databases were searched, with details on the number of articles found on each shown above. All non-duplicate articles found underwent title and abstract review, with ten retrieved for full text review. Five of these were dropped at full review, with five included in the final review.

2.4.2 Details of excluded and included studies

2.4.2.1 Excluded reviews

The reasons for excluding studies is given in the Table 2-1 with specific reference to the PICOD (Population, Intervention, Comparator, Outcome, Design) inclusion criteria the review failed where applicable.

Table 2-1 Excluded reviews, primary care incentives, details

Study	PICO	Reason for exclusion
Gosden, 1999[38]	D	The focus was not on high level study design research evidence: Majority of included studies were case control.
Gosden, 2001[39]	N/A	This study repeats the findings of the Cochrane reviews previously undertaken by the authors, which are included in the review[36 40].
Peckham, 2010[41]	N/A	The article is a discussion paper which broadly discusses the literature rather than analysing it systematically
Vahidi, 2013[42]	D	The focus was not on high level study design research evidence: The majority of the primary literature were cohort and comparative studies
Van Herck, 2010[43]	D	The focus was not on high level study design research evidence: Included any primary evaluation study published in a peer reviewed journal

2.4.2.2 Included reviews

Key points concerning the conduct of the included studies are summarised in Table 2-2.

Table 2-2: Included reviews, primary care incentives, details

Author	Databases searched	Search Period	Interventions considered	Selection criteria / Specific clinical focus	Studies included
Giuffrida, 1999[36]	Medline, Embase, EconLit, Health Star, HMIS, Cochrane library	1966-1997	Target payments	-RCT's, CBA, ITS -Had to meet EPOC criteria	2
Gosden, 2000[40]	Embase, ISI Social Science Index, EconLit, Health Star, HMIS, Cochrane library, Medline	1966-1997	Capitation, salary, FFS	-RCT's, CBA, ITS -Had to meet EPOC criteria	8
Petersen, 2006[44]	PubMed	1980-2005	Any financial incentive designed to improve care	-RCT's, CBA -Studies had to assess the use of financial incentives as the independent variable and a measure of quality as the dependent variable	17
Scott, 2011[45]	Medline, PsychInfo, Embase, Cumulative Index to Nursing and Allied Health, EconLit, PAIS, Cochrane library	2000-2009	P4P, Salary, FFS, Mixed, Capitation	-RCT's, Quasi RCT's, CBA, ITS Had to meet EPOC criteria	7
Town,	EconLit, Business Source	1966-	Any financial	-RCT's	6

2005[46]	Premier, PsychInfo, Medline, Cochrane library	2002	incentive	Had to look at the incentive independently of any other service interventions -Could not be simply for participation in a study -Limited to preventative care	

Study design: CBA=Controlled Before and After, ITS=Interrupted Times Series, RCT= Randomised Control Trials

Cochrane Library refers to a search of any of the Cochrane databases; HMIS=Health Management Information System

EPOC criteria refers to Cochrane Effective Practice and Organisation of Care criteria

2.4.3 Included reviews discussion

The systematic reviews search period includes years from 1966-2009; a large number of search platforms and literature databases. At the same time there is a concentration on certain databases and consequently there is duplication in the reviews returned and examined.

Three of the included reviews are from the Cochrane collaboration. The earliest by Gosden, 2000, looks at capitation, FFS and salary[40]. The most recent by Scott, 2011[45], looks at all forms of physician payment with the exception of target payments which are covered in the third Cochrane review by Giuffrida, 1999[36].

Town focuses on a specific area, the effect of incentives on provider preventative care[46]. Petersen is one of the earlier studies and the only one to limit its search to one database, PubMed[44].

2.4.4 Quality of included reviews

The methodological quality of included studies was assessed using a reduced form of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist[47]. Ten of the 27 checklist points specific to systematic review and pertinent to this review of reviews were selected. These are summarised in the text below and reported in Table 2-3 with full details of the checklist specification provided in Appendix 2.

1. Provide a structured study summary in the abstract
2. Describe the rationale for the review in the context of existing work
3. Specify study inclusion criteria
4. Describe all information sources
5. Present the full search strategy for at least one database
6. State the process of screening and selecting studies
7. Specify any assessment of risk of bias that may have affected the cumulative evidence
8. Summarise the main findings of the studies
9. Discuss limitations at the study, outcome and review level
10. Give a general interpretation of the findings in the context of other evidence

Table 2-3: PRISMA checklist, included systematic reviews

Author	Abstract	Rationale	Criteria	Sources	Search	Selection	Bias	Findings	Limitations	Interpretation	Total/10
Giuffrida, 1999[36]	√	√	√	√	√	√	√	√	√	√	10
Gosden, 2000[40]	√	√	√	√	√	√	√	√	√	√	10
Scott, 2011[45]	√	√	√	√	√	√	√	√	√	√	10
Petersen, 2006[44]	√	√	√	√		√	√	√	√	√	9
Town, 2005[46]		√	√	√				√	√	√	6

Three of the five included reviews met all the criteria, with one study deficient in one of the criteria. The review in question was by Petersen and was deficient with regard to reporting fully replicable details of their search strategy for at least one database[44]. This may have been available from the authors, but was not made available within the paper itself or within supporting material referenced within the paper. The remaining review by Town was deficient in four of the criteria, meaning that it has a number of methodological quality issues when judged against this checklist[46]. Common with the review by Petersen it does not provide fully replicable details for one of its searches. In addition there is not a structured summary in the abstract; the review does not assess the risk of bias within their included studies; and they fail to produce a PRISMA diagram or description of how documents were screened and filtered to final inclusion[46].

2.5 Primary study level reporting

None of the systematic reviews conducted any form of meta-analysis using their included studies. This is unsurprising given that the outcome measures in the primary literature were heterogeneous, making it impossible to meta-analyse their

results. The reviews, the Cochrane reviews in particular, focus on reporting the findings of each included study. There is some attempt to draw together findings in terms of common themes but this is limited in its success as the study findings are so disparate. Therefore in this section the main findings of the individual studies covered in the systematic reviews that met the inclusion criteria are reported, and evaluated, in the context of this chapter objective.

The findings of the studies are reported in two separate tables in the following pages. Table 2-4 relates to studies where the comparison is between alternative payment systems, this includes all the non P4P studies. Since P4P incentives are either compared against the absence of P4P in within study designs or a control who did not receive the incentives in a between study design, the comparator is always against non P4P and not an alternative form of incentive. Therefore the results for P4P, the largest body of work, are reported separately in Table 2-5. Studies looking at target payments are also included under the P4P classification since, as discussed earlier; these can also be viewed as performance based incentives. Target payments take various forms but all forms reward performance to a pre specified outcome.

Table 2-4: Included studies, primary care incentives, non P4P

Study	Intervention	Comparator	Context	Measure	Outcome
Davidson, 1992[48]	Capitation	FFS / Capitation	Children's Medicaid Programme, USA	Physician patient visits during study	Regression coefficients (effect on utilisation): Capitation:+0.59 ($p \leq 0.05$); FFS +0.9 ($p \leq 0.01$)
				Referrals to specialists and hospitals during study	Regression coefficients (effect on utilisation): Capitation:-0.25 ($p \leq 0.01$); FFS +0.03 ($p > 0.05$)
				Hospitalisations during study	Regression coefficients (effect on utilisation): Capitation:-0.02 ($p \leq 0.05$); FFS +0.01 ($p > 0.05$)
Gosden, 2003[49]	Fixed Salary	Capitation	UK NHS	Cervical cytology	Mean difference (PMS-GMS) after 1 year =0.38 (95% CI, -14.9 to 15.67) not significant
				Childhood immunisation	Mean difference (PMS-GMS) after 1 year =-1.08 (95% CI, -17.95 to 15.80) not significant
				Pre-school booster	Mean difference (PMS-GMS) after 1 year =-3.08 (95% CI, -9.63 to 3.47) not significant
Hickson, 1987[50]	Salary	FFS	USA: Resident paediatricians, Continuity clinics	Emergency visits	Numbers per patient (9 months):0.22 v 0.12, $p<0.01$
				Scheduled visits	Numbers per patient (9 months): 2.83 v 3.69, $p<0.01$
Hutchison,	Capitation	FFS	Ontario, Canada	Hospital days per	3 years post intervention compared to 1 year before;

1996[51]				1000 patients	additional 3 hospital days/1000, p=0.774
Krasnik, 1990[52]	Partial FFS	Capitation	Copenhagen county, Denmark	Face to face contacts	+7.2% after 6 months, -0.5% after 12 months*
				Telephone contacts	+10% after 6 months; +11% after 12 months*
				Prescription renewals	-9% after 6 months; -27.4% after 12 months*
				Hospital referrals	-9.7% after 6 months; -33.7% after 12 months*
				Specialist referrals	-9.3% after 6 months; -21.2% after 12 months*
				Diagnostic services	+32.8% after 6 months; +52.2% after 12 months*
				Curative services	+88.6% after 6 months; +79.8% after 12 months*

*Change in absolute numbers over a one week period per 1000 patients compared to 6 months prior to the intervention

Table 2-5: Included studies, primary care incentives, P4P

Study	Context	Measure	Outcome
An, 2008[53]	Fairview Physicians Associates, USA	Mean % of smokers referred to a telephone counselling service	11.4% in the intervention group; 4.2% control group, over 10 months (p=0.001)
Fairbrother, 1999[54]	New York, USA	Increase in up to date Immunisations in children under 3 years old over 8 months	+25.3% Bonus & feedback (p<0.01), +4.3% Enhanced FFS & Feedback, +12.6% Feedback only, +6.1% Control
Fairbrother, 2001[55]	New York, USA	Increase in up to date Immunisations in children under 3 years old over 1 year	+5.9% Bonus & feedback(p<0.05), +7.4% Enhanced FFS & Feedback (p<0.01), -2.5% Control
Grady, 1997[56]	Massachusetts, USA	Annual mammography compliance in females over 50	Absolute increase from baseline over a 1 year period: Intervention 17.9% Control 13.7% (no significant dif.)
Hillman, 1998[57]	Philadelphia, USA	Compliance with cancer screening guidelines:	Absolute change from baseline over 18 month period Intervention: 26.3% Control: 26.4% (no significant dif.)
Hillman, 1999[58]	Philadelphia, USA	Preventative care in under 6 year olds	Absolute change from baseline over 18 month period Intervention: 17.2% Control: 11.3% (no significant dif.)
Kouides, 1998[59]	New York State, USA	Influenza vaccination rate for eligible patients	Median practice specific improvement over 1 year: + 10.3% in intervention group; +3.5% in the control group, p=0.03
Mullen, 2010[60]	Independent Health Association,	Cervical screening rate	Mean difference in absolute change from baseline Int group 1 (QIP) = -0.043% (p>0.1)

	California USA		Int group 2 (IHA1+QIP2) =3.5% (p<0.01) Int group 3 (IHA2+QIP2) =6.01% (p<0.01)
		Mammography screening rate	Mean difference in absolute change from baseline Int group 1 (QIP) = -1.067% (p>0.1) Int group 2 (IHA1+QIP2) =0.12% (p>0.1) Int group 3 (IHA2+QIP2) = 1.29% (p>0.1)
		HbA1c testing	Mean difference in absolute change from baseline Int group 1 (QIP) = 1.357% (p>0.1) Int group 2 (IHA1+QIP2) = -3.756% (p<0.1) Int group 3 (IHA2+QIP2) =1.916% (p>0.1)
		Childhood immunisation	Mean difference in absolute change from baseline Int group 1 (QIP) = 3.155% (p<0.05) Int group 2 (IHA1+QIP2) =2.078% (p<0.1)
Ritchie, 1992[61]	Grampians, Scotland	Pre-school immunisation rates	Practices achieving 95% coverage +50%; 90% +20% ┼
		Primary school immunisation rates	Practices achieving 95% coverage +41%; 90% +41% ┼
Rosenthal, 2005[62]	PacifiCare Health Systems, USA	Cervical cancer screening	1 year difference in differences compared to control: 3.6% (p=0.02)
		Mammography	1 year difference in differences compared to control: 1.7% (p=0.13)

		Haemoglobin A1c testing	1 year difference in differences compared to control: 0% (p=0.5)
Roski, 2003[63]	Mid West USA	7-day sustained abstinence from smoking	Int1 (incentive) = 22.4% Int2 (registry) = 21.7% Cont = 19.2% (between groups, p=0.2)
Twardella, 2007[64]	Germany	Smoking abstinence at 12 months	Int1 (Training + Incentive)= 3% (p=0.75)# Int2 (Training + Medication)= 12% (p=0.046)# Int3 (TI + TM)= 15% (p=0.02)# Control (usual care)= 3%
Young, 2007[65]	Rochester, NY, USA	Haemoglobin A1c testing	Difference in changes in adherence rate (Pre2 to Pre3 vs Pre3 to Post1) = 0.9% (Non-significant)
		Urinalysis	Difference in changes in adherence rate (Pre2 to Pre3 vs Pre3 to Post1) = 1.7% (Non-significant)
		Lipoprotein density level (LDL)	Difference in changes in adherence rate (Pre2 to Pre3 vs Pre3 to Post1) = 0.3% (Non-significant)
		Eye examination	Difference in changes in adherence rate (Pre2 to Pre3 vs Pre3 to Post1) = 5.1% (p<0.001)

┼ Over a 21 month period. Results not significant when adjusted for previous time trend #compared to the control group

2.5.1 Primary study details

2.5.1.1 Non P4P

The individual studies cover the UK, USA, Canada and Denmark. Davidson compares primary care physicians (PCPs) paid by age adjusted capitation to two FFS groups, one paid the normal Medicaid fee and the other a new higher rate; double the existing rate; all of whom participated in the Children's Medical Programme[48]. The capitation group had to set aside $25 a month to provide services they did not directly provide. If they spent below this budget they received a 40% bonus, if they exceeded it they were liable to a charge of up to $2000 per child and 25% of their budget. Gosden looked at the effect of changes introduced in the NHS (Primary Care) Act of 1997 which provided the opportunity for GP's to move from a capitation General Medical Services (GMS) contract to locally arranged salaried employment to provide Personal Medical Services (PMS)[49]. Hickson compared resident paediatric PCPs allocated to FFS and fixed salary in Ontario, Canada[50]. Krasnik looked at the effect of introducing fees in a capitation system compared to a control group of PCPs already paid under a mixed capitation FFS system in Stockholm[52].

2.5.1.2 P4P

Included studies looking specifically at P4P are drawn from the UK, USA and German health care systems. An et al, looked at the effect of financial incentives, $5000 for the first 50 patients, $25 for each patient thereafter; on referrals of smokers to a telephone counselling service by clinics in Minnesota, USA[53].

Grady's study is concerned with a bonus scheme that paid physicians $50 if their referral rate of women over 50 for mammography was over 50%.

Kouides considered target payments made to 54 general practices based in Monroe County, New York state[59]. The intervention paid an additional 10% on top of the normal $8 fee for each influenza vaccination made over the target rate of 70%, and 20% for each vaccination over an 85% target.

Mullen examines the effects on 172 eligible medical groups of two schemes[60]. One implemented by PacifiCare Health Systems, known as the Quality Incentive Programme (QIP). The second a much larger scheme, which soon overshadowed the former, was a P4P scheme operated by the Independent Health Association (IHA), a

not for profit coalition of six HMO's in California. The analysis was broken down further following an updated version of the QIP and different stages of adoption of the IHA (Year 1 or 2). The study uses difference in difference analysis, an econometric technique which compares the marginal effect difference within patients in the intervention and control group, assuming fixed time effects. In this instance the control group were comparable medical groups from the Pacific North West who were also in contracts with PacifiCare. Rosenthal also considered the HMO, PacifiCare Health Systems, to look at the effect of quarterly bonus payments, administered through the QIP programme in 134 medical groups on rates of cervical cancer and mammography screening, and HbA1c testing[62]. The control group were 33 medical groups from the Pacific North West, whose results were reported contemporaneously. Difference in differences analysis was again used to compare differences in marginal attainment between the intervention and control group.

Ritchie looked at the Grampian region of Scotland covering 95 General Practices and 313 GP's, immediately before and after the 1990 GP contract changes[61]. These changes led to practices being given a lump sum payment if they immunised 70% of their eligible population and a higher lump sum if that figure was 90% or more.

Roski examines the effects of incentives paid to 25 clinics in the Mid-West of the USA for meeting pre-set targets for recording smoking status and offering smoking cessation advice (intervention 1)[63]. Ten of these were also provided with access to a centralised patient registry and intervention system (intervention 2).

Twardella considers the effect of an incentive offered to German GP's of $130 for every patient they recruited who smoked more than 10 cigarettes a day and who were smoke free at 12 months follow up[64]. The study was split into four arms with a control group of 20 practices offering usual care; 21 in the incentive scheme (Int1); 21 whom could spend up to the amount of the incentive on offering free prescriptions for drugs proven effective in smoking cessation (Int2); and a fourth arm who had the flexibility to spend the money on offering free prescriptions or keep it, in whatever proportion they wished (Int3). All three intervention groups were offered training on methods of delivering smoking cessation.

There are two studies by Fairbrother, both looking at the effectiveness of incentives in the area of infant immunisation in New York, USA[54 55]. The earlier of the two

included studies looks at the effect of a cash bonus and enhanced FFS, both combined with feedback, compared to a group that received feedback only and a control group. Physicians were randomly allocated to one of the four groups. Those in the cash bonus group were paid $1000 for practice wide increases from baseline of 20%, $2500 for 40% improvement, and $5000 for an 80% up to date coverage rate[54]. The enhanced FFS intervention group received quarterly payments of $5 for each vaccination given within 30 days of it coming due, and $15 for each visit at which more than 1 vaccine was due and all were administered. The effectiveness of the interventions was measured 4 and 8 months from baseline. The later study was intended to follow the first, but was delayed by 12 months[55]. In this study the effectiveness of the intervention was measured at 4, 8 and 12 months from baseline, and there was a number of other changes: In the bonus group the threshold for the $1000 payment increased to 30% and for the $2500 payment to 45%; while a new payment of $7500 was introduced for a 90% up to date coverage rate. In addition the feedback only group was removed; details for the enhanced FFS group however did remain the same. Of the 60 physicians who took part in the first study 8 were lost to the second study, and 12 physicians were newly recruited to take their place.

Two studies are also included by Hillman, concerning different clinical areas[57 58]. The earliest looks at incentives paid for compliance with cancer screening guidelines which paid half yearly bonuses for good performers[57]. Based on aggregate scores and improvement in scores the best performing 3 sites received a bonus equal to 20% of capitation for all females over 50 on their list, with the 3 next best performers receiving a 10% of capitation bonus for females over 50 on their list. The more recent study was concerned with preventative care in children under 6 years of age[58]. Based upon a minimum attainment of 20% on all quality indicators relating to immunisations and visits; the three highest performing practices received an extra 20% on top of their capitation rate for the under 6 population; and the three next best an extra 10%.

Young et al, use an interrupted times series analysis, to look at the effect of an incentive scheme paid to 334 PCP's in the Rochester Individual Practice Association, based on their performance on diabetes targets relative to other PCP's in the study group[65]. To fund the incentive scheme each member had 5% of their fees with-held but would have 50-150% of that returned based on their relative

performance. The targets were for HbA1c test, Urinalysis, Lipoprotein density levels and eye examinations.

2.5.2 Discussion of findings

There is a large heterogeneity of interventions considered and outcome measures used, which makes meta-analyses impractical. No strong evidence was found for a specific form of physician incentive. Incentives which reward activity or outcomes were generally found to lead to desired improvements on those measures. However these improvements were measured using process measures principally so it is uncertain if they had any clinical benefits.

Overall none of the primary studies looked at the effectiveness of incentives in primary care to evidence linked clinical outcomes. Instead the focus was largely on process measures in primary care and where secondary care outcomes were reported they were general measures of hospitalisation. None of the P4P studies looked at clinical outcomes, and very few used surrogate outcome measures.

As with the reporting in the tables the discussion is split down into P4P studies and non P4P studies with an additional overall section which draws findings from all the studies for shared clinical areas.

2.5.2.1 Non P4P studies

Three of the five non P4P studies compare FFS and capitation. These largely found more activity on the part of FFS physicians in routine and preventative primary care. Davidson found a statistically significant higher rate of physician visits in the higher FFS intervention compared to the usual FFS rate control group, ($p \leq 0.01$). Differences in referrals to specialists and hospital, and hospitalisations during the intervention period, were however insignificant. In comparison those paid by capitation registered a smaller increase in visits relative to higher rate FFS physicians, but one that was still statistically significant, ($p \leq 0.05$), and a statistically significant decrease in referrals and hospitalisations. Hence the authors concluded that there was no disadvantage to patients from the switch to capitation[48].

Hutchison finds no statistically significant difference in hospital days per 1000 patients between the two payment methods[51]. Krasnik, a lower rate of prescribing by FFS physicians but a higher rate of face to face consultations, after 6 months only;

and in referrals, diagnostic, and curative services, after both 6 and 12 months[52]. These results were not tested for statistical significance.

The remaining two non P4P studies included salaried physicians, comparing them against capitation and FFS remunerated physicians. Gosden finds no significant difference in cervical cytology, childhood immunisation and preschool booster rates between capitation and fixed salary. Hickson that paediatric physicians paid by FFS carried out more scheduled visits per patient ($p<0.01$) and fewer emergency visits ($p<0.01$) than their capitation remunerated counterparts; the latter the authors attribute to a better continuity of care by FFS physicians[50]. The comparison between the results of these two studies would seem to confirm the hypothesis that activity based incentive payments do generate greater physician activity. This may have some clinical benefits but with no clinical outcomes reported it is not possible to confirm this from these studies.

2.5.2.2 P4P studies

The evidence generally showed an improvement in outcomes but these were insignificant in the majority of studies, and so overall not conclusive. An, finds a much higher rate of referrals to a smoking counselling service for those PCP's offered incentives, a result which is statistically significant ($p=0.001$)[53]. Mullen shows significant positive increases in mean absolute screening percentages for those physicians paid by the two P4P incentive schemes as opposed to those paid by a single P4P scheme the Quality Incentive Programme (QIP), and from later variants of the same schemes ($p<0.01$). However the results on the other measures are less encouraging with those paid by the QIP alone having higher mean absolute changes in childhood immunisation ($p<0.05$) than those paid by combinations of the QIP and Independent Health Association (IHA) ($p<0.1$), and QIP alone outperforming IHA1 and QIP2 on HbA1c[60]. Looking at similar outcomes to Mullen, Rosenthal finds significant improvements on cervical cancer screening in the intervention group compared to the control ($p=0.02$), but not in mammography screening ($p=0.13$), and HbA1c testing ($p=0.5$)[62].

Roski looks at the effect of incentives on smoking recording and cessation, with and without a patient registry, compared to a control[63]. Rates of abstinence, intention to quit and use of a counselling programme were higher in the intervention groups but lower for the use of pharmaceutical aids and cessation assistance. On the main

outcome measure 7 day sustained smoking abstinence Roski found no significant difference between the intervention and control groups (p=0.02). The only significant difference found was with regard to a higher percentage of the registry population reporting having used any counselling service for smoking cessation (p=0.001). Twardella also examines the effects of incentives on smoking cessation[64]. This showed no impact of the incentives relative to a control group, but noticeable difference in rates when the GP had the option to spend some (p=0.02) or all (p=0.05) of the incentive on offering free prescriptions for proven pharmaceuticals in that area.

Grady looks at the annual change in mammography compliance and finds no significant difference in this figure in the intervention group relative to the control[56]. In the study by Ritchie the number of practices achieving at least a 95% coverage figure for primary school immunisation increased by 50% over a 21 month period following the intervention, and by 20% for the at least a 90% coverage[61]. For pre-school immunisation the numbers meeting at least a 95% and 90% coverage, both increased by 41%. However after fitting a logistic regression model the authors found no significant difference to trend following the introduction of targets payments.

There are two studies by Fairbrother, both looking at the same target payment scheme on childhood immunisation rates.[54 55]. In the earlier study the bonus group increased coverage significantly more than the control group ($p<0.05$). However this was from a much lower baseline than the enhanced FFS group, 29.1% compared to 46.2%. From a much higher baseline in the second study period the increase in up to date coverage for the bonus group was more modest but still significant ($p<0.05$), albeit less than that experienced in the enhanced FFS group ($p<0.01$). When the incentive was withdrawn for the 12 months between the two studies, coverage did not fall back and remained similar to the rates seen at the end of the first incentive period. While this was only a short period it suggests that targets may change behaviour even after they are withdrawn.

Finally Young finds no significant difference in adherence rates, allowing for time trends, of a P4P scheme based on relative performance and paid for by withholding existing monies, on all measures, with the exception of eye examinations[65].

2.5.2.3 Overall findings

A number of studies have a shared focus on the same clinical measures which presents the opportunity to find agreement and disagreement in their results and attempt to weigh up the wider effect of physician incentives. One of these is childhood immunisation, which five studies consider[49 54 55 60 61]. Three of these considered target payments, with two coming from the same author, examining broadly the same scheme in New York State[54 55], the remaining study considered a scheme in the UK[49]. All showed an increase from baseline, the studies by Fairbrother showed this to be significant compared to the control group[54 55], the study by Ritchie found it to be insignificant when existing trends were taken into consideration[49]. In the study by Ritchie the majority of physicians met the required levels to trigger payments, while in the Fairbrother studies, based on overall attainment figures one target was met in the first period, and none in the second period. Despite this the effect was significant in both periods compared to the control group. It is worth noting that the Fairbrother studies did not adjust for baseline attainment, or trend as in the study by Ritchie, so it possible that these effects may not have been significant had they done so.

The study by Mullen also considers the impact on childhood immunisation in a pay for performance setting[60]. This showed, counter intuitively and in contrast to the other clinical areas considered, that greater incentives available through participating in multiple schemes had a less significant impact than participating in the single QIP scheme ($p<0.05$). This may be a service that physicians are minded to do regardless of specific incentives which would explain this result. Finally the study by Gosden looks at the impact of moving from capitation to fixed salary payments on childhood immunisation rates[49]. Unsurprisingly, since neither payment method incentives productivity nor improved outcomes, the results were not significantly different.

Three studies looked at cervical cancer screening[49 60 62]. In the study by Gosden the change from capitation to fixed salary again had no impact on screening rates at the 95% significance level[49]. Mullen found a significant increase in the marginal mean difference for those participating in multiple incentive schemes compared to the control group ($p<0.01$), but an insignificant difference between the single incentive and control groups ($p<0.1$)[60]. This results suggest that the financial size of incentives do play a role contrary to their findings for childhood immunisation[60]. Equally

Rosenthal finds a significant increase in cervical cancer screening rate in those receiving P4P incentives compared to a control group (p=0.02)[62].

Three studies, all P4P, consider HbA1c testing as one of their outcomes[60 62 65]. Mullen finds a fall in the marginal rate of testing in those who adopted the IHA in its first year (IHA1) and the second version of QIP (QIP2) but an increase in those joining in the second year of the IHA incentive scheme, IHA2+QIP2. However, none of these changes were significantly different to baseline rates (p>0.1). Rosenthal finds the same increase from baseline in the intervention and control group (2.1%), though only in the intervention group is this increase significant (p=0.02)[62]. Finally Young who uses an interrupted time series to attempt to adjust for trend finds no significant increase in HbA1c testing following the introduction of P4P, when the existing trend is taken into consideration (p>0.05).

Two P4P studies look at smoking cessation services[63 64]. Both the study by Roski and Twardella suggest that incentives alone had no impact on smoking cessation. However Twardella found that giving individual physician's discretion on how some or all of that financial incentive was spent did have a significant impact on smoking abstinence in patients.

Overall comparing across the different interventions these results suggest that payment methods that do not distinguish by performance such as capitation and fixed salary had no impact on levels of activity. P4P does have a positive impact on the whole, but one that is not consistently statistically significant.

2.6 Implications for practice

The primary studies have identified a number of implications for practice. These are listed below:

- The studies did not find conclusively in favour of any particular form of financial incentive. They generally showed that PCP's responded to greater financial incentives with better performance or more activity in those areas which attracted incentives. It is uncertain from the studies included whether this was at the expense of non-incentivised areas, however given that a physician's time is limited it is plausible to assume that there was may have been some level of resource substitution.

- Using existing money to pay for the incentive may have a negative impact on its effectiveness[64], while giving individual physician's greater autonomy in how financial incentives can be spent to meet targets appeared to improve attainment on those targets[64 65].
- All studies that have looked at the impact of changes in physician payment have done so over a relatively short time period. Incentives where they have been put in place have been limited in their duration. Possibly due to this the focus has been on process measures with no consideration of hard clinical outcomes or endpoints.
- There is evidence that physician response to incentives may be short lived, with normal activity levels resumed after an initial spurt in activity when the incentive is first introduced. In the study by Krasnik, FFS generated more activity compared to a capitation system in areas that attracted the fee, but differences dissipated over time[52]. This was consistent with the target income hypothesis explored in the Krasnik paper. This hypothesis suggests that a physician's supply of labour and behaviour is determined by their motivation to maintain a target level of income. Hence in the short term changes to incentives can lead to significant changes in behaviour and activity by the physician, while they learn to adapt to the new circumstances. In the longer term however, once that learning process is complete, activity and behaviour settles around that level needed to maintain their target income. On the opposite side evidence was found which showed that performance did not drop to levels seen before the introduction of the incentive, in this case target payments, when they were temporarily withdrawn[54 55]
- Mixed incentives schemes by mitigating the perceived weakness or perverse incentives in a specific form of incentivisation, may lead to more desirable outcomes. In the study by Hickson fixed salary PCP's undertook fewer routine visits and had higher emergency admissions in secondary care[50]. This could be indicative of the incentive in such a system to pass workload onto other parts of the system. However in the study by Davidson looking at capitation which has a similar incentive to pass costs onto other parts or the system; when this was combined with an incentive to limit referral to outside services, those referrals were lower[48].

2.7 Implications for future research

This review identified a lack of high level study evidence on the effects of incentives in primary care, with 5 reviews returned and 18 separate studies meeting the review study design criteria. This reflects the difficulties of doing such research in the area of physician payment, where randomisation is difficult and concealment impossible. In such circumstances future research should consider quasi randomised, and other lower level study designs in any review of the evidence.

In terms of the future research on physician incentives there is a need for research which looks at the impact of primary care incentives on condition specific outcomes or endpoints, and considers the effects of incentives over an extended duration. This need is most significant for P4P given the absence of clinical outcomes and lack of surrogate outcomes identified by this research.

2.8 Conclusion

This chapter has identified gaps in the research evidence on the effects of physician payment in primary care. Among these is an absence of studies looking at clinical endpoints and the impact of incentives over an extended duration. Another issue is that outcomes have tended to be general process measures, specifically in the P4P literature, with no link to evidence.

It is important that the relationship between the intervention and outcome measure is supported by high level research evidence, to ensure that clinical benefits are demonstrable. The next chapter aims to ensure this process is followed in the thesis by examining the clinical evidence for QOF targets to select those which have a high level evidence base through to condition specific hospital admissions, and have been in place for a sufficient duration to ensure any effect is measurable. This ensures that if the financial incentive is effective in improving clinical outcomes that impact will be evident in the outcome variable.

Chapter 3 The evidence base for QOF clinical targets and feasibility of extracting linked hospital admissions

3.1 Objective

To select from the QOF clinical domain, exemplar targets whose clinical benefits have been evidenced in high level study design research, and can be modelled using CPRD and HES data. This will involve:

1. Using NICE and SIGN clinical guidelines to find which targets from the QOF guidance 2009/10 had the highest level of clinical evidence
2. From those targets which have the highest clinical evidence selecting those for whom clinical benefits can best be measured and extracted using CPRD and HES data.

3.2 Background

When the New General Medical Services contract was drawn up in 2003 the 'Quality Standards', which included the QOF clinical targets were drawn up by an independent panel of experts based on evidence and current professional practice[66]. The contract was the result of consultations between GP's represented by the General Practitioners Committee of the BMA and the government of the day. They took place against a backdrop of difficulties in recruiting and retaining GP's, a determination to address it and the money to do so. Under such circumstances it is perhaps not surprising that the details on how the expert panel agreed on the initial targets is lacking, although rationale has been provided in QOF guidance. The over-riding aim may have been to make the profession more attractive financially, rather than a focus on 'quality outcomes'. There has been criticism of the design of the QOF incentive scheme by among others, Ashworth et al, who raised concerns around the usefulness of using QOF targets to determine improvements in patient care, and questioned whether the QOF was more concerned with paying for recording than paying for performance[18]. In 2005 and 2007 the QOF was reviewed by an expert panel of appointed primary care academics supported by a group of clinicians, using feedback from calls for evidence sent out to a cohort of Primary Care Trusts[7]. Not until the introduction of QOF NICE committee in 2009 did the policy of evaluating

existing QOF targets and piloting new targets become rigorous, independent, and transparent.

Hence it is often unclear as to why specific targets were chosen and how their wording and the specific payment thresholds were decided upon. For this reason the decision was made to go through the QOF guidance 2009/10, to discover which, of all the targets in the clinical domain, had high level clinical evidence. Then based on that evidence determine which of those could be effectively measured and examined using CPRD and HES data. This was to ensure that targets chosen for analysis would produce substantive and measurable benefits that would be picked up in any analysis. The previous chapter highlighted a number of deficiencies in existing research on incentives in primary care. Specifically the absence of clinical outcome measures, a short intervention or follow up period, and heterogeneity in outcome measures. The approach taken in this chapter seeks to ensure that those deficiencies are addressed in this research. By ensuring that targets have been stable or in place for a significant portion of the QOF intervention period, it will be possible to examine their longer term effects. In linking their effect into clinical outcomes supported by high level evidence their effectiveness will be clearly demonstrable. By basing outcome measurement on clinical evidence only, this removes the scope to self select those measures, and reduces the potential for bias in their selection.

3.3 Methodology

In order to ensure that any targets selected had high level clinical evidence and that those benefits could be clearly measured in the primary and secondary care datasets a three stage selection process was undertaken:

Stage 1; an appraisal of the QOF 2009/10 guidance to remove those targets which had no measurable clinical outcomes, or possible link to clinical outcomes

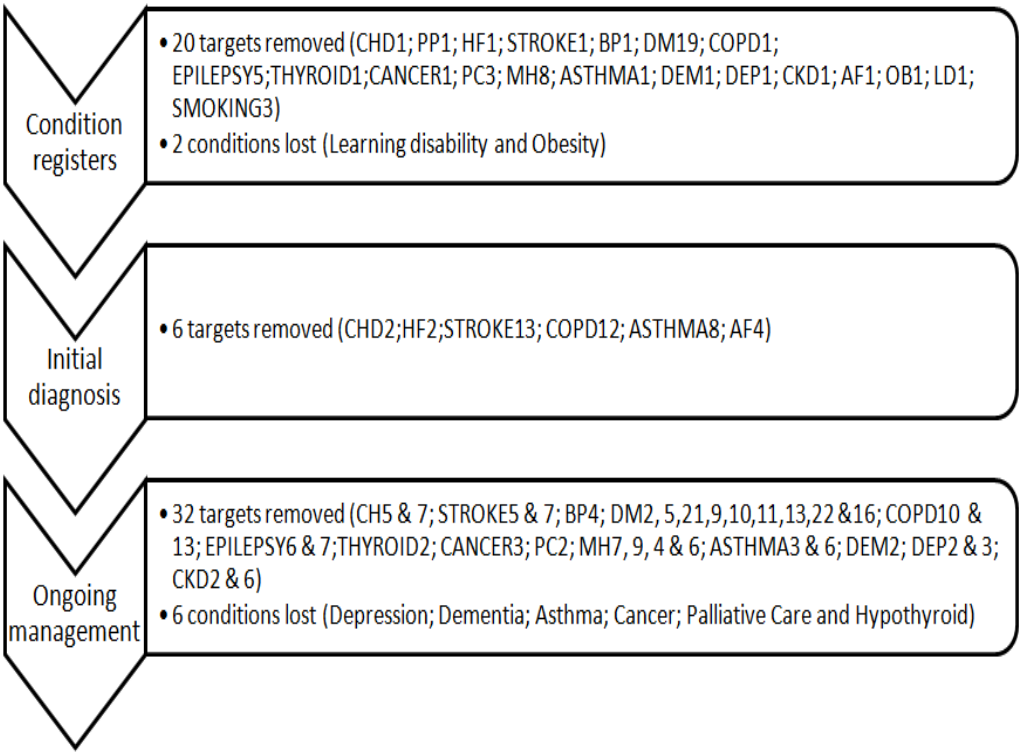
Stage 2; a review of the NICE and SIGN clinical guideline evidence to find which targets, of those remaining, had the highest level of clinical evidence.

Stage 3; an examination of the suitability of the remaining targets with regards to their longevity, stability, and whether their benefits could be measured and linked across CPRD and HES data.

3.3.1 Stage 1: Review of the QOF clinical guidance

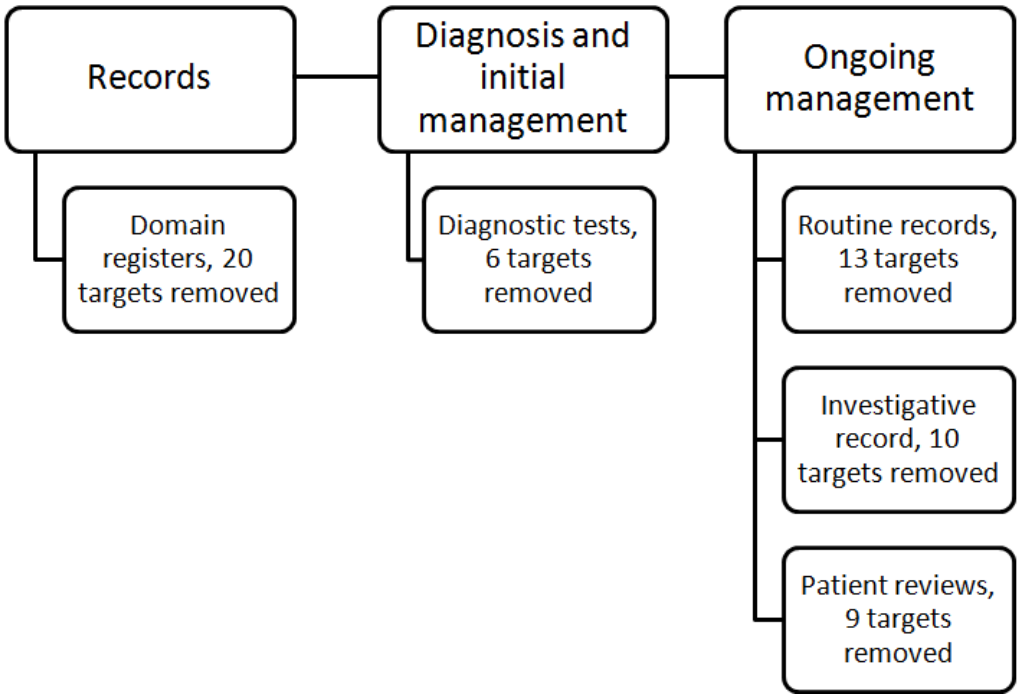
The 2009/10 guidance was chosen as there were no changes made in 2010/11 so it represented the most up to date for the study period. Within the QOF for that year there were 19 clinical conditions, a smoking domain, and 86 targets in total. Each condition follows a similar flow of targets, starting with an initial condition register, which is used as the denominator for a number of subsequent targets. The following section, not present in all conditions, is termed 'diagnosis' or 'initial management' which relate to confirmatory tests for the disease or tests to determine its specific form. Finally there is an 'ongoing management' section which contains the bulk of the QOF targets. Details of the targets and conditions removed at this stage are given in Figures 3-1 and 3-2.

Figure 3-1 QOF targets removed by guidance section



This process is further broken down in Figure 3-2.

Figure 3-2 QOF target types removed by section, more detail



The first targets removed were those within the 'Records' section of the disease indicators that referred to the practice producing a register for each respective condition. These numbered twenty in total and for two domains were the only targets, resulting in their complete removal. They incentive practices to maintain condition registers for their patients enabling them to be treated promptly, however any benefits will be dependent on subsequent management and treatment. The next targets removed were those subsequent to the disease registers, within the diagnosis and initial management section where present, which referred to a diagnostic test to confirm the presence of the disease or tighten its definition. One example of this is HF2 which rewarded the practice if heart failure had been confirmed by echocardiogram or specialist assessment. Again the benefits of these targets are determined by how the information is utilised; and since this is not stipulated in these targets it would be extremely difficult to measure their impact. Six targets were removed at this point in the process

The next area examined were those that came into the 'Ongoing management' section of the disease domain, which contains the bulk of the targets and relates to how the patient's condition is managed. Within this section three groups of targets have been removed. The first of these are referred to as 'routine records'. These are targets for recording that an event has taken place without specifying the outcome,

for surrogate measures like blood pressure or cholesterol measurement. These precede targets which stipulate what level those measures need to meet, and it may be possible to combine them in a single target. While clinical benefits may result from these 'routine records', for instance they incentive GP's to monitor blood pressure in all patients and not just those likely to meet the target level which may improve outcomes in those patients; those benefits are difficult to quantify and attribute. For this reason they were therefore removed.

Following on from these there are what have been termed investigative records. These are again targets which refer to a record of a test but are for investigative and preventative reasons and disease specific. As before while there may be some benefits from these targets, they are difficult to quantify or are picked up in subsequent targets. So for instance, in diabetes there is a target for the percentage of patients with a record of micro-albuminuria testing in the previous 15 months, DM13. In some instances these will lead into a subsequent target to treat those patients based on its results; such as DM15 which rewards prescriptions of ACE inhibitors to diabetic patients with micro-albuminuria. There were ten targets fitting this classification all of whom were removed.

Finally there were targets that have been categorised as 'patient reviews'. These refer to targets which involve a review of the patient's condition; or an agreed care plan. A third of these are in mental health and they are present in dementia also, both conditions where the individual may be dependent or reliant on others or outside services to administer aspects of their care. Therefore the consent and opinion of those additional interested individuals or bodies is sought, as well as that of the individual, to review present care and agree future care. An example in Mental Health is MH6 which refers to the percentage of patients with a documented comprehensive care plan agreed between the individual, their family and, or their carer. The other aspect to this 'patient reviews' category are reviews of the condition for diseases where care will need to be amended, scaled up or back depending upon responses to previous treatments and changes in the severity of it. As targets in this category only refer to a process and make no stipulation on an expected outcome it is difficult to attribute any subsequent improvements to them. Hence nine targets fitting this description were removed. At the end of this review of the QOF guidance 58 targets had been removed, leaving 28 targets remaining.

3.3.2 Stage 2: Best practice guidelines

The next stage was to consult national clinical guidelines to review the level of clinical evidence for the remaining targets. NICE and SIGN commissioned guidelines were the first areas consulted to find clinical evidence for QOF targets. If guidance for targets were missing from these areas, individual or joint guidance from disease specific organisations was consulted. Both NICE and SIGN use the same ranking methodology to score the research evidence and generate recommendations. NICE classifications and their corresponding level of evidence is shown in Figure 3-3, taken from the Atrial Fibrillation NICE guideline.[67] The same guidance is produced in a slightly adapted format within the SIGN guidelines and replicated in the guidance produced by the British Hypertension Society (BHS)[68]. In this chapter and throughout the thesis the term 'high level evidence' is used to refer to level 1+ and 1++ study designs

Figure 3-3: NICE, Evidence to recommendations, AF guidelines

Levels of evidence		Classification of recommendations	
Level	Type of evidence	Class	Evidence
1++	High-quality meta-analysis (MA), systematic reviews (SR) of randomised controlled trials (RCTs), or RCTs with a very low risk of bias.	A	Level 1++ and directly applicable to the target population or level 1+ and directly applicable to the target population AND consistency of results. Evidence from NICE technology appraisal.
1+	Well-conducted MA, SR or RCTs, or RCTs with a low risk of bias.		
1–	MA, SR of RCTs, or RCTs with a high risk of bias.	Not used as a basis for making a recommendation.	
2++	High-quality SR of case-control or cohort studies. High-quality case-control or cohort studies with a very low risk of confounding, bias or chance and a high probability that the relationship is causal.	B	Level 2++, directly applicable to the target population and demonstrating overall consistency of results. or extrapolated evidence from 1++ or 1+.
2+	Well-conducted case-control or cohort studies with a low risk of confounding, bias or chance and a moderate probability that the relationship is causal.		
2–	Case-control or cohort studies with a high risk of confounding, bias or chance and a significant risk that the relationship is not causal	Not used as a basis for making a recommendation.	
3	Non-analytic studies (for example case reports, case series).	C	Level 2+, directly applicable to the target population and demonstrating overall consistency of results or extrapolated evidence from 2++.
4	Expert opinion, formal consensus.	D	Level 3 or 4 or extrapolated from 2+ or formal consensus.
		D (GPP)	A good practice point (GPP) is a recommendation based on the experience of the GDG.

This study was primarily interested in Grade A recommendations, level 1+ and 1++

evidence, or their equivalent, though lower levels are reported. In the past few years NICE guideline evidence has adopted the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) toolbox to rank the quality of evidence. They then produce 'recommendations' or 'key priorities for implementation' for those areas which have the highest ranking evidence. In the GRADE ranking system this equates to a strong recommendation which describes circumstances where they are very certain that the benefits of the treatment do or do not outweigh its risks and burden[69 70]. Where the GRADE criteria have been applied in making the recommendation an 'R' is placed before the referenced source. It is not possible to make direct comparisons between the two ranking tools. GRADE limits itself to just two levels of recommendation, strong and weak, which are determined by the scale of the treatment's benefit relative to its risk and financial costs. This is a very different approach to the A-D class of recommendation used in earlier guideline evidence, most notably with regards to cost, where there was no explicit requirement to consider the costs implications of any recommendation. However in terms of the level of clinical evidence, a GRADE recommendation would reasonably equate to no lower than level 2++/Class B evidence. On other occasions no evidence could be found for the target. In these cases only the guideline source is given, with the reason for the lack of a level of recommendation provided in the summary column.

Outlined in Table 3-1 is a summary of the documents examined and information gathered on the clinical effectiveness of the remaining 28 targets. Where targets were common across conditions and guidance was shared or reasoning similar, the evidence is presented for the group of conditions and targets collectively, rather than separately.

Table 3-1 QOF targets, source and level of evidence

QOF Indicator(s) 2009 summary	Clinical guideline(s) summary	Level/ source of evidence
CHD; Stroke; Hypertension: Patients in whom	Individuals with established cardiovascular disease with sustained systolic blood pressure greater than 140mmHg and/or diastolic greater than 90mmHg	

the last blood pressure reading, measured in the previous 15 months (9 months for Hypertension) is 150/90 or less. For CKD a blood pressure of 140/85 or less. For diabetes a blood pressure of 145/85 or less. CHD6; STROKE6; BP5; DM12; CKD3.	should be considered for blood pressure lowering drug therapy. In patients aged over 80 with treated hypertension, a blood pressure target of 150/90 or below should be aimed for. 140/90 in those aged under 80. The same levels were referenced and recommended in the most recent Angina guideline. In non-diabetics with hypertension, the optimal BP treatment goals are <140/85 and the minimum acceptable level <150/90. In diabetic patients the optimal BP target should be <135/80 with a minimum acceptable (audit standard) of <140/80 Patients with CKD and proteinuria should have a target maximum systolic blood pressure of 130mmHg In patients with CKD aim for a systolic BP target range of 120-139, diastolic BP<90. Diabetic patients with cardiovascular disease should have a diastolic BP target<=80 mmHg; and a systolic BP<130mmHg. In Stroke the target should be less than 140/85, 130/80 and below if they have diabetes.	A[71] SIGN 1+[72] NICE R[72]NICE R[73]NICE B[68] BHS B[68] BHS A[74]SIGN A?[75]NICE A[76] SIGN D[76]SIGN D[77]NICE
Smoking cessation advice to patients with any or any combination of CHD, Stroke, Hypertension, Diabetes, COPD, CKD, Asthma,	There is limited high quality evidence for the effectiveness of offering advice and behavioural support from NHS Stop Smoking Services. Pharmacy led interventions have been shown to have a significant validated impact on quit rates. Nicotine replacement therapies or bupropion should be used as part of a smoking cessation programme to augment professional advice and increase long term abstinence rates.	1 x 1++ study[78] NICE A[78] NICE A[71] NICE

Schizophrenia, Bipolar disorder or other pyschoses. SMOKING4	No guideline evidence was found specifically related to the effectiveness of smoking cessation advice in patients with schizophrenia, bipolar disorders or other psychoses.	NICE[79 80]
CHD9. Patients with CHD prescribed an aspirin, alternative anti-platelet therapy, or an anti-coagulant (unless a contraindication of side-effects are recorded)	Aspirin should be offered to all patients after an MI and be continued indefinitely. Clopidogrel should be offered as a first line monotherapy after an MI. Clopidogrel, in combination with low dose aspirin, is recommended for use in the management of non ST segment elevation acute coronary syndrome in people who are at moderate to high risk of MI or death. Aspirin (75mg) a day should be routinely given and continued long term in patients with diabetes and CHD. Consider 75mg daily dosage of aspirin in people with stable angina, taking into account the risk of bleeding and co-morbidities.	A[81] NICE A[81] NICE A[81] NICE A[71] NICE R[73] NICE
CHD10. Patients with CHD who are currently treated with a beta blocker	Patients with clinical MI should be maintained on long term beta blocker therapy. Beta blockers should be used as the first line therapy for the relief of symptoms of stable angina.	A[71] NICE A[82] SIGN
CHD11. Patients with a history of MI (diagnosed after 1 April 2003) treated with an ACE inhibitor or Angiotensin II	After an MI all patients with preserved left ventricular function or left ventricular dysfunction (LVD) should continue with an ACE inhibitor indefinitely, whether or not they have symptoms of heart failure. In post MI patients who have to discontinue an ACE inhibitor because of intolerance or allergy, an ARB should be substituted.	A[81] NICE A[81] NICE

antagonist.		
HF3. Patients with a current diagnosis of heart failure due to LVD treated with an ACE inhibitor or ARB, who can tolerate therapy and for whom there is no contra-indication	ACE inhibitors should be considered in patients with all NYHA classes of heart failure due to LVD. Patients with chronic heart failure due to LVD alone, or heart failure, LVD or both following MI, who are intolerant of ACE inhibitors should be considered for an ARB	A[83] SIGN A[83] SIGN
HF4. Patients with a current diagnosis of HF due to LVD treated with an ACEI or ARB, who are currently treated with a beta blocker licensed for heart failure	All patients with heart failure due to LVD of all NYHA functional classes should be started on beta blocker therapy as soon as their condition is stable. Offer both ACE inhibitors and beta-blockers licensed for heart failure to all patients with heart failure due to LVD. Use clinical judgement when deciding which drug to start first	A[83] SIGN 1++[84] NICE
STROKE12. Patients with a stroke shown to be non-haemorrhagic, or a history of TIA, who receive an anti-platelet agent, or an anti-	Low dose aspirin (75mg daily) and dipyridamole (200mg modified release daily) should be prescribed after ischaemic stroke or TIA. Clopidogrel (75mg daily) monotherapy should be considered as an alternative to combination aspirin and dipyridamole after ischaemic stroke or TIA.	A[77] NICE A[77] NICE

coagulant		
DM23, 24 and 25. Patients with diabetes in whom the last HbA1c is 7, 8 or 9 or less respectively.	HbA1c should be controlled, ideally around 7%, in all diabetic patients to prevent the onset and progression of diabetic eye disease. An HbA1c target of 7.5% is reasonable to reduce the risk of microvascular disease, with a target of 6.5% at diagnosis. Targets should be based on individual circumstances in order to balance benefits and harms, in particular hypoglycaemia and weight gain.	A[76] SIGN A[76] SIGN
DM15. Diabetes patients with a diagnosis of proteinuria or micro-albuminuria who are treated with ACE inhibitors or A2 antagonists.	Both type 1 and 2 diabetic patients with micro-albuminuria should be treated with ACE inhibitors (or ARB's in type 2 patients only), irrespective of blood pressure. ACE inhibitors and/or ARB's should be used as agents of choice in diabetic patients with Chronic Kidney Disease and proteinuria.	A[76] SIGN A[76] SIGN
EPILEPSY8. Patients aged over 18 and on drug treatment for epilepsy who have been seizure free for the last 12 months	It is difficult to draw specific recommendations from the clinical guidelines for this target. Firstly there is evidence that if someone has been seizure free for 2 or more years, the odds of a seizure are sufficiently low that treatment can be withdrawn. Treatment whether using mono or combination drug therapy has to be patient specific, and at the patients consent. Side effects can outweigh any treatment effects in terms of seizures avoided. Risk of seizure are highest in the under 16's and the over 59. Future seizure risk depends on seizure type, neurological deficit and epilepsy syndrome[85].	[85] NICE
MH5. Patients	No evidence found within the schizophrenia	[79] [80] NICE

on lithium therapy with a record of lithium levels in the therapeutic range	guideline[80]. Lithium is a key priority recommendation in bipolar patients dependent on co-morbidities, response to previous treatment, patient preference and history of adherence. It should only be used if symptoms are not severe because of its slow onset. There is a lack of evidence around what is a therapeutic range. The three placebo controlled RCT's reported in the guideline adopted a minimum plasma level of 6mmol/l or more. UK laboratories use a lower limit of 4-6mmol/l or less[79]	
CK5. CKD patients with hypertension and proteinuria who are treated with an ACEI or ARB	ACE inhibitors and/or ARB's should be used as agents of choice in CKD patients with proteinuria regardless of diabetes status. Offer ACEI or ARB to non-diabetic CKD patients with hypertension	A[74] SIGN R[75] NICE
AF3. Patients with atrial fibrillation who are treated with anti-coagulants or platelets	In patients with permanent AF where antithrombotic therapy is given to prevent strokes and/or thromboembolism, adjusted warfarin should be given. Where warfarin is not appropriate aspirin should be given. In AF patients who are either post stroke or have a TIA, warfarin is the most effective thromboprophylactic agent.	A[67] NICE B[67] NICE A[67] NICE
PP2. People with hypertension diagnosed after 1 April 2009 who are given lifestyle advice on increasing	Diets low in total and saturated fats should be recommended to reduce CV risk. People with hypertension should be advised to reduce their salt intake as much as possible to reduce BP. Physical activity of at least moderate intensity is recommended for the whole population.	A[71]NICE A[71] NICE B[71] NICE

physical activity; smoking; alcohol and diet	All people who smoke should be advised to quit and offered support to do so. Patients may be advised that light to moderate alcohol consumption may be protective against the development of CHD	B[71] NICE B[71] NICE

3.3.2.1 Evidence discussion

3.3.2.1.1 Blood pressure

It is clear that blood pressure control has high level clinical evidence in all the conditions where it constitutes a target. What is usually less certain is the evidence behind the specified level in the QOF, with the QOF target tending to be less stringent than those recommended by the guidelines. In the case of Hypertension, Stroke and CHD the QOF sets a target of 150/90mmHg or less. However the guideline evidence argues that the QOF level should be considered the minimum acceptable, an 'audit' level, which is suitable only for hypertensive patients over 80[72]. The evidence available for a level below that 'audit level' however, of 140/90 for patients under 80 years of age in the NICE guidance[72]; and of 140/85 from the BHS is sub Grade A standard[68].

For diabetes the QOF target for blood pressures is 145/85. Grade A level evidence, was found for a diastolic blood pressure of 80mmHg or below in diabetic patients with CVD[76]. Weak, Grade D level evidence was found for a systolic blood pressure of 130mmHg or below in the same patient group[76]. The BHS meanwhile recommends a target level of 140/80 or less citing Grade B level evidence. None of which is in exact agreement with the QOF target.

Like diabetes, CKD has a more stringent target than the other conditions, in this instance it is140/85. There is some contradiction within the guidelines regarding a target level, with NICE recommending a 140/90 or below level which appears to be based on high level evidence[75], and SIGN, 130mmHg and below for systolic blood pressure with high level evidence[74]. Indeed the NICE guidelines for CKD recognise the difficulties in setting a diastolic and systolic blood pressure target in this condition in tandem, as treatments given to maintain systolic blood pressure in an optimal range, result in diastolic blood pressure falling below its optimal range. For this reason they recommend a systolic range of 120-139 mmHg and a diastolic target

of 90mmHg or less. This recommendation appeared to be based on the highest level clinical evidence although it couldn't be determined with certainty[75].

Overall it would have to be concluded that while there is high level evidence for controlling blood pressure, there is uncertainty around specific levels that it should be controlled to. That while the QOF targets are probably not stringent enough they represent a good compromise based on uncertainty over what the exact level should be.

3.3.2.1.2 Cholesterol control

In all the conditions where cholesterol control is a target, the level is set at 5mmol/l or less. Once again there is high level evidence for the policy of control, but in this case greater certainty, and high level evidence as to what level should be aimed for. That is for a level of 4mmol/l or less, one that is more ambitious than that set by the QOF, which the guidelines recommend as an 'audit level'; a target that all patients should reach and we can measure progress to[86 87]. While the clinical evidence argues that 4mmol/l is the optimal level, and doctors should consider increasing statin dosage to meet it; it also acknowledges that less than half of patients will meet it[86]. Given this fact, and despite questions over the level at which it is set; the level set by the QOF can be seen as a target which incentives provision for all patients. One which practices should aim to ensure all its patients meet and a significant number progress beyond.

3.3.2.1.3 Glycaemic control in diabetes patients

Within the QOF there were 3 targets for HbA1c levels; for levels at 9 or less, 8 or less and 7 or less. There is grade A level evidence concerning the control of glycaemic levels in diabetic patients ideally to 7%, though with some flexibility to reflect individual patient circumstances[76]. This range suggests lowering HbA1c to 6.5 in those patients who can tolerate treatment to that level, and 7.5 as a reasonable level to reduce the risk of microvascular disease[76]. The logic behind having three targets for this one measure in the QOF is clearly to reward GP's for progressively lowering a patient's glycaemic level; while at the same time providing sufficient incentive to treat patients who are unlikely to reach lower target thresholds. However it was felt that based on the strongest clinical evidence there was certainly no evidence for the higher level target, particularly given that the 8 or less target provided some room to progressively lower glycaemia down to its optimal range. In

which case it may be better to allocate the points from the 9 or less target to the remaining two; to provide a greater incentive for GP's to bring patients down towards the optimal range. For this reason DM25, HbA1c of 9 or less was dropped.

3.3.2.1.4 Prescription and immunisation targets

In all targets which refer to prescribing ACE inhibitors or ARB's; aspirin, anticoagulants or platelets there is high level supportive clinical evidence. In the case of Heart Failure there is an additional target, HF4, relating to the prescription of ACE or ARB and beta blockers. The evidence statements in the Heart Failure clinical guidance showed high level evidence for the prescription of both ACEI and beta blockers using an RCT study design comparing the drug class against placebo[84]. Within the NICE HF guideline there was also RCT studies which sought to determine which of the two drugs when used in combination, should be initiated first[84]. The evidence was inconclusive and that judgement was left to the individual clinician in the guideline recommendation. What was not shown in the NICE guideline however was the evidence for ACE inhibitor or ARB plus beta blockers over ACE or ARB alone; or for that matter beta blockers compared to ACE inhibitors or ARB. So in this instance the evidence for this target, as worded, was unable to be determined from the clinical guidance. For this reason this target was dropped and only HF3 was taken into the next stage. There was strong evidence for prescribing beta blockers to CHD patients, and for the QOF target as worded, CHD10[71 82].

3.3.2.1.5 Mental health, epilepsy and influenza immunisation

For mental health and epilepsy the evidence was conditional on so many patient factors and based on a number of provisions that it is near impossible to frame a target which covers all these eventualities. Given this the guidance on what constitutes a lithium therapeutic range, in the mental health target is possibly deliberately vague, as it needs to reflect individual patient characteristics. The therapeutic range currently used by UK laboratories and referenced in the guidance, is lower than that adopted in the few RCT's to have taken place in this area[79]. Due to the lack of robust evidence for this target it had to be dropped. With epilepsy there was also great uncertainty around the target. Linking drug therapy and seizure free period is dependent on a range of patient factors including type of seizures, side effects and drug tolerance and therefore will not be appropriate to all[85]. For this

reason this target was also removed. Finally no clinical evidence to support the inclusion of influenza immunisation targets in the QOF was found except a reference to it being a current recommendation of the Department of Health and Joint Committee on Vaccination and Immunisation, contained within the QOF guidance itself. These targets as a result were also dropped.

3.3.2.1.6 Primary prevention

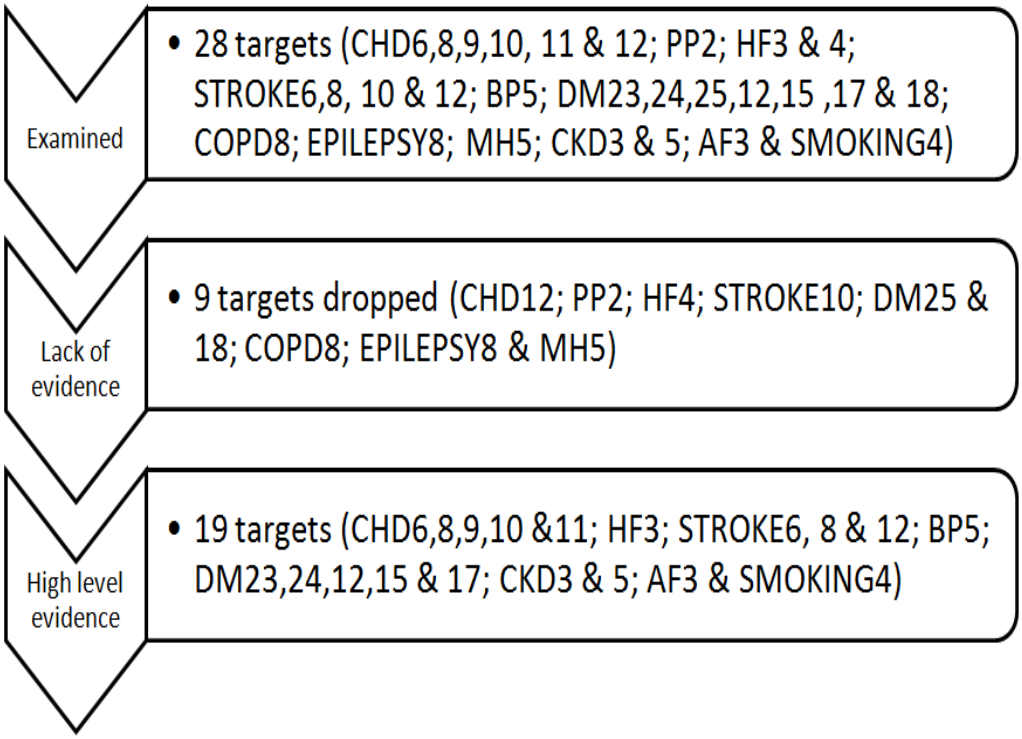
In cardiovascular disease primary prevention, PP2, there is high level evidence that reducing salt and fat intake reduces cardiovascular complications. However this evidence relates to controlled environments where individuals are compelled to do so[71]. There is lower level evidence around the protective qualities of alcohol within moderated limits and smoking cessation advice both applied to the whole population[71]. What could not be found was evidence for the effectiveness of offering lifestyle advice alone outside of controlled environments where there was some form of compulsion on participants to follow it. It is difficult to map lifestyle guidance to clinical outcomes. This target also replicates the smoking cessation target for hypertensive patients. Given these reasons the evidence for this target was considered to be uncertain and it was removed.

3.3.2.1.7 Smoking cessation

The health benefits of ceasing smoking are well established, even in those patients who been heavy smokers for long periods[88]. There is high level evidence for smoking cessations services delivered through NHS stop smoking services and pharmacy led interventions, but weaker evidence for the policy generally[71 78]. Additionally there is high level evidence for the effectiveness of nicotine replacement therapies in patients with cardiovascular and respiratory disease, particularly when used in conjunction with cessation services[78]. However the target is not explicit in its wording as to what form advice should take, or which therapies should be applied, so its validity is very open to interpretation. Nonetheless as the target refers to specialist services, which have an evidence base, and as it gives the GP the opportunity to recommend nicotine replacement therapies as part of a smoking cessation programme; it was considered to be in keeping with the high level evidence cited.

The targets starting this process and remaining at its conclusion are shown in Figure 3-4 below.

Figure 3-4: Stage 2 summary



3.3.3 Stage 3: Targets with high level clinical evidence

3.3.3.1 Objective

To select from the 19 targets found to have high level clinical evidence, those which could be effectively analysed using CPRD and HES data. The following factors were considered when making the decision:

1. The targets longevity
2. The stability of the target
3. The size of the patient population treated by the target
4. The ability to extract data on clinical outcomes in HES

3.3.3.2 Description

Nineteen targets had a high level of clinical evidence; covering seven disease conditions, as well as a smoking cessation target which included multiple disease conditions. Faced with a range of possible QOF targets the decision was whether to focus on a number of individual targets across a range of conditions, one's that were common across a range of conditions such as the Blood Pressure and Serum Cholesterol monitoring, or focus on one or more disease condition. To reach a final decision in a logical manner a number of points important to the analysis were

checked against each target. These were to do with longevity of the target, namely the number of QOF years in the study in which the target featured. The target's stability, in terms of the number of changes to the points awarded and payment thresholds during the QOF study period. Patient numbers affected by the target and the ability to measure the expected clinical benefits of the target, and attribute it to the target, using HES data.

Since their introduction QOF targets have undergone a number of changes with targets being dropped, new ones introduced; payment thresholds, wording and points amended. Within the QOF every time a target has changed its wording the initial target has been dropped and a new one created with a new target number. In this thesis the 2009/10 wording and numbering is used as this is the most up to date in the study period. Earlier versions of the QOF have been visited to see if the target exists in those and if not whether its wording is comparable. If the target number occurred in previous QOF guidelines or its 2009/10 wording was directly comparable to earlier versions, this would count towards the longevity of the target, which is a maximum of 7 years for the study period. However each change in wording, points and payment threshold would be reflected in the stability criteria. To measure the impact and benefits of a target with confidence the target should preferably be in the QOF for a long period but not have changed over that period, reflected in a high longevity and low stability figure.

This study draws from a very large dataset where small sample sizes are unlikely to be an issue. Nonetheless it is still important to consider the size of the potential population, as while it may not be a factor that excludes targets it is one which could make analysis potentially more difficult and therefore be a deciding factor when choosing which targets to analyse. For this purpose the prevalence of the condition is shown for whole condition targets, or where it refers to a subset of that population, the approximate population prevalence for that subset. Details on how these figures were calculated are covered in detail in Appendix 3.

The fourth consideration was whether the outcome could be measured in HES, and clearly attributed to the target. This was of prime importance since as shown in the previous chapter there is an absence of studies looking at the effect of incentives in primary care on clinical outcomes and the overall aim of the thesis is to find evidence for the effectiveness of QOF targets. To address this deficiency and meet

the thesis aim it was necessary to establish that the clinical outcome for the target, that had been shown to supported by high level evidence in the previous stages; could be clearly linked to the QOF target in the datasets.

The results of applying these criteria are shown in Tables 3-2 to 3-8, along with the expected health improvements. The smoking cessation target is not included in the tables below as using these criteria would give misleading results regarding its suitability. This is a target that applies to multiple conditions and was initially listed, when the QOF was first introduced, as a separate target in each of the core chronic conditions to which it applies today. In 2006 smoking as a condition in its own right was introduced into the QOF and those separate targets merged in to one. In 2008, smokers from the mental health condition group were added. So despite being around since the beginning of the QOF for certain conditions, it would have a low figure for longevity due to the high number of changes the target has incurred. Even applying the target in the context of a specific QOF condition, despite essentially remaining the same, a misleadingly high stability figure would result given all the changes mentioned. The expected improvement in clinical outcomes are also common across all respiratory and cardiovascular diseases from successfully quitting smoking and relate to reduced mortality, hospitalisations, morbidity, cardiovascular and respiratory complications.

3.3.3.3 Analysis

Table 3-2 CHD Secondary prevention high level evidenced targets

Target	Longevity	Stability	Population %	Clinical outcomes
CHD6 BP 150/90 or less	7	2	3.4%	Reduced risk of major and non-fatal cardiovascular complications; all cause and cardiovascular mortality[72]
CHD8 Cholesterol 5mmol/l or less	7	2	3.4%	Reduced risk of all-cause mortality, cardiovascular mortality, CHD mortality and fatal MI. Also for rates of non-fatal stroke, TIA, MI and

				unstable angina[86]
CHD9 Anti-platelets or coagulants	7	1	3.4%	Reduced risk of death and cardiovascular events, non-fatal MI or stroke[71 81].
CHD10 Beta blockers	7	1	3.4%	Reduced risk of death, all-cause mortality, cardiovascular-cause mortality, non-fatal MI[73 81].
CHD11 MI: ACEI or AII	7	1	0.54%	Reduction in all-cause mortality, and recurrent MI[81 73]
All of the CHD targets have been in place since the beginning of the QOF period and have also been very stable. The whole population targets have a large prospective population, but for the MI subgroup it is substantially smaller. There is a clear link between the targets and expected benefits which are extractable using CPRD and linked HES data				

Table 3-3 Heart Failure high level evidenced targets

Target	Longevity	Stability	Population %	Clinical outcomes
HF3 ACEI or ARB	5	0	0.4%	Reduced risk of hospital admissions due to heart failure or related complications and in mean level all-cause mortality[84]
This target was present for 5 years of the QOF study period and underwent no changes during that period. The potential patient population is low but the expected benefits are clearly related to the condition.				

Table 3-4 Atrial Fibrillation high level evidenced targets

Target	Longevity	Stability	Population %	Clinical outcomes
AF3 Anti-platelets or coagulants	5	1	1.4%	Reduced risk of thromboembolic complications, ischaemic stroke and vascular events (MI, stroke, haemorrhage, vascular death)[67]
The target was present for 5 years of the study period and had a single change. The potential patient population is low in relation to other conditions. In terms of expected benefits there are potential problems in being able to attribute the outcome to the target using the datasets since the outcomes manifest themselves in other conditions such as Stroke and CHD.				

Table 3-5 Chronic Kidney Disease high level evidenced targets

Target	Longevity	Stability	Population %	Clinical outcomes
CKD3 BP 140/85 or less	5	0	4.2%	Slows the deterioration of glomerular filtration rate and reduce proteinuria; reducing cardiovascular risk[74]
CKD5 Hypertension & proteinuria: ACEI or ARB	3	1	0.21%	Reduced risk of all-cause mortality, proteinuria, and end stage renal disease. Reductions in the rate of disease progression[74].
CKD5 is present for only 3 years of the QOF study period and underwent a change during it. The prospective patient population is small; though the expected benefits can be related to the target. CKD3 is present for 5 years and has been stable. The population prevalence appears to be adequate though the outcomes that manifest themselves in reduced cardiovascular risk may be hard to link in the HES data.				

Table 3-6 Diabetes high level evidenced targets

Target	Longevity	Stability	Population %	Clinical outcomes
DM23 HbA1c, 7 or less	2	0	5.3%	Reduced risk of micro-albuminuria and retinopathy[76].
DM24 HbA1c, 8 or less	2	0	5.3%	Reduced risk of micro-albuminuria and retinopathy[76].
DM12 BP, 145/85 or less	7	2	5.3%	Reduced risk of macro-vascular and microvascular disease. Reduction in major cardiovascular events[71 76]
DM15 Proteinuria or micro-albuminuria : ACE or AII	7	1	0.59%	Reduced risk of arterial events (mortality or stroke) and MI[76]
DM17 Cholesterol, 5mmol/l or less	7	1	5.3%	Reduced risk of all-cause mortality, cardiovascular mortality, CHD mortality and fatal MI. Also for rates of non-fatal stroke, TIA, MI and unstable angina[76 86]
DM23 and DM24 have a low longevity figure which would make analysis very problematic. All the targets have been stable, DM12 the least with 2 changes but over a 7 year period. There is no issue with the percentages for the relevant population for all of the targets with the potential exception of DM15. In terms of outcomes for the remaining targets outside of DM23 & 24, it may be difficult to attribute these to a diabetes QOF target as they are cardiovascular events and not metabolic complications				

Table 3-7 Hypertension high level evidenced targets

Target	Longevity	Stability	Population %	Clinical outcomes
BP5 BP, 150/90 or less	7	2	13.4%	Reduced risk of major and non-fatal cardiovascular complications; all cause and cardiovascular mortality[71 72]
This target covers the whole of the QOF study period, and was largely stable. The prospective study population is large. The clinical outcomes are broad, and would apply to a host of clinical treatments. This is further complicated by the fact that the target itself is shared across a number of cardiovascular and metabolic conditions which would make it difficult to isolate the effects of this specific target.				

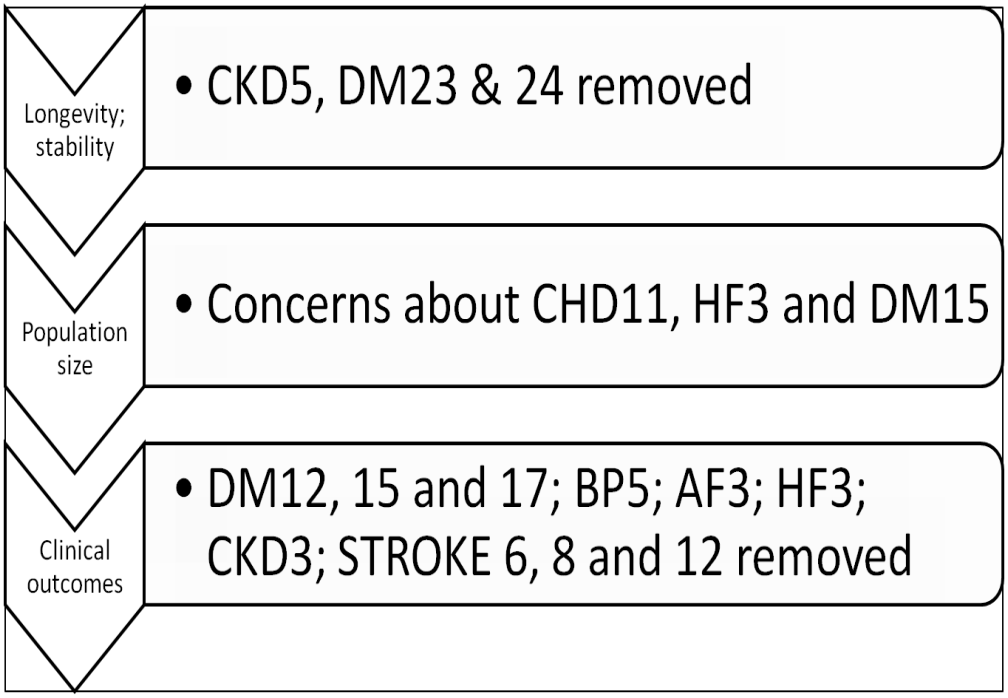
Table 3-8 Stroke and Transient Ischaemic Attack high level evidenced targets

Target	Longevity	Stability	Population %	Clinical outcomes
STROKE6 BP, 150/90 or less	7	1	1.7%	Reduced risk of major and non-fatal cardiovascular complications and all-cause mortality[71 77]
STROKE8 Cholesterol, 5mmol/l or less	7	1	1.7%	Reduced risk of all-cause mortality, cardiovascular mortality, CHD mortality and fatal MI. Also for rates of non-fatal stroke, TIA, MI and unstable angina[77 86]
STROKE12 Anti-platelets or coags	7	2	1.7%	Significant reduction in symptomatic pulmonary embolism or recurrent stroke.[77]
All the targets have longevity, stability and are sufficiently prevalent. There may be issues with attributing the outcomes of the first two STIA to those targets as they are common cardiovascular outcomes shared with a number of QOF targets.				

3.3.3.4 Results

The results of this section are summarised in the figure 3-5 below with details of targets dropped and those where concerns were raised against each criteria. Years and changes are merged as years proved to be a much bigger factor in decisions; with changes in one instance adding weight to the decision to drop a target and making an additional target questionable.

Figure 3-5 Stage 3 Summary



3.3.3.4.1 Longevity and stability

A number of CHD, Stroke, Hypertension and certain Diabetes targets have been present throughout the QOF study period. As had glycaemic control for diabetic patients; however this target had seen repeated changes such that the present targets could not be directly compared against previous versions. With only 2 years of data these targets had to be dropped as this would not be sufficiently long to pick up an effect. Even applying earlier versions it would be difficult to apportion an effect of treatment on outcomes as the treatment target has constantly moved. Focusing on years still, CKD5's 3 years of follow up is probably insufficient for benefits to be reflected. In that short period it also had a change to its points awarded, and the target applied to a small proportion of the population. Due to all these concerns this target was also dropped.

For the remaining targets there are no concerns as in most cases these targets have been present throughout the QOF study period. None of these therefore escaped the changes made to the QOF in 2006, which saw reform of payment thresholds, among other changes. However the substance of the targets has remained essentially the same. They have also been no more than two changes to the stability terms of the target. The remainder were introduced in 2006 so have 5 years of data but no changes over that period, AF3 aside.

3.3.3.4.2 Prospective population size

In the case of prevalence this obviously favours the common chronic conditions, hypertension in particular, raises potential issues for Heart Failure and certain targets which consider a subset of the condition population; namely CHD11 and DM15. However this is being used for information only at this stage, not to specifically exclude targets, but as an additional source of information and potential deciding factor in any final decision.

3.3.3.4.3 Access to clinical outcomes in HES

The final column looks at the expected benefits of the target, taken from the best practice guidelines. Within HES data the study is interested in the primary diagnosis coding given to the patient upon hospitalisation. For all the remaining targets an expected common benefit was a reduction in cardiovascular complications. Repeatedly mentioned among these were both fatal and non fatal MI, CHD and Stroke. It is important when using pre-collected data like CPRD and HES where patients are not being observed that events in either can be linked. Extraction of relevant data is facilitated when outcomes are condition specific and where there is a clear progression from the care and condition to any outcomes. In this area CHD secondary prevention stood out. Following diagnosis, if the condition is well managed in primary care, we expect to see fewer ischaemic heart disease admissions. There is a clear causal pathway with poor care leading to more complications, and more severe forms of the condition, with each progressive form having a separate outcome code.

With the diabetes targets, STROKE6 and 8, AF3 and CKD3 the link is not so clear. Complications manifest themselves in other conditions and then coding and extractability within HES becomes an issue. Firstly because there is not the certainty we had with CHD that the condition caused the outcome. Secondly there will also be

concerns about the data, as there could be doubts that the underlying cause or the presenting condition formed the primary diagnosis. For BP5 while the outcomes are clearly linked to hypertension management, the issues are with the fact that the outcomes are broad and related to a number of other clinical treatments and that the target is replicated across a number of QOF conditions. Hence there would be real difficulties in finding and thereby linking outcomes specifically related to this target in HES.

3.3.3.5 Summary

Going through all the criteria the targets remaining are the CHD targets, HF3, and STROKE12. For HF3 the percentage of the study population that would be covered under this target is low enough to question the practicality of analysing it, and its absence for two years of the study period is also a negative. STOKE12 is strong on longevity and stability with outcomes that are clearly linked to the target. However again this is a target which is common in a number of other targets, involves a drug that is given widely, and most likely in addition to a host of other drugs with similar benefits, which could make direct linkages to secondary outcome measures problematic. CHD was a strong performer on all the criteria set. In particular it was very strong with regard to the ability to measure and relate clinical outcomes in HES. An additional benefit, given the number of CHD targets going into this final stage and performing well on the criteria used, is that it offers the opportunity to select a large part of the QOF targets for a single disease domain. This offers the opportunity to consider the effects of a package of care for a single disease group, which is how the QOF works in practice; rather than considering separate aspects of care across multiple conditions.

3.3.3.5.1 Smoking

Smoking is a predictor of poor health generally, and a significant risk factor for all cardiovascular diseases, CHD and Stoke in particular[88-90]. The benefits of ceasing to smoke are therefore significant and smoking cessation services have proven effective in helping patients to quit[78]. It was however omitted from this section as the criteria used would give a misleading result on the stability and longevity of the targets and prejudice its selection. Following the appraisal of the other targets a decision now had to be made on this target. The options available were to apply it to CHD only, in effect applying the pre 2006 version of the target. Or alternatively apply it to all

conditions that were included in Stage 3 or apply the target as worded, which would include additional conditions. Given these options the natural solution was to apply it to CHD, keeping with the focus of the other targets.

3.4 Conclusion

The objective of this chapter was to find which targets had the highest level of clinical evidence through a two stage process; and then in the third and final stage determine which of those it would be possible to analyse using CPRD and HES linked data. In this third stage, the CHD targets performed consistently well on all the criteria set. They possessed stability, longevity and clear linkage to Secondary care outcomes. The breadth of the targets presents the opportunity to consider a whole package of care, and to consider a QOF disease domain almost in its entirety. With a number of these targets shared by other conditions and common outcomes there is also the opportunity to draw lessons for other conditions. No other condition or combination of targets offered this opportunity. On the basis of this, the Secondary prevention of CHD, QOF targets, were chosen as the area of research in this PhD.

The next chapter will undertake a literature review to uncover what research has been undertaken to date on the effect of physician incentives on hard clinical outcomes and endpoints. Since CHD is now the exemplar condition, this will provide the reviews focus.

Chapter 4 The effectiveness of physician incentives on CHD clinical outcomes

4.1 Objective

To examine the effectiveness of financial incentives on hard, and surrogate, CHD clinical outcomes.

4.2 Introduction

In the previous chapter CHD QOF targets were selected as the exemplar targets following a review of the clinical guideline evidence. The focus now narrows to CHD and this chapter looks at the literature regarding the effectiveness of physician incentives on CHD clinical outcomes.

The review in Chapter 2 revealed a lack of high level study design evidence on incentives. It was felt that if this review limited itself to high level study design evidence, given its focus on CHD, few if any studies would be returned. The study design criteria were therefore relaxed to include any study design. Chapter 2 also identified a lack of clinical outcomes which this review sought to address.

This review set out to find evidence principally for the effectiveness of P4P incentives on clinical outcomes. This was to find out if research, similar to that which this thesis would undertake on the QOF, had previously been undertaken. Given that the distinction between primary and secondary care is not as clearly made in other health systems as it is in the UK, and these studies could provide useful information for this research, incentives in secondary care were also included. It was also felt that limiting the search to P4P would find few studies, and that evidence on other forms of incentives could provide useful information on the effectiveness of incentives more generally. Most importantly in the context of this research it would provide some means of comparison for any final results on the QOF, which lacks a control group or comparator.

This chapter is interested in surrogate, and more specifically, hard CHD clinical outcomes. Hard clinical outcomes are defined as a significant clinical event where the individual has a serious or end stage form of disease. An example for CHD patients would be a myocardial infarction (MI). Surrogate clinical outcomes relate to measures, generally measured in primary care, that have an evidenced relationship

with a clinical outcome, and hence form a mapped measure. An example would be cholesterol control and CVD risk. Surgery to resolve CHD complications in this context was considered to be a hard clinical outcome, as it was assumed that this would predominantly be undertaken in patients who had more serious forms of CHD, and could therefore be considered to be a significant clinical event. However since surgical rates are also related to access to services and the physicians ability to induce demand, to ensure these had clinical benefits and hence were clinically necessary, hard clinical endpoints such as mortality had to be reported in those studies[91]. To ensure that outcomes were the result of incentives and not caused by differences in patient groups covered by different forms of incentives, some attempt had to be made to match patients on key socio-demographic variables. This could take place within patients, such as before and after study designs on the same patients, or between comparable patients from different settings, e.g. states or countries.

This chapter is looking at the impact of physician financial incentives on clinical outcomes measured at the patient level. As with Chapter 2, it is not interested in physician level outcome measures in judging the effectiveness of those incentives, even though the incentives impact on the physician and not the patient. The reason for this is that while physicians may benefit financially from these incentives, be directed to act in a certain manner, or provide certain treatments as a result of their introduction; they also have an impact on the patient. Due to the manner in which healthcare is organised it is often via the physician that policy makers seek to influence patient outcomes. It is this direction of effect that this thesis seeks to explore, and hence patient level outcomes are its focus.

4.3 Background

Given that this review is looking at incentives in primary and secondary care the search terms have been broadened to reflect this. It also means that this review will cover different forms of service delivery, and new terminology, that it is important to understand. These are broken down into three categories, national and local government backed insurance schemes; private insurance backed; and managed care organisations. The predominant form of incentivisation operating in each is also discussed.

4.3.1 National and local government backed insurance schemes

Examples of centrally funded and administered schemes are the NHS in the UK and Medicare and Veteran Affairs (VA) in the USA. Medicare provides care to everyone aged 65 and over, as well as those under that age with certain disabilities. Physicians treating Medicare patients are principally paid on a fee for service (FFS) basis. The VA provides a range of benefits, one of which is healthcare to former members of the Armed Forces. Physicians are directly employed by the Department for Veterans Affairs and are therefore paid on a salaried basis

In countries like the USA there is also state and national government involvement in health care. An example of which is Medicaid in the USA. This is a means tested limited provision for low income families, funded at the federal and state level. Dependent on state of residence patients may be expected to make co-payments. Around 70% of physicians who treat Medicaid patients are paid on a capitation basis, the remainder are paid on a FFS basis.

4.3.2 Private insurance backed schemes

Private health insurance schemes are most common in the USA but are also used in Western Europe and Australia. These operate on the same basis as any insurance scheme where individuals pay a premium based on their level of coverage, risk and the size of deductible or co-payment. Physicians who work in the private insurance sector are generally paid on a FFS basis

4.3.3 Managed care organisations

The best known of these are Health Maintenance Organisations (HMOs), in the USA. These liaise with health care providers to provide care for their members on a prepaid basis. Physicians agree to provide a set level of care as laid out in the HMO guidelines in return for a guaranteed stream of income. Individuals take out insurance with HMOs through their workplace or on the open market. They are similar to the NHS in that they employ physicians in primary care to act as gatekeepers to secondary care services in order to control costs. Physicians are paid through various forms of incentivisation depending upon their contractual relationship with the managed care organisation. The most common in HMOs is capitation.

4.4 Inclusion criteria

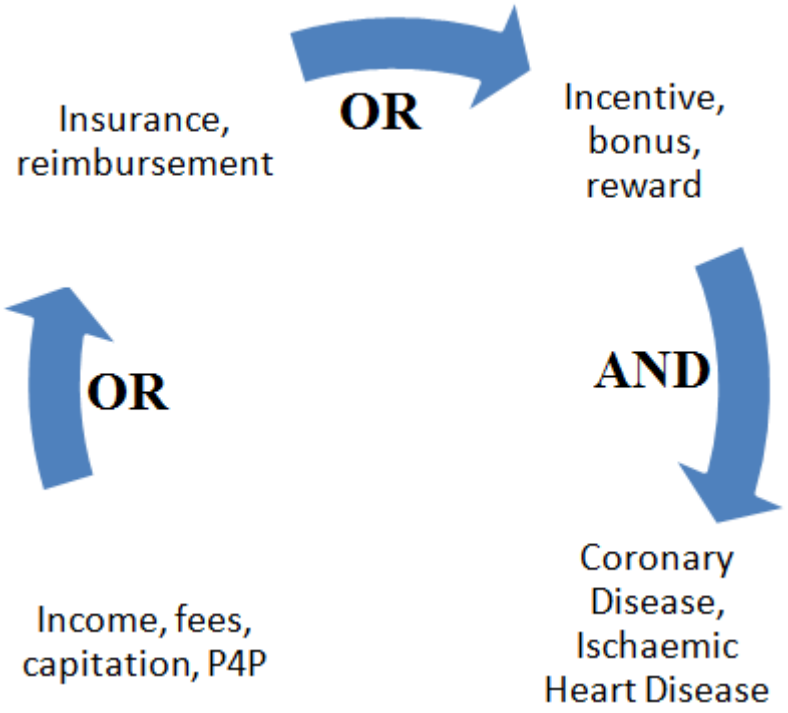
The article had to be written in English, or translated into English. An attempt had to be made in the study to match individuals in the intervention and control arms on key socio-demographic variables with adjustments made either at baseline or in final results. In addition to these stipulations patients were included if they met the following PICO criteria:

1. Population: Patients who had a known CHD diagnosis or were admitted with complications arising from CHD
2. Population: Where the patients had co-existing conditions, the results of the intervention on CHD were reported separately or could be disaggregated from any final results
3. Intervention: Any form of physician financial incentive
4. Comparator: Within comparison (before and after) or between comparison against comparable populations in different settings (e.g. Countries, States) where different physician incentives operated.
5. Outcome: Condition specific hard and surrogate clinical outcomes reported at the patient level.

4.5 Search Terms

The main search terms are summarised below and specified in full for each search engine in Appendix 4. Just looking at the QOF or P4P would severely limit this search hence search terms covered all forms of incentive. To capture CHD clinical outcomes, Mesh heading for CHD were included and matched to incentive search terms using the AND command.

Figure 4-1 CHD Physician Incentives Search Terms

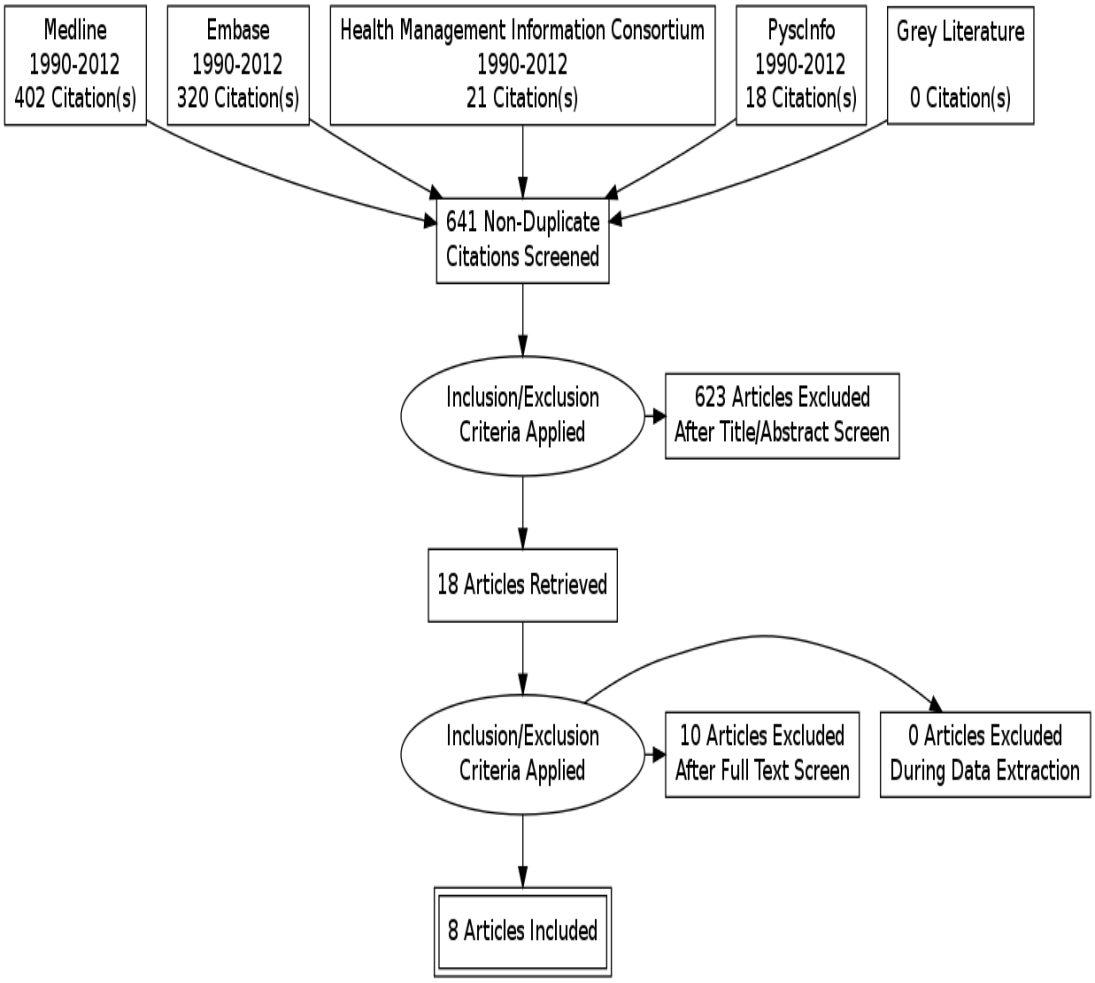


4.6 Search Methodology

The search used the OvidSP search platform, from which the following databases were searched: Embase, Medline, HMIC and PyscINFO. Further to this a grey literature search was undertaken of the websites belonging to the following institutions: The Commonwealth Fund; Nuffield Trust; and The Kings Fund.

The PRISMA diagram in figure 4-2 explains how the search was conducted. In order to capture primary and secondary care studies, generic terms for physicians were used when included. As a result over 700 articles were sourced by this search, as these terms picked up all articles where payments or incentives and CHD were mentioned in the abstract. 118 of the articles were duplicates leaving 643 articles whose abstracts were screened to see if they met the inclusion criteria, of which 18 full articles were reviewed. The main reasons for dropping articles at the abstract review stage were that the intervention was directed at the patient; there was no control or comparator; the intervention was not financial; the study did not look at clinical outcomes; and the patient did not have CHD. After reading these documents, a further 10 were dropped, leaving 8 journal articles which fully met the inclusion criteria.

Figure 4-2 CHD Physician Incentives PRISMA diagram



4.6.1 Articles excluded after full text screen

Reasons for excluding 10 of 18 articles at the full text screen stage are given in table 4-1 along with the specific PICO inclusion criteria they failed to meet. Author refers to the cited primary author.

Table 4-1 CHD Physician Payment Articles, Reasons for study exclusion

Author	PICOD	Explanation for exclusion
Assaf, 1993[92]	O	Does not measure clinical outcomes ....Looks at the effect of a change in a payment regime from FFS to a prospective payment scheme on use of diagnostic codes. It is unclear however whether changes to coding indicates changes in activity or choice of codes in response to financial incentives.

Brown, 2004[93]	O	Does not report clinical endpoints....Looks at differences in rates of surgical procedures between HMO and FFS patients. These are reported in isolation so there is no means of determining if they are a result of poor condition management and/or had clinical benefits.
Choudhry, 2011[94]	I	The intervention was not directed at the physician.....Looks at patient co-payments
Heidenreich, 2002[95]	I	The intervention was not financial....Looks at the effects of concentration of HMOs on patient care. This measure is applied to the intervention and control, and acts as a proxy for competition.
Hilleman, 2004[96]	I	The intervention was not financial....The intervention examined was a physician prompt system
Jackevicius, 2008[97]	I	The intervention was not financial...The intervention examined was changes to administrative rules.
Jha, 2012[98]	P	CHD population could not be disaggregated...CHD was one of a number of conditions incentivised under the P4P scheme examined.
Nelson, 1998[99]	P O	CHD population could not be disaggregated and outcome not CHD specific.... Looks at the effects of patient demographics, chronic conditions and physician incentives on resource utilization, measured as numbers of all hospitalisations
Ryan, 2009[100]	C	There was no comparator or control arm...Looked at Hospital Care, a pay for reporting scheme for hospitals who wish to treat patients under the Medicare payment scheme. However no control or comparator was used to assess the effectiveness of the scheme.

Shen, 2003[101]	I	The intervention was not financial....Refers to the effects of financial pressures measured by HMO penetration and prospective payment levels, rather than a comparison of payment methods

4.7 Included studies

Details of the included studies are summarised in Table 4-2. Composite measures refer to any qualitative outcome or outcomes composed from other outcome measures.

Table 4-2 CHD Physician Payment Included Studies Details

	Carlisle	Glickman	Luft	Paone	Petersen	Sada	Starr	Young
INTERVENTION INCENTIVE								
FFS			√			√		
Salaried				√	√		√	
P4P		√						
Mixed	√							
Capitation								√
COMPARISON INCENTIVE								
FFS	√			√	√		√	√
Mixed			√			√		
Non P4P		√						
Capitation						√		
Uninsured						√		
OUTCOME MEASURES								
In hospital mortality		√		√		√	√	√
30 day mortality measures	√		√					√
1 year and above mortality measures	√				√			

Composite measures	√	√						
Length of stay measures				√				
Angiography rates						√		√

Four, half of the included studies are direct comparisons between HMOs (mixed and salaried) and FFS. Of the remaining half, one looks at P4P, one compares FFS to Salaried, one capitation to FFS, and the remaining study compares a number of physician incentives. Further details on specific studies are set out in the tables below. All of the studies used routinely or pre collected data to analyse the effects of financial interventions, none collected data specifically for their study:

Carlisle et al, 1992[102]

Population	Patients over 65 admitted with a principal diagnosis of Acute Myocardial Infarction (AMI)
Intervention	Mixed- salaried and capitation
Comparator	FFS
Outcome	1. 30 day mortality rate following admission 2. 180 day mortality rate following admission 3. Quality of process score (mean)
Setting	Secondary Care
Context	USA, 1 July 1985-1 July 1986
Data source	Hospital medical records – pre collected data
Notes	1. Three HMOs operating in separate states, were selected from a national consortium of twelve. A final sample fully meeting the selection criteria involving 369 admissions to 27 hospitals were used 2. The FFS sample was taken from a previous study and included 1119 patients drawn from 297 hospitals in five states.

	3. Outcomes were adjusted for sickness at admission. This incorporated items from the Killip and Norris measures of the severity of AMI, the APACHE II severity of illness scale and measures of co-morbidity. 4. The quality of process score is an algorithm incorporating 93 specific process of care measures, reviewed by an expert panel to see if they were clinically acceptable and could be reliably recorded. It includes measures on physician and nurse performance and appropriate use of tests, procedures and pharmaceuticals, as well as intensive care or telemetry.

Glickman et al, 2007[103]

Population	Patients with acute non-ST segment elevation myocardial infarction admitted to 500 CRUSADE reporting hospitals
Intervention	P4P
Comparator	Absence of P4P
Outcome	1. Incentivised composite measure 2. Non incentivised composite 3. In hospital mortality
Setting	Secondary Care
Context	USA, 1 July 2003 and 30 June 2006
Data source	Patients enrolled in the ACC/AHA CRUSADE national quality improvement initiative – pre collected data
Notes	1. 105,383 patients included in the study 2. Can rapid risk stratification of unstable angina patients with early implementation of the American College of Cardiology/American Heart Association (ACC/AHA) guidelines (CRUSADE) 3. Hospital Quality Incentive Demonstration (HQID) is a P4P

	scheme which provided a bonus to hospitals that met Centers for Medicare and Medicaid (CMS) measures across a range of clinical conditions including AMI. Hospitals in the top 20% on these measures received a reimbursement bonus, while poorly performing hospitals risked being excluded from future contracts with Medicare and Medicaid. 4. Incentivised composite based on CMS measures (Aspirin at arrival; aspirin at discharge; ACEI or ARB for LVD; smoking cessation counselling for active or recent smokers; beta blocker at arrival and beta blocker at discharge). Non incentivised (Glycoprotein IIb/IIa inhibitor use; Clopidogrel at discharge; any heparin use; lipid lowering medication; dietary modificiation counselling; referral for cardiac rehabilitation; ECG within 10 minutes; cardiac catheterization within 48 hours)

Luft, 2003[104]

Population	All Medicare beneficiaries (aged over 65) admitted with AMI
Intervention	FFS
Comparator	Mixed-salaried and capitation
Outcome	1. 30 day standardised mortality ratio (SMR)-Model A 2. 30 day standardised mortality ratio (SMR)-Model B
Setting	Secondary Care
Context	California, USA, 1994-96
Data source	California Hospital Outcome Project – pre collected data
Notes	1. Data is taken from hospital submitted discharge abstracts 2. The two models were produced by the UCLA with 30 day mortality as the dependent variable. Model A potentially under compensates for true case mix differences as it only includes

	variables (co-morbidities etc.) known prior to admission. Model B additionally includes variables revealed during admission (e.g. shock, other co-morbidities); potentially over-compensating for risk differences by including variables that may be complications of care.

Paone et al, 1995[105]

Population	Patients who received a Coronary Artery Bypass Graft (CABG)
Intervention	Salaried
Comparator	FFS
Outcome	1. In hospital mortality rate 2. Mean ICU length of stay 3. Total hospital length of stay
Setting	Secondary Care
Context	Henry Ford Hospital, USA, 1 Jan 1990 to 31 Jan 94
Data source	Computerized cardiac surgical database, routinely collected
Notes	1. The Henry Ford hospital is the sole provider of cardiac services to patients in the HAP and provides services to the community at large using alternative health insurance arrangements 2. 569 HAP (all Medicare beneficiaries excluded) compared to 225 FFS patients

Petersen et al, 2003[106]

Population	Men, 65 and over, who had had an AMI within the preceding 8 weeks
Intervention	Salaried
Comparator	FFS

Outcome	1. Angiography (Model 1) 2. Angiography (Model 2) 3. 1 year mortality (Model 1) 4. 1 year mortality (Model 2)
Setting	Secondary Care
Context	USA, Jan 1994 to September 1995
Data source	Various routinely/ pre-collected data sources
Notes	1. To ensure diagnostic and not therapeutic emergency primary angioplasty were compared; only patients who became eligible for angiography more than 12 hours after the onset of symptoms but before discharge were included. 2. Patients considered to have a clinical need for angiography if they were in class I of the ACC-AHA guidelines 3. Model 1 adjusted for age, race, BMI, presence of morbidities and other risk factors. Model 2 adjusted additionally for the availability of angiography and cardiac surgery on site 4. Medicare records collected from 7 US states for male patients discharged between 1 Feb 1994 and 30 July 1995. 5. Veterans Affairs (VA) male patients discharge from non-psychiatric VA hospitals between 1 Jan 1994 and 30 Sept 1995 6. 1665 males from 81 facilities (VA population). 19, 305 males patients from 1530 non-federal acute care hospitals (Medicare)

Sada et al, 1998[107]

Population	Patients under 65 enrolled in the National Registry of Myocardial Infarction (NRMI)
Intervention	FFS

Comparator	Mixed
Outcome	1. Non-discretionary angiography rates in high risk patients 2. In hospital mortality
Setting	Secondary Care
Context	USA, June 1994 to October 1995
Data source	NMRI: Pre- collected data from 1482 hospitals (26% of all medical/surgical hospitals)
Notes	1. The study population totalled 17,600 patients; 10,498 FFS (59.6%), 3,273 HMO (18.6%), 1,354 Medicaid (7.7%) and 2,475 uninsured patients (14%). 2. Uninsured were those without identifiable insurance 3. Whether treatment was deemed discretionary/non-discretionary and patients were at a high/low risk was based on ACC/AHA guidelines

Starr et al, 1996[108]

Population	Patients admitted for CABG surgery at St Vincent Hospital and Medical Centre
Intervention	FFS
Comparator	Salaried
Outcome	1. Operative mortality
Setting	Secondary Care
Context	Oregon, USA, January 1985 to March 1994
Data source	St Vincent hospital database and Keiser Permanente – routinely collected

Notes	1. The final regression model included 3686 (43%) of the 8483 qualifying patients. No significant differences were observed between those excluded (based on one or more missing risk factors) and those included

Young and Cohen, 1992[109]

Population	Patients who had an emergency admission, with a principal diagnosis of AMI.
Intervention	Capitation
Comparator	FFS
Outcome	1. Inpatient mortality 2. 30 day post admission mortality 3. 30 day post discharge mortality 4. Arteriography 5. CABG 6. Angioplasty
Setting	Secondary Care
Context	Massachusetts, USA, 1987
Data source	Various routinely/ pre- collected data sources
Notes	1. Diagnosis based on ICD 9 codes 410-410.9 2. The study population contained 4,033 patients; 3755 privately insured; 278 Medicaid 3. Results adjusted for differences in clinical and demographic characteristics and type of hospital where treatment occurred

4.7.1 Quality assessment of included studies

A reduced form of the Downs and Black ranking scale was used to assess the quality of the included studies[110]. This particular ranking scale was chosen because it is

widely used, seen as a gold standard and was comparable to or better than equivalent ranking scales[111]. The Downs and Black rating scale contains 27 questions in 3 categories; Reporting; External validity; and Internal validity. Questions were chosen that reflected the nature of the studies being evaluated, to ensure any deficiencies evident would be due to poor relative study design. Consequently questions relating to blinding were omitted as the studies were interested in how physicians responded to incentives therefore blinding was impractical. Likewise questions relating to compliance with the intervention were also not relevant as compliance with a financial incentive is assumed. The checklists which were retained and deemed suitable were in the areas of study design, reporting, controlling for confounders and potential for bias. While the intervention was directed at the physician the concern was with how this impacted on patient outcomes. Hence when assessing outcomes and judging study quality, the concern was to ensure patients were comparable between the intervention and control or comparator arms. The assumption in this review and the published literature is that physicians as a group are broadly similar in both arms and hence any difference in outcomes is due to responses to incentives and payment mechanisms by the physician, and not difference in their clinical knowledge, training or practices. Details regarding the ten criteria are summarised below as interpreted in relation to the studies evaluated; with further details on the Down and Black checklist provided in Appendix 5. Table 4-3 details how each of the included studies performed in relation to these. The author column refers to the principal author.

1. Are the main outcomes to be measured clearly described in the Introduction or Methods section?
2. Are the characteristics of the patients included in the study clearly described?
3. Are the main findings of the study clearly described?
4. Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?
5. Is the intervention and control group followed up for the same length of time?
6. Were the statistical tests used to assess the main outcomes appropriate?
7. Were the main outcome measures used accurate (valid and reliable)?

8. Were patients in the intervention and control arm selected from the same hospital, comparable care settings, joint registry or reporting schemes?
9. Were the study periods in the intervention and control group matched?
10. Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?

Table 4-3 Down and Black Checklist

Author	1	2	3	4	5	6	7	8	9	10
Carlisle	√	√	√		√	√	√	√	√	
Glickman	√	√	√	√	√	√	√	√	√	√
Luft					√	√	√	√	√	√
Paone	√	√	√	√	√	√	√	√	√	√
Petersen	√	√	√		√	√	√			√
Sada	√	√	√	√	√	√	√	√	√	√
Starr	√	√	√	√	√	√	√	√	√	√
Young	√	√	√		√	√	√	√	√	√

4.7.2 Quality assessment discussion

Ticks show that the checklist criteria were met, a score of 1 on the ranking criteria; while shading shows it was missed, a zero score. Generally the studies performed well on these measures and were well conducted observational studies, with a clear description of data sources, patient selection criteria, analyses and results, with appropriate adjustments for confounders. Nonetheless there were significant concerns about two studies, namely Luft and Petersen[104 106]. Both of these failed to give a full breakdown of the p values for their main outcome measures. In the study by Petersen they are not reported at all; in Luft they are reported for certain significance levels. The Luft article performs poorly on all reporting measures[104]. The main outcomes are not described until the results section. There are no output tables, results are presented within text and are difficult to extract and relate at times.

Nor is there a breakdown of the key patient characteristics in the intervention and control bar general details in the text, though reference is made to earlier work where this information may be available. Petersen performs poorly on internal validity as the study periods in the intervention and control group differ, albeit by only 3 months. It is also not certain if the intervention and control group are matched sufficiently, as the VA study sample used national data after the number of VA hospitals in the 7 states from which the Medicare population was selected, was found to be too small to provide adequate analytic power.

Outside of these studies both Young and Carlisle fail to give a full breakdown of p values for their results, though do report them at the 1% and 5% significance levels[109]. More seriously in the case of Carlisle it is unclear whether the results are adjusted for the significantly higher number of females in the FFS study arm[102]. It is possible that differences in risks between the genders could be captured in the adjustments for sickness at admission but it is not possible to be certain on this.

4.8 Results

All eight studies are from the USA and secondary, hospital care. None linked primary care incentives to clinical outcomes, and all used administrative databases. Results from the studies are shown in Table 4-4, and are broken down by clinical measures and forms of reimbursement.

Table 4-4: Included studies, physician incentives CHD, results

SURGERY				
Study	**Outcome**	**Intervention**	**Comparator**	**Results**
Petersen	Clinically needed Angiography	Salaried	FFS	Adjusted Odds Ratio Model 1: 0.75 (0.57-0.96) Model 2: 1.02 (0.82-1.26)
Sada	Angiography: High risk patients, non-discretionary	FFS	Capitation	OR 3.13 (95% CI, 2.80-3.85)#
Sada	Angiography: High risk patients, non-discretionary	FFS	Mixed	OR 1.45 (95% CI, 1.19-1.75)#
Sada	Angiography: High risk patients, non-discretionary	FFS	Uninsured	OR 2.08 (95% CI, 1.69-2.52)#
Young	Arteriography utilisation	Capitation	FFS	OR 0.38 (95% CI, 0.23-0.78)
Young	CABG utilisation	Capitation	FFS	OR 0.27 (95% CI, 0.21-0.65)
Young	Angioplasty utilisation	Capitation	FFS	OR 0.28 (95% CI, 0.24-0.63)
MORTALITY				
Carlisle	Mortality: 30 day post admission	Mixed	FFS	Mixed: 23.8%; FFS: 24.2% ; p>0.05
Carlisle	Mortality: 180 day following admission	Mixed	FFS	Mixed: 35.3%; FFS: 35.4%; p>0.05
Glickman	Mortality: In hospital	P4P	Non P4P	P4P: 4.87% (2003); 3.93% (2006); OR 0.91 (95% CI

				0.84-0.99) Control: 5.03% (2003); 4.06% (2006); OR 0.97 (95% CI 0.94-0.99) P (comparison) = 0.21
Luft	Standardised Mortality Rate	FFS	Mixed	Model A: 1.047 (p<0.01) Model B: 1.018 (p, non-significant)
Paone	Mortality: In hospital	Salaried	FFS	Salaried 1.9%; FFS 2.2%; p=0.794
Petersen	Mortality: 1 year	Salaried	FFS	Adjusted Odds Ratio Model 1: 1.10 (0.92-1.33) Model 2: 1.08 (0.89-1.28)
Sada	Mortality: In hospital	Capitation	FFS	OR 1.55 (95% CI, 1.19-2.01)* OR 1.40 (95% CI, 1.04-1.87)±
Starr	Operative Mortality	FFS	Salaried	p=0.059ǂ
Young	Mortality: In hospital	Capitation	FFS	OR 1.91 (1.28-2.95)
Young	Mortality: 30 day post admission	Capitation	FFS	OR 1.92 (95% CI, 1.33-2.95)
Young	Mortality: 30 day post discharge	Capitation	FFS	OR 1.90 (95% CI, 1.31-2.89)
OTHER MEASURES				
Carlisle	Quality of process score	Mixed	FFS	Mixed: 0.54; FFS: 0.23; P<0.01
Glickman	Incentivised composite	P4P	Non P4P	P4P: 87% (2003); 94.2% (2006); OR (95% CI) 1.23 (1.15-1.30) Control: 88% (2003); 93.6% (2006); OR (95% CI)

				1.17 (1.14-1.20) P (comparison) = 0.16
Glickman	Non incentivised composite	P4P	Non P4P	P4P: 59.5% (2003); 73.1% (2006); OR (95% CI) 1.09 (1.05-1.14) Control: 60.5% (2003); 68.6% (2006); OR (95% CI) 1.08 (1.06-1.09) P (comparison) = 0.49
Paone	Mean ICU stay	Salaried	FFS	Salaried: 2.6 +/-0.3 days FFS: 2.3 +/- 0.3 days p=0.734
Paone	Total length of hospital stay	Salaried	FFS	Salaried: 9.8 +/- 0.8 days FFS: 8.6 +/- 0.6 days p=0.911

*After adjustment of model for variables present on admission

± After adjustment of the model for differences in treatments

ǂ FFS as a risk factor # Adjusting for different hospital and demographic characteristics between groups

OR= Odds Ratio

As the focus was on the effect of incentives on clinical outcomes, measures non clinical such as length of stay are only reported where these were additionally included. Likewise studies which look at surgical procedures were only included if clinical endpoints were also included, and vice versa. For all of these measures there is a potential for bias, as reporting is selective and hence results may not be generalisable. It is worth noting that issues such as this are not limited to this review and reflect issues with research in this area, some of which were identified in Chapter 2.

4.8.1 Surgical procedures

Petersen et al find significantly lower odds of clinically needed angiography, based on the patients being in class 1 of the ACC-AHA guideline in VA patient treated under a salaried system compared to FFS, OR 0.75, 95% CI 0.57-0.96, when adjusting for patient risk factors only[106]. However when this figure is adjusted for the on-site availability of cardiac procedures (Model 2) there is no significant difference, OR 1.02, 95% CI 0.82-1.26.

Sada looks at non-discretionary use of angiography in patients at high risk of future cardiac events patients, specified as those in whom it would be most likely to confer a survival benefit, which was lower in all other forms of health payments relative to FFS[107]. High risk FFS patients were found to have over 3 times higher odds of receiving non-discretionary angiography than high risk Medicaid patients where physicians were paid by capitation (OR 3.13, 95% CI, 2.80-3.85). Sada also finds that high risk FFS patients have 45% higher odds than their HMO counterparts (mixed incentives) and more than twice the odds of equivalent uninsured patients to receive non-discretionary angiography: HMO (OR 1.45. 95% CI, 1.19-1.75); Uninsured OR (2.08. 95% CI, 1.69-2.52).

Young and Cohen consider the number of arteriography, CABG and angioplasty procedures[109]. No distinction is made between clinically necessary and unnecessary interventions, although all patients were emergency admissions suggesting their needs would be high. They found much lower odds of utilisation for all three procedures in patients covered by capitation compared to FFS; Arteriography, OR 0.38, 95% CI, 0.23-0.78; CABG, OR 0.27, 95% CI, 0.21-0.65; Angioplasty, OR 0.28, 95% CI, 0.24-0.63.

4.8.2 Mortality

While HMOs comprising a mix of salaried and capitation generally had lower rates of mortality than FFS the differences were largely insignificant[102 104 105 108]. Only in the study by Luft and then in his under specified model did mixed incentive HMO patients have significantly lower rates of mortality ($p \leq 0.05$)[104]. There are two studies that compare capitation in the Medicaid system against FFS[107 109]. In both studies the odds of death is significantly higher for patients in the capitation system, by 40% or 55%, depending on what adjustments are made, in the study by Sada[107], and around 90% on the 3 measures used by Young[109]. Petersen compares the salaried Veterans Affairs health care system against FFS Medicare patients[106]. The odds ratios adjusted for risk factors show a higher rate of death within 1 year of an AMI in VA patients, Model 1, 1.10 (0.92-1.33); which falls slightly when adjusted for access to onsite cardiac surgery and angiography, Model 2, 1.08 (0.89-1.28). However in both instances confidence intervals are wide and cross 1, meaning that the results are not statistically significant.

Glickman looks at the before and after effects of the introduction of P4P in an incentivised and non-incentivised group. This showed that mortality fell in both groups over the study period, OR P4P 0.91 (0.84-0.99), OR Control 0.97 (0.94-0.99), a level of change that was not significantly different between the two groups, $p=0.21$[103].

4.8.3 Other measures

The studies by Glickman and Carlisle looked at measures of quality of patient care which related to following correct care protocols. In the paper by Carlisle a mixed salaried and capitation HMO group performed significantly better than FFS on a quality of process score (0.54 vs 0.23, $p<0.01$) which included 93 separate measures of care based on clinician performance, and the appropriate use of tests, facilities, procedures and drugs[102]. Glickman considers non incentivised and incentivised targets in the presence and absence of P4P[103]. For the incentivised targets, performance over the study period improved in the control group, from 88% to 93.6%, but by a greater amount in hospitals participating in P4P, from 87% to 94.2%. There was however no statistically significant change in the rate of change between the intervention and control ($p=0.16$). In terms of non-incentivised targets

the results were the same with the P4P group enjoying a bigger increase relative to the control but again the difference between the two was non-significant (p=0.49).

Paone finds a slightly higher mean ICU stay (2.6 days vs 2.3 days) and total length of hospital stay (9.8 days vs 8.6 days) in patients from a salaried HMO compared to FFS, though neither result was significant, p=0.794 and 0.734 respectively[105].

4.9 Discussion

Overall the literature reveals the complexities of disaggregating effects and drawing lessons in this area of health care research. Not only is the researcher faced with the challenge of dealing with patient differences which are inherent in the different forms of service delivery but there are different incentives operating at the organisational and the individual physician level. Nonetheless there are still a number of significant findings and general results than can be drawn. These are broken down into findings, limitations, lessons for practice, and future research recommendations

4.9.1 Findings

The literature on the effects of incentives on CHD clinical outcomes was all USA based and looked at Secondary care. No studies were found that linked primary care incentives into CHD clinical outcomes as proposed by this thesis. Only one study was found that looked at P4P. This looked at similar targets to those selected for this study but as a composite measure and in patients admitted to hospital with acute non ST MI. This found no statistically significant impact of P4P on incentivised measures, non-incentivised measures, or in hospital mortality. Hence while P4P was ineffective in this context, it did not lead to resource substitution. In addition to these principal findings a number of incidental findings were made, which did not relate to the research question but are of wider interest.

In line with economic theory FFS tended to lead to greater resource utilisation, or conversely alternative payment methods which do not pay by activity lead to lower, or under, resource utilisation. However there is evidence that access to services, and greater physician and patient choice may be partly responsible for this result rather than it being wholly a physician response to FFS to induce demand. Individual studies themselves factored this into their outcome measures by only looking at clinically necessary angiography, so largely discounting any unnecessary surgery

from their results. Additionally the study by Petersen showed that once the results were adjusted for the availability of onsite cardiac facilities, VA patients had slightly higher odds of receiving clinically needed angiography than FFS, though the result was not significant[106]. The VA healthcare system concentrates care in large regionalised hospitals similar to teaching hospitals in the UK. In comparison FFS offers both the physician and patient greater choice and access to services.

Big discrepancies in the utilisation of clinically necessary surgery in the Medicaid capitation system and among the uninsured, compared to FFS, were associated with significantly higher mortality rates. However despite lower surgical utilisation in all forms of HMOs, there were no significant differences in mortality rates between FFS and HMOs. This suggests that there is some level of surgical intervention that while clinical necessary had no impact on mortality. It is possible that these patients had some quality of life benefit, and surgery therefore conferred some benefit which was not picked up in mortality. Or conversely that HMO's, by presenting the physician with greater resource constraints, made them more judicious in their use of surgery. Hence it may have avoided over investigation and the use of invasive surgery in patients where surgery was harmful or the benefits were marginal. Medicaid appears to place too many resource constraints on physician's and access restrictions on patients leading to significantly worse outcomes, with FFS perhaps placing too little, leading to unnecessary investigation and intervention. HMOs may represent a compromise between the two which does not excessively constrain physician discretion, but still restricts the incentive to intervene and induce demand.

A final finding leading on from this is that overall HMOs performed much better than their resource constrained counterparts and at worst performed as good as but generally better than FFS. These have an element of limited choice and cost controls found in the Medicaid capitation funded systems but provision of care is far less restricted. They are similar to an NHS model of care, with physician's providing a primary and secondary care interface and managing access to Secondary care. However there is far greater competition in this system than in the NHS and not the clear separation of primary and secondary care. With care in both sectors provided under these plans there is the incentive for both to work together to ensure efficiencies and be more innovative.

4.9.2 Limitations

There are a number of limitations in the included literature. Firstly all the studies examine routinely or pre collected data, none use data specifically generated for the purposes of their study. This presents possible data quality issues and also increases the potential that findings show correlation and not causality and that data was used to fit rather than determine results. With the exception of Starr who examines the effects of the intervention over a nine year period, all of the studies examine the effects of the intervention over a relatively short study period. Clinically significant endpoints such as mortality are also measured over short time spans, of no more than 1 year, usually less, when surgery in particular may provide benefits that extend to periods way beyond that. All the studies look at Secondary care only; none looked at the effect of physician incentives in primary care on clinical outcomes. With the exception of Glickman none considered the impact of P4P, or looked at surrogate or prescribing measures which have relevance to the QOF. However this was done so in the context of hospital care, in patient's admitted with AMI.

4.10 Implications for practice

In both this Chapter and in Chapter 2, no evidence has been uncovered for the clinical benefits of P4P. Hence despite the increasing preference for it, and significant invest in it in UK primary care, there appears to be little evidence supporting it, whether we look at process measures or clinically significant outcomes.

Resource constraints which place excessive restrictions on access to services do have significant impacts on clinical endpoints. After adjusting for socio-demographic factors Medicaid patients were significantly more likely to die and significantly less likely to receive clinically necessary cardiac surgery. This could be due to the managed care programmes which these programmes use placing excessive gate keeping demands on physicians undermining their ability to treat; and co-payments deterring this economically deprived group from seeking primary preventative care, leading to poor secondary care outcomes.

Some form of physician gate keeping role, resource limitations and competition to contain costs while maintaining quality would appear to be necessary to limit over investigation and treatment. These are characteristics of the HMO managed care

system that may account for the fact that it secured at worst as good clinical endpoints for patients despite fewer surgical interventions.

4.11 Implications for future research

This review has focused on clinical outcomes. FFS not only place lower resource constraints on physicians but also offer patients the greatest access to services. An effect of this may be greater demands by patients for services including surgery. This works suggests this does not always impact on longevity, but there may be quality of life benefits not considered, which patients who have the means are willing to pay for.

With only one study identified on P4P there is a clear need for research to assess its effectiveness on clinical outcomes, certainly in the area of CHD care. With Glickman only following up the intervention for 3 years, future research would benefit from using a longer period of follow up to determine if effects are related to duration of the intervention.

Research is needed which spans primary and secondary care to look at the benefits of integrated working between the two. This should seek to determine if improved or more intensive condition management in primary care impacts on secondary care activity and clinical outcomes.

4.12 Summary

The focus of Chapter 2 was on the effect of physician payment in Primary care. This review relaxed the study design criteria in that review, extended the search for literature into Secondary care, while limiting the focus to CHD. As a result this thesis has considered the impact of physician payment through primary and into secondary care and clinical outcomes, the latter albeit for CHD only. This is important for this research as it is what it intends to do for the QOF. The logic being that a focus on process measures and arguably surrogate outcomes misses the real concern of whether targets and changes in physician incentives impact on important clinical outcomes. While no single piece of research has been found which does this in the area of CHD care, this is important, as it shows a clear need for the work being undertaken in this thesis. The QOF represent a significant investment in primary care. Most significantly in a global budgeted service it consumes resources that could otherwise be used in other areas of the health service. Hence it is important

that there is evidence that investment in the QOF have led to improved clinical outcomes and consequently fewer demands being placed on secondary care. Following the Health and Social Care Act (2012) primary care physicians now find themselves in control of even greater shares of health expenditure and responsible for purchasing services. Consequently it is more important than ever to demonstrate that allocating greater resources to primary care leads to positive clinical outcomes to justify this change.

4.13 Conclusion

This review has revealed a paucity of work on the effect of incentives on clinical outcomes, and consequently the need for further research. There was no evidence found for the clinical benefits of P4P, and hence justification for the choice of this means in delivering the QOF over alternatives.

In terms of this thesis it has highlighted the need for further research which considers the effect of incentives in primary care on evidence linked hospital admissions. The next chapter will examine P4P from a UK perspective, and examine one of the biggest financial investments in P4P in the world, the QOF. This will seek to find research on the QOF and CHD to determine how its effectiveness has been demonstrated and reported.

Chapter 5 CHD in the QOF

5.1 Objective

To determine what research exists on CHD in the QOF, what it shows, and what this research contributes to it. This will involve:

1. A structured literature search to identify the literature looking at the impact of QOF CHD targets.
2. Evaluating the research methods and outcome measures used.
3. Appraising the research in terms of what has been found and where it is deficient

5.2 Background

Chapter 2 examined the systematic review evidence for the effectiveness of different forms of physician payment in primary care. In Chapter 4 the focus moved to CHD and the effect of different forms of physician incentive on clinical outcomes. To date due to the inclusion criteria imposed, this thesis has not looked at the effects of the QOF on CHD outcomes. This chapter sets out to address this and explore what evidence has been found for the effectiveness of the QOF on CHD care. This will complete the review of the effects of incentives on patient care. The specific question for this review is how has the QOF impacted on CHD care? Answering this question will reveal what type of research has been undertaken, how the QOF's impact has been measured and what this research will contribute to existing knowledge in this field.

5.3 Introduction

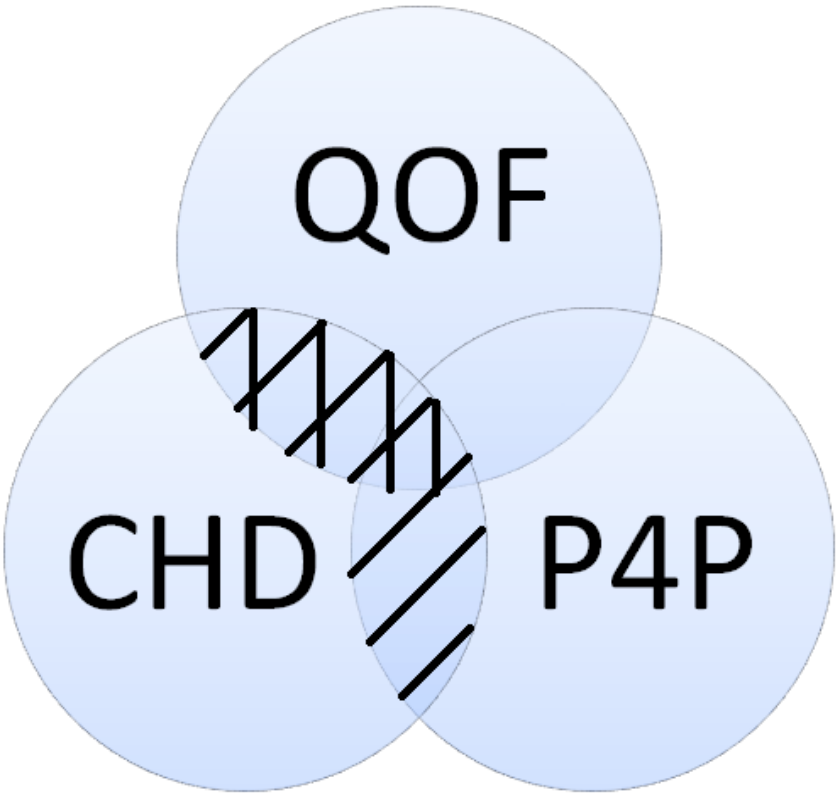
No pilot studies were published prior to the QOF so there is no body of evidence supporting its introduction. With near universal adoption there is no control group to evaluate its effectiveness. All of this makes research on the QOF challenging, and one where the choice over study design will be limited. While the QOF has been in place for nearly a decade, increasing the scope for research into its effects, it is still an area of research which is relatively young and evolving. In order to inform and place this research it was important to uncover what research had been published on QOF CHD care to date. Given the nature of the intervention, and in order to capture the literature fully, no stipulations were placed on study design

5.4 Inclusion criteria

1. Population: Patients diagnosed with CHD
2. Population: Where patients were treated under other QOF conditions, the results for CHD were reported separately.
3. Intervention: QOF P4P scheme.
4. Comparators: Had to involve a minimum of two time points at least one of which had to be post the QOF's introduction.
5. Outcomes: Had to measure performance on QOF targets, measures that mapped to QOF targets, or clinical outcomes.
6. Outcomes: Had to consider the impact of the QOF at the patient level.
7. Outcomes: Where composite measures are used they need to be composed in their entirety of QOF targets

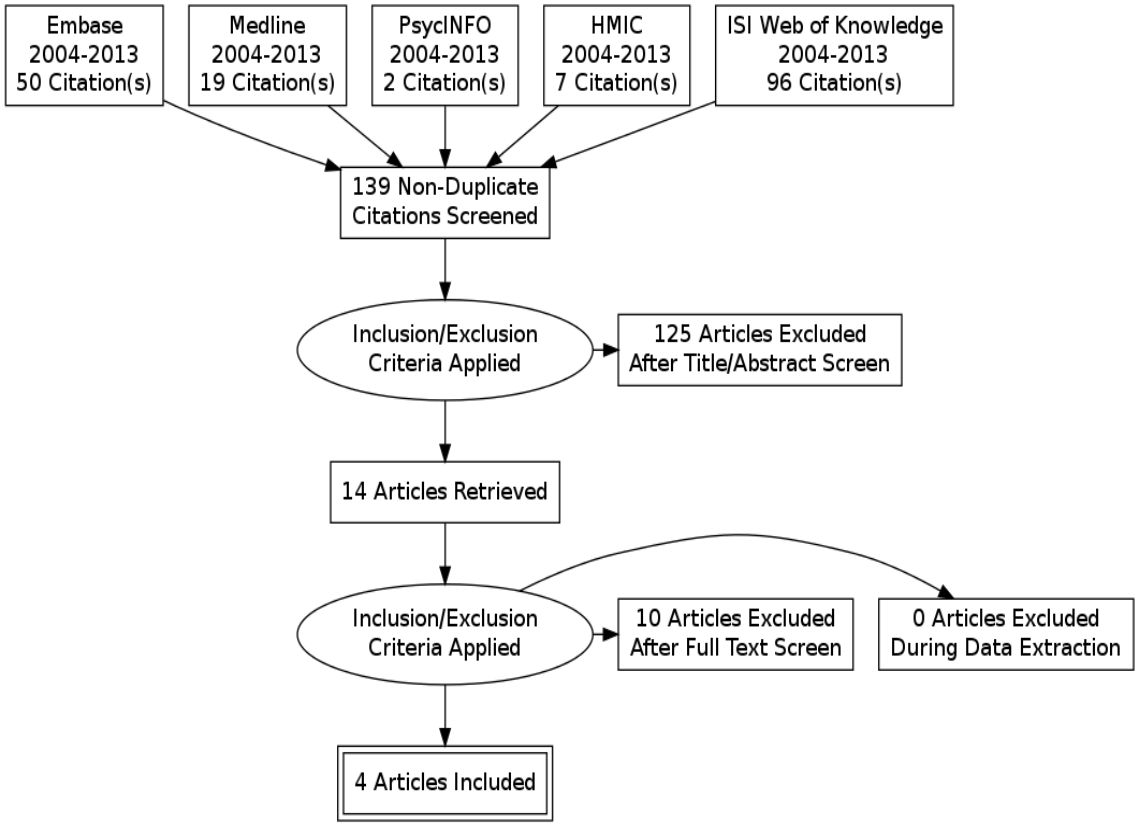
The search used the OvidSP and ISI Web of Knowledge search platforms. Within OvidSP the following databases were searched: Embase, Medline, HMIC and PyscINFO. Details of the full search terms used for each database are detailed in Appendix 6, and summarised in the Venn diagram below.

Figure 5-1: CHD QOF Search terms



The review was interested in articles which looked at the QOF in relation to CHD, so primarily with articles which contained these two search terms. Pay for performance (P4P) is the main descriptive term used to describe physician incentive schemes which pay by result or outcome, of which the QOF is an example. Hence this was also included as a search term in order to ensure any study on the QOF, which included pay for performance in preference, and in the absence of QOF, in its search fields, was also included. Whether or not such a study had a QOF focus was determined at the screening stage. Hence while all studies falling within the overlaps shown with shading were returned by the search, only those in the cross hatched area in Figure 5-1, were included.

Figure 5-2: CHD QOF PRISMA diagram



The search yielded 174 published articles, of which 35 were duplicates. Title and abstract screening of the non-duplicate articles using the exclusion and inclusion criteria led to the removal of 125 of those articles. Fourteen full journal articles were retrieved and reviewed. A further ten were removed at the full article review phase, with four considered to meet the inclusion criteria in full.

5.5 Articles excluded after full review

Table 5-1 outlines the reasons behind the exclusion of studies at the full text review stage, as well as which PICO criteria the study failed to meet.

Table 5-1: CHD QOF, studies excluded at full text review

Author(s)	PICO	Description of study and reasons for exclusion
Bottle, 2008[112]	P	Patients' CHD status uncertain..... Regression analysis is used to look at the relationship between QOF CHD targets and secondary care outcomes measured in HES. However it is not linking those events directly to CHD patients rather to all patients in the PCT adjusting for QOF CHD scores and other variables.
Campbell, 2007[113]	O	The outcome measure used is a composite measure that included non QOF CHD targets.....This composite mean score measure is calculated in 1998 and 2003 and the trend growth rate between those two points is used to predict attainment in 2005. Actual attainment in 2005 then determines how the QOF affected attainment relative to a predicted trend. However the majority of the indicators included in the composite score are not present in the QOF; and in instances where the indicator matches the QOF target, very often the measurement qualifying period differs from that in the QOF
Campbell, 2009[12]	O	The outcome measure used is a composite measure that included non QOF targets.....This study builds upon the study reported above following access to new data for 2007. This is used to compare trend growth rates between 2005 and 2007 with that between 1998 and 2000.
Crawley, 2009[114]	P	Patients' CHD status not clinically verified.....Uses Health Survey for England data to look at the effect of occupational class on QOF attainment. Patients were selected if they

		declared a doctor confirmed diagnosis of CHD, however this was not medically verified.
Kiran, 2010[115]	P	Patients' CHD status uncertain.....Relates CHD hospital admissions and deaths to a CHD quality measure which includes Stroke, Hypertension, Diabetes and Smoking domain QOF targets in addition to CHD at a practice level. It was not possible to be certain that the patients admitted had CHD or were QOF CHD qualifying
Levene, 2010[116]	P	Patients' CHD status uncertain.....The article is concerned with CHD mortality for each year of a three year period but does not make QOF CHD care its main focus. Rather it examines those rates in relation to the proportion of maximum points achieved in CHD, diabetes, hypertension, blood pressure, and serum cholesterol control; as well as a host of other practice characteristics. In final models CHD is dropped entirely as a predictor for CHD admissions.
McLean, 2007[117]	O	Outcomes are not measured at the patient level.....Looks at the effect of practice location in terms of distance to urban settlement on attainment on QOF targets. Potential wider patient level factors behind the relationship are not explored, nor any adjustments made, meaning the relationship is merely a practice level one.
Purdy, 2010[13]	P I	Uncertainty over whether patient had CHD and was treated under the QOF incentive scheme.....The study uses HES data to look at emergency hospital admissions for MI or Angina, over a 12 month QOF period, 2004/05. It then runs a model with these as the dependent variable, with a number of explanatory variables including national CHD prevalence and QOF targets score. The study is not looking at the QOF or CHD specifically as it is possible that the patients in HES

		neither had CHD, nor were treated by the QOF.
Ryan, 2011	O	Outcomes were not measured at the patient level...Aggregates QOF CHD targets into two composite measures; surrogate outcomes and process measures. The relationship between improvements in attainment on the composite process and surrogate outcome measures is then explored. This is however conducted at an English national level and not the patient level.
Strong, 2006[118]	P O	Outcomes were analysed at a whole practice level and the study does not focus on CHD patients..... The study looks at the correlation between CHD QOF target scores and deprivation. As deprivation is measured at the practice level, there is no direct link being made between deprivation within CHD patients and their QOF scores. Nor can this be disaggregated from the final results.

5.6 Articles included after full review

Two of the studies focus purely on blood pressure and look at systolic and diastolic blood pressure levels rather than attainment to the QOF threshold[119 120]. Both of these use data from the same source, namely family practices in Wandsworth PCT, London, as does the study by Millet[121]. The remaining study collects its data from Scottish practices who participated in SPICE, a clinical effectiveness programme developed by the Royal College of General Practitioners (Scotland), and whom used the General Practice Administrative Software System (GPASS)[122]. All of the studies which used data provided by Wandsworth PCT, conducted sub group analysis by ethnicity, with Murray performing additional subgroup analysis by gender[120]. In the study by Millet only sub group results are reported and whole population figures are not[121]. Furthermore, Millett uses a 140/80 threshold for BP control, which is more exacting than the threshold used in the QOF. Three of the four included studies compare just two time points in their results[120-122], though one of these had data to enable them to compare substantially more[120]. The remaining study compares trend rates in a QOF period against those in a pre QOF period[119]. None of the studies used

data specifically generated for the purposes of their study. Further details on each of the four included studies are provided below:

Lee et al, 2011[119]

Population	Patients over 18, diagnosed with CHD, Stroke or Hypertension registered with 29 family practices in Wandsworth in 2007
Study design	Retrospective cohort
Method of Analysis	Segmented regression analysis of interrupted time series
Study period	2000-2007
Subgroup(s)	Ethnicity: White, Black and Asian
Outcome(s)	1. Mean Systolic BP change 2. Mean Diastolic BP change 3. Mean Cholesterol level change
Limitations	• A small sample size of 1,753 CHD patients, meant that some of the results did not achieve statistical due to a lack of power
Notes	• Unbalanced panel data, collected retrospectively from patients registered in 2007

McGovern et al, 2008[122]

Population	Patients with a CHD Read code selected from 310 practices in Scotland who used the GPASS administrative system and participated in SPICE
Study design	Serial cross section
Method of Analysis	Binary logistic regression model
Study period	March 2004 & March 2005
Sub group(s)	Female, Age 75+, Most deprived quintile: compared to Males, Aged 65 and below, least deprived quintile, respectively
Outcome(s)	1. Angina patient exercise test/specialist assessment 2. Smoking status recorded 3. Smokers cessation advice 4. Blood pressure recorded

	5. Blood pressure 150/90 or less 6. Cholesterol recorded 7. Cholesterol≤5mmol/l 8. Anti-platelet or anti-coagulant therapy recorded 9. Blocker therapy recorded 10. ACE inhibitor recorded 11. Influenza vaccination recorded
Limitations	• There is a near 50% rise in the CHD population in March 2005 compared to 2004 despite a slight fall in the database population, which cannot be explained entirely as a QOF incentive effect to better record and diagnose.
Notes	• **S**cottish **P**rogramme for **I**mproving **C**linical **E**ffectiveness is part of a clinical effectiveness programme developed by the Royal College of General Practitioners (Scotland) • **G**eneral **P**ractice **A**dministrative **S**oftware **S**ystem is used by approximately 36% of all practices in Scotland

Millett et al, 2009[121]

Population	Patients with CHD registered with 32 general practices in Wandsworth, London.
Study design	Cross sectional survey
Method of Analysis	Multivariate logistic regression
Study period	June 2003-Oct 2003 and Nov 2005-Jan 2006
Sub group(s)	Ethnicity: White, Black and South Asian
Outcome(s)	1. Angina patient exercise test/specialist assessment 2. Smoking status recorded 3. Smokers cessation advice 4. Blood pressure recorded 5. Blood pressure 140/80mmHg or less 6. Cholesterol recorded 7. Cholesterol≤5mmol/l 8. Aspirin prescribed 9. Blocker therapy prescribed

	10. ACE inhibitor prescribed
Limitations	• None
Notes	• Odds ratios were adjusted for age, gender, deprivation and practice level clustering

Murray et al, 2010[120]

Population	Patients with CHD selected from 29 General Practices in Wandsworth, London
Study design	Longitudinal
Method of Analysis	Unpaired two tailed t tests of mean values
Study period	1998 and 2007
Sub group(s)	Ethnicity: White; Black and South Asian Gender: Male or Female
Outcome(s)	1. Systolic BP mean difference 2. Diastolic BP mean difference 3. Cholesterol level mean difference
Limitations	• Although the study collects 10 years of data the results only report differences between 1998 and 2007
Notes	• Patient records were retrospectively collected for all patients aged 18 years and above registered on 31st December 2007, for each of the previous 10 years from January 1998 to December 2007.

5.7 Table of results

The results from the included studies are reported in Table 5-2, broken down by group, and target or clinical area. Whole population figures are shown in bold and italics. Significance levels where reported in the studies are included in the table.

Table 5-2: CHD QOF Included study results

Study	Comparators, Metrics used	Group	Target	Result
Lee	Trend Change 2005-07 vs 2000-03 mmHg	***Whole population***	***Systolic BP***	***-0.53 (-1.09, 0.02)***
		Black		0.63 (-1.28, 2.53)
		White		-0.36 (-1.02, 0.29)
		South Asian		-1.77 (-3.05, -0.50)**
	Trend Change 2005-07 vs 2000-03 mmHg	***Whole population***	***Diastolic BP***	***0.32 (0, 0.64)***
		Black		0.83 (-0.24, 1.90)
		White		0.13 (-0.25, 0.51)
		South Asian		0.91 (0.2, 1.63)*
	Trend Change 2005-07 vs 2000-03 mmol/l	***Whole population***	***Cholesterol***	***0.02 (-0.01, 0.05)***
		Black		0.14 (0.04, 0.25)**
		White		0.01 (-0.03, 0.05)
		South Asian		0.01 (-0.06, 0.08)
McGovern	March 2004 v March 2005 Percentage of patients meeting the target	***Whole population***	***Exercise testing ǂ***	***63.9% v 66.2%****
			Smoking status recorded	***69.5% v 95.7%****
			Smoking cessation advice	***81% v 96.2%****
			BP recorded	***75.7% v 97.2%****
			BP≤150/90mmHg	***79.3% v 80%****

			Cholesterol recorded	***44.1% v 85.8%****	
			Cholesterol≤ 5mmol/l	***86.3% v 75.5%****	
			Anti-platelet or coagulant	***65.8% v 90.3%****	
			Beta blocker	***42.6% v 70%****	
			ACE inhibitor ±	***66.4% v 77.9%****	
			Influenza vaccination	***57.4% v 85.5%***	
	March 2004 & March 2005 Odds Ratio	Female[1]	Exercise testing ǂ	0.93 (0.84-1.04)	0.87 (0.81-0.93)
			Smoking status recorded	1.15 (1.09-1.21)	1.06 (0.98-1.15)
			Smoking cessation advice	1.10 (0.99-1.21)	1.19 (1.05-1.34)
McGovern	March 2004 & March 2005 Odds Ratio	Female[1]	BP recorded	0.92 (0.87-0.97)	0.89 (0.82-0.97)
			BP≤ 150/90mmHg	0.84 (0.79-0.89)	0.84 (0.80-0.87)
			Cholesterol recorded	0.87 (0.83-0.92)	0.82 (0.78-0.87)
			Cholesterol≤ 5mmol/l	0.57 (0.53-0.62)	0.54 (0.51-0.56)
			Anti-platelet or coagulant	0.89 (0.85-0.93)	0.73 (0.70-0.77)
			Beta blocker	0.85 (0.81-0.88)	0.81 (0.79-0.84)
			ACE inhibitor ±	0.67 (0.56-0.81)	0.88 (0.79-0.98)
			Influenza vaccination	1 (0.96-1.04)	0.91 (0.88-0.94)
	March 2004 & March 2005 Odds Ratio	Age 75+[2]	Exercise testing ǂ	0.41 (0.35-0.49)	0.37 (0.34-0.40)
			Smoking status recorded	0.54 (0.48-0.60)	0.99 (0.89-1.15)

			Smoking cessation advice	0.56 (0.48-0.66)	0.44 (0.36-0.53)
			BP recorded	0.37 (0.32-0.42)	0.73 (0.65-0.81)
			BP≤ 150/90mmHg	0.90 (0.84-0.97)	0.86 (0.81-0.90)
			Cholesterol recorded	0.45 (0.41-0.49)	0.73 (0.68-0.78)
			Cholesterol≤ 5mmol/l	1.39 (1.22-1.58)	1.19 (1.12-1.26)
			Anti-platelet or coagulant	0.60 (0.55-0.67)	0.97 (0.91-1.03)
			Beta blocker	0.39 (0.36-0.42)	0.53 (0.51-0.56)
			ACE inhibitor ±	0.56 (0.45-0.69)	0.50 (0.43-0.57)
			Influenza vaccination	1.44(1.28-1.61)	2.85(2.67-3.04)
	March 2004 & March 2005 Odds Ratio	Most deprived 20% [3]	Exercise testing ǂ	0.76 (0.57-1.02)	0.84 (0.61-1.14)
			Smoking status recorded	1.04 (0.86-1.26)	0.78 (0.62-0.99)
			Smoking cessation advice	1.35 (0.92-1.97)	0.83 (0.60-1.14)
			BP recorded	0.95 (0.74-1.20)	0.59 (0.45-0.78)
			BP≤ 150/90mmHg	1.05 (0.89-1.24)	1.06 (0.93-1.21)
			Cholesterol recorded	1 (0.77-1.31)	0.79 (0.57-1.08)
			Cholesterol≤ 5mmol/l	1.31 (0.83-2.06)	1.31 (0.98-1.76)
			Anti-platelet or coagulant	1.11 (0.95-1.28)	1.14 (1-1.22)
McGovern	March 2004 & March 2005 Odds Ratio	Most deprived 20% [3]	Beta blocker	0.87 (0.77-0.97)	0.84 (0.76-0.92)
			ACE inhibitor ±	1.64 (1.18-2.28)	1.67 (1.34-2.10)

			Influenza vaccination	0.92 (0.78-1.09) 0.83 (0.74-0.94)
Millett	Adjusted Odds Ratio 2003:2005	White	Smoking status recorded	3.6 (2.5-5.4)
		Black		3.5 (2.2-5.7)
		South Asian		5.3 (3-9.5)
		White	Cholesterol measured	1.6 (1.2-2.1)
		Black		2.1 (0.9-4.9)
		South Asian		2.6 (0.9-7.7)
		White	BP measured	2.9 (2-4.2)
		Black		3.0 (1.1-8.6)
		South Asian		6.3 (2.2-18.3)
		White	Aspirin prescribed	1.6 (1.3-1.9)
		Black		2.1 (1.4-3.3)
		South Asian		1.8 (0.9-3.6)
		White	Beta blocker prescribed	1.3 (1.1-1.5)
		Black		1.3 (1-1.7)
		South Asian		1.2 (0.9-1.7)
		White	ACE inhibitor ±	1.4 (1.1-1.7)
		Black		2.1 (1-4.6)
		South Asian		1.4 (0.7-2.7)

		White	Cholesterol≤5mmol/l	1.5 (1.3-1.8)
		Black		1.7 (1.1-2.5)
		South Asian		1.2 (0.9-1.7)
		White	BP≤140/80mmHg	1.3 (1-1.6)
		Black		1.7 (1-2.7)
		South Asian		1.5 (1-2.2)
Murray	Difference in mean: 1998-2007 mmHg	***All Male patients***	***Systolic BP***	***-6.25 P<0.001***
		All Female patients		***-9.43 P<0.001***
		White Male		-5.69 P<0.001
		White Female		-10.11 P<0.001
		Black Male		-8.00 P=0.120
		Black Female		-9.45 P=0.134
		South Asian Male		-7.11 P=0.049
		South Asian Female		-4.27 P=0.376
		All Male patients	***Diastolic BP***	***-6.73 P<0.001***
		All Female patients		***-6.31 P<0.001***
		White Male		-5.59 P<0.001
		White Female		-6.55 P<0.001
		Black Male		-15.62 P<0.001

		Black Female		-4.77 P=0.196
		South Asian Male		-7.47 P<0.001
		South Asian Female		-7.19 P=0.008
		All Male patients	***Cholesterol***	***-0.87 P<0.001***
		All Female patients		***-0.90 P<0.001***
		White Male		-0.94 P<0.001
		White Female		-0.81 P<0.001
		Black Male		-0.20 P=0.656
		Black Female		-1.08 P=0.007
		South Asian Male		-0.70 P=0.006
		South Asian Female		-1.08 P=0.036
** P≤ 0.01 * P≤ 0.05 [1] Compared to males [2] Compared to under 65s [3] Compared to least deprived 20%				
ǂ Patients with angina ± Patients with a history of myocardial infarction 95% confidence intervals shown in parentheses ( )				
Whole population results shown in bold and italics Subgroup results shown in normal font				

5.8 Included studies discussion

None of the studies reported hard clinical outcomes or endpoints. Studies examined the impact of incentives over a short period, often using as little as two time points. Only one study adjusted results for existing trends and only one used data that was not from an administrative dataset collected by Wandsworth PCT, an inner London borough. In terms of informing this thesis, and answering the research question, it has identified a clear need for the research being undertaken by this thesis. No previous research has examined the effectiveness of QOF targets on evidence linked hospital admissions at the individual patient level, or tentatively researched this area.

Results are discussed in detail by target area below:

5.8.1 Blood Pressure

Lee does not examine the QOF target but looks at the underlying measures, systolic and diastolic blood pressure, examining changes between 2005 and 2007 relative to trend between 2000 and 2003[119]. For the whole population systolic blood pressure was lower than trend by 0.53 mmHg, though this result was not significant at a 95% significance level. For all subgroups with the exception of black ethnicities the decrease was below trend, though only in South Asians was it significantly so, $p \leq 0.01$. For diastolic blood pressure, at the whole population level and for each subgroup, levels were higher than they would have been had they followed the pre QOF trend, though only in South Asians was it significantly so, $p \leq 0.05$.

Murray also looks at the underlying blood pressure measures rather than attainment to QOF thresholds[120]. Though in this instance the comparison is between 2007 and 1998 and the figures are reported for males and females at the population level with subgroup analysis by ethnicity and gender. Systolic blood pressure fell between 1998 and 2007 by 6.25mmHg in males and by 9.43mmHg in females, $p \leq 0.001$ in both instances. Within the different ethnic groups systolic blood pressure was lower in 2007 compared to 1998, though these results were not significant at a 95% confidence level for both male and female black ethnicities and South Asian females. Diastolic blood pressure fell by 6.73mmHg in all male patients and by 6.31mmHg in all female patients between 1998 and 2006, $p \leq 0.001$. It fell in all male and female ethnic groups with a 99% significance level, with the exception of black females, -4.77%, $p=0.196$.

At the population level McGovern finds a large increase in BP recording immediately following the QOF, from 75.7% in March 2004 to 97.2% in March 2005, which was statistically significant, $p \leq 0.05$[122]. In terms of patients meeting the BP≤150/90 target the effect is less dramatic but still significant, rising to 80% from 79.3%, $p \leq 0.05$. When examining sub group populations the odds of a female CHD patient having their blood pressure recorded was around 10% lower than males; and controlled around 15% lower, in both the pre and post QOF study period. Patients aged 75 and above had lower odds of having their blood pressure recorded and controlled in both periods though this improved noticeably in the first year of the QOF for recording but was marginally lower for control, relative to those 65 and under. The most deprived had a 5% lower odds of having their blood pressure recorded prior to the QOF but a 5% higher odds of having it controlled at or below 150/90 relative to the least deprived. While the odds of blood pressure being controlled remained similar in the first year of the QOF the odds for having blood pressure recorded declined noticeably to 41% lower than that in the least deprived.

Millett compares odds ratios for 2003 and 2005 adjusted for age, gender, deprivation and practice level clustering[121]. No whole population level results were provided. The adjusted odds ratio was higher in all ethnic groups for blood pressure recording in 2005 relative to 2003. In all these results the 95% confidence interval did not cross one. For blood pressure control Millett uses a more demanding target than that adopted by the QOF, BP≤140/80. For all ethnic groups the adjusted odds ratio is greater than 1; 1.3 White, 1.7 Black and 1.5 South Asian. However for all groups the 95% confidence includes 1, suggesting no significant difference at a 95% significance level

5.8.3 Cholesterol

At the population level Lee finds that cholesterol levels were higher between 2005 and 2007 then they would have been expected to be if they had followed the trend rate experienced between 2000 and 2003, by 0.02mmol/l[119]. However this result was not significant. In terms of ethnicity subgroups cholesterol levels were higher in 2005 than they would have been had they followed the 2000-2003 trend, though only in the Black ethnic group were they significantly so, 0.14mmol/l, $p \leq 0.01$.

Murray also looks at cholesterol levels rather than the QOF thresholds, comparing mean differences between 1998 and 2007[120]. Mean levels are lower in 2007 for both

males and females, by 0.87mmol/l and 0.90mmol/l respectively, p≤0.001. When these are further broken down into ethnicity subgroups the results are all significant at a 95% significance level with the exception of Black males, -0.2mmol/l, p=0.656.

McGovern finds a large and statistically significant increase in the number of patients with a cholesterol record after the first year of the QOF, 85.8% (March 2005) compared to 44.1% (March 2004), p≤0.05[122]. However he also finds a drop in the percentage having their cholesterol controlled to 5mmol/l or less, from 86.3% in March 2004 to 75.5% in March 2005, p≤0.05. In terms of subgroups females had 13% lower odds of having cholesterol recorded and 43% lower odds of meeting the QOF target in March 2004, relative to males. Both of these figures worsened slightly in the QOF's first year. Over 75's had 55% lower odds of having cholesterol recorded but 39% higher odds of meeting the QOF threshold in March 2004, relative to those aged 65 and below. In the first year of the QOF these figures had risen to 27% lower odds and fallen to 19% higher odds, respectively. For the most deprived the odds of meeting the QOF target were 31% higher in both periods relative to the least deprived. In terms of having a cholesterol level record, there were equal odds in March 2004, but by the end of the QOF's first year the most deprived were 21% less likely to have their cholesterol recorded.

Millett measures cholesterol recording and control to the QOF threshold, comparing adjusted odds of meeting the target in 2003 against 2005[121]. In all ethnic groups the adjusted odds of meeting both targets increased over that period. The 95% confidence intervals around the adjusted odds ratio remained above 1, for the white and black ethnic groups, but for the South Asian ethnic group the 95% confidence interval fell below 1 on both measures, recorded (0.9-7.7), 5mmol/l or less (0.9-1.7).

5.8.4 Smoking

The adjusted odds ratio for having smoking status recorded was over 3 times higher in the white and black ethnic groups and over 7 times higher in South Asians in 2005 compared to 2003, in the study by Millett[121].

McGovern finds that the percentage of patients with a record of smoking status and given smoking cessation advice both climbed above 95% in the year immediately following the QOF's introduction, from 69.5% for recording and 81% for cessation advice in March 2004[122]. Females had higher odds of smoking status recording and smoking cessation advice in both study years. However whereas those odds fell

relative to males for smoking recording from 1.15 to 1.06, they increased for cessation advice from 1.19 to 1.28. Those aged 75 and above saw a large increase in the odds of having smoking status recorded from nearly half that in those aged 65 and below to near equal odds. The odds of being offered smoking cessation advice however fell back from 44% lower to 56% lower. The least deprived went from having 4% higher odds of having smoking status record, though with 95% confidence intervals crossing 1, to having 22% lower odds in the QOF's first year. Smoking cessation also fell back from 35% higher odds, though again with 95% confidence intervals crossing 1, to 22% lower odds

5.8.5 Prescriptions

Both McGovern and Millett look at CHD QOF prescription targets. In the study by McGovern the percentage of the population meeting the QOF prescription targets increased in the first year of the QOF, by over 25% for anti-platelets to 90.3%, near 28% for beta blockers to 70%, and 11.5% to 77.9% for ACE inhibitors[122]. Those aged 75 and above had lower odds of meeting all QOF prescription targets relative to those aged 65 and below, though odds improved to near equal for anti-platelets, to 47% lower for beta blockers but fell further behind for ACE inhibitors to 50% lower odds. Females also had lower odds of receiving QOF prescribed drugs than males in both periods, though this time with odds for anti-platelets and beta-blockers falling, but odds for ACE inhibitors increasing, in the year following the QOF's introduction. The most deprived had higher odds of receiving anti-platelets and ACE inhibitors prior to the QOF, 11% and 64% higher respectively, with both odds increasing marginally in the first year.

Millet finds that patients of all ethnicities had higher adjusted odds of receiving all incentivised drugs in 2005 compared to 2003[121]. However for beta blockers and ACE inhibitors in the non-white population, and aspirin in South Asians, the 95% confidence interval included 1

5.9 Limitations

There are a number of limitations in the included studies which are split down into two broad headings; generalisability and strength of evidence.

Three quarters of the included studies used data from Wandsworth PCT[119-121]. This is an inner city area of South West London, affluent by London standards, urban with

no rurality, with a relatively young and very ethnically diverse population. As such its population is not representative of the UK as a whole. Furthermore the qualifying patient numbers in each study from this region were small, ranging from 913 to just over 3,100. Based on a relatively small sample from a small geographical area of London, and on one PCT, it is difficult to draw service wide lessons from the results. McGovern uses a larger sample, one covering around 36% of practices in Scotland[122]. However it is not clear if these are representative of Scotland practices in general. Nor whether given the size of Scotland relative to the UK as whole and differences in service organisation across the different nations, that these results would be applicable across the whole of the UK. Two of the studies use the measures underlying the QOF rather than the specific QOF targets thresholds[119 120]. While they show what is happening in the measures that map to the QOF targets which one would expect to broadly predict attainment on the targets, they can nonetheless hide significant between patient variation.

In terms of strength of evidence, small sample sizes in most of the studies raises concerns around the statistical power of reported outcomes. This issue is indeed recognised by Lee, who reports that a number of results did not reach statistical significance due to a lack of statistical power. Most of the studies have very limited data points in their analysis. In the case of McGovern, Millett and Murray they are simply comparing two points in time which make it impossible to account for any underlying trends which may be influencing the results additional to a QOF intervention effect. This is particularly noticeable in the study by Murray that looks at absolute differences in mean figures for readings taken in 1998 and 2007. While this includes a period in the QOF the significant time difference between the two dates and the earliest date and the introduction of the QOF in 2004, means that it is difficult to attribute differences to the QOF. It is unclear in the study by Murray why, when they apparently had ten years of longitudinal data, they used these two time points only. Nor why they did not make greater use of their data to examine trends and report further on pre QOF and QOF trends that they refer briefly to.

5.10 Implications for practice

The evidence base on the clinical effects of CHD QOF targets is extremely limited and extends simply to analysis of attainment on the QOF targets, with a reliance on a small number of routinely collected datasets. Analysis rarely extends beyond a

comparison between two points in time measures with little consideration of time trends. As such there is no strong evidence of clinical benefits of either the targets or using P4P to deliver improvements in patient outcome.

5.11 Implications for future research

This research area is significantly underdeveloped and has a number of deficiencies which future research should attempt to address. Specifically there is a need for more ambitious research which extends beyond simply looking at attainment on QOF targets or their mapped measures. This should look at the effectiveness of QOF targets on relevant clinical outcomes, ideally over longer periods which can account for trends. Studies need to be sufficiently powered to pick up any effects, and extend beyond the data sources used by the studies in this review.

5.12 Summary

The objective of this literature review was to determine the impact of the QOF on CHD care and the scope and depth of research undertaken to date. On the whole it has shown an improvement in performance to CHD QOF targets and on their mapped measures. What has not been established is to what extent, if any, the QOF has been responsible for these improvements; in part due to the lack of a control group but also due to a lack of follow up in the included studies. Nor has any evidence been found for the impact of QOF targets on relevant clinical outcomes.

5.13 Conclusion

This literature review has identified a clear need for further research in this area. This research should address the limitations of the research identified in this review. Having focused purely on QOF targets or their mapped measures the question of whether the QOF has impacted on relevant clinical outcomes is still one that future research needs to address. To account for underlying trends as well as any cumulative benefits from repeated compliance; the research should involve repeated patient observations where possible and a sufficient period of follow up to pick up any intervention effects. These are items that this research will seek to address in the remaining chapters. The next chapter will describe the databases used in this study and how study variables were created using the information extracted from them.

Chapter 6 Methodology: Using CPRD and HES data to model the QOF

6.1 Objective

To explain how CPRD and HES data can be used to construct variables relevant in the modelling of the QOF. This will involve:

1. Briefly describing how the CHD patient cohort was defined in CPRD and linked into HES data
2. Describing the CPRD file structure and how it can be used to define and extract data to construct CHD QOF targets and other explanatory variables
3. Defining secondary care outcome measures and describing how they can be identified using the HES data file structure
4. Describing how data was quality assured and how missing or unreliable data were dealt with

6.2 Defining the CHD study population

The starting point for the study was to define the study population. This took place in the CPRD dataset before the patients were linked into HES. Patients were selected on the basis that they had a CHD Secondary prevention Read code recognised for reimbursement purposes by QOF Business rules version 17. Qualifying Read codes are detailed in Appendix 7. Those Read codes needed to be recorded at any time during the study period, 01/01/2000 to 01/04/2011. In order to have a starting population a second define was undertaken, this time allowing Read codes prior to the study start date to qualify, but only for those patients who had a Read code during the study. This process did not change the patients in the population nor the study start date, it simply ensured that on the first day of the study there were patients already 'recruited' into the study.

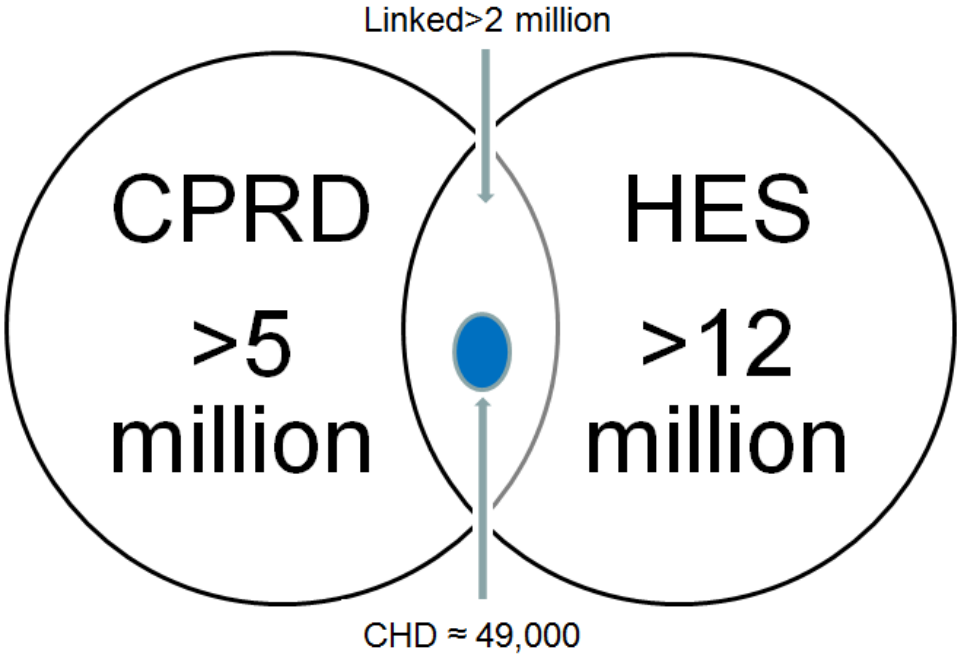
6.2.1 Defining the study population: CPRD and HES linkages

As the study was concerned with the impact of QOF care on secondary care outcomes, the patient cohort defined in CPRD had to be linked into HES. This reduced the CPRD cohort since not all CPRD practices have consented to HES linkages and HES collects data on the English NHS only, limiting the study population to patients in English practices. The CPRD population was linked into

HES in two stages. The first tranche of HES linked data which was made available to the study in August 2012 defined the HES linked population. This HES upload included data up to the end of December 2010, contained 304 consenting practices and over 4.5 million patient events. In March 2013 a new upload was received which included HES linked data up to the end of March 2012 and contained 374 consenting practices. Had this data been used in the study it would have increased the existing study population over the course of the study by just under 12,000 before any data cleansing, and added an additional QOF year to the study. In the absence of time to make changes, the new HES upload was used to take the HES data up to the end of March 2011 to fit in with the end of the QOF year, for the previously defined patient population. Hence, in effect, the new HES upload was used to find the secondary care events of interest; but only for existing patients who had been linked using the previous upload.

The way the two databases interact and where the patient population fits in is illustrated in the Venn diagram, Figure 6-1. The figures represent approximations of the active population, i.e. both up to standard and contributing data in any given year of the study rather than total number of patient records on each.

Figure 6-1CPRD and HES linkages



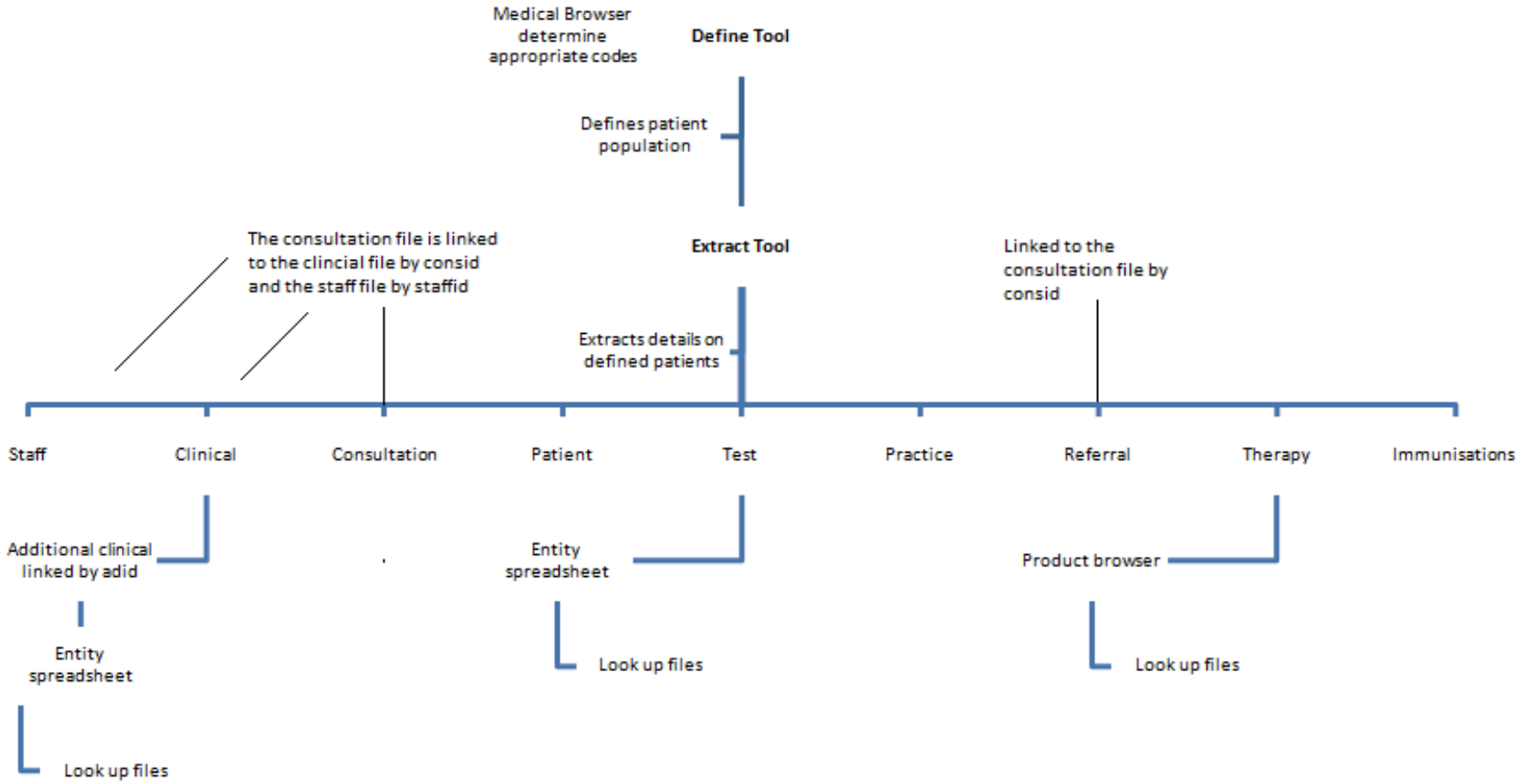
CPRD and HES are stand alone databases, illustrated by the two circles above. So long as a practice on CPRD consents to linkages its patients at that point become

HES linked. Of the over 5 million active and up to standard patients within CPRD this applied to approximately 2.5 million of its population, roughly 2.2 million of whom had been matched in HES; across 304 separate practices (figures based on the December 2010 uplift). HES only covers the English NHS, and contains inpatient data only. A patient in a HES linked practice has their data captured in HES by matching a combination of their date of birth, gender, NHS number and postcode. For the HES link upload used in this study to determine the qualifying patients, nearly 5 million patients qualified for linkages and 3.9 million had an inpatient admission matched to them. In the study that population is restricted further to patients who have a CHD Read coding in CPRD and are HES linked, illustrated by the shaded oval in the overlap between the HES and CPRD circles.

6.2.2 Defining and extracting the study population using CPRD

The CPRD file structure that was used to define, extract and then create variables is illustrated in diagrammatic form in Figure 6-2 on the following page. Explanation of the process and diagram are given in subsequent pages.

Figure 6-2 CPRD file structure: Diagramatic representation



Within the QOF, Read codes are used to classify conditions and treatments. In CPRD the comparable information is recorded using medical codes (*medcode*). The first task in ascertaining the patient population was to translate CHD Read codes into medcodes using the Medical browser provided by CPRD. Version 1.3.2, November 2011, was used in this study.

With GPRD medical codes it is then possible to perform a search for patients on the database who have an entry for one or more of those medcodes, by feeding them into the Define tool. Up to 2 inclusion criteria and 4 exclusion criteria can be specified in the define tool, by inputting the medcodes of interest in each. A time period for the search is specified, which was the study period in this instance and a minimum and maximum age range for qualifying patients can be entered, though none was for this study. The CPRD database is updated monthly using information submitted by practices. These updates not only add to information on patients based on activity since the last database build but can include updates and additions to information submitted for previous periods. The January 2012 database build was used for all data defines. Only events in the registration period were selected meaning they had to take place in the patient's current registration period. The period within which codes are acceptable are also selected at this point from one of three options: any in the study period; any before the study end; and first ever in study period. In this study first ever in the study period were first selected, as this was the point from which patients became eligible to participate, namely from the point of their first CHD code in the study period. Subsequent to this however, to ensure a non-zero study start population, the any before the study end option was selected, which produces a list of patients with the codes of interest at any point prior to the study end date. If the previously defined patients had a CHD coding prior to the study start date, they entered the study at its start, 01/01/2000, providing it met quality control markers.

Within the define browser there are 5 fields (tick box options) which generate files on different aspects of care. The first four: Referral, Test, Clinical and Immunisation look at patient activity with respect to consultation episodes, referrals for outside services, clinical tests and immunisations delivered in surgery. The last one, Therapy, cannot be selected at the same time as Referral, Test and Clinical, and vice versa; as they search on product codes, namely for medical prescriptions issued to

the patient. When the define process is completed, text files are produced from each of the four files listing the date when that medical code was recorded, and the identifier for the patient it was recorded against (*patid*). In addition separate log files give details of the search, numbers satisfying each inclusion and exclusion criteria, and a list of patients meeting the inclusion criteria.

The define tool simply produces a list of patients with an index date denoting the date when the one of the specified medcodes of interest was first recorded. To gather further information on those patients the Extract browser is then used. This uses the patient lists from one of the inclusion or exclusion files generated in the define process to extract further details on those patients. At this stage it is also possible to get therapy details on patients who were selected on the basis of medical codes. Likewise it is possible to find clinical details on patients selected on the basis that they had been given a certain therapy, for instance a prescription for anti-coagulants. If choosing this route the product browser is used to search for relevant codes as opposed to the medical browser.

The extract tool has ten tick box fields which generate separate files for each of those fields, shown on the horizontal line in Figure 6-1, and the additional clinical file which is shown beneath. While additional clinical is a separate file it has to be linked to the clinical file using the additional identifier variable, *adid*. These ten tick box fields are further separated into a number of further tick box fields specific to that file, making it possible to specify the content of the file. In conducting this research all files and fields were extracted, meaning that all CPRD database details on the study population were extracted. In addition to the Define and Extract browser there are two additional browsers, Refine and Transform. These were not used in the course of this work and are therefore not elaborated on further.

6.2.3 What form CPRD data takes and how files are linked

The extract tool generates ten files: Clinical, Additional Clinical Details, Consultation, Immunisation, Patient, Practice, Referral, Staff, Test and Therapy. Different aspects of patient care, staff and practice details are held on these files. Depending on the nature and scope of the research being undertaken it may be necessary to link across the various files to build up a bigger picture on the unit of analysis. In CPRD there are 3 possible units of analysis, the patient through the unique patient identifier, *patid*, the practice through their unique identifier, *pracid*, or

alternatively an individual member of staff through *staffid*. In the case of this study the data was aggregated and analysed at the patient level. Although the QOF rewards practice attainment, the incentivised treatments are delivered, and their impact felt, at the individual patient level. Hence for this study which was concerned with the clinical impact of physician incentives, the patient represented the logical unit of measurement. In addition to these unit identifiers there are also a number of shared identifying variables that enable entries to be linked across files. A summary of the nature of the information contained on the separate files is specified below:

- The patient file contains patient specific information such as sex, date of birth, and registration date.
- The Practice file provides information about the practice where the patient was treated such as its geographical location, and time since it has been up to standard for CPRD data collection purposes.
- The Staff file records the gender and role of practice staff.
- The Clinical file details all medical events including symptoms, signs and diagnoses. When further details regarding these events were recorded in the structured areas of the GP's software these are shown in the Additional Clinical Details file. These details are linked by an additional identifier (*adid*) unique to the clinical event and recorded in up to seven data fields whose details are determined by consulting the *entity* data file.
- The Consultation file contains details on the type of consultation as entered by the practice staff according to a pre-determined list. This primarily details the type and duration of the consultation. It is linked to the staff file by means of the *staffid* and where further clinical events occur as a result of the consultation, to the clinical file, by a unique consultation identifier *consid*.
- The Immunisation file contains details on all immunisations recorded on the GP vision system including product details, where it was administered, the reasons and method.
- The Referral file details any patient referrals to outside care agencies, normally to secondary care centres, generated by the different staff members and recorded on their systems.
- The Test file details all tests requested on behalf of the patient by the practice staff. This information is given in numbered fields, 4, 7 or 8 data fields in length;

whose values are determined by its numbered *enttype*, which in consultation with the entity file and various look up variables gives details on the tests undertaken.

- The Therapy file gives details on all prescriptions issued to individual patients by individual practice staff.

6.2.4 Using CPRD data to create QOF variables: An example

The QOF target in question determined which file needed to be consulted. Naturally the first task was to build a CHD patient population that would act as the denominator in the majority of targets. Patients entered the study population at their first recorded CHD Read code date, in the case of this study it was specified that it needed to be during the study period. So long as that date was in the patient's current registration period and when the practice was classed as up to standard, meaning of research quality, this date was acceptable. Otherwise the first coding following practice up to standard was used. The patient remained in the CHD study population up until the study end or the earliest of their transfer out date or death date. If the transfer out or death date occurred at any point in a QOF year, 1st April to 31st March the following year, the patients were classed as QOF qualifying for that QOF year so attainment on QOF targets was measured. Likewise the patient immediately entered the CHD denominator and QOF population at first CHD coding, so formed part of the CHD population from the first QOF year they entered. This is in keeping with QOF business rules, though exceptions can be made for patients who became QOF qualifying in the last three months of a QOF year and did not meet a QOF target. It should be noted that the same targets and rules applied to the whole study period, including years prior to the QOF, so in the context of this study QOF years refers to an accounting period when QOF targets were applied, and not necessarily that the QOF P4P scheme was in existence.

Having established the denominator the next task was to determine the numerator for each target. This differed for each target, and brief outlines of how each was determined are given later on in this chapter. For purposes of illustration the QOF target, CHD8, last measured total cholesterol, measured in the previous 15 months, of 5mmol/l or less, is used. The first point of reference to determine where this is recorded in CPRD is the *entity* spreadsheet. Searching on the medical term of interest in the description field determines which CPRD file the information is contained in, and what coding, *enttype*, is used to record it. There is further information on this

spreadsheet which, in conjunction with CPRD lookup files, details what is recorded in the file data fields for that *enttype*. Among these is the actual reading or result and the unit of measurement used.

The *entity* spreadsheet reveals that serum cholesterol can be found in the test file by searching for the value 163 in the *entity* variable. It also specifies what look up file needs to be consulted to determine units of measure and which field in the test file this information is recorded in. In this instance the SUM look up file needs to be consulted to determine the unit of measure recorded in the data3 field. In near 100% of instances where a unit of measure was recorded, mmol/l was used, which related directly to the QOF wording. Where data is entered an *eventdate*, the date the event happened, should be recorded. In its rare absence the *sysdate*, system date, was used in its place which is the date the event was entered on the surgery IT systems from which CPRD collects its data. This usually, though not always, matches the former. These dates were then placed into QOF years using the same criteria as that applied to the denominator. The wording of the QOF target specifies three components that must be met for the serum cholesterol test result to be QOF qualifying. Firstly that the reading should be 5mmol/l or less; secondly that the last recorded measure in that QOF year should meet that criteria; and thirdly that the last reading can be at any point in the preceding 15 months. In order to determine these were met all readings over the past 15 months prior to the end of the QOF year were selected. These were then ranked in date order, and if the last recorded reading over that period was 5mmol/l or less the patient met the QOF criteria, and formed part of the practice numerator to determine QOF achievement. Both patient and practice level performance was recorded, though the analysis took place at the patient level, so the study was more concerned with whether or not the patient met the target rather than the practice's overall performance.

6.3 How is HES linked data organised?

The lowest level of activity in HES is an episode. An episode in this sense refers to each time a patient receives a medical intervention. A patient may receive a number of episodes of care with the same consultant in the course of a routine examination or surgery, and once the last episode in the process is administered by the consultant this constitutes a finished consultant episode. The patient may then be seen by different consultants or health professionals including the initial consultant in

between, generating a number of finished consultant episodes, before finally being discharged from that hospital. This process of care constitutes a spell. Should the patient then be discharged from their first hospital immediately to a different NHS hospital or other NHS funded provider for continued care, the recording process continues in the order discussed. The whole process of care from the time when the patient enters the health system to the time that they are discharged back into the community is referred to as a continuous inpatient spell. This consists of one or more spell's, which refers to the care received in one hospital or health setting, which in turn consists of one or more finished consultant episodes, which are a record of care with each consultant.

As with CPRD, HES data is split into files. In HES files are primarily broken down into episodes and hospitalisations, the latter referring to spells of care. Episodes are uniquely identified by an *epikey* variable and spells by a spell number, *spno* variable. As well as ICD 10 codes, OPCS 4 procedure codes are recorded by episode in a procedure file, and further files are provided on periods of augmented care, time spent in critical care and maternity care, where applicable. All of these files can be related to one another through the patient identifier and one or a combination of episode key or spell number whichever is most applicable to the care recorded

6.4 QOF targets examined in this study

A description of how CPRD was used to model the QOF target CHD8 has already been described in detail. Table 6-1 goes through the wording of the remaining targets and the rule set followed to model these in CPRD. For guidance in drawing up these rule sets and determining compliance with QOF targets, QOF Business Rules were used. Although the QOF was not introduced until April 2004, the same rules were applied for periods prior to that for means of comparison. Full details of the medical and Read codes used are given in Appendix 7. In all instances the denominator was the whole study population unless specified otherwise. Exclusions were not applicable; and in all cases it was assumed that missing data meant the event had not taken place and the target was therefore unmet, except where specified. Table 6-1 details the rules followed for QOF targets only, with details on the remaining variables provided in Appendix 8, along with the expected relationship of all the variables with the study outcome measures.

Table 6-1 QOF study variables ruleset

CHD6	The percentage of patients with CHD in whom the last blood pressure reading (measured in the last 15 months) is 150/90mmHg or less
Variable name	BP
CPRD ruleset	A last recorded blood pressure at or below 150/90 measured at any point in the QOF year or 3 months prior to it, recorded in the additional file against an enttype of 1.
CHD9	The percentage of patients with CHD with a record in the last 15 months that aspirin, an alternative anti-platelet therapy, or an anticoagulant is being taken
Variable name	Anti
CPRD ruleset	At least one prescription for aspirin, an anti-platelet or anti-coagulant (recorded in the therapy file using bnfcodes 15, 196, 912, 918 and 919); during a QOF year and the 3 months prior.
CHD10	The percentage of patients with CHD who are currently treated with a beta blocker (unless a contraindication or side effects are recorded)
Variable name	BB
CPRD ruleset	At least one prescription for a beta-blocker in the last six months of the QOF year recorded in the therapy file using bnfcode 18.
CHD11	The percentage of patients with a history of myocardial infarction (diagnosed after 1 April 2003) who are currently treated with an ACE inhibitor or Angiotensin II antagonist
Variable name	Acea2q
CPRD ruleset	Pre QOF period: A Myocardial Infarction Read code entered within

	CPRD after 1 April 1999, and at least one prescription for an ACE inhibitor or Angiotension II antagonist (recorded using bnfcodes 5 and 142), in the last six months of a QOF year QOF period: A Myocardial infarction Read code, entered within CPRD after 1 April 2003 and at least one prescription for an ACE inhibitor or Angiotension II antagonist (recorded using bnfcodes 5 and 142), in the last six months of a QOF year
Denominator	All patients with an MI code entered on or after 01/04/1999 prior to 01/04/2004. After this date all patients with an MI code on or after 01/04/2003.
Exclusions	Patients with a MI Read code prior to 01/04/1999 in the pre QOF period Patients with a MI Read code prior to 01/04/2004 in QOF period
Related to **Smoking3**	The percentage of patients with CHD whose notes contain a record of smoking status in the previous 15 months
Variable name	Smokstat
CPRD ruleset	A smoking medcode recorded in the clinical file in the QOF year and the three months prior. These are broken down into the categories set out in that QOF ruleset: never, ex and current. Some patients had multiple records over a QOF year, and could move between smoking states. Where multiple records existed, the codes were selected in the following preference order, a current smoking record in the first instance, then ex and finally never.
Related to **Smoking4**	The percentage of patients with CHD 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
Variable name	Smcess

CPRD ruleset	A smoking cessation medcode recorded in the clinical or referral file during a QOF financial year, and the three months prior, matched against a current smoker code recorded at any point in the same 15 month period, in the smokstat variable.
Denominator	All patients with a current smoker code
Exclusions	All patients with no record or a non-current smoker code

6.5 HES outcomes of interest

The study was concerned with secondary care admissions for complications arising out of existent CHD, which could be indicative of poor condition management in primary care. These outcomes were ICD10 codes for Ischaemic Heart Disease, I20 to I25. Codes within this grouping for myocardial infarction (I21 and I23) had to form the primary diagnosis during a hospitalisation. The remaining codes within this grouping (I20, I22, I24 and I25), are for first or early stage forms of CHD, and for that reason not all codes were selected, and for those that were, the hospitalisations had to be emergency admissions. The reason being that non-emergency admissions for these codes may not be indicative of poor management in primary care, as they could be pre planned and therefore represent the contrary. The details of all the ICD10 codes used are detailed in Appendix 7, pages 279 and 280, and the rule set used to create the outcome variables detailed below. Again the denominator is the CHD patient population, and missing data signifies that a hospital admission of interest did not take place in all instances with additional points listed where appropriate. The hospitalisation had to occur after CHD coding, when the patient's practice was up to standard and before the earliest of the patient's transfer out or death date. Not meeting these criteria meant the event was excluded from analysis. Other exclusions are reported where applicable

Table 6-2 HES study variables ruleset

Variable	HES Ruleset
miadmis	Refers to a hospitalisation where an MI code formed the primary diagnosis. These were recorded in the primary diagnosis across a hospitalisation HES file in the *ICD_Primary* variable. Any codes in the I21, I22 or I23 groupings recorded in that field qualified.
emang	Relates to an emergency hospitalisation where specific ICD10 I20, I24 and I25 codes formed the primary diagnosis. These were recorded in the primary diagnosis across a hospitalisation HES file in the *ICD_Primary* variable. To select only emergency admissions this file had to be merged by patient id (*patid*) and spell number (*spno*) to the same event record in the HES Episodes file. In this file the method of admission variable (*admimeth*) gives details on how the patient was admitted. The following codes constituted an emergency admission: 21, Emergency via A&E; 22, Emergency via GP; 23, Emergency via bed bureau; 24, Emergency via consultant outpatient clinic; and 28, Emergency other means. If the hospital admission had one of these codes the record was retained.
Exclusions	• Primary diagnoses by hospitalisation which were not an emergency admission
alladmis	This was created by adding the emang variable to the miadmis variable. It represents a count of all admissions for all secondary care outcomes of interest
alladmisx	This has been created from the *alladmis* variable, adding additional MI events. For a full explanation of this variable see 6.7.1

6.5.1 Measuring outcomes, definitions and impact on numbers

Work by Herrett et al looks at the differences in numbers of acute myocardial infarctions (AMI) in patients when measured in CPRD, HES, and the Myocardial Ischaemia National Audit Project (MINAP)[123]. Office for National Statistics (ONS) cause specific mortality data was additionally used as a confirmatory and cross referencing source where applicable. All of these sources with the exception of MINAP were available to this study. While this meant that this study could not adopt the approach taken by Herrett et al, this was the smallest of the three databases, and therefore its impact was relatively minor in comparison. It also did not prevent analysis of the impact of broadening the search and definition of MI.

The process used by Herrett et al was as follows: In CPRD acute MI's were identified by searching on condition specific Read codes, in HES by the primary diagnosis in the first episode of a hospital admission using ICD10 codes, and MINAP using internationally agreed definition markers. While there was significant shared coverage between the three datasets, the different databases and definitions nonetheless revealed differing numbers of acute MI's. MINAP uses the most stringent criteria of the three and had the fewest cases. CPRD where the criteria for recording an event was the least stringent had the most, with HES falling between the two. This poses potential problems when relying on just one of these datasets as events may be missed or overestimated. Over the course of the Herret study, 01 Jan 2003 to 31 March 2009, for instance, 4588 potential acute MI's were recorded in CPRD but neither of the other two databases, 2352 in HES alone, and 1634 in MINAP alone.

Lessons from this paper were applied to this study to add potential additional MI's to the *alladmis* variable. A new variable, *alladmisx*, was created and the following process undertaken to find these additional events:

- First MI codes at or following first CHD Read coding were extracted from the CPRD clinical file
- If these codes fell within 30 days of discharge following a hospital admission recorded in the HES primary diagnosis by hospitalisation file they were retained.
- Providing there was no other admission in that 30 day period and that this admission was not given one of the study ICD 10 codes of interest and therefore

already recorded in the *alladmis* outcome variable: This was considered to be a potentially missed MI hospitalisation

This process was a more cautious approach than that applied by Herrett, but reduced the possibility of over estimation brought about by potentially counting routine and repeat entries. Only the first recorded MI code was used as only in these circumstances could it be proposed with reasonable confidence that a hospital admission prior had been the result of an MI, even if it did not form the primary diagnosis. Subsequent codes recorded in CPRD could simply represent routine entries following consultation, and where they followed a hospital admission where MI was not the primary diagnosis would more likely represent secondary and tertiary diagnoses. Over the study period this added a further 2942 events (9.3%) to those picked up in the study outcome specification, *alladmis* variable.

6.6 Missing data and data issues

If data was missing on the patient population during their study qualifying period, it was assumed that the activity did not take place. There is the possibility that no record rather than indicating the absence of an event; reflected it simply not being recorded. This is of course impossible to ascertain or control for. However there are a number of factors which mitigate this and certainly make it unlikely. Firstly all of the patients were in practices that were up to standard, meaning that their data collection and entry practices were considered to be at a research standard. The data related to prescriptions and important test measures that practices would need to have a record of to ensure the patient's condition was being effectively monitored and treated. In a chronically ill patient population this would be even more of an imperative. Finally, certainly since the QOF, but also partially in schemes prior, there was a financial incentive to ensure that such events were recorded when they took place. Hence it is reasonable to assume that the absence of data meant that the patient was not receiving the care examined.

Inaccurate or incorrectly entered data appeared to be a rarity, and as there were usually multiple records for each QOF target event in each year, it did not necessarily prejudice any final results. However human error is impossible to eliminate and for the purposes of the study, cut offs were generated outside of which the data was considered medically unreliable or extremely unlikely. These were a

serum cholesterol of greater than 30mmol/l; systolic blood pressure below 70mmHg and at 250mmHg and above; diastolic blood pressure below 40mmHg and at 250mmHg and above. Such incidences were very rare and accounted for less than 1% of observations. Where they did occur the reading was classed as invalid and the next available valid reading which fell within the plausible range was used.

In rare instances, 0.0015% of total clinical records, the date an event occurred was missing in CPRD, however in all instances a system date was present. The system date refers to the date the event was recorded on the GP practice I.T system. In general where both were recorded the two were in agreement, however it is possible for one to deviate from the other, usually by a couple days but very rarely significantly so. For instance in situations where events are entered retrospectively by a practice administrator, using a GP's notes; or there is a long delay in getting a test result. However such occurrences are rare and where the date of event was missing the system date was used in its place.

There were 83 patients within the population who were first diagnosed with CHD at less than 30 years of age. These were a tiny percentage (0.1%) and were most likely patients who suffered from congenital heart disease that had been wrongly entered as a CHD Read code. There was also the possibility that date of birth had been entered incorrectly for these patients meaning their age was incorrect. Due to the small probability of actual CHD in the under 30 population the following cautious approach was taken: All patients who had a first CHD diagnosis under 18 years of age were dropped from the study; and of the remaining aged 18 to 30 population, patients were dropped if they were present on the database for 1 year only. It was felt that if no data had been entered for that patient following CHD coding, indicating it was not being treated, it was very likely that the patient did not have CHD and therefore the code had been entered incorrectly. This led to a loss of 25 patients and one of the myocardial infarction outcomes. As a minimum of two patient observations are required for panel data analysis these patients would have dropped out of data analysis regardless, so this had no impact on final results.

Over 11% of individuals had a record in HES for one of the study outcomes prior to a QOF recognisable diagnosis of CHD in CPRD. A much smaller percentage had multiple outcomes prior to study entry. These were on occasions prior to the registration period specified above, in which case they may have occurred

subsequent to a CHD code outside the qualifying period in this study. However in the majority of instances they were the first time CHD became recognisable in those patients and thus recorded in CPRD. In all instances these were retained and placed in the pre-study outcome variable, *Hospre* (See Appendix 8).

An individual's participation in the study during their qualifying period was ended at the earlier of their transfer out or death date. It was noted that these dates differed in around 40% of instances where both occurred. On average the difference was 18 days, and in 98% of cases the difference was less than 60 days, meaning there will have been minimal if any impact on the panel. However in 0.7% of cases (0.2% of total data) the difference was over 6 months and up to 16 years. In these instances qualifying years of data may have been lost to the study.

6.7 Summary

The quality and accuracy of data was assured throughout the study using controls put in place by CPRD as well as by conditions placed by this study. Hence at the define stage only data in the individual's current registration period was admissible, namely during their current registration period and when their practice was considered to be up to a research standard. This was confirmed and assured by this study when the data was placed into a panel data structure for analysis. These two processes ensured the data complied with CPRD's research standard markers. The same markers were applied to the HES data using the same conventions, both by using variables placed in HES to act as quality markers (*hes_start* and *hes_end*) and by confirmatory checks using the base CPRD variables that were used to create those quality markers.

Aside from these there was little in the way of amendments to the data or need to question it, and the data was seen to be of a very high standard. A number of minimal adjustments were made, which did not affect data integrity as their frequency was very small, and as they often applied to data with multiple records, did not change the data used. These are detailed below:

1. Controls were placed on permissible blood pressure readings and cholesterol but these only applied to a small percentage of the total records, less than 1%.
2. A number of patients were given a CHD Read code from birth to the age of 30, which were most likely incorrect codes for congenital heart failure. These accounted for less than 0.1% of the CHD study population however.

3. A very small percentage of clinical entries (0.0015%) had no date of event recorded. In these instances the system date was always present, and used in its place, though it should be noted that these are not always in agreement.

In addition to these there were some general issues with how the QOF targets were applied in this study. Specifically

1. Exception reporting codes were not applied as the data was concerned with the effect of treatment and outcomes at the individual patient level, and these have greater significance for practice attainment
2. Equally the practice attainment variable, *pracatt*, did not adjust for exception reporting
3. In this study an individual became immediately QOF qualifying upon entry and hence was included in that QOF. In contrast patients entering in the last 3 months of a QOF year can be exempted from targets.

The chapter noted work undertaken by Herrett et al, which showed differences in numbers of AMI's recorded in CPRD, HES and MINAP[123]. The latter was not available to this study, so this was explored for HES and CPRD only, with ONS data also acting as a further reference point, as it did in the work by Herrett. An approach was taken which utilised HES and CPRD to pick out potential missed AMI's. This was a more formal approach than that undertaken by Herrett et al, which was deemed more suitable to this study and reduced the risk of over-estimation of events. Over the duration of the study this increased the number of outcomes by 9.3%, which are recorded along with existing outcome events in a new outcome variable, *alladmisx*.

6.8 Conclusion

This chapter has described how CPRD and HES were used to define the study population and create the variables for use in subsequent analyses. The following chapter will describe and analyse those variables, and construct an econometric model to examine the relationship between the quality of CHD care given to patients in primary care and its impact on relevant clinical outcomes.

Chapter 7 Data Description and Analysis

7.1 Objective

The objective of this chapter is to describe the data, introduce econometrics, specify an econometric model and undertake analysis of it. This will be conducted in the following steps:

1. An overview of econometrics
2. Model specification
3. Descriptive analysis of the data
4. Model estimation and discussion of results

7.2 Background

The data extracted by this thesis has been placed into a panel data format. The unit of analysis is the individual CHD patient. Panel data combines two dimensions, in this thesis cross sectional and time series.

Table 7-1 Panel data structure, a two patient example over 4 time periods

patid	*Year*	*alladmis*	*BP*	*pracatt*
1096	1	0	0	0.8
1096	2	0	1	0.82
1096	3	2	1	0.85
1096	4	0	0	0.87
35769	1			0.68
35769	2	0	1	0.7
35769	3	1	1	0.77
35769	4	1	1	0.73

An example of the panel data format used in this study is presented for hypothetical patients in Table 7-1. The outcome (also referred to as dependent) variable, *alladmis*, in this thesis is the number of primary diagnosis hospital admissions or emergency admissions for selected ICD10 I20-I25 codes. *BP* and *pracatt* are explanatory variables (covariates) which are believed to have some causal relationship with the outcome variable; the extent of which is captured by their coefficients, denoted and

referred to as their beta, β. *BP* shows whether the individual met the QOF BP target and *pracatt*, relates to practice attainment on QOF targets. Since the study includes a number of explanatory variables, key among which are the QOF target variables, it is a form of multivariate analysis. A final point with regard to this illustration is to do with the type of panel used in the study. As individual 35769 does not contribute observations in year 1 the dataset is called an unbalanced panel. These points will be developed further below.

7.2.1 What is econometrics?

Economics is a social science which attempts to explain how individuals, firms, and governments go about allocating scarce resources to meet unlimited wants. Econometrics is a branch of economics which applies mathematical and statistical techniques to economic phenomena arising from this basic economic principle.

As with any social science there are uncertainties and assumptions in any economic theory. This is because economic theories are applied to uncontrolled 'real' world settings and it is therefore impossible to control for all the influences that can impact on human behaviour. As a result people do not behave in the same predictable manner, nor indeed always act rationally; making modelling problematic. Econometrics applies statistical methods to economic problems to attempt to design causal regressions which explain observed economic relationships and events, using data from uncontrolled settings. It differs to statistics which is more generally applied to natural sciences, controlled settings, and seeks to predict behaviour or relationships, rather than precisely explaining them; looking for patterns, rather than causes. Different methodologies have consequently developed to fit these contexts. The approach taken by econometricians is to develop a model based on economic theory, which then guides data collection. Then based on limitations in the data available within real world settings such as; vaguely defined, missing or uncollectable data; to develop approaches to get as much information as possible from that data to explain the outcome of interest. This differs from medical statistics for instance where data is collected, generally in controlled settings, and then fitted into models based on their ability to predict the outcome of interest. One of the respects in which econometrics is in regards to model specification and the treatment of the error term. This is demonstrated in the simple linear equation (1) below:

$$Y = \alpha + \beta_1 X + \varepsilon \qquad (1)$$

This error term or disturbance term, ε, is a random variable which captures the variation in the outcome variable, Y, not captured by the included explanatory variables. Since econometrics operates in a world where variables are not observed or cannot be collected, the error term plays an important role, and methods such as fixed effects are more frequently used. The inclusion of this error term has implications in panel data which will be discussed subsequently.

7.2.2 More on panel data

A panel in economics refers to a group of economic agents be they individuals or higher level aggregates such as households, firms or countries. Panel data combines two dimensions, usually a cross sectional and a time series or longitudinal dimension. It does this by repeatedly analysing a cross section of individuals over a given period of time. If the same individuals are present throughout the study it is termed a balanced panel. If individuals enter, and/or, exit at different time periods, it is described as an unbalanced panel. A panel data regression model is written in functional form below:

$$Y_{it} = \alpha + \beta_1 X_{1it} + \cdots + \beta_k X_{kit} + \varepsilon_{it} \qquad (2)$$

The model parameters seen are the constant term, α, and the coefficients,β_1,...,β_k, indicating the influence that the explanatory variables have on *Y*.

Subscript *i* refers to individual *i* = 1,....,*N*, where there are a total of *N* individuals included in the panel, and time *t* = 1,...,*T* refers to the time period of which there are a total of *T* in the panel (in this study *N*=81,201 and *T*=11). The modelled variable *Y* is commonly referred to as the dependent variable, X_1 to X_k as the explanatory variables and ε denotes the unobserved error term. There are a number of assumptions made concerning the unobserved error term. These are:

$E(\varepsilon \,|X_i\,) = 0$ Zero conditional mean: The expected value of the error term is zero for any given value of the explanatory variables (3)

$Var\,(\varepsilon\,|X_i\,) = \sigma_\varepsilon^2 \;\forall\; i$ Constant variance: The variance of the error term is the same for all values taken by the explanatory variables (4)

$Cov(\varepsilon_t, \varepsilon_s) = 0 \;\forall\; t \neq s$ Correlation of error observations is zero: The error terms are independent and not related to each other, meaning there is no correlation over time, with t and s denoting different time periods. (5)

7.2.2.1 Advantages of panel data

Panel data offers significant advantages over data measured in cross section alone or in time series alone. These are[124] [125]:

1. By combining time series and cross sectional observations panel data creates a larger number of data points ($N*T$) enabling a more information rich analysis. This has a number of potential benefits including reduced variability and collinearity, and larger degrees of freedom. This produces more efficient estimators, and allows for stronger inference.
2. By studying repeated cross sections, panel data is particularly well suited to study the dynamics of change. This could be individual level choices such as the decision to quit smoking or the effects of periods of unemployment. Alternatively the dynamics of interest could be the effect of higher, national for instance, level changes, such as national minimum wage rates, on individual level units.
3. Panel data allows greater flexibility in modelling differences in behaviours among individuals even when the model is not fully specified and there is unobserved heterogeneity. In particular panel data allows us to deal with omitted variables (unobserved or non-measurable) which are correlated with the included explanatory variables and can result in biased estimators in other forms of econometric analysis. Assuming that the omitted variable does not change over time (time invariant values) and has same effect over time (time invariant effects) fixed effects panel data analysis provides a means of dealing with their omission. Examples of variables fitting this description are gender, race, education level and innate ability.
4. By observing the same unit across time, panel data enables the researcher to examine temporal effects, namely the effects of variables in previous periods on outcomes in the present period.

7.2.2.2 Fixed and random effects

For panel data regression analysis the choice of modelling option depends on model specification and assumptions. Most frequent are assumptions concerning fixed and random effects.

The different modelling assumptions change the functional form in (2) depending on the presentation of the error term, ε. To explain these the error term presented needs

to be examined further along with assumptions regarding its relationship with the explanatory variables.

$$\varepsilon_{it} = v_i + u_{it} \qquad (6)$$

This is done in (6) above, which breaks the error term into two component parts. The first is an unobserved heterogeneity or fixed effects component, v_i, which varies with the individual but is fixed over time. The second what is called an idiosyncratic or time varying error, u_{it} , represents unobserved factors that change over time and affect the individuals outcome variable. This is a completely random component that is not associated with any of the cross sectional units, namely across individual entities. These can equally be described as between variation, namely between individual units within the panel and within variation, which considers the variation within the same unit over time.

The idiosyncratic component of the error term, u_{it} , does not create problems as it is assumed to vary randomly over time and is therefore not correlated with any of the included explanatory variables. Hence it is treated the same in both fixed and random effects, retained within and captured by the error term. Where they differ is with respect to their treatment of the unobserved fixed effects component, v_i.

Random effects analysis assumes that the fixed effects component is not correlated with any of the explanatory variables, $Cov(X_{it}, v_i) = 0$. This assumption does not mean that the individual effects are identical for every observation rather that their values are random with no association with the observed values of the explanatory variables. If this assumption does not hold random effects estimation produces biased estimates.

Fixed effects analysis however relaxes this assumption and allows the fixed effects component to be correlated with the included explanatory variables. In this sense fixed effects views, v_i , as an omitted variable that is correlated with the included explanatory variables. Fixed effects deals with this by including v_i , in the constant term, α, so that the intercept varies across each unit of observation and/or time.

$$Y_{it} = (\alpha + v_i) + \beta_1 X_{1it} + \cdots + \beta_k X_{kit} + u_{it} \qquad (7)$$

This fixed effects error component captured in the intercept can then be dealt with in a number of different ways. The dummy variable approach introduces a dummy

variable for each unit of observation so that each has an individual specific intercept. This can however be extremely burdensome when there are a large number of individual units (large *N*) in the panel. An alternative is the first difference transformation, which involves subtracting the lag value for each variable from its present value. This differences away the v_i from (7) as it is fixed with respect to time *t*. A final method is referred to as the fixed effects estimator, which is the most common method used in applied econometric analysis. Most statistical packages, Stata included, have specific commands which carry out this operation, making appropriate adjustments to the degrees of freedom to calculate accurate standard errors. This involves time de-meaning the data by subtracting the mean value for all observed values for each individual unit of analysis from each of its observed values. It is explained in functional form in (8) and (9) below:

$$(Y_{it} - \bar{Y}_i) = \beta_0(v_i - \bar{v}_i) + \beta_1(X_{it} - \bar{X}_i) + (u_{it} - \bar{u}_i) \qquad (8)$$

$$\tilde{Y}_{it} = \beta_1 \tilde{X}_{it} + \tilde{u}_{it} \qquad (9)$$

Time demeaning deals with any unobserved factors by removing from the regression the components that are fixed over time. By assumption that will be the entire amount of the unobserved fixed effects, v_i .

7.2.2.3 When to use fixed or random effects?

Fixed effects are usually the preferred method in econometrics, compared to random effects in medical statistics, since it cannot often be assumed that models are fully specified and hence there may be omitted variables potentially correlated with included variables. These omitted variables will be captured in the error term. Fixed effects, assuming those omitted variables have time invariant effects, namely are fixed over the study period, is a way of dealing with them. While fixed effects removes the time invariant component of the error term it comes at a cost. Not only does it remove the fixed effect errors but any variable that is fixed over time. Consequently variables such as gender are removed completely. Hence the decision on which method to use comes down to a number of considerations:

1. Whether or not time invariant explanatory variables are important in the model
2. How well the model is specified and hence likely there are missing variables which are potentially correlated with included explanatory variables
3. The level of within subject variation in time variant explanatory variables

4. Whether it can be assumed that the missing variables have time invariant values and effects

If the model is well specified, such that, missing variables and correlation of the explanatory variables with the error term is unlikely, random effects is the preferred option. It includes more data, uses fewer degrees of freedom and produces more efficient estimators. Likewise if it is suspected that there may be omitted variable bias, but time constant variables are important in the model, fixed effects would not be suitable. In this instance potentially biased estimators would be a necessary trade off and attempts would need to be made to adjust standard errors for correlation within group. Outside of these instances fixed effects are preferable as it more likely than not that there are unobserved variables potentially correlated with the error term. Fixed effects analysis however is not a panacea as it can be ineffective in dealing with omitted variable bias. Referring back to (6), fixed effects analysis assumes that the unobserved fixed effects component of the error term is constant over time, in other words that it has time invariant effects. In random effects v_i seen in (6) is in effect replaced by v_{it} , and the latter is assumed to vary randomly.

7.2.2.4 Designing an econometric model

When building an econometric model, the model should be underpinned by economic theory. This should guide the choice of variables and model specification. In this study, economic theory would argue that utility maximising GP's will have responded to QOF financial incentives and provided better care in those areas which are incentivised. Those improvements in care in response to incentives will be detected by QOF target variables, which have a causal relationship with the study outcome measure. Hence the QOF targets will be included in any analysis as they are its focus, as will key confounding variables such as age and gender. Outside of these, variables will be included if they are expected to have a causal relationship with the study outcome, are well specified and add to the model.

7.3 Model specification

In this section the model is briefly specified in a simplified functional form and issues arising out of the use of a count outcome variable in a panel data setting, as well as potential solutions, are discussed. The model is specified as follows:

$$Y = f(X_1 \ldots . X_m, U, \beta_1 \ldots . . \beta_k)$$

Where Y refers to either a hospital admission where MI formed the primary diagnosis, or an emergency hospital admission where selected IHD codes formed the primary diagnosis. Namely the *alladmis* or *alladmisx* variables

The X variables are the explanatory variables. These are broken down into three groups which describe where these fit in the explanation of the dependent variable. The text in italics refers to the shortened variable name, where one was necessary, used in the datasets and referred to at various point in the text.

(Quality of patient care variables) = QOF targets {Blood pressure control≤ 150/90mmHg (*BP),* Cholesterol controlled≤ 5mmol/l (*Chol*), The prescription of Aspirin, anti-platelets or coagulations (*Anti),* The prescription of beta blocker (*BB),* The prescription of ACE or A2 to MI patients (*Acea2q*), Smoking cessation advice to current smokers (*Smcess*)}, QOF related target (*ACE)*

(Patient risk factors) =*age*; *gender*; Co-morbidities {Heart Failure (*HF*), Diabetes (*Diab*), Atrial Fibrillation (*AF*), Stroke and Transient Ischaemic Attack (*STIA*), Hypertension (*Hyper*), Rheumatoid Arthritis (*RA*), Chronic Kidney Disease *(CKD)*}; SHA Region *(Region),* Angina outcome pre-admission *(Angpre),* MI pre-admission *(MIpre)*; Outcome pre-admission *(Hospre)*; Index of Multiple Deprivation 2007 (*IMD)*; Current smoker (*Smoker*)

(Practice level factors) = Practice size measured in quintiles *(sizeq)*; Practice GP workload measured in quintiles *(workldq)*; Practice level CHD QOF attainment *(pracatt)*

U is the error or disturbance term which is created when the explanatory variables included in the model do not fully explain the value of the dependent variable

The β's are the parameters for the explanatory variables.

7.3.1 Dependent variable

Both the *alladmis* and *alladmisx* variables will be used in the same model specifications to see if results differ between the two. It is expected that results will not differ significantly between the two as the numbers of outcomes recorded in both are similar. In final reporting and model specification the *alladmis* variable will be used unless the results from using *alladmisx* are more significant and in keeping with a priori expectations

7.3.2 Specification of the outcome variable

The HES outcome measures are all count variables. There are a number of issues concerning the use of count models and panel data that are discussed below along with possible solutions which will be followed in any analysis. Full details regarding the different regression models are detailed in Appendix 9

7.3.2.1 Over dispersion

The Poisson model assumes that the conditional mean and the conditional variance of the outcome variable are equal. Hence where the data deviates substantially from that assumption results can be unreliable. In those instances the negative binomial will be used which allows the variance to adjust independently of the mean.

7.3.2.2 Excess zeros

An excess of zeros leads to significant over dispersion immediately discounting the Poisson Model. The negative binomial is a possible alternative in these circumstances but it too may give misleading results when there are a large number of zeros in count models. To overcome this, all, or an excess of zeros, need to be accounted for separately within the data. Two approaches have been developed to do this, zero inflation models and hurdle models. Which of these is chosen depends on the nature of the data being modelled and any underlying theoretical assumptions. Zero inflated models assume there are two kinds of zeros, referred to as excess zeros, and true zeros. Excess zeros are due to specific structures in the data, which mean that certain observations will always generate a zero. For example in a study looking at the number of fish caught individuals who did not fish will always generate a zero. True zeros on the other hand represent true sample zeros, namely those individuals who were fishing but caught no fish. A hurdle model splits the sample down into two groups, those who did and did and did not have the outcome. A binary model is used to model whether an outcome takes place, and a truncated count model is used for those who 'clear' the hurdle and have a positive count. In essence the hurdle model treats all zeros as true zeros but assumes those who do have an outcome are different to those who do not and hence models them separately.

In the case of this study the vast majority, if not all zeros are expected to be true zeros as their CHD has been clinically verified, it is looking at clinically significant outcomes, and the data is quality assured. Hence a zero inflated approach would not

be appropriate. It is very reasonable on the other hand to assume that patients who have an event differ intrinsically from those who have not and can therefore be modelled separately. This will be explored within the data and where appropriate and feasible a hurdle specification will be developed.

7.3.2.3 The Hurdle Model

As mentioned the hurdle model consists of a binary model which predicts the outcome of interest taking place and passage into a truncated count model. Once an outcome has occurred those events are measured in a count model, truncated from below to remove the excess of zeros and deal with over dispersion. Consequently the count model component of the count model starts at 1 removing all the excess zeros. These then provide the reset starting point for the count model. Hurdle models have been used in a variety of areas including health care for cross sectional data[126]. However their implementation in panel data has been undertaken and explored to a much lesser extent, and no facilities to date have been developed in Stata to apply them. Hence any application will be explorative, utilise existing commands and be adaptive to the specific requirements of the analysis.

7.3.2.4 Alternative to the hurdle model

An alternative approach to dealing with excess zeros in a count model, besides a pragmatic approach of using the negative binomial and accepting there maybe bias in the estimators, is to dichotomise the data, i.e. replace counts greater than one with one. In instances where counts greater than one are a rarity this does not alter the outcome variable significantly and therefore this approach is justifiable and results will be similarly interpretable. It also represents a much simpler methodological approach to over dispersion thereby adding less complication and assumption.

7.4 Data description

This section descriptively analyses the outcome and key explanatory variables, as well as the surrogate measures that mapped to those variables. The reason for doing this is to get an understanding of what the individuals who make up the study population are like in terms of their demographics, socio-economic status, health status, and QOF attainment. It will also provide an insight into what was happening to those variables and this study population over the period of analysis. This will flag up any concerns with regard to how representative the study population is, and any

issues concerning variable construct. These details are ignored in any regression results, and considered in isolation can lead to their misinterpretation. This section providers some context for those results and to their generalisability.

This section is split into 4 parts, which are: population dynamics; demographic and socio-economic measures; QOF targets and mapped measures; outcome measures; and panel data measures:

7.4.1 Study population dynamics

Table 7-2 Study population dynamics

Year	Total Population	New cases	Exiting
2000/01	25,983	25,983	0
2001/02	33,992	9,172	1,163
2002/03	40,797	8,838	2,032
2003/04	44,929	6,922	2,791
2004/05	47,081	5,488	3,336
2005/06	48,356	4,672	3,397
2006/07	49,308	4,437	3,484
2007/08	49,910	4,238	3,637
2008/09	50,031	3,914	3,792
2009/10	50,064	3,711	3,679
2010/11	48,973	3,650	4,741

Table 7-2 provides details on the dynamics of the study population in terms of number of patients entering and exiting with each year. All figures, outside of the first year which ran from 1st January 2000 to 31st March 2001, relate to an annual QOF reporting period, 1st April to 31th March of the following year. Starting in the first column the figures show the total population at the year-end for each respective year. The new cases column refers to patients with a first recorded CHD code that year, and who therefore joined the study population during that year. Numbers exiting was calculated by subtracting the previous year population from the present year and then subtracting that figure from the number for new cases. Consequently the figure for the first year is zero despite the fact that patients died and transferred out that year. This is due to the way the study population was calculated. Any patient

present at the beginning of a QOF year qualified, and hence was counted, for the whole of that year regardless of whether or not they died or transferred out part way through it.

Strictly speaking therefore the exiting column shows the number of patients who left the study population in the year prior, and who were consequently absent in the year shown. Total population represents the numbers of patients present at the beginning of the QOF year plus any newly diagnosed cases that year, as a patient joined the study population and became QOF qualifying immediately upon diagnosis. For the first year, 2000/01, the two figures tally, and represent patients who entered the study immediately due to a CHD coding prior to the study start and those who had a code during that year.

7.4.2 Key patient socio-demographic measures

7.4.2.1 Age and gender

Table 7-3 Age gender statistics for the study population

	Age		Number & % Male	
Year	Mean Age (Years)	SD (Years)	N	%
2000/01	69.30	11.09	15,199	58.5
2001/02	69.60	11.23	19,921	58.6
2002/03	69.93	11.29	23,910	58.6
2003/04	70.25	11.34	26,463	58.9
2004/05	70.53	11.38	27,897	59.3
2005/06	70.83	11.37	28,780	59.5
2006/07	71.09	11.44	29,525	59.9
2007/08	71.37	11.43	30,051	60.2
2008/09	71.59	11.48	30,374	60.7
2009/10	71.84	11.49	30,570	61.1
2010/11	72.06	11.55	30,087	61.4

Table 7-3 shows the mean age and its standard deviation (SD) for the study population over the study period, as well as the number and percentage of that population who were male, over that period. The figures show that the study population aged over the course of the study, by on average nearly 3 years. At the

same time the percentage of the study population that was male increased, with males making up 3% more of the study population by the end of the study period compared to its start. These figures are generally reflective of national statistics. While statistics on the mean age of CHD diagnosis were not found, incidence rates for Angina, the first symptoms of CHD were. These show that at the English national level, the population from which the study is drawn, the incidence of Angina is very rare in those under 45 (3.2 per 100,000 patient years in 2011)[20]. The incidence rate starts to rise significantly thereafter to 44.7 in those aged 45-54, 109.3 in those aged 55-64, peaking at 151.5 in the 65-74 age group before dropping to 98.3 in those aged 75 plus; all figure per 100,000 patient years, 2011[20]. These results are generally supportive of the age profile of the study population, as the mean age coincides with Angina's highest incidence. While the figures may count multiple incidences, these are generally first symptoms. The ageing of the population is to be expected in a panel where the same patients are followed for multiple periods, and in the context of improvements in treatment and increasing longevity.

Since males are more likely to suffer from CHD than females[127]; develop CHD 7-10 years earlier than females[128]; and have a higher incidence rate of Angina[20]: The higher percentage of males in the study population is expected. Their increasing presence could be indicative of improved care over the study period; as such a situation would arise if patients were living longer with CHD and being diagnosed earlier. Under such a scenario as males make up a higher proportion of the population they would benefit in greater numbers from any falls in the mortality rate over the study period. These patients would be supplemented by proportionately higher numbers of males if diagnosis improved over the study period causing the effect seen. Indeed it would be evident but less so if new cases came in at a steady rate as existing one's died at a slower rate.

7.4.2.2 Morbidity

Details on the numbers and prevalence of selected co-morbidities in the CHD study population are explored in this subsection. The selected co-morbidities, Stroke and TIA (*STIA);* Diabetes *(Diab)*; Atrial Fibrillation *(AF);* Chronic Kidney Disease *(CKD)*; Hypertension *(Hyper)*; Heart Failure *(HF);* and Rheumatoid Arthritis *(RA)* are all atherosclerotic diseases, or risk factors for cardiovascular disease, with a possible or probable impact on the study outcome. The effects of co-morbidities are

explored further in Appendix 9, which looks at how the co-morbidity free component of the study population differs from the total study population. This provides some insight into how an experimental controlled trial population would appear in comparison to the study population on a number of study variables.

Table 7-4: Selected co-morbidities, numbers by study year

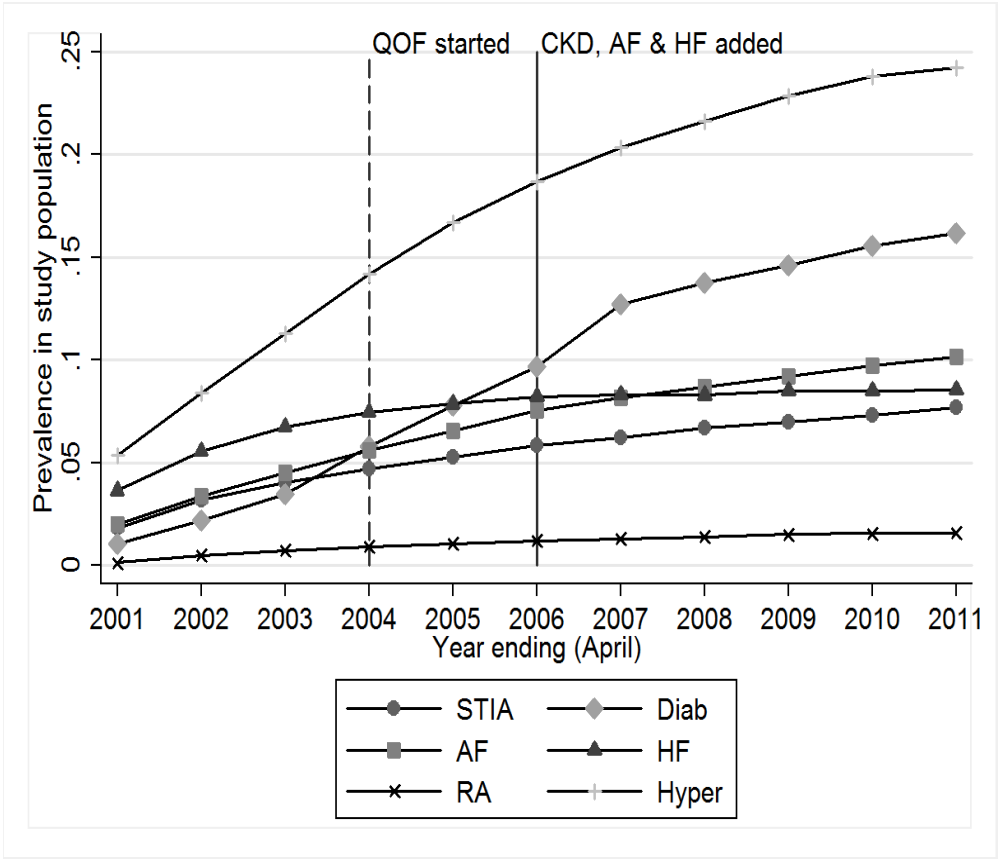
	CKD	STIA	AF	Diab	HF	Hyper	RA
2000/01	8	473	516	270	948	1397	33
2001/02	17	1080	1152	747	1890	2857	161
2002/03	37	1650	1841	1416	2749	4608	289
2003/04	70	2120	2524	2598	3354	6372	404
2004/05	168	2484	3083	3655	3699	7859	496
2005/06	621	2817	3639	4691	3971	9035	576
2006/07	9300	3053	4025	6266	4099	10042	641
2007/08	11908	3333	4333	6858	4138	10791	693
2008/09	12735	3487	4601	7321	4258	11444	750
2009/10	13063	3655	4868	7780	4253	11916	773
2010/11	12749	3765	4962	7922	4191	11883	761

In Table 7-4 all the co-morbidities increase at a faster rate than the CHD population. This could be caused by a number of factors. Firstly the initial population started off with no co-morbidities as the pre-study period was not searched to find codes prior to the study start, so there will have been a period of catch up. Secondly as patients spent longer in the study, they aged, had lived with CHD longer and were therefore more likely to develop co-morbidities[129]. Finally the QOF incentive scheme may have played a part as besides payments for establishing condition registers, patients with the co-morbidities shown, *RA* aside, could attract greater rewards for equal effort, as a number of targets were shared across multiple conditions. The effect of the introduction of the QOF on the recording and coding of a condition is most starkly demonstrated for *CKD*. This is a condition which was barely recorded prior to the QOF's introduction, started to increase noticeably thereafter; and increased around 15 fold in its first year of introduction into the QOF alone. By the end of the study period it was the most common co-morbidity in the study population. It is difficult to trust the figures in light of its extraordinary rise. There is however some

support for the pattern evident in these figures. Studies have shown a high prevalence of the condition, particularly in the elderly, which would account for its high prevalence following its introduction in the QOF[130]. At the same time there is an acknowledgement of difficulties in diagnosing CKD, and differences in opinion as to how to do so, and hence wide variations in diagnosis rates which could partly account for the pattern pre QOF entry[130 131]. Nonetheless it makes it difficult if not impossible to include it in the context of this study.

The figures in the Table 7-4 are also presented in Figure 7-1, though in this instance in terms of their prevalence in the CHD population. Prevalence is calculated as the number of people with a coding for the condition divided by the total number of patients in the study population that year. *CKD* is omitted due to the issues raised previously which make comparisons over the study period difficult. Vertical lines show when the QOF was introduced and when conditions of interest not present at the start of the QOF were added.

Figure 7-1 Prevalence of co-morbidities in the CHD study population



Every co-morbidity increases in prevalence over the course of the study period. This is not surprising for the reasons given previously. All the conditions are assumed to

be chronic, meaning that once diagnosed the patient had the disease throughout the study period. This is logical though for Atrial Fibrillation it is possible, albeit unusual, for it to be resolved, making this assumption false[132].

Lastly, co-morbidities in terms of numbers co-existing are shown in Table 7-5, with *RA* and *CKD* omitted from the figures. This shows that throughout the study most patients had none of the selected co-morbidities. However as the study progressed patients became increasingly co-morbid and moved into multi-morbidity states.

Table 7-5: No, co, and multi morbidity numbers

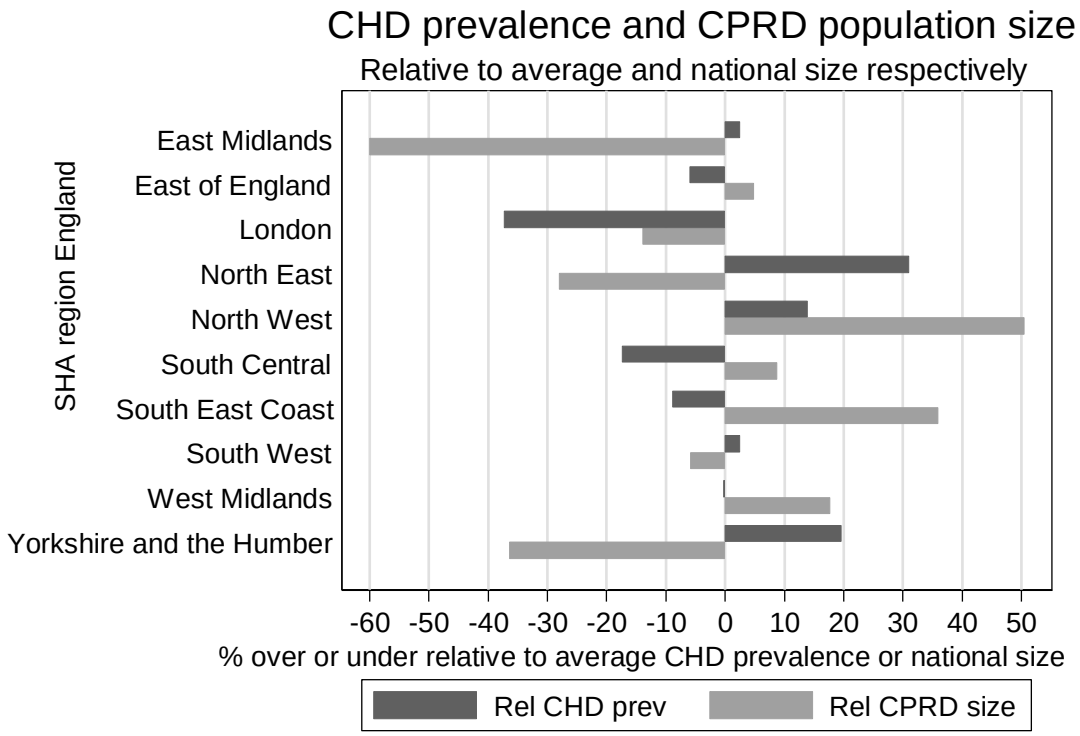
	Number of co-morbidities					
Year	0	1	2	3	4	5
2000/01	22,651	3,073	246	13	0	0
2001/02	27,134	6,064	722	70	2	0
2002/03	30,239	9,008	1,400	144	6	0
2003/04	30,814	11,580	2,237	278	20	0
2004/05	30,264	13,365	2,978	437	37	0
2005/06	29,326	14,660	3,677	634	58	1
2006/07	28,121	15,883	4,407	806	85	6
2007/08	27,470	16,567	4,847	920	98	8
2008/09	26,627	16,992	5,252	1,032	121	7
2009/10	25,921	17,280	5,560	1,150	143	10
2010/11	24,884	17,049	5,633	1,233	161	13

7.4.2.3 SHA Region

Figure 7-2 shows two things. Firstly as denoted by the Rel CHD prev bars, how much greater or less prevalent CHD is in that region in percentage terms, relative to the English national average in 2010/11[20]. Secondly as denoted by the Rel CPRD size bars the amount under or over represented the region is in the study population, in percentage terms, in relation to its population size based on 2011 census figures[133]. From this graph it is evident that the North East, North West and Yorkshire and the Humber SHAs have a much higher prevalence of CHD in relation to the national average. Conversely London has a much lower prevalence. In terms of the study population the North West and South East Coast are significantly over represented while the East Midlands and Yorkshire and the Humber are significantly

under represented. Even allowing for the higher prevalence of CHD in the North West this region remains significantly over represented in the study. Furthermore on this basis the under-representation of the North East, East Midlands and Yorkshire and the Humber would be considered worse still. While the South East Coast SHA would become increasingly over represented and the South Central SHA could be considered significantly over represented.

Figure 7-2: English regional SHA: CHD prevalence and CPRD population



7.4.2.4 Deprivation-Index of Multiple Deprivation

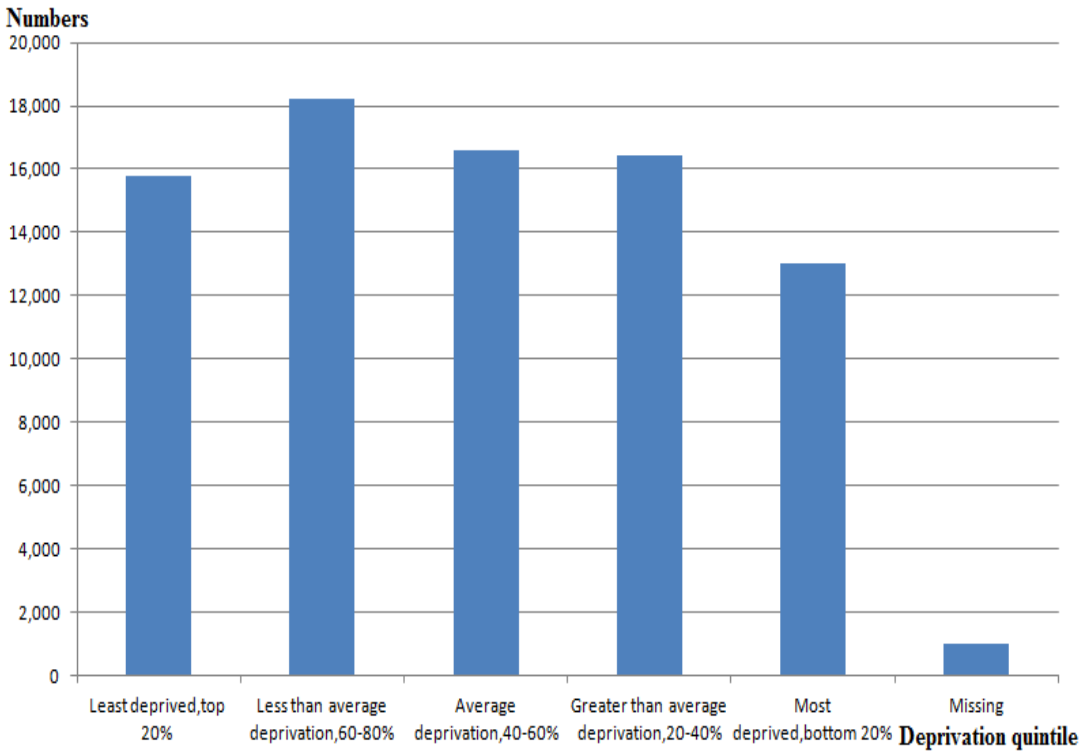
IMD was generated using the 2007 deprivation algorithms based on the most up to date information on the patient population at the time of request, March 2012. These are once only generated measures but were assumed to apply to the whole of the patient's participation in the panel.

If there were equal representation across the deprivation quintiles there would be approximately 16,204 (20%) in each column. While this is not the case the deprivation quintiles are fairly evenly represented. However there is noticeable under-representation of people from the lowest quintile, who account for just over 16% of the total population. In their place there is an over representation of those in

the less than average 60-80% quintile, accounting for nearly 22.5% of the population. A relatively small proportion of the population, 1.25%, had no IMD 2007 data and would need to be factored into those figures.

As these patients have been included into the study on the basis of existing CHD however they are not representative of the population as a whole. Since the main risk factors for CHD; an unhealthy diet, lack of exercise, obesity and smoking are more common in individuals in lower socio-economic groups; they would be expected to have a greater representation than those from less deprived quintiles. Previous research has also shown a higher prevalence of CHD in more deprived grouping using, individual and neighbourhood level data in Sweden[134] [135]. A link between higher deprivation and greater incidence of CHD has further been found at a local authority, practice, and electoral ward level in the UK[118] [136] [137]. This would be evident in Figure 7-3 by bars increasing in height from left to right, as deprivation quintile increases. Hence the study population on this measure is unlikely to be representative of the UK CHD population as a whole.

Figure 7-3: IMD 2007 deprivation numbers by quintile in the study population



7.4.3 QOF target variables and mapped variables

The following subsection looks at attainment on evidence based targets in the QOF over the study period and the underlying measures that mapped to those variables.

7.4.3.1 Diastolic and systolic blood pressure

Mean diastolic and systolic blood pressure level fell in most years, a finding supported by previous research discussed in Chapter 5[119 120]. These show that based on mean levels over a 15 month period, on average patients were meeting the QOF BP≤150/90mmHg target throughout the study period, comfortably so. Whilst in order to meet the BP QOF target the last reading over the 15 month period would have to be at or below that level, it nonetheless puts these figures in a context. Encouragingly the mean figures also met the more demanding levels suggested by the clinical guidelines of 140mmHg or below systolic and 85-90mmHg or below diastolic, for most of the study period[68 72].

Table 7-6 Mean systolic and diastolic blood pressure over the study period

BP	Systolic		Diastolic	
Year	Mean (mmHg)	SD (mmHg)	Mean (mmHg)	SD (mmHg)
2000/01	143.19	18.19	80.25	8.45
2001/02	141.03	18.76	78.63	9.05
2002/03	140.09	18.41	77.91	9.14
2003/04	138.96	17.96	77.25	8.99
2004/05	137.36	16.94	76.48	8.72
2005/06	136.21	16.35	75.98	8.67
2006/07	135.12	16.01	75.42	8.68
2007/08	134.61	15.87	75.12	8.72
2008/09	134.11	15.65	74.88	8.58
2009/10	133.83	15.63	74.63	8.64
2010/11	133.67	15.14	74.55	8.42

7.4.3.2 Total cholesterol

Figures represent the mean of all reading taken over a 15 month period, for all patients in the study population in the respective years, as well as their standard deviations. Mean cholesterol levels have been falling for the majority of the study period, though they appear to have reached a plateau. This conforms to the trend

found in previous research discussed in Chapter 5[119] [120]. While these figures do not take into account QOF business rules, they do indicate trends, based on which the mean patient reading, averaged over the QOF accounting period, met the QOF standard from 2002/03 onwards. They also show that based on yearly means, the average patient did not meet the more demanding target evidenced in the clinical guidelines, of 4 mmol/l or less, at any point over the study period[86] [87]. This suggests this may have been an unrealistic target, had it been implemented.

Table 7-7 Mean cholesterol levels over the study period

	Cholesterol levels	
Year	Mean (mmol/l)	SD (mmol/l)
2000/01	5.40	1.06
2001/02	5.04	1.05
2002/03	4.86	1.04
2003/04	4.69	1.01
2004/05	4.57	0.99
2005/06	4.42	0.96
2006/07	4.32	0.95
2007/08	4.28	0.95
2008/09	4.26	0.97
2009/10	4.26	0.98
2010/11	4.25	0.99

7.4.3.3 Smoking status

Table 7-8 shows the rates of the different smoking states in terms of their percentages relative to the total patient population in each year. The different smoking states map to the categories set out in the QOF business rule, and follow the codes set out in that guidance. While the recording of smoking status was a QOF target, as with the other recording targets such as blood pressure and cholesterol, the clinical benefits of interest to this study related to subsequent targets which used these records, in this instance smoking cessation advice. For most of the pre QOF study period, the majority of patients had no record of smoking status. The imminent and actual introduction of QOF incentives to record smoking status appears to have

had an impact on this figure. Since the QOF's introduction the no record figure, the missing column, has held reasonably constant in a range of 24% to 29%. The proportion of current smokers has been stable around the 10.8% to 12% mark.

Table 7-8 Smoking status, recorded and carried forward

Year	Never	Ex	Current	Missing	CurrentR	MissingR
2000/01	4.9%	6.2%	4.1%	84.9%	4.1%	84.9%
2001/02	5.3%	4.6%	2.9%	87.2%	5.0%	78.9%
2002/03	11.1%	11.8%	5.3%	71.8%	7.6%	61.6%
2003/04	22.1%	27.4%	9.8%	40.8%	11.5%	30.3%
2004/05	27.4%	35.9%	11.8%	24.9%	13.2%	12.5%
2005/06	25.0%	36.7%	11.6%	26.7%	13.2%	8.9%
2006/07	25.0%	37.3%	11.1%	26.6%	13.0%	8.0%
2007/08	24.9%	38.3%	11.1%	25.6%	12.8%	7.5%
2008/09	25.1%	38.9%	10.9%	25.0%	12.7%	6.8%
2009/10	25.2%	35.1%	10.8%	28.9%	12.7%	6.7%
2010/11	27.1%	36.7%	12.0%	24.2%	13.2%	5.1%

The figures relate to the QOF target, so smoking status had to be entered every year to qualify for payments. The same terms were applied within CPRD when modelling the QOF. The last two columns CurrentR and MissingR show what happens to those figures when missing values are replaced by records from previous periods, where one existed. In terms of current smokers the effect is more noticeable prior to the QOF, but in the QOF period would add around 1% to 2% to the population who are current smokers. The more noticeable impact is with respect to the two Missing columns. The difference between them represents the impact of carrying forward any of the smoking states. This approach would reduce the population with no record significantly, more noticeably as the study progressed. Clearly most of these would have a non-smoking status, and would therefore not be included in the QOF target were this approach adopted. The figures for current smoker used in the study are supported, certainly in the QOF period, by ONS survey data which shows a smoking rate in males and females aged 60 and over, of between 12 to 17% during the study period[138]. In a CHD population this figure would be expected to be lower.

7.4.3.3 QOF target measures

Table 7-9: Performance on study CHD QOF targets

	BP≤150/90mmHg		Chol≤5mmol/l		Anti plats or coags		Beta blocker		ACEI / A2		Smoking cessation	
Year	N	%	N	%	N	%	N	%	N	%	N	%
2000/01	14,241	54.8%	6,223	24.0%	9,862	38.0%	11,537	44.4%	1,851	49.4%	256	24.3%
2001/02	18,700	55.0%	9,713	28.6%	13,528	39.8%	15,824	46.6%	3,530	57.2%	202	20.3%
2002/03	24,510	60.1%	14,922	36.6%	19,528	47.9%	19,609	48.1%	5,268	61.8%	644	29.7%
2003/04	29,928	66.6%	21,294	47.4%	24,989	55.6%	22,744	50.6%	7,046	65.8%	2,748	62.5%
2004/05	35,753	75.9%	28,628	60.8%	29,699	63.1%	24,926	52.9%	4,697	71.4%	4,711	84.8%
2005/06	37,701	78.0%	31,309	64.7%	32,525	67.3%	26,212	54.2%	6,471	73.6%	4,883	86.8%
2006/07	38,886	78.8%	32,777	66.5%	34,267	69.5%	26,950	54.7%	8,112	75.5%	4,661	85.0%
2007/08	40,029	80.2%	33,732	67.6%	35,795	71.7%	27,628	55.4%	9,483	76.1%	4,564	82.2%
2008/09	40,086	80.1%	33,450	66.9%	36,673	73.3%	28,136	56.2%	10,629	76.6%	4,578	83.8%
2009/10	39,956	79.9%	32,535	65.0%	37,229	74.4%	28,238	56.4%	11,413	75.9%	4,473	82.6%
2010/11	41,391	84.6%	34,564	70.6%	38,789	79.2%	27,589	56.3%	11,934	74.9%	4,854	82.6%

Figure 7-4: Study Smoking cessation and surrogate outcome QOF targets attainment

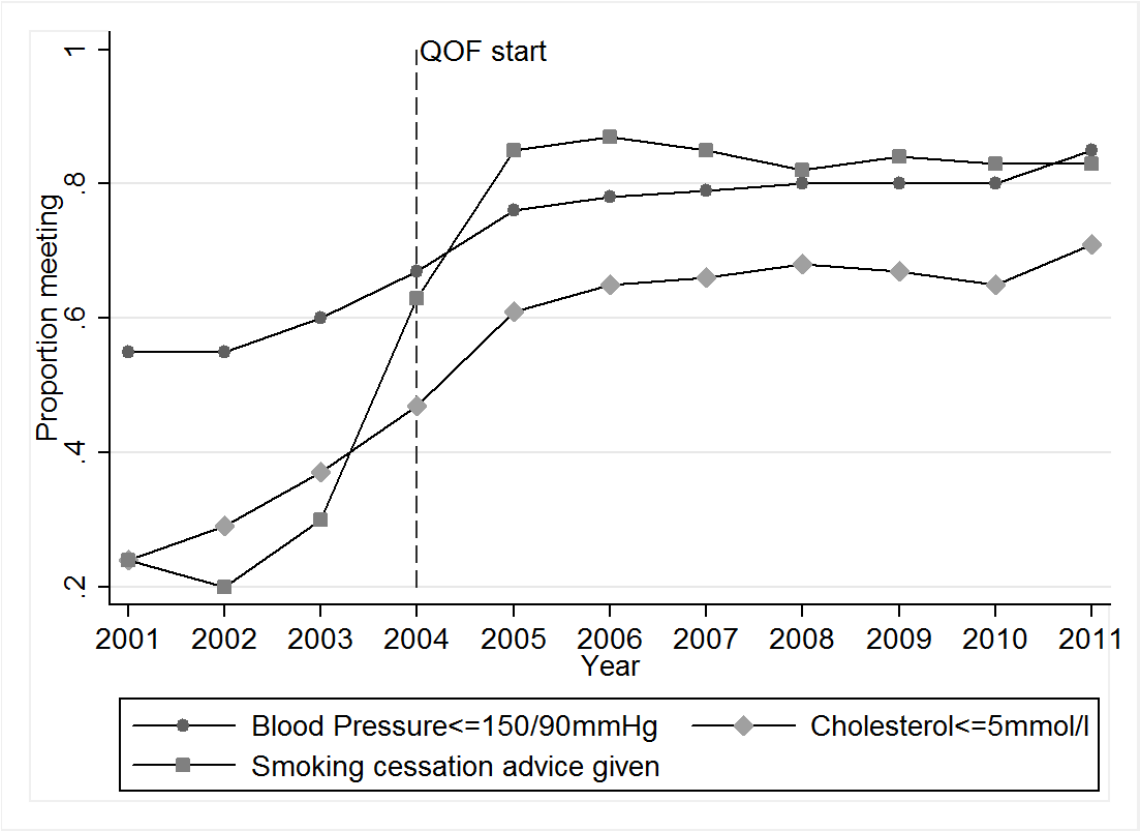


Figure 7-5: Study QOF CHD Prescription targets attainment

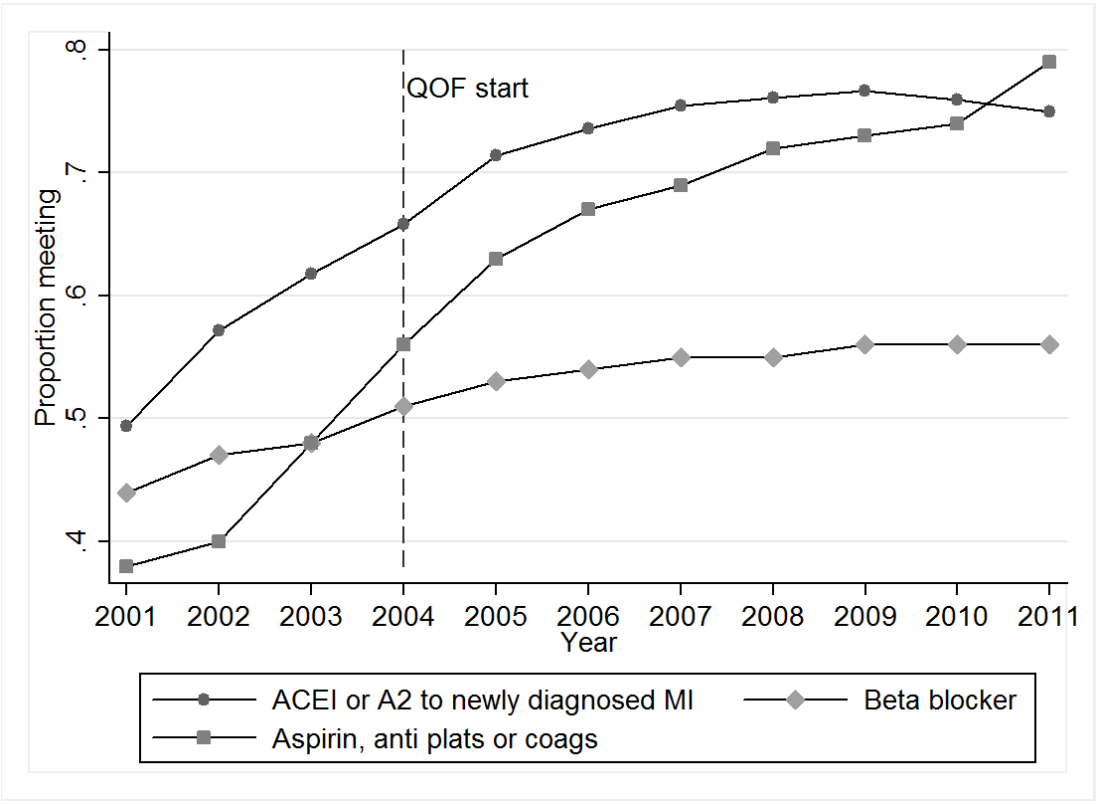


Table 7-9; Figures 7-4 and 7-5 show attainment on the QOF CHD targets that were included in this thesis.

The QOF variables reported are all binary variables, with zero representing a patient who missed the target or had no record, and 1, a patient who met the target. Numbers and percentages of patients meeting the targets are reported in the table, while proportions of the study population meeting the target are shown in the graphs. These measures are equivalent.

Table 7-9 and Figures 7-4 and 7-5 show improvements over the study period in patients meeting the targets. However a large amount of that improvement occurred prior to the QOF and the QOF built on an existing upward trend. That trend was maintained for all targets in the early years of the QOF though by 2006 a number of targets had reached a plateau. Only in the case of the aspirin, anti-platelet or coagulant target did attainment increase throughout the study period. In the final year of the study this target and that for blood pressure and cholesterol had a noticeable jump in attainment. Without further data it is impossible to determine if this is a blip before the resumption of the static or downward trend or represents the beginning of a new upward one. Taken as a whole these figures are indicative of national trends for the whole CHD target group shown in Figure 1-1.

When placed in the context of their starting point, those targets which had lower attainment at the study start saw the biggest improvements over the study periods. While the BP target saw a much more modest increase in relation to some of the other targets, it nonetheless started and ended the study as the target which had the highest level of attainment. The prescription of beta blockers had the lowest level of achievement by the study end due to it experiencing the smallest rise over the study period.

While informative the results will not be the same as those reported by the HSCIC on QOF attainment[139]. Firstly they are reported at a patient rather than a practice level, hence upper and lower thresholds have not been applied. Secondly exception reporting has not been taken into consideration as again this has an impact at the practice and not the patient level.

7.4.4 Secondary care outcome measures

As the outcome is a count variable Figure 7-6 shows the number of events as a percentage of the patients in the study population, by each study year (01 April to 31 March the following year). Hence the actual percentage of patients having an event will be lower still, albeit marginally. This highlights the fact that any admission, more so an MI hospitalisation, was rare within the study population and fell throughout the study period. Indeed even at its peak in the first year of the study, less than 10% of the population had any admission. This could present issues in any subsequent analysis, as an excessive number of zeros in a count model, no admissions in this model, can bias any results.

Figure 7-6 Secondary care all and MI admissions as a % of study population

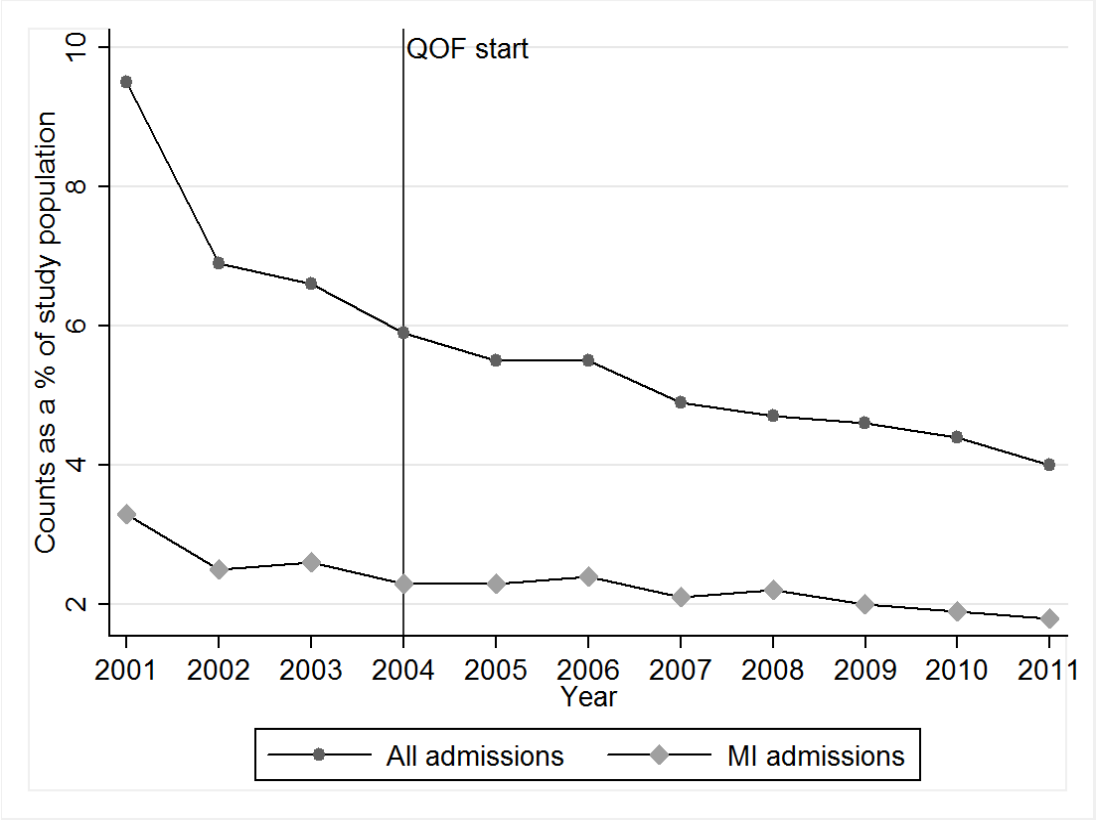


Table 7-10 looks at the figures for patients over their whole study period. These show that 75% of the study population had no secondary care outcomes of interest over the whole of their study period. Fewer than 20% had one admission, meaning that just over 5% of the population had more than 1 admission during their time in the panel.

Table 7-10: Hospital admissions over the study period

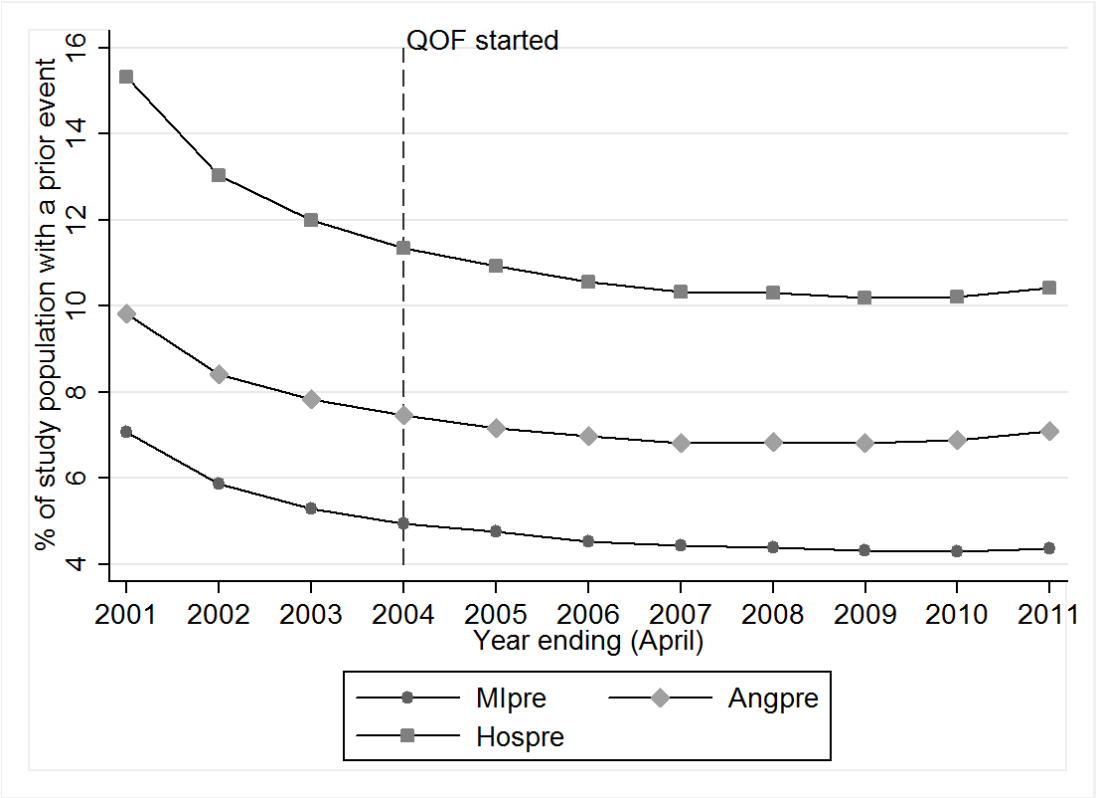
Admissions during the study period	Total	%
No admissions during the study	60,822	75%
1 admission during the study	15,847	19.6%
More than 1 admission during the study	4,353	5.4%

7.4.4.1 Outcomes prior to study participation

Patients did not enter the study population unless they had CHD diagnosed using a Read code recognised by the QOF, and met certain requirements to ensure any data recorded on them was up to a research standard. Prior to this however a number of patients had one or more of the study outcomes of interest, which were not included in the outcome variable as they did not occur after the patient qualified for the QOF in this study period. In the vast majority of cases these events led to an immediate CHD diagnosis, suggesting that a study outcome was the first time CHD was evident in these patients. A small number had an event following a first QOF recognised CHD diagnosis, however as this fell outside the study period, and since those patients did not join until the study start; these have been counted as pre study events, as they constituted a risk factor for outcomes in the study period. These will all have entered in year 1 of the study and hence the higher figure for that year needs to be set in that context. Small percentages of the study population also had multiple events, and were diagnosed some time later. Whether a patient had a study outcome prior to their participation in it is shown by the three variables in Figure 7-7 which relate to the outcome classifications discussed in the previous chapter. While these variables are by right count variables, they have been dichotomised in this study.

Figure 7-7 shows the percentage of patients in the study population that year, who had an outcome event (MI, emergency angina or either) prior to entry into it. Hence patients were included for each year they spent in the study in this figure, so long as they had one of the outcome events prior. The lines shows that the percentage of patients who had any of these outcomes measures declined over the study period, reaching a plateau around 2007.

Figure 7-7 Percentage of patients present in study who had an outcome prior to entry



7.4.5 Panel data analysis

As previously mentioned the data is a repeated sample on the same individuals over time, referred to as panel data. The panel is unbalanced because the population is not static with patients exiting and joining the study at different points in the study period.

7.4.5.1 Panel participation breakdown

Table 7-11 gives a breakdown of the most frequent patterns observed of the 81,022 patients who participated in the study. It shows for which of the eleven years of the study period they were present in the panel, with a 1 denoting presence, going from 2000/01 on the far left of the pattern column to 2010/11 on the right.

Table 7-11: Individual level panel participation, frequency and patterns

Freq.	Percent	Cum.	Pattern
12549	15.49	15.49	11111111111
4765	5.88	21.37	..111111111
4635	5.72	27.09	.1111111111
3991	4.93	32.02	...11111111
3650	4.50	36.52	1
3373	4.16	40.68	1111111
3366	4.15	44.84	11
3230	3.99	48.83	111
3195	3.94	52.77	1111
38268	47.23	100.00	(other patterns)
81022	100.00		XXXXXXXXXXX

This shows that the most frequently observed pattern is for a patient to spend all eleven years of the study period in the panel, though this accounts for just less than 15.5% of the study population. Around 32% of the study population spent 8 or more consecutive years on the panel as evident by the cumulative figure four rows down. The largest category is, 'other patterns', which are all the permutations other than those displayed. The ones of greatest interest within these are those with only 1 year of data as these will be lost entirely if a single lag is used. Individuals with 2 years of data in those circumstances would be lost in fixed effects analyses, as with only 1 year of data, they would have no means of measuring within variance. For the permutations shown these groups accounts for 8.65% of the study population, but other patterns not shown will certainly add to that figure.

7.4.5.2 Within, between and overall study variation

One of the benefits of using panel data is the introduction of time effects through the repeated sampling of the same individuals, enabling examination of the effects of a state change in that individual on a specified outcome. This makes any results potentially more powerful than say a cross sectional study design which uses one sample of individuals or compares different samples in the case of a repeated cross section. However for this benefit to be realised individuals have to change state. In

table 7-12 this is explored for the main explanatory variables that were not fixed over the study period.

Table 7-12: Study QOF variables variance: overall, between and within percentages

	Overall	Between	Within
Variable	Percent	Percent	Percent
BP: No	26.2	73.75	41
BP: Yes	73.8	92.05	75.79
Chol: No	42.96	88.75	55.23
Chol: Yes	57.04	78.92	64.6
BB: No	47	68.2	72.16
BB: Yes	53	66.69	76.16
Anti: No	36.07	61.52	57.43
Anti: Yes	63.93	82.33	78.55
Acea2q: No	34.06	53.35	73.17
Acea2q: Yes	65.94	71.61	85.13
Never smoked	21.59	40.41	47.98
Ex smoker	30.13	47.47	57.01
Current smoker	9.72	16.85	55.07
No record	38.56	87.72	50.47

Figures are presented for the overall population, between panels in the study and within the same individuals. The No and Yes refer to whether the individual missed or met the target, respectively; or in the case of the Smokstat variable whether they had a record of the specified smoking state or not. The overall column shows the percentages that did and did not meet the target for all the observations, person years, in the panel data. These can be viewed as pooled figures since they are a global figure for all patients. As these are expressed as percentages the figures sum to 100. Taking the BP figure for instance this shows that for 26.2% of all observations on this variable, patients failed to meet the target, and in 73.8% they did. The between column shows the percentage of patients who spent some time in that state. Using the figures for BP these show that 73.75% of patients failed to meet the target at some point during their time in the panel, while 92% of patients met the target for at least one of the periods they spent in the panel. Finally the within figures show,

conditional on the individual having a Yes or a No; the percentage of their observations that were a Yes or a No. Hence among those individuals who ever met the BP target, over 75% of their observations met the target. While those individuals who failed to meet the target did so for 41% of all their observations.

Overall the figures show variability among the patients and most importantly for fixed effects panel analysis, within patients over time. They also highlight a high level of state persistence, particularly for the prescription targets. In the case of the prescription of ACE or A2 post MI over 85% of patients who received an ACE did so for all their observations in the panel. This is to be expected as these were relatively easy to meet and it is important for chronically ill patients to be given medication in order to manage their condition, even more so if they have had an MI. However patients who did not meet the ACE inhibitor target, did not meet it for over 73% of all their observations. This is a smaller number but still a large figure, suggesting that a section of qualifying patients did not receive the recommend care for a large proportion of their time in the study. It is possible that these individuals were exempted from the QOF for some or all of those periods. For all the QOF targets patients were more likely to change from not meeting the target to meeting it, than to go from meeting it to not meeting it, as evidenced by the higher percentages of Yes in the within column relative to No.

For smoking status the figures show a high percentage of patients who did not have a record for all observations. Over 87% of patients had no record at some point in their study period and of those who had no record; over half of their observations showed no record. The most important figures are for current smokers which potentially offer some clues into the effectiveness of smoking cessation. Interpreting these and the figures for other states are however made more difficult by the high level on non-recording. They show that current smokers accounted for fewer than 10% of all observations; just under 17% of the study population had a current smoker status at any point in the study period; and those who smoked had a current smoking record for over 55% of their observations. This suggests patients may have been quitting but there was relatively high persistence among smokers during the study period.

This is explored further in Table 7-13 which shows transitions between the different smoking states. These show some movement from current smoker and ex-smoker to never smoked, which is impossible and therefore points to inaccuracies in data

recording. Thankfully the percentages are small, for smokers certainly, at 0.6%. The figures show that over 72% of smokers continued to smoke in the next period, with a similarly high figure for ex-smokers. Around 10% of patients who smoked were ex-smokers in the next period, pointing to some patient's choosing to cease smoking but still a persistently high percentage that did not. The small percentage of ex-smokers moving into the current smoking state, around 2.5%, and high percentage of ex-smokers who were ex-smokers in the next period, 75.7%, both point to patients successfully quitting and not taking up smoking again.

Table 7-13: Transitions between smoking states, percentage rates

	Never	Ex	Current	No record
Never	65.68%	6.19%	0.25%	27.88%
Ex	4.67%	75.69%	2.48%	17.17%
Current	0.6%	10.43%	72.7%	16.27%
No record	21.8%	22.33%	7.12%	48.76%

Finally the variation within and between patients on the outcome measure is explored in Table 7-14 for the all hospital admissions, *alladmis*, variable. The row of particular significance and the one which dominates the table is for those with no admission. These accounted for over 94% of all patient data years, and over 98% of patients had no admissions for at least one of their years in the study. Of those who ever had no admission, nearly 94.5% of their years in the study had no admission.

Table 7-14: All admissions, overall, between and within study variation

	Overall		Between		Within
alladmis	Frequency	Percent	Frequency	Percent	Percent
0	462756	94.55	79592	98.24	94.47
1	21675	4.43	17578	21.7	26.67
2	3706	0.76	3447	4.25	24.7
3	892	0.18	849	1.05	24.05
4	238	0.05	226	0.28	24.68
5	88	0.02	81	0.1	22.61
6	33	0.01	33	0.04	20.78
7	12	0	12	0.01	17.9
8	9	0	9	0.01	16.76
9	9	0	9	0.01	14.08
10	4	0	4	0	15.8
12	1	0	1	0	20
13	1	0	1	0	9.09

7.4.6 Key descriptive findings

The key descriptive findings from this section which inform this thesis are:

1. The population aged and the co-morbidities measured figured more prominently in the study population over the course of the study, as it became increasingly co-morbid and multi-morbid. These are atherosclerotic risk factors, expected to aggravate CHD and therefore will be importance variables in panel data models.
2. While the number of patients in the study population with a CKD coding increased nearly nine fold in the period prior to the QOF, it was still rarely recorded. The first two years of the QOF saw big increases, over 200 fold, in CKD records, but it was not until CKD's introduction into the QOF in 2006/07 that it became a highly prevalent; indeed the most prevalent by the study end, co-morbidity in the study population. This makes it difficult to include in the context of this study.
3. Mean systolic and diastolic BP, and cholesterol levels fell during most years in the study, in line with trends evidenced in previous research on QOF qualifying CHD patients.

4. The study population has an over-representation relative to its size nationally of individuals from the North West and South Central SHA's in particular, and is under-represented most significantly in the East Midlands and London SHA. Taking into consideration regional differences in CHD prevalence did not correct for this and in some instances made it worse.
5. Smoking status was poorly recorded particularly in the early years of the study when around 85% of individuals had no record. This situation improved significantly with, and immediately prior to, the introduction of the QOF. In the context of this study the percentage of current smokers, fell within a range of 10-12%, around 2-5% lower than that nationally in those aged 60 and over.
6. The percentage of individuals meeting the QOF targets increased over the study period, and for most years in the study, conforming to national trends.
7. The number of secondary care outcomes relative to the study population fell throughout the study period, conforming to national trends.
8. Over 75% of individuals had none of the study outcomes during their participation in the panel, and nearly 95% of those who did not have an outcome in one of their years in the study did not have an outcome for all of their years in the study. This means there will be a large amount of zeros in the outcome count variable with implications for analysis.
9. All of the QOF targets, specifically the prescription targets, had high percentage values for within variance meaning that there was little variation in whether the individual met or missed the target. This will have implications for fixed effects analysis.

7.5 Econometric analysis

7.5.1 Objective

This section conducts econometric panel data analyses to examine the impact of QOF CHD targets, and other relevant variables, on CHD clinical outcomes at an individual patient level.

7.5.2 Methods

The following structure will be used in the construct of economic panel data models:

1. There are a number of key explanatory variables that will be included in all regressions. These are all the QOF targets that apply to the whole CHD group as

they are the focus of the study, practice attainment, age and co-morbidities with the exception of CKD (see point 4). The non QOF *ACE* variable will be included in all regressions outside of those for the MI subgroup. ACE inhibitors have a number of cardiac benefits that may be realised in a wider CHD population and not just those with recently diagnosed MI.

2. The subgroup QOF targets, smoking cessation (*Smcess*) and ACE inhibitors, which apply only to individuals who smoke or with newly diagnosed MI respectively, will be included in separate subgroup analyses.
3. Time invariant variables, whose values do not change during the study: *gender, IMD, region, Angpre, MIpre, Hospre* will be included where model specification allows
4. Where there are alternative specifications of the same measure (covariates and dependent variables), e.g. whether someone is a current smoker, the specification used in the model will be chosen on the basis of: That which either is most significant, behaves according to prior expectations (suggesting a better specification), or where differences are negligible; that which there is most confidence in.
5. Due to concerns over the surge in use of CKD QOF codes follow the QOF's introduction and more so following CKD's introduction into the QOF, the decision on whether or not to include it will be based on preliminary model analysis
6. The remaining practice variables outside of practice attainment, namely size and workload will be included if there is evidence of significance in preliminary analysis
7. Random and fixed effects models will both be utilised. Random effects are important as a number of key variables such as *gender* and *IMD* are fixed over time. There is also a high level of individual time invariance in a number of the QOF targets variables. Fixed effects are important as models are rarely fully specified and therefore there are likely to be variables omitted from the model which are potentially correlated with included explanatory variables. Both specifications will have their results reported to enable comparison, and appropriate tests will be conducted to determine which specification best fits the model.

8. Population averaging will also be conducted, averaged at the individual level, using robust standard errors. These correct for heterogeneity in the variance of the error term.

Data analysis will follow a four stage process. These will start from initially specifying the model before moving into panel data regressions which will explore the impact of alternative specifications of the outcome variable. They are set out below:

Stage 1 Model specification

Stage 2 Analysis using panel count data models

Stage 3 Analysis of count data using a panel hurdle model

Stage 4 Logit panel data analysis using a binary transformed outcome measure

7.5.3 Stage 1 Model specification

7.5.3.1 Objective

1. To confirm the existence of time effects within the data to justify the use of panel data analysis
2. In the presence of the year variable and the key study variables to determine which outcome measure should be used and explanatory variables included

7.5.3.2 Background

Given the scale of the data under investigation cross sectional analysis provides a quick overview of the relationship between the explanatory and dependent variables. In the presence of time dummies these relationships are expected in large part to be borne out in more computationally complex models. It therefore provides an opportunity not only to flag up initial talking points and concerns but also importantly to guide model specification for subsequent stages

7.5.3.3 Methodology

A model was built around the QOF targets selected in chapter 3. All additional covariates were selected on the basis that they were expected to have an important explanatory relationship with the outcome (dependent) variable which contains the clinical outcomes identified in chapter 3. Further details about all these explanatory variables and their expected relationship with the dependent variable are given in

Appendix 8. In defining the model the process was constrained by the availability of data within the CPRD and HES datasets, and the quality of data collection in those databases. This meant that a covariate that ideally would be in the model, ethnicity, could not be due to a large amount of missing data. Other variables were constrained by the unit and frequency of measurement. Hence while data shows that death rates from CHD, and MI more specifically, do vary by strategic health authority in England, and that there is a general north-south divide with the northern SHA's having higher rates; this unit of measure hides significant pockets of variation in deaths from CHD[20]. As such while this variable could have an important explanatory role for differences in outcomes between individuals, this was dependent on those individuals being generally representative of those regions, of which we had no indication or control over. The results for the IMD variable, presented in figure 7-3, suggest that if this result is due to regional variations in deprivation, that the study population may not be representative.

Other variables such as workload and size were constructed by this study. In addition alternative specifications were attempted on a number of variables and the outcome measure. This section therefore undertook preliminary regressions to determine if these variables had an independent relationship with the study outcome, in the presence of key study variables, which would warrant their inclusion in the model. Or in the case of variables with alternative specifications, determine which, in instances if any, should be included.

The key variables included in all the analyses are specified below. The additional variables were added separately in the presence of those key variables, with the exception of *Smoker2* which was entered in the place of *Smoker*. The results presented in the tables show the regression coefficients for the individual variables in the presence of the key variables only. Negative binomial regressions were run as the outcome variables were over dispersed, meaning their variance was greater than their mean. The choice of the negative binomial over the Poisson was also verified by examining the value of the dispersion parameter, α, and the likelihood test on it, both of which confirmed over dispersion and the choice of the negative binomial.

Dependent (outcome) variable: *alladmis* (All ICD10, I21 and I23 codes and selected I20, 24 and 25 codes; for full breakdown of codes used see Appendix 7, page 279

and 280) ; *alladmisx* (including additional MI Read codes, see page 143 and 144 for details)

Key covariates (explanatory variables) included: *age; gender;* QOF variables {Blood pressure controlled≤ 150/90mmHg (*BP),* Cholesterol controlled≤ 5mmol/l (*Chol*), The prescription of Aspirin, anti-platelets or coagulations (*Anti),* The prescription of beta blocker (*BB),* The prescription of ACE or A2 to MI patients (*Acea2q*), Smoking cessation advice to current smokers (*Smcess*)}, QOF related target (*ACE),* Index of Multiple Deprivation 2007 *(IMD); Years;* Co-morbidities {Heart Failure (*HF*), Diabetes (*Diab*), Atrial Fibrillation (*AF*), Stroke and Transient Ischaemic Attack (*STIA*), Hypertension (*Hyper*), Rheumatoid Arthritis (*RA*)}*;* current smoker (*Smoker);* Practice level CHD QOF attainment *(pracatt)*

Additional covariates tested: *Smoker2,* SHA Region *(Region),* Angina outcome pre-admission *(Angpre),* MI pre-admission *(MIpre),* Outcome pre-admission *(Hospre),* Chronic Kidney Disease *(CKD),* Practice size measured in quintiles *(sizeq),* Practice GP workload measured in quintiles *(workldq)*

7.5.3.4 Discussion and results

7.5.3.4.1 Years

The first stage analyses ignored the data's panel structure and examined it as a single cross section. This was in order to study the impact of the year variable in the presence of the key variables and other possible explanatory variables to explore potential model specifications. Since panel data analysis considers time and individual unit effects, as denoted in equation (2), these effects play an important part in any panel analysis. A lack of time variance would favour a simple pooled regression, as the analysis would simplify down to an individual cross section if time played no part in outcomes. This preliminary analysis also provided an opportunity to see if there were any differences in results between the *alladmis* and *alladmisx* outcome measures.

The results are presented in a series of tables on the following pages, starting with the effect of time (years) on the outcome variables and key explanatory variables. The coefficients show log values, however since these are logs of counts, their interpretation is more difficult. Nonetheless they still provide an indication of effect,

with a negative value showing the variable reduces the outcomes, a positive the opposite, and the larger the value the bigger the effect.

Table 7-15 Effect of time period on study outcome measures

Year	alladmis	St. Error	alladmisx	St. Error
2001/02	-0.38***	(0.033)	-0.39***	(0.031)
2002/03	-0.37***	(0.033)	-0.38***	(0.031)
2003/04	-0.33***	(0.039)	-0.33***	(0.038)
2004/05	-0.28***	(0.048)	-0.27***	(0.046)
2005/06	-0.28***	(0.049)	-0.28***	(0.047)
2006/07	-0.41***	(0.050)	-0.40***	(0.048)
2007/08	-0.48***	(0.051)	-0.47***	(0.048)
2008/09	-0.55***	(0.051)	-0.56***	(0.049)
2009/10	-0.64***	(0.051)	-0.64***	(0.048)
2010/11	-0.73***	(0.053)	-0.72***	(0.051)

*** p<0.001, ** p<0.01, * p<0.05

Table 7-15 clearly shows that time, measured in years (01 April to 31 March the following year), had a statistically significant negative effect on the outcome variables, i.e. reduced their numbers, relative to the base year 2000/01, holding the key variables constant. In other words, adjusting for all the key variables, the outcome variable was declining over the study period. CHD statistics reported in Chapter 1 show supporting evidence for a declining trend in deaths and admissions for the study outcomes over the period of analysis[20]. This demonstrates the importance of considering time effects as cross sectional analyses could wrongly attribute the relationship shown to other included variables. The results suggest that there were large declines in relation to the base year in 2001/02, but that improvement stalled in the early years of the QOF, before figures improved again.

Overall they point to the importance of time on the study outcome measures, an effect included in panel data analysis. This is done by setting the data by time, and also in this study with respect to the individual, which allows for the effect of time to be examined in different contexts within this study. Firstly through introducing lags on variables it is possible to examine temporal effects and determine the impact of previous period variable values on present period outcomes. More generally looking at the individual unit level, in this case the patient, since the data is set by time,

depending on the method of analysis it is possible to see what happens within the individual or between individuals. Hence allowing for temporal effects the data tells us what the impact is when an explanatory variables changes value on the outcome at the individual level using within effects analysis, and what is the difference in outcomes between individuals when an explanatory variable changes value, in between effects analysis. In this setting what is happening to the outcome variable over the study period does not affect the validity of the results, since we are concerned with what is happening to outcomes within patients over time, or the effect of changes in covariate values between individuals, allowing for time effects.

7.5.3.4.2 Deprivation

The results for deprivation, one of the key explanatory variables, are shown in Table 7-16. Once again the results are very similar for both specifications of the outcome variable. Compared to the least deprived these results show that moving up the deprivation quintiles increases the expected number of admissions, increasingly so, with each increase in deprivation quintile. Hence being in the most deprived quintile increases the expected number of admissions relative to the next quintile, those with greater than average deprivation, and so on and so forth. All results are statistically significant, $p<0.001$. This results suggests that the more deprived the individual is, based on IMD 2007 metrics, the higher their rates of the study outcomes.

Table 7-16 Effect of deprivation quintile on study outcome measures

IMD 2007	alladmis	St. Error	alladmisx	St. Error
Less than average deprivation,60-80%	0.081***	(0.021)	0.082***	(0.020)
Average deprivation,40-60%	0.092***	(0.021)	0.082***	(0.020)
Greater than average deprivation,20-40%	0.17***	(0.021)	0.16***	(0.020)
Most deprived, bottom 20%	0.23***	(0.022)	0.22***	(0.021)
Missing	0.095	(0.059)	0.053	(0.057)

Compared to least deprived 20% Missing accounted for 1.22% of obs *** p<0.001, ** p<0.01, * p<0.05

7.5.3.4.3 Region

For the region variable which is based on the 10 SHAs there is no natural order to its categorisation. The order used was that in which it was placed by CPRD, with the North East used as the baseline comparator. Using other regions as the baseline

either due to their expected affluence or deprivation, did not give improved results. Furthermore it can be argued that the North East performs the role well given its expected relative deprivation in relation to the other regions. The coefficients for this variable, holding the key variables constant, are shown in Table 7-17.

Table 7-17 Effect of region on study outcome measures

SHA Region	alladmis	St. Error	alladmisx	St. Error
North West	0.11**	(0.038)	0.12***	(0.037)
Yorkshire & The Humber	0.047	(0.044)	0.046	(0.042)
East Midlands	0.080	(0.051)	0.055	(0.048)
West Midlands	0.090*	(0.040)	0.099*	(0.038)
East of England	0.14***	(0.041)	0.13***	(0.039)
South West	0.14***	(0.041)	0.12**	(0.039)
South Central	0.16***	(0.042)	0.12**	(0.040)
London	0.29***	(0.042)	0.29***	(0.040)
South East Coast	0.26***	(0.042)	0.23***	(0.040)

Compared to the North East *** p<0.001, ** p<0.01, * p<0.05

Surprisingly therefore the coefficients show that other regions are associated with an expected increase in admissions relative to the North East. This effect is highly significant in the Southern regions producing a reverse of the usual north south divide. When interpreting region in this way it is assumed that it is a measure of regional deprivation but as the figures are for SHA, they can hide significant variation and are possibly a blunt tool in this respect. At the same time it may be replicating the work done by the *IMD* variable and therefore there may be correlation between the two giving misleading results. This was not apparent in the *IMD* variable which remained highly significant with regions included, though the coefficients were smaller and the staged increases through the quintiles not as marked. Therefore the regression was run with *IMD* absent. In this instance the sign of the coefficients did not change but their size reduced while their standard errors increased. As a result the coefficients for the East of England and South Central were no longer significant and the significance level for the South West reduced to 5%. This did not suggest high correlation, which was confirmed by further analysis which showed a correlation coefficient between the two variables of, -0.2224. What this perhaps highlights is that due to the broad geographical categories there is

potentially a relationship between region and deprivation for certain patients accounting for its effects on that variable. However since the categories contain pockets of affluence and deprivation it does not form a good measure of relative deprivation across the regions of England. It may be capturing regional differences in access to and quality of health and wider social services. Or the fact that relative to living costs incentives in these regions will have been less generous than the North East. While its interpretation is presently unclear, the significant differences between some regions are worth exploring further to see if consistencies occur and patterns emerge, which could aid interpretation.

7.5.3.4.4 Sizeq and workldq

These were included as quintiles as the coefficients were near zero and insignificant when they were included as non-categorised variables. Both are attempts to capture the impact of practice level factors on individual patient outcomes.

Table 7-18 Effect of size quintile and workload quintile variables on *alladmis* outcome variable

Category	Workload	St. Error	Practice size	St. Error
Bottom 20%-40%	0.012	(0.021)	-0.012	(0.021)
Mid, 40-60%	0.025	(0.021)	-0.063**	(0.021)
Top 60-80%	0.038	(0.021)	-0.030	(0.021)
Highest 20%	0.077***	(0.021)	-0.0078	(0.021)

Compared to bottom 20% *** p<0.001, ** p<0.01, * p<0.05

The results are largely non-significant in both instances with the exception of those practices with the highest workload and practices in the mid range for size. The coefficients are however the correct size and increase in magnitude as we go through the quintiles as expected, for workload certainly. Results are presented with respect to the *alladmis* outcome variable only, though the coefficients for the *alladmisx* variable were similar, and their statistical significance identical. These results suggest that while it may not be evident in each quintile, relative practice size and workload does have an effect on the study outcome, which merits exploration.

7.5.3.4.5 Remaining variables

The results from the remainder of the variables are shown in the Table 7-19.

Table 7-19 Remaining variables, effect on outcome measures

VARIABLES	alladmis	St. Error	alladmisx	St. Error
BP≤ 150/90mmHg	-0.11***	(0.016)	-0.19***	(0.015)
Anti coag, platelet or aspirin	0.62***	(0.016)	0.57***	(0.015)
BB prescribed	0.38***	(0.014)	0.36***	(0.013)
Cholesterol ≤ 5 mmol/l	-0.30***	(0.015)	-0.38***	(0.014)
ACE or A2	0.48***	(0.014)	0.48***	(0.014)
age	0.0039***	(0.00062)	0.0028***	(0.00059)
gender	0.091***	(0.014)	0.11***	(0.013)
Heart Failure	0.64***	(0.021)	0.66***	(0.020)
Atrial Fibrillation	0.21***	(0.023)	0.22***	(0.022)
Rheumatoid Arthritis	0.12*	(0.060)	0.15**	(0.057)
Chronic Kidney Disease	0.039	(0.023)	0.033	(0.022)
Hypertension	-0.066***	(0.018)	-0.081***	(0.017)
Stroke TIA	0.16***	(0.027)	0.18***	(0.026)
Diabetes	0.047*	(0.022)	0.039	(0.021)
Practice CHD attainment	-0.013***	(0.001)	-0.016***	(0.001)
Pre study period MI admis	0.21***	(0.027)	0.19***	(0.026)
Pre study period IHD admis	0.93***	(0.020)	0.85***	(0.019)
Current smoker	0.11***	(0.023)	0.078***	(0.022)
Current smoker forward ±	0.067**	(0.022)	0.022	(0.021)
Pre study hospital admission┼	0.77***	(0.017)	0.70***	(0.017)

*** p<0.001, ** p<0.01, * p<0.05 ┼ When used in place of pre study MI and IHD admissions Observations 489,384
± Carries forward previous smoking state where information is missing, used in place of current smoker variable

Firstly with regard to the different versions of the explanatory variables (*alladmis* and *alladmisx*), results in terms of coefficients and significance are very similar; as has been evident in the other tables produced. Looking further into the results the coefficients for all the co-morbidities are postive and significant as expected with the exception of CKD, perhaps unsurprisingly due to issues raised previously. The QOF outcome measures BP and Cholesterol are the expected sign, namely meeting the target reduced the number of admissions and significant but all the prescription targets are positive suggesting that receiving these increased admissions. This relationship will have to be explored further within panel data analysis to look at

potential reasons. Having an outcome event prior to the patients participation in the QOF was a postive and significant predictor for future outcomes. This could support some form of hurdle structure. This was the case whether the variables were specified separately by MI or Angina or merged to create the *Hospre* variable, placing it on the same basis as the outcome variable. Different specifications of the smoking variable were included, one using codes recorded in CPRD, Smoker, and the second, *Smoker2*, by carrying a smoking record forward into subsequent periods, if details were missing. Both of these variables had the expected sign and were significant.

7.5.3.4.6 Structure of the count variable

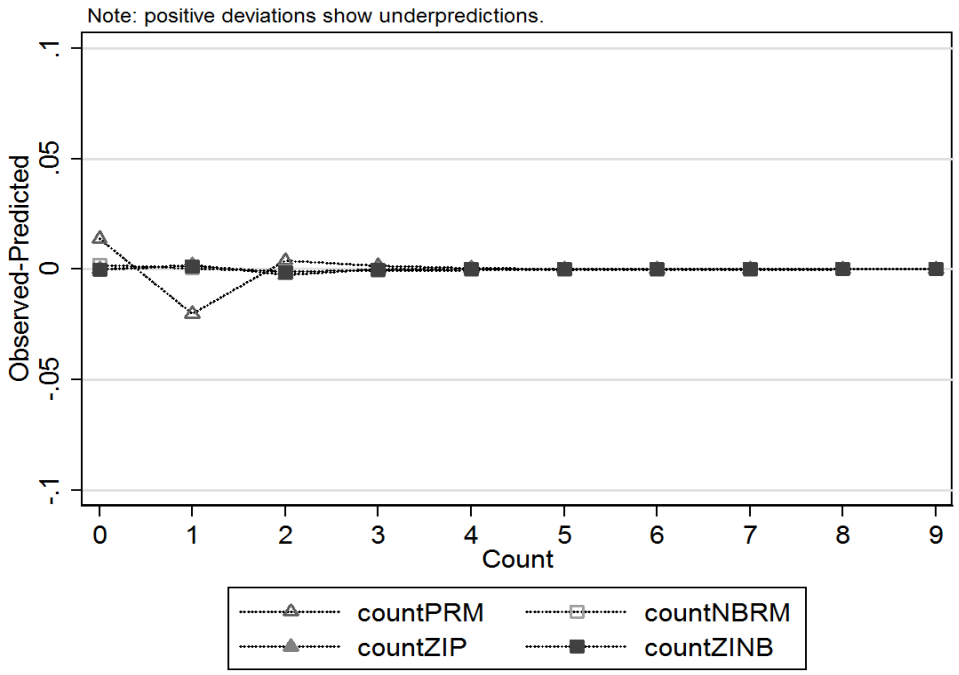
The final analysis in this stage looked at various model specifications incorporating the key explanatory variables and the *alladmis* outcome measure using the countfit command in Stata. This commands examines whether a Poisson or negative binomial is the best fit for the model and whether zero inflated models should be used. This command does not look at hurdle models but the results can be used more generally to see if the model has an excess of zero outcomes which some form of excess zero correction could improve upon. The results are also not applied to a panel structure though years were included in the model specification. The summary results of the output are shown in Table 7-20 and Figure 7-8.

The results in Table 7-20 show a strong preference for the negative binomial (countNBRM) over the Poisson (countPRM) confirming the results from other checks, and for a zero inflation correction. The effectiveness of the different model specifications in fitting the data are shown in Figure 7-8 below that table. This shows that the negative binomial is better at predicting the value of the count variable *alladmis* than the Poisson for values between 0 and 2, with the Poisson under-predicting zeros and overpredicting ones. Overall both the non zero inflated negative binomial and the zero inflated models for both specifications appear to be a good fit for the data when interpreting the graph only. However the results from Table 7-20 argue strongly for a zero inflated approach and the use of a zero inflated negative binomial (ZINB) in preference to a zero inflated Poisson (ZIP).

Table 7-20 Preferred model specification

countPRM	BIC=-6.160e+06	AIC= 0.513	Prefer	Over	Evidence
vs countNBRM	BIC=-6.176e+06	dif= 15671.430	NBRM	PRM	Very strong
	AIC= 0.481	dif= 0.032	NBRM	PRM	
	LRX2=15684.531	prob= 0.000	NBRM	PRM	p=0.000
vs countZIP	BIC=-6.179e+06	dif= 18879.724	ZIP	PRM	Very strong
	AIC= 0.474	dif= 0.039	ZIP	PRM	
	Vuong= 36.433	prob= 0.000	ZIP	PRM	p=0.000
vs countZINB	BIC=-6.180e+06	dif= 19723.409	ZINB	PRM	Very strong
	AIC= 0.473	dif= 0.041	ZINB	PRM	
countNBRM	BIC=-6.176e+06	AIC= 0.481	Prefer	Over	Evidence
vs countZIP	BIC=-6.179e+06	dif= 3208.294	ZIP	NBRM	Very strong
	AIC= 0.474	dif= 0.007	ZIP	NBRM	
vs countZINB	BIC=-6.180e+06	dif= 4051.979	ZINB	NBRM	Very strong
	AIC= 0.473	dif= 0.009	ZINB	NBRM	
	Vuong= 2.955	prob= 0.002	ZINB	NBRM	p=0.002
countZIP	BIC=-6.179e+06	AIC= 0.474	Prefer	Over	Evidence
vs countZINB	BIC=-6.180e+06	dif= 843.685	ZINB	ZIP	Very strong
	AIC= 0.473	dif= 0.002	ZINB	ZIP	
	LRX2= 856.786	prob= 0.000	ZINB	ZIP	p=0.000

Figure 7-8 Predictive accuracy of different model specifications



7.5.3.5 Summary and conclusion

There is no evidence with respect to any of the variables examined to use the *alladmisx* outcome variable in preference to *alladmis*. Due to this it was decided to focus on the original *alladmis* variable for subsequent analysis as this contains known hospital admissions for the codings of interest, and there is no additional benefits from including further possible but uncertain events. Likewise nor does *Smoker2* improve upon *Smoker* in terms of its fit with other variables and relationship with the outcome variable. For this reason *Smoker* is retained for subsequent analysis as this uses only actual recorded entries. *CKD* is not included, not simply on the lack of significance in this analysis but also due to wider concerns with the consistency of its coding over the study period. *Hospre*, which incorporates both the emergency angina and MI pre-study outcome measures variables will be used rather than splitting it down into its component variables. The analysis confirmed that both are significant in increasing the count of the outcome variable in the study period, an effect that was captured in this one variable. *Region*, *sizeq* and *workldq* are also included. These were found to be significant for certain regions and quintiles, in relation to the base measure, and therefore may play some role in explaining study outcome. For brevity and clarity of presentation only those categories that differ significantly, $p \leq 0.05$, from the baseline comparator will be reported in output tables.

Therefore the model to be used in all subsequent analysis is as follows.

For whole population analyses:

Dependent variable: Original outcome measure (*alladmis)* in count models, or the binary derivation of the same outcome measure, in binary models (*Admis)*

Explanatory variables (covariates): QOF variables {Blood pressure control≤ 150/90mmHg (*BP),* Cholesterol control≤ 5mmol/l, The prescription of Aspirin, anti-platelets or coagulations (*Anti),* The prescription of beta blocker (*BB)*}; QOF related variable, the prescription of ACE inhibitors or A2 antagonists (*ACE*); Practice size measured in quintiles (*sizeq);* Practice GP workload measured in quintiles (*workldq)*; Index of Multiple Deprivation 2007 (*IMD*); Current smoker (*Smoker*); SHA region (*Region);* Co-morbidities {Heart Failure (*HF*), Diabetes (*Diab*), Atrial Fibrillation

(*AF*), Stroke and Transient Ischaemic Attack (*STIA*), Hypertension (*Hyper*), Rheumatoid Arthritis (*RA*)}

In those analyses looking at the MI subgroup:

The additional QOF variable will be included, the prescription of ACE or A2 antagonists (*Acea2q*), which is applied in the QOF to that subgroup. The *ACE* variable which applies that target to the whole of the CHD population (*ACE*) will be removed in these analyses.

In those analyses looking at the smoking subgroup:

The QOF target of offering smoking cessation advice to smoker will be included (*Smcess*), and the current smoker variable (*Smoker*) removed, as there will be collinearity with the added variable.

7.5.4 Stage 2: Analysis using count panel regression analysis

7.5.4.1 Introduction

A benefit of regression methods that consider time is the ability to look at temporal effects through the imposition of time lags. The downside of such an approach is that data is inevitably lost as patients with only one year of data are removed from the analysis when lags are introduced. Introducing lags of QOF explanatory variables makes intuitive sense in the context of this study. Patients could join the study at any point in a QOF year following a first recorded CHD QOF code. Likewise they were assumed to be QOF qualifying for a full year so long as they were present at the beginning of that QOF year. In both cases lags become important, though for different reasons. When joining not only may there not be enough time to measure an impact depending on the timing of their first coding but naturally there is a period of treatment initiation and then compliance before the intervention has an impact on the condition and reaps clinical benefits. Likewise as the outcome can occur at any point in the QOF year, and hence prior to the administration of the evidence based QOF targets, depending upon its timing, treatment in the previous year may be more pertinent to the outcome in that year.

Regression without time lags were run, these produced very similar results to those reported in Stage 1 as expected. The more interesting results are for those in which lags were introduced that are presented below. Random effects, fixed effects and

population average with robust standard errors were performed on the same model specification, dropping time invariant variables for the fixed effects analysis. Any variable with a suffix -1, refers to its one year lagged value. Results for deprivation and practice size are shown in relation to the lowest quintile, the least deprived and smallest practices respectively. SHA region and practice GP workload were also included but values are only shown for those regions and quintiles where the results were signifcant. N represents the number of patients who were included in that regression and T, the number of time periods. In all tables the evidence based QOF variables, identified in chapter 3, are shown at the top of the output table followed by age and gender, and then the remaining study explanatory variables.

7.5.4.2 Whole population targets

The results for whole population targets are shown in Table 7-21. The coefficients show incident rate ratios (IRR), which are easier to interpret than the previous results which showed the difference between the logs of expected counts. For those coefficients a positive figure indicated that the variable increased the number of counts, a negative that it decreased the number of counts, though the magnitude of those effects is difficult to interpret when expressed in logs. For IRR's shown in the Table 7-21, a figure greater than one means that a unit increase in the variable increases the rate of counts, those less than one that it reduced them. Hence looking at the coefficient for gender for random effects this shows that holding other variables constant being male compared to female is expected to increase the rate of *alladmis* counts by a factor of 1.086. IRRs are presented for all analyses where the outcome variable is specified as a count variable.

Table 7-21 Whole population QOF targets, Negative Binomial

Covariates	Random Eff IRR	St. Error	Pop Ave IRR	St. Error	Fixed Eff IRR	St. Error
Blood Press	1.042*	0.020	1.049*	0.021	1.042	0.024
Blood Press-1	0.915***	0.017	0.905***	0.017	0.932***	0.020
Anti-coags	2.868***	0.071	2.835***	0.081	2.446***	0.074
Anti-coags-1	0.612***	0.013	0.611***	0.015	0.623***	0.016
Beta block	1.943***	0.048	1.863***	0.068	1.929***	0.054
Beta block-1	0.627***	0.015	0.657***	0.024	0.660***	0.018
Cholesterol	0.917***	0.016	0.913***	0.017	0.956*	0.020
Cholesterol-1	0.830***	0.015	0.839***	0.015	0.872***	0.018
age	1.015***	0.001	1.014***	0.001	1.001	0.003
gender	1.086***	0.022	1.106***	0.024		
Heart Failure	2.098***	0.055	1.987***	0.059	2.020***	0.093
Atrial Fibril	1.285***	0.038	1.242***	0.040	1.482***	0.074
Rheum Arth	1.137	0.089	1.109	0.100	1.439*	0.242
Hypertension	0.902***	0.022	0.902***	0.023	1.072	0.062
Stroke TIA	1.213***	0.041	1.195***	0.043	1.226**	0.077
Diabetes	1.121***	0.031	1.119***	0.035	1.111*	0.053
ACE	1.804***	0.043	1.700***	0.060	1.665***	0.045
ACE-1	0.711***	0.017	0.735***	0.025	0.622***	0.017
Practice CHD attainment	0.986***	0.001	0.987***	0.007	0.996***	0.009
Outcome pre-admission	2.532***	0.063	2.480***	0.069		
Smoker	1.199***	0.034	1.184***	0.037	1.058	0.043
Practice size 60-80% quint	0.956	0.028	0.970	0.032	0.802**	0.065
Practice size largest 20%	0.957	0.029	0.981	0.032	0.827*	0.074
IMD 20-40% quintile	1.131***	0.035	1.140***	0.037		
IMD 40-60% quintile	1.158***	0.036	1.173***	0.040		
IMD 60-80% quintile	1.225***	0.038	1.276***	0.044		
Most deprived quintile	1.447***	0.049	1.455***	0.053		
N/T	72661/10		72661/10		12709/10	

*** p<0.001, ** p<0.01, * p<0.05 Region and workldq also included in models but effects were insignificant, size quintiles only reported where significant

The results in terms of their size and significance are similar across the different model specifications. Clearly introducing a lag makes a difference and all QOF

targets, and the additional ACE prescription target applied to the whole population, reduce rates of admissions, significantly so, $p<0.0001$, when lagged by one year. This effect is most noticeable for aspirin, anti-platelets and coagulants, where meeting the target in the previous period is seen to reduce the rates of the outcome by nearly 40% (IRR=0.611 to 0.623). With the exception of the cholesterol target the results suggest that meeting the QOF target increases the rate of admissions in the present year. This relationship is harder to interpret. It could be that immediately following an admission the patient is more intensively managed and hence meets the targets and this effect is being captured in the same year panel. Meeting the target in the previous year, in this instance, has a more causal relationship with events subsequently, which is being accurately captured by the lag.

All co-morbidities are significant with the exception of rheumatoid arthritis, and increase rates of admission with the exception of hypertension, which only does so in fixed effects analysis. Increasing practice GP workload, size, as well as SHA region were insignificant in these analyses, and hence are not produced in the table. Having an admission for the outcomes of interest prior to entry into the study increased the rates of admissions by a factor of around 2.5 in non-fixed effects analyses (IRR=2.480 to 2.532, $p<0.001$). Current smoking also increased rates of outcomes as expected by near 20% in random effects (IRR=1.199, $p<0.001$) and population averaging (IRR=1.184, $p<0.001$), and 5% within patient using fixed effects analysis (IRR=1.058), a result that was non-significant. There was a significant and linear stepped increase in rates of outcomes as the patients moved into more deprived quintiles, with the most deprived quintile having around a 45% increase in rates of outcomes (IRR=1.447 random effects; 1.455 population average, $p<0.001$). This variable was not included in fixed effects analysis as it was fixed over the study period for the individual patient, and therefore time and would be removed by fixed effects adjustments. Practice CHD QOF attainment significantly reduces rates of admissions ($p<0.001$), by 0.4% for each 1% increase in attainment in fixed effects analysis (IRR=0.996), and around 1.4% for each 1% increase in attainment in random effects (IRR=0.986) and population averaged analyses (IRR=0.987).

A Hausman test was conducted which looks at the differences in the variance of the fixed and random effect estimators to indicate which specification is most efficient.

This test came out in favour of fixed effects estimation whose results are reported, time constant variables aside. The fixed effects coefficients need to be interpreted differently as they consider only within individual variation and not that between individuals which is also included in the random and population average coefficients. Hence gender which does not differ within an individual is lost in fixed effects analysis, while in the random and population averaged analysis the coefficient refers simply to the differences between male and female patients in the study. Alternatively for a variable such as Heart Failure (*HF*) which does change within patients during the study, the fixed effects regression coefficient shows that having an *HF* diagnosis significantly increases the individual rate of admissions by just over 2 fold (IRR=2.020, $p<0.001$). The random effects coefficient on the other hand shows that an individual diagnosis of *HF*, and an individual with *HF* compared to one without, significantly increases the rate of outcomes by a factor of nearly 2.1 (IRR=2.098, $p<0.001$). Encouragingly in fixed effects the results are similar and the QOF targets remain significant with a lag, as does prescriptions of ACE or A2. The results for co-morbidities are also similar although *RA* is now a significant risk factor (IRR=1.439, $p<0.05$) and hypertension is associated with an increase in rates (IRR=1.072), though is insignificant. Age also becomes insignificant in fixed effects (IRR=1.001), while patients in larger practices (largest two quintiles) have around a 20% lower rate of admissions compared to the smallest practice size quintile, $p<0.01$. A 1% increase in practice attainment reduces rates of admissions by 0.4% for the individual. These results are however based on a much smaller dataset involving 12,709 patients, with large numbers of observations and patients lost, in addition to those through lags, to individual time invariance in included time variant variables, predominantly the QOF targets.

7.5.4.3 Sub group CHD targets: MI

The coefficients for the MI subgroup targets largely mirror those for the whole population, see Table 7-22. The ACE inhibitor or A2 prescription QOF target applied to its designated QOF population is significant in all models when lagged, reducing rates of admission by around 27% (IRR=0.726, $p<0.001$) in fixed effects analysis, and around 15% in non-fixed effects analyses (IRR=0.854, $p<0.001$ random effects; 0.881, $P<0.05$ population averaged) in the subsequent period, when met in the present period. The reduction in incident rate ratios is lower than that seen

when the target was applied to the whole population. Rates do not increase incrementally with each increase in deprivation as witnessed in the whole study population, with there now being no significant difference in rates of outcomes between the least deprived 20% and those in the least deprived 20-40%. However more generally increased deprivation is seen to increase rates of admissions, by around 33% for the most deprived individuals compared to the least deprived (IRR=1.324-1.329, $p<0.001$). This variable is fixed over the study period and therefore could not be run in fixed effects analyses, and the impact of changes in an individual's deprivation status on the study outcome could not be examined. In fixed effects besides Heart Failure (IRR, 2.131) and Hypertension (IRR, 1.391) none of the co-morbidities increase rates of outcomes significantly. Hypertension is no longer significant outside of fixed effects analysis (IRR=1.39,$p<0.05$), the reverse of the findings found in the whole population, suggesting that in this population having hypertension does not increase admissions between individuals, but it does within a patient who goes from being non-hypertensive to hypertensive. While none of the size quintiles are significant in any of the analyses, individuals in practices where GP workload was in the lowest 20-40% quintile had significantly higher rates of outcomes than those in the lowest workload quintile in all analyses, of between 12% and 20% (IRR=1.119-1.120, $p<0.05$). The only remaining difference between this group and the whole study population is that there is no evidence for male patients being at higher risk of having an outcome than female patients in the MI subgroup.

Table 7-22 MI subgroup, Negative Binomial

Covariates	Random Eff IRR	St. Error	Pop Ave IRR	St. Error	Fixed Eff IRR	St. Error
Blood Press	1.038	0.043	1.048	0.045	0.984	0.056
Blood Press-1	0.852***	0.033	0.834***	0.034	0.882***	0.047
Anti-coags	2.277***	0.127	2.220***	0.136	1.652***	0.145
Anti-coags-1	0.647***	0.029	0.658***	0.031	0.630***	0.040
Beta block	1.584***	0.081	1.597***	0.111	1.722***	0.121
Beta block-1	0.683***	0.034	0.691***	0.048	0.715***	0.047
Cholesterol	0.816***	0.029	0.823***	0.030	0.854***	0.044
Cholesterol-1	0.782***	0.029	0.781***	0.030	0.812***	0.040
ACE/A2	1.081	0.054	1.056	0.066	1.090	0.076
ACE/A2-1	0.854**	0.042	0.881*	0.054	0.726***	0.050
age	1.009***	0.002	1.009***	0.002	0.973**	0.008
gender	0.990	0.039	1.016	0.042		
Heart Failure	1.880***	0.091	1.770***	0.086	2.131***	0.258
Atrial Fibril	1.158*	0.068	1.160***	0.072	1.562	0.234
Rheum Arth	1.214	0.174	1.152	0.151	1.753	0.884
Hypertension	0.970	0.045	0.933	0.045	1.391*	0.230
Stroke TIA	1.211**	0.082	1.238**	0.088	1.247	0.213
Diabetes	1.179**	0.064	1.175**	0.068	1.060	0.146
Workld: lowest 20-40% quintile	1.120*	0.061	1.119*	0.063	1.196*	0.100
IMD 20-40% quintile	1.056	0.062	1.075	0.065		
IMD 40-60% quintile	1.185**	0.070	1.191**	0.074		
IMD 60-80% quintile	1.128*	0.068	1.154*	0.075		
Most deprived quintile	1.324***	0.086	1.329***	0.090		
Practice CHD attainment	0.990***	0.001	0.991***	0.001	0.985***	0.003
Outcome pre-admission	2.392***	0.122	2.334***	0.125		
Smoker	1.232***	0.064	1.244***	0.069	1.102	0.098
N/T	26744/10		26744/10		2997/10	

Region, workldq and sizeq reported where significant. *** p<0.001, ** p<0.01, * p<0.05

7.5.4.4 Smoking subgroup, Smoking cessation target

The results for the smoking subgroup are shown in Table 7-23. Of specific relevance to this group, smoking cessation is seen to reduce rates of each admission by around

25%, a result that is significant in all models (IRR=0.728-0.745, p<0.01). However in this instance in the present period and not when lagged. With the exception of the blood pressure target, all the remaining QOF targets are seen to significantly reduce rates of outcomes after a lag, as does the ACE inhibitor target when applied to this population. Higher practice attainment significantly reduces the rates of patient admissions in random effects and population averaged models (IRR=0.983-0.984, p<0.001), though this effect is not significant in fixed effects analysis (IRR=0.992). Having an outcome prior to admission remains a significant factor in increasing rates of admission, more than doubling those rates (IRR=2.350-2.514, p<0.001). Being registered in a mid-sized GP practice significantly reduces rates of outcomes by nearly 30% compared to the smallest 20% of practices in non-fixed effects analyses (IRR=0.707-0.715, p<0.05). Males have significantly higher rates of outcomes relative to females, by around 30% in non-fixed effects analyses (IRR=1.277-1.317, p<0.05)

The only co-morbidity that consistently significantly increases rates of outcomes across all models is Heart Failure, by around 2 fold in non-fixed effects analyses (IRR=2.068-2.118, p<0.01) and over 3 fold in fixed effects analyses (IRR=3.088, p<0.001). Diabetes (IRR=1.2, p<0.05) and Atrial Fibrillation (IRR=1.429, p<0.01) are additionally significant in random effects and Hypertension in fixed effects (IRR=2.496, p<0.01). Rheumatoid Arthritis is seen to reduce rates in non-fixed effects analysis but increase them four fold in fixed showing great variability in effect for this variable, though none of the coefficients are significant.

Deprivation has an insignificant effect on outcomes except in the most deprived where it increases rates by around 40% (IRR=1.147-1.149, p<0.05). Age significantly reduces rates of admissions by around 1% outside of fixed effects, where they increased with age, but insignificantly so. Again when the Hausman test was performed it showed a preference for fixed effects analysis.

Table 7-23 Smoking subgroup, Negative binomial

Covariates	Random Eff IRR	St. Error	Pop Ave IRR	St. Error	Fixed Eff IRR	St. Error
Blood Press	1.119	0.096	1.148	0.103	1.260*	0.148
Blood Press-1	0.910	0.071	0.905	0.070	1.003	0.105
Anti-coags	3.156***	0.336	3.045***	0.362	2.520***	0.353
Anti-coags-1	0.626***	0.057	0.640***	0.062	0.706**	0.093
Beta block	2.141***	0.221	1.953***	0.329	1.922***	0.252
Beta block-1	0.476***	0.048	0.513***	0.085	0.692**	0.087
Cholesterol	0.800**	0.053	0.804**	0.055	0.797*	0.071
Cholesterol-1	0.802**	0.052	0.805**	0.053	0.836*	0.073
Smoke cess	0.728***	0.057	0.737***	0.061	0.745**	0.078
Smoke cess-1	1.046	0.085	1.062	0.085	0.926	0.097
age	0.991*	0.003	0.990**	0.004	1.003	0.015
gender	1.277**	0.097	1.317**	0.104		
Heart Failure	2.118**	0.236	2.068***	0.229	3.088***	0.782
Atrial Fibril	1.429**	0.182	1.302	0.181	1.480	0.360
Rheum Arth	0.719	0.245	0.636	0.185	4.905	5.843
Hypertension	1.103	0.090	1.138	0.099	2.496**	0.764
Stroke TIA	1.227	0.156	1.280	0.179	1.083	0.290
Diabetes	1.200*	0.109	1.150	0.101	1.051	0.224
ACE	2.136***	0.217	2.023***	0.279	2.020***	0.264
ACE-1	0.619***	0.060	0.641**	0.083	0.536***	0.073
IMD 20-40% quintile	1.175	0.175	1.199	0.188		
IMD 40-60% quintile	1.190	0.174	1.222	0.191		
IMD 60-80% quintile	1.161	0.164	1.199	0.177		
Most deprived quintile	1.417*	0.201	1.418*	0.216		
Practice CHD attainment	0.984***	0.004	0.983***	0.004	0.992	0.007
Outcome pre-admission	2.514***	0.211	2.350***	0.193		
Size: Mid Quintile 40-60%	0.715**	0.084	0.707**	0.083		
N/T	8696/10		8696/10		870 /10	

Region, workldq and sizeq reported where significant. *** p<0.001, ** p<0.01, * p<0.05

7.5.5 Stage 3 Hurdle model

In this part of the analysis only the whole study population targets were analysed as patient numbers in sub group analysis were insufficient to allow fixed effects analysis to converge on a result and small sample sizes gave volatile and potentially misleading results in all models. As discussed this approach splits the study population into two components. Firstly a binary 'at risk' model, which the whole population participates in. In this 'at risk' model the coefficients show the relationship between the variable and the individual having an outcome. For those who have an event the hurdle is crossed and they participate in a count model, where the coefficients represent the effect of the variable on the number of events. The hurdle approach deals with potential over-dispersion issues arising from excess zeros.

7.5.5.1 Binary component

In a panel data setting for a binary outcome variable, the choice was between using logit or probit regression. The logit was used in preference to the probit as this has a fixed effects option in Stata, and enables the production of odds ratios which are easy to interpret. Besides these factors the two distributions are near identical with the logit having slightly longer tails, which may be more accommodating to the large number of zeros in the dataset. Nonetheless the two models generally give very similar results. The results for the 'at risk' logit model are shown in the Table 7-24. In this instance the results for fixed effects are not included as the regression did not converge. As the outcome variable has a binary specification the coefficients show odds ratios (OR). These are interpreted similar to IRRs, in that a figure above one shows that the variable increases the outcome measure, while those below one, that it decreases them. However since the outcome variable is simply showing whether an outcome occurs or not, the coefficients represent the impact that the variables have on the odds of an outcome occurring. As opposed to IRRs, which show the impact of the outcome variable on the rates of outcomes.

Table 7-24 Whole population QOF targets, Hurdle model, binary component

Covariates	Random Effects Odds Ratio	St. Error	Pop Ave Odds Ratio	St. Error
Blood Press	1.034	0.026	1.035	0.025
Blood Press-1	0.929**	0.022	0.919***	0.021
Anti-coags	3.177***	0.101	3.026***	0.094
Anti-coags-1	0.526***	0.015	0.523***	0.015
Beta blockers	2.651***	0.089	2.592***	0.110
Beta blockers-1	0.466***	0.015	0.465***	0.019
Cholesterol	0.944*	0.022	0.952*	0.021
Cholesterol-1	0.804***	0.019	0.791***	0.018
age	1.021***	0.001	1.018***	0.001
gender	1.168***	0.028	1.149***	0.024
Heart Failure	2.240***	0.088	1.923***	0.061
Atrial Fibrillation	1.313***	0.051	1.217***	0.043
Rheumatoid Arth	1.049	0.106	1.034	0.094
Hypertension	0.890***	0.026	0.902***	0.024
Stroke TIA	1.240***	0.053	1.192***	0.045
Diabetes	1.089*	0.039	1.069*	0.034
ACE	2.300***	0.074	2.239***	0.090
ACE-1	0.595***	0.019	0.592***	0.023
Practice CHD attainment	0.985***	0.001	0.983***	0.001
Outcome pre-admission	2.307***	0.083	1.996***	0.053
Smoker	1.236***	0.045	1.219***	0.039
IMD 20-40% quintile	1.150***	0.042	1.128***	0.036
IMD 40-60% quintile	1.141***	0.042	1.120***	0.037
IMD 60-80% quintile	1.216***	0.046	1.184***	0.039
Most deprived quintile	1.483***	0.060	1.403***	0.049
Yorkshire &The Humber			0.877*	0.055
East of England			0.883*	0.052
N/T	65,236/10		65,236/10	

Region reported where significant. *** p<0.001, ** p<0.01, * p<0.05 Sizeq and workldq also included in models but effects were insignificant NS= Not significant

The results are in keeping with those found in count model specifications. Meeting the QOF targets reduces the odds of having an admission when the targets are lagged. These results are statistically significant, most at a 99.9% significance level. Co-morbidities significantly increase odds of having an outcome with the exception

of hypertension which reduces odds between patients, namely an individual with hypertension has lower odds of having an outcome compared to one without (OR=0.890-0.902, p<0.001), and Rheumatoid Arthritis where the increase is insignificant. Smoking increases odds by over 20% (OR=1.219-1.236, p<0.001), while having an admission for one of the chosen outcomes prior to CHD diagnoses is associated with around a doubling of odds of a study admission in both models (OR=1.996-2.307, p<0.001). For each percentage increase in the individual's practice attainment on the selected CHD QOF targets, the odds of an outcome fall by around 1.5% (OR=0.983-0.985, p<0.001). Both size and workload were insignificant risk factors for having an admission and are not reported in the table. So too was region in the random effects model, though in the population averaged model, Yorkshire and the Humber (OR=0.877, p<0.05), and the East of England (OR=0.883, p<0.05), had around 12% lower odds of an admission compared to the North East.

7.5.5.2 Count component

The results for those who had an admission for an outcome of interest; the count element of the hurdle model, are shown in Table 7-25. In these results there is some commonality with the binary model, certainly in the direction of effects for the majority of variables. Meeting QOF targets is still important in how many admissions an individual has when lagged, namely if they met the target in the previous period they have lower rates of outcomes in the present period in all models. Meeting the cholesterol target in the present period also reduces rates by around 8% (IRR=0.912-0.923, p<0.001), suggesting more immediacy with this target, though this figure is not as great as its lagged value (IRR=0.786-0.843, p<0.001). Being male in this instance reduces rates, though only by around 1%, and only in the random effects model (IRR=0.987, p<0.05). Four regions (West Midlands, South West, London, and South East Coast) are associated with significantly lower rates of admissions in the random effects model. For the West Midlands, London and South East Coast rates of admission are additionally significantly lower in the population averaged model relative to the North East. The effects size range from a 15% reduction in rates of admissions (London, random effects, IRR=0.853, p<0.01) to a 10% reduction (South East Coast, population averaged, IRR=0.901, p<0.05). As these are in the population who had an event this

may be indicative of better access to, and response by, not only primary health care services but also secondary care and wider social services. There is mixed results with regard to co-morbidities. Rheumatoid Arthritis and Hypertension are insignificant in all models, Stroke is significant only in random effects (IR=1.077, $p<0.05$), and Diabetes (IRR=1.203, $p<0.01$) and Atrial Fibrillation (IRR=1.384, $p<0.001$) in fixed effects. Only Heart Failure is significant in increasing the odds of admission in all models. Higher practice CHD QOF attainment continues to significantly reduce rates of admissions, by around 2% for each 1% increase in non-fixed effects analyses, and by 0.3% with each 1% increase in fixed effects analysis. Smoking increases them in all models, by between 11.4% (population averaged, $p<0.001$) and 25% (fixed effects, $p<0.001$).

Increases in age are shown to reduce rates of outcomes in fixed effects analysis (IRR=0.71, $p<0.001$), an effect also seen in fixed effects in the MI subgroup, and non-fixed effects analyses in the smoking subgroup. It would appear that for individuals that could be considered to be in poorer health groups, namely those with recently diagnosed MI; or who had a study outcome during the study period, increasing age has a within patient protective effect. A possible explanation is that this is due to a 'survivor effect' in these instances. In the case of those who had an outcome or with recently diagnosed MI, the risk of having an event are often highest immediately following an event. Hence if the individual survives beyond that immediate period their chance of having an outcome decreases; resulting in the finding that age reduces rates.

For smoking the comparison is in non-fixed effects, primarily between patients. In this instance the more chronically ill smokers may be quitting and exiting the smoking population, meaning that the smoking population is relatively healthier in comparison. In that scenario increasing age would be seen to reduce rates of outcomes between patients, in a smoking subgroup. Smokers were certainly younger on average than non-smokers, by nearly 8 years, mean 63.9 years compared to 71.8, which could point to this part of the study population being in better health.

Table 7-25 Whole population QOF targets, Hurdle model, count component

Covariates	Random Eff IRR	St. Error	Pop Ave IRR	St. Error	Fixed Eff IRR	St. Error
Blood Press	0.997	0.019	1.001	0.017	1.013	0.026
Blood Press-1	0.836***	0.015	0.838***	0.013	0.892***	0.021
Anti-coags	2.025***	0.048	1.957***	0.043	1.442***	0.048
Anti-coags-1	0.588***	0.012	0.599***	0.011	0.594***	0.016
Beta blockers	1.541***	0.033	1.478***	0.035	1.624***	0.051
Beta blockers-1	0.658***	0.014	0.695***	0.016	0.725***	0.020
Cholesterol	0.912***	0.016	0.923***	0.014	0.916***	0.021
Cholesterol-1	0.786***	0.013	0.803***	0.012	0.843***	0.019
age	1.009***	0.001	1.007***	0.001	0.710***	0.006
gender	0.987*	0.019	1.000	0.019		
Heart Failure	1.306***	0.031	1.293***	0.033	1.851***	0.106
Atrial Fibril	1.038	0.028	1.032	0.029	1.384***	0.089
Rheum Arth	1.011	0.074	0.992	0.078	0.791	0.206
Hypertension	0.959	0.022	0.957	0.022	1.078	0.089
Stroke TIA	1.077*	0.034	1.056	0.033	0.997	0.086
Diabetes	1.031	0.027	1.019	0.028	1.203**	0.075
Size: Largest 20%	0.908**	0.026	0.918**	0.027		
ACE	1.324***	0.028	1.279	0.029	1.323***	0.040
ACE-1	0.618***	0.013	0.649***	0.014	0.608***	0.017
IMD 20-40% quintile	1.054	0.031	1.056	0.030		
IMD 40-60% quintile	1.092**	0.033	1.087**	0.032		
IMD 60-80% quintile	1.109**	0.033	1.118***	0.033		
Most deprived quintile	1.198***	0.039	1.187***	0.037		
West Midlands	0.879*	0.047	0.898*	0.042		
South West	0.879*	0.047				
London	0.853**	0.048	0.887*	0.045		
South East Coast	0.869*	0.050	0.901*	0.046		
Practice CHD attainment	0.980***	0.001	0.981***	0.001	0.997**	0.001
Outcome pre-admission	1.404***	0.031	1.384***	0.032		
Smoker	1.124***	0.030	1.114***	0.029	1.241***	0.056
N/T	18998/10		18998/10		10759/10	

Region, workldq and sizeq reported where significant. *** p<0.001, ** p<0.01, * p<0.05

7.5.6 Stage 4 Logit panel data analysis using a binary transformed outcome measure

7.5.6.1 Whole population

It was shown in the data description section that there is a large dominance of zero outcomes within the data, amounting to 94.55% of all outcomes. Individuals having one event account for 4.43% of all outcomes, meaning that counts of greater than one account for just over 1%. Applying count models with such a high proportion of zeros may produce biased estimators. One of the solutions is to apply a form of hurdle to the data which was done previously. However this led to issues of convergence, meaning that subgroup targets had to be excluded entirely and fixed effects did not converge in the whole study binary component. An alternative was to dichotomise the outcome measure. This approach alters a small portion of outcomes, removes issues concerning zero inflation pertinent to count models and enables regressions to be completed on a greater proportion of the study data. To this effect a new variable, Admis, was created which replaced all counts greater than 1 with 1. As in the hurdle model the logit was used in preference to the probit, for the reasons discussed previously. Hausman tests were performed where possible throughout this stage and showed a preference for fixed effects. The results from this analysis are presented in Table 7-26.

Table 7-26 Whole population QOF targets, Logit

Covariates	Random Eff Odds Ratio	St. Error	Pop Ave Odds Ratio	St. Error	Fixed Eff Odds Ratio	St. Error
Blood Press	1.038	0.023	1.039	0.020	1.038	0.027
Blood Press-1	0.914***	0.019	0.929***	0.017	0.922**	0.022
Anti-coags	3.197***	0.089	2.824***	0.076	2.612***	0.090
Anti-coags-1	0.570***	0.014	0.605***	0.014	0.579***	0.017
Beta blockers	2.102***	0.059	1.986***	0.069	2.124***	0.069
Beta blockers-1	0.588***	0.016	0.606***	0.021	0.616***	0.020
Cholesterol	0.910***	0.018	0.915***	0.016	0.956	0.023
Cholesterol-1	0.812***	0.016	0.828***	0.015	0.862***	0.020
age	1.018***	0.001	1.015***	0.001	1.015*	0.007
gender	1.099***	0.024	1.086***	0.022		
Heart Failure	2.295***	0.069	1.965***	0.053	2.322***	0.133
Atrial Fibril	1.330***	0.044	1.251***	0.038	1.635***	0.100
Rheum Arth	1.136	0.100	1.144	0.096	1.373	0.291
Hypertension	0.902***	0.024	0.912***	0.022	1.208*	0.089
Stroke TIA	1.237***	0.047	1.196***	0.040	1.223**	0.093
Diabetes	1.132***	0.035	1.118***	0.031	1.092	0.063
ACE	1.939***	0.053	1.819***	0.059	1.790***	0.055
ACE-1	0.676***	0.018	0.697***	0.022	0.574***	0.018
IMD 20-40% quintile	1.139***	0.039	1.128***	0.034		
IMD 40-60% quintile	1.166***	0.040	1.146***	0.036		
IMD 60-80% quintile	1.244***	0.043	1.223***	0.038		
Most deprived quintile	1.506***	0.056	1.426***	0.047		
Practice CHD attainment	0.984***	0.001	0.986***	0.001	0.994***	0.001
Outcome pre-admission	2.762***	0.078	2.423***	0.060		
Smoker	1.220***	0.038	1.189***	0.034	1.057	0.050
N/T	72661/10		72661/10		12546/10	

Region, workldq and sizeq reported where significant. *** p<0.001, ** p<0.01, * p<0.05

They show a general consistency in results among the different forms of analyses. All the QOF targets significantly reduce odds ($p<0.001$) when that variable is lagged, as does the ACE inhibitor/A2 prescription target when applied to the whole CHD population. Cholesterol was also found to reduce odds of an admission in the present period (OR=0.91-0.915, $p<0.001$) outside of fixed effects analysis. Each one

year increase in age increases the odds of having an outcome; by 1.5% in fixed effects and population averaged models, and 1.8% in random effects. Being male increases the odds of an admission by near 1% in models where the variable is present (OR=1.086-1.099, p<0.001). The odds of having an outcome increase as the deprivation quintiles increase, in a clear stepped manner, relative to the least deprived. With the exception of Rheumatoid Arthritis in all models and Diabetes in fixed effects (OR=1.092), the effect of co-morbidities are significant. The odds in all cases are increased by the presence of the included co-morbidities with the exception of hypertension in the random effects and population average specifications where it is seen to reduce odds between and within patients by 9-10% (p<0.001). Each 1% increase in practice attainment reduces the odds of having an outcome, by around 1.5% (p<0.001) in between effects analyses and 0.6% in within, fixed effects analysis (p<0.001). Smoking increases odds, by a range of 5% in fixed effects (not significant) to 22% in random effects (p<0.001). Having one of the study outcomes of interest prior to entry into the study increased odds of having an outcome during the study by around 2.5 times, in non-fixed effects analyses (OR=2.423-2.762, p<0.001) where the variable could be included. GP workload, SHA region and practice size were included in all analyses but were not significant in any of their categories relative to their baseline comparator.

7.5.6.2 MI subgroup

Regression results for the MI subgroup are shown in Table 7-27, and are similar to those reported in Table 7-26 for the whole study population. In comparison to those figures QOF targets continue to be significant when lagged, including the prescription of ACE or A2 to recently diagnosed MI patients. Hypertension is however no longer significant except in fixed effects where it continues to increase odds, though by a much bigger magnitude of over 86% within patients (OR=1.864, p<0.01). Atrial Fibrillation is no longer significant in population averaged analysis (OR=1.126). Region continues to be an insignificant predictor of events, as are the size quintiles in this subgroup, while workload increases odds in the 20-40% quintile (OR, 1.112 to 1.271, p<0.05). Diabetes is no longer significant in fixed effects analysis (OR=1.195) and neither is smoking (OR=1.125). Males have lower odds of an outcome in non-fixed effects analysis, but the effect is insignificant. In terms of deprivation there is no longer a clear increase in odds of an outcome as we move into

Table 7-27 MI subgroup, Logit model

Covariates	Random Odds Ratio	St. Error	Pop Ave Odds Ratio	St. Error	Fixed Odds Ratio	St. Error
Blood Press	1.016	0.048	1.019	0.042	0.941	0.062
Blood Press-1	0.842***	0.037	0.864***	0.033	0.870*	0.053
Anti-coags	2.537***	0.160	2.261***	0.130	1.719***	0.176
Anti-coags-1	0.600***	0.031	0.639***	0.029	0.610***	0.046
Beta blockers	1.662***	0.098	1.581***	0.107	1.853***	0.155
Beta blockers-1	0.654***	0.038	0.682***	0.045	0.670***	0.053
Cholesterol	0.797***	0.033	0.820***	0.030	0.807***	0.048
Cholesterol-1	0.768***	0.032	0.795***	0.030	0.845**	0.048
ACE/A2	1.079	0.062	1.067	0.064	1.017	0.083
ACE/A2-1	0.842**	0.048	0.871*	0.052	0.708***	0.058
age	1.010***	0.002	1.008*	0.002	0.902***	0.017
gender	0.977	0.044	0.983	0.039		
Heart Failure	2.064***	0.114	1.821***	0.085	2.713***	0.423
Atrial Fibril	1.152*	0.077	1.126	0.065	1.659**	0.317
Rheum Arth	1.302	0.211	1.268	0.174	2.081	1.468
Hypertension	0.990	0.052	0.986	0.045	1.864**	0.427
Stroke TIA	1.221**	0.094	1.195**	0.079	1.593*	0.341
Diabetes	1.200**	0.074	1.178**	0.063	1.195	0.216
Workld: lowest 20-40%	1.131*	0.070	1.112*	0.059	1.271*	0.123
IMD 20-40% quintile	1.062	0.070	1.051	0.059		
IMD 40-60% quintile	1.211**	0.081	1.178**	0.068		
IMD 60-80% quintile	1.138	0.078	1.124*	0.066		
Most deprived quintile	1.378***	0.101	1.311***	0.082		
Practice CHD attainment	0.989***	0.002	0.991***	0.001	0.986***	0.004
Outcome pre-admission	2.704***	0.160	2.335***	0.115		
Smoker	1.233***	0.073	1.218***	0.063	1.125	0.120
N/T	26744/10		26744/10		2895/10	

Region, workldq and sizeq reported where significant. *** p<0.001, ** p<0.01, * p<0.05

more deprived quintiles. Being in the least deprived 20-40% quintile no longer increases odds of an outcome in random effects and population average models, nor

does being in the most deprived 20-40% in random effects relative to the least deprived quintile. Having an outcome prior to entry into the study is still a large and significant predictor of an admission during the study, more than doubling odds (OR=2.335-2.704, p<0.001). Increases in practice attainment continue to significantly reduce odds of an outcome (OR=0.986 to 0.991, p<0.001). In fixed effects analysis age is again shown to significantly reduce odds of an admission in those with recently diagnosed MI (OR=0.902, p<0.001).

7.5.6.3 Current smokers-smoking cessation

In the non-fixed effects analyses just less than 11% of the total patient population are involved, in fixed effects that figure is down to just over 1%, making results from fixed effects less generalisable. Results are shown in Table 7-28.

The results are very similar to those presented for the negative binomial count model presented in stage 2 for this population. Coefficients for age again show this variable to significantly reduce, in this instance the odds of an outcome, in random effects models by near 1% (OR=0.991, p<0.001).

The BP QOF target is only significant in fixed effect analysis increasing odds in the immediate period (OR=1.326, p<0.05), though not when lags were introduced. Meeting the cholesterol target reduces odds of an outcome in both the immediate period and the subsequent period, though the lagged effect was not significant in fixed effects analysis. The remaining QOF targets are seen to significantly reduce the odds of an outcome, but only when lagged.

The results concerning co-morbidities are mixed. Heart Failure increases odds of an outcome by near double in population averaged (OR=1.917) analysis to over treble in fixed effects analysis (OR=3.225), p<0.001. Atrial Fibrillation increases odds by around 1.5 times (OR=1.5 random effects, p<0.01; 1.397 population averaged, p<0.05) but this is not significant in the fixed effects model. Rheumatoid Arthritis is insignificant in all models and Hypertension only in fixed effects analysis where it increases odds within patients by over 2.5 times (OR=2.763, p<0.05). Stroke is insignificant in all analyses and diabetes only in random effects where it increases odds within and between patients by 23% (OR=1.232, p<0.05).

Table 7-28 Smoking subgroup, Logit model

Covariates	Random Eff Odds Ratio	St. Error	Pop Ave Odds Ratio	St. Error	Fixed Eff Odds Ratio	St. Error
Blood Press	1.114	0.108	1.098	0.093	1.326*	0.178
Blood Press-1	0.880	0.078	0.895	0.068	1.017	0.122
Anti-coags	3.637***	0.443	3.164***	0.372	2.617***	0.409
Anti-coags-1	0.560***	0.060	0.595***	0.058	0.661**	0.100
Beta blockers	2.507***	0.305	2.376***	0.353	2.324***	0.369
Beta blockers-1	0.409***	0.049	0.430***	0.063	0.586**	0.092
Cholesterol	0.776**	0.058	0.807**	0.054	0.781*	0.080
Cholesterol-1	0.793**	0.058	0.811**	0.051	0.861	0.088
Smoke cess	0.702***	0.063	0.744***	0.059	0.732*	0.089
Smoke cess-1	1.061	0.098	1.079	0.086	0.913	0.108
age	0.991*	0.004	0.992*	0.004	1.013	0.029
gender	1.289**	0.110	1.240**	0.091		
Heart Failure	2.277***	0.291	1.917**	0.202	3.225***	0.960
Atrial Fibril	1.500**	0.221	1.397*	0.186	1.580	0.519
Rheum Arth	0.764	0.280	0.761	0.234	5.671	7.573
Hypertension	1.084	0.100	1.082	0.085	2.763*	1.108
Stroke TIA	1.203	0.173	1.140	0.143	1.104	0.377
Diabetes	1.232*	0.126	1.167	0.101	0.979	0.243
ACE	2.430***	0.284	2.242***	0.298	2.281***	0.349
ACE-1	0.578***	0.065	0.610***	0.077	0.540***	0.089
IMD 20-40% quintile	1.227	0.204	1.211	0.177		
IMD 40-60% quintile	1.235	0.202	1.199	0.174		
IMD 60-80% quintile	1.251	0.197	1.219	0.168		
Most deprived quintile	1.520*	0.242	1.426*	0.201		
Practice CHD attainment	0.981***	0.005	0.984***	0.004	0.990	0.008
Outcome pre-admission	2.783***	0.269	2.376***	0.189		
Size: Mid Quintile 40-60%	0.717*	0.093	0.731*	0.081		
N/T	8696/10		8696/10		850/10	

Region, workldq and sizeq reported where significant. *** p<0.001, ** p<0.01, * p<0.05

Receiving an ACE inhibitor or A2 antagonist reduces odds of having an outcome in the smoking population when the variable is lagged in all analyses, by between 40

and 45% (OR=0.54 to 0.610, p<0.001). Deprivation is no longer a significant predictor of an outcome with the exception of the most deprived quintile where odds are increased by over 40% (OR= 1.426, p<0.05) in the population averaged and 50% in random effects analyses (OR=1.52, p<0.05), relative to the least deprived. Percentage increases in practice attainment continue to decrease odds of an outcome, by around 2% in random effects (OR=0.981, p<0.001) and population averaged analyses (OR=0.984, p<0.001); though the variable is no longer significant in fixed effects. Having an outcome prior to entry into the study remains a significant cause of admissions during the study increasing odds by around 2.5 times (OR=2.376 to 2.783, p<0.001). Smoking cessation is seen to have an immediate effect, reducing odds in all analyses by nearly 30% (OR=0.702 to 0.744 non-fixed effects, p<0.001; 0.732 fixed effects, p<0.05). However the variable is not significant when lagged. This may point to the long run ineffectiveness of this target, as it suggests that while it may have resulted in immediate benefits, possibly brought about by individuals quitting smoking, in the subsequent period it had no significant impact. This may be due to patients resuming smoking once the intervention was withdrawn or because, in this chronically ill population, many of whom had co-morbidities, there were no longer term benefits in terms of reduced adverse outcomes from quitting smoking. It should also be noted that there may be wider issues when interpreting figures for this subgroup due to the way the data is set up. This is because individuals only feature in this population if they are current smokers, so those who were recorded as non or ex-smokers would not feature. Hence when interpreting lags we are only looking at patients who continued to smoke and naturally the target would be ineffective in respect of that population. What is not known is how representative this patient population is of the wider population who received smoking cessation advice. Finally being in a mid-sized practice is seen to reduce the odds of an admission relative to the smallest practice size quintile, by up to 28% in non-fixed effects analysis (OR=0.717 to 0.731, p<0.05).

7.5.7 Summary

Table 7-29 shows a summary of the main results for the regression coefficients on analyses undertaken on the whole study population. The exception to this being the QOF target variables for the prescription of ACE or A2 to MI patients (*Acea2q*) and smoking cessation advice to current smokes (*Smcess*), which applied to their relevant

subgroups only, detailed in the table. The coefficients on the remaining variables are not shown for these subgroups though they have been detailed previously. Also omitted are the coefficients for SHA region, practice size and GP workload variables as these were largely insignificant. Only the lags of the QOF target variables are reported as only these were significant, with the exception of the cholesterol target, which also had an immediate impact.

Table 7-30 shows the regression coefficients from the hurdle model analyses. Again SHA region, practice size and GP workload were excluded for the same reasons given above. Subgroup target variables are not shown as they were not analysed in the hurdle model. The results in the binary hurdle represent the odds of a patient having an outcome and the count hurdle the rates of outcomes in those who had an outcome. They are not intended for comparison as they show different risk groups, with those in the count model assumed to be higher risk patients as they have had an outcome in the study.

These tables shown a general consistency in the size and significance of the key variables, certainly the QOF variables, and subgroup specific targets where included across all analyses. For the binary and count model specifications for the whole study population, the results are very similar for all variables. In these two analyses co-morbidities, Rheumatoid Arthritis and Hypertension aside, significantly increase the odds of the study outcome. These results are largely but not wholly replicated in the hurdle model. Further discussion follows after Table 7-31.

Table 7-29 Whole study binary and count models results summary

	Count (Incident rate ratio)			Binary (Odds ratio)		
Covariates	Random	Pop Ave	Fixed	Random	Pop Ave	Fixed
Blood Press-1	0.915***	0.905***	0.932***	0.914***	0.929***	0.922***
Anti-coags-1	0.612***	0.611***	0.623***	0.570***	0.605***	0.579***
Beta blocks-1	0.627***	0.657***	0.660***	0.588***	0.606***	0.616***
Cholesterol	0.917***	0.913***	0.956*	0.910***	0.915***	0.956
Cholesterol-1	0.830***	0.839***	0.872***	0.812***	0.828***	0.862***
ACE/A2-1┼	0.854**	0.881*	0.726***	0.842***	0.871*	0.708***
Smoke cess₸	0.728***	0.737***	0.745**	0.702***	0.744***	0.732*
age	1.015***	1.014***	1.001	1.018***	1.015***	1.015*
gender	1.086***	1.106***	N/A	1.099***	1.086***	N/A
Heart Failure	2.098***	1.987***	2.020***	2.295***	1.965***	2.322***
Atrial Fibrilation	1.285***	1.242***	1.482***	1.330***	1.251***	1.635***
Rheum Arthritis	1.137	1.109	1.439*	1.136	1.144	1.373
Hypertens	0.902***	0.902***	1.072	0.902***	0.912***	1.208*
Stroke TIA	1.213***	1.915***	1.226**	1.237***	1.196***	1.223**
Diabetes	1.121***	1.119***	1.111*	1.132***	1.118***	1.092
ACE-1	0.711***	0.735***	0.622***	0.676***	0.697***	0.574***
Practice CHD attain	0.986***	0.987***	0.996***	0.984***	0.986***	0.994***
Outcome pre-admis	2.532***	2.480***	N/A	2.762***	2.423***	N/A
Smoker	1.199***	1.184***	1.058	1.220***	1.189***	1.057
IMD 20-40% quintile	1.131***	1.140***	N/A	1.139***	1.128***	N/A
IMD 40-60% quintile	1.158***	1.173***	N/A	1.166***	1.146***	N/A
IMD 60-80% quintile	1.225***	1.276***	N/A	1.244***	1.223***	N/A
Most deprived quintile	1.447***	1.455***	N/A	1.506***	1.426***	N/A

*** p<0.001, ** p<0.01, * p<0.05 ₸Current smokers sub group ┼MI subgroup

Table 7-30 Hurdle model results summary

	Binary Hurdle (OR)		Count Hurdle (IRR)		
Covariates	Random	Pop Ave	Random	Pop Ave	Fixed
Blood Press-1	0.929**	0.919***	0.836***	0.838***	0.892***
Anti-coags-1	0.526***	0.523***	0.588***	0.599***	0.594***
Beta blocks-1	0.466***	0.465***	0.658***	0.695***	0.725***
Cholesterol	0.944*	0.952*	0.912***	0.923***	0.916***
Cholesterol-1	0.804***	0.791***	0.786***	0.803***	0.843***
age	1.021***	1.018***	1.009***	1.007***	0.710***
gender	1.168***	1.149***	0.987*	1	N/A
Heart Failure	2.240***	1.923***	1.306***	1.293***	1.851***
Atrial Fibril	1.313***	1.217***	1.038	1.032	1.384***
Rheum Arth	1.049	1.034	1.011	0.992	0.791
Hypertension	0.890***	0.902***	0.959	0.957	1.078
Stroke TIA	1.240***	1.192***	1.077*	1.056	0.997
Diabetes	1.089*	1.069*	1.031	1.019	1.203**
ACE-1	0.595***	0.592***	0.618***	0.649***	0.608***
Practice CHD attainment	0.985***	0.983***	0.980***	0.981***	0.997**
Outcome pre-admission	2.307***	1.996***	1.404***	1.384***	N/A
Smoker	1.236***	1.219***	1.124***	1.114***	1.241***
IMD 20-40% quintile	1.150***	1.128***	1.054	1.056	N/A
IMD 40-60% quintile	1.141***	1.120***	1.092**	1.087**	N/A
IMD 60-80% quintile	1.216***	1.184***	1.109**	1.118***	N/A
Most deprived quintile	1.483***	1.403***	1.198***	1.187***	N/A
*** p<0.001, ** p<0.01, * p<0.05					

The results for all the regressions are summarised in the Table 7-31 for the key and consistently included variables. A number of variables from this list are excluded: Rheumatoid Arthritis is absent as this was significant in the negative binomial fixed effects model only. Practice size, practice GP workload and the SHA region variables are also excluded as they showed no consistency in effect and it was

therefore not practical to include them. Subgroup specific targets, ACE inhibitors in recently diagnosed MI patients and the smoking cessation target are not included as they applied only to their specific context, though did significantly reduce rates or odds in all instances when included. The results are broken down into the whole and subgroup populations as well as to the different models analysed in Stage 2-4; negative binomial, hurdle and logit models. Results are only reported if the coefficient was significant at a 95% or greater level ($p \leq 0.05$). The +/- reflects the direction of the effect of the explanatory variable on the outcome measure, as the explanatory variable increased or went from zero to one. The absence of results for fixed effects analysis in the binary component of the hurdle model reflects the fact that the regression in this instance could not converge on a result.

Table 7-31 incorporates all the pertinent information on all analyses and is therefore information heavy. In interpreting it therefore there are a number of general pointers. As the table shows instances (regressions) where the variable was significant ($p \leq 0.05$), the more entries there are against a variable the more consistently significant it was. Changes in sign show inconsistency in the direction of effect, hence regular changes, particularly in the same regression type or outcome variable specification, show that the variable had an inconsistent effect and was unstable in this respect. This may be indicative of a badly specified variable, and certainly makes it difficult to interpret its impact. To avoid confusion where the absence of detail could be interpreted as the absence of any effect, those variables that could not be included in fixed effects analyses are denoted by a not applicable, N/A, entry. Hence when looking for variables which play a significant role in predicting an outcome or outcomes, a greater number of entries and consistency on the sign recorded, indicate this. A quick examination of the table reveals importantly that the lagged QOF variables were all important explanatory variables.

Table 7-31 Summary of results for all models

Variable	Random Effects	Population Average	Fixed Effects
Age	+NB①② -NB③ +B +C +L①② -L③	+NB①② -NB③ +B +C +L①② -L③	-NB② -C +L① -L②
Male	+NB①③ +B -C +L①③	+NB①③ +B +L①③	N/A
BP≤150/90 lagged	-NB①② -C -L①②	-NB①② -B -C -L①②	-NB①② -C -L①②
Chol≤5mmol/l lagged	-NB①②③ -B -C -L①②③	-NB①②③ -B -C -L①②③	-NB①②③ -C -L①②
Anti plats/coags lagged	-NB①②③ -B -C -L①②③	-NB①②③ -B -C -L①②③	-NB①②③ -C -L①②③
Betablockers lagged	-NB①②③ -B -C -L①②③	-NB①②③ -B -C -L①②③	-NB①②③ -C -L①②③
ACE inhibitors lagged	-NB①③ -B -C -L①③	-NB①③ -B -L①③	-NB①③ -C -L①③
Heart Failure	+NB①②③ +B +C +L①②③	+NB①②③ +B +C +L①②③	+NB①②③ +C +L①②③
Atrial Fibrillation	+NB①②③ +B +L①②③	+NB①② +B +L①③	+NB① +C +L①②
Hypertension	-NB① -B -L①	-NB① -B -L①	+NB②③ +L①②③
Stroke TIA	+NB①② +B +L①②	+NB①② +B +L①②	+NB① +L①②
Diabetes	+NB①②③ +B +L①②③	+NB①② +B +L①②	+NB① +C
Practice CHD attainment	-NB①②③ -B -C -L①②③	-NB①②③ -B -C -L①②③	-NB①② -C -L①②
Least deprived 20-40%	+NB① +B +L①	+NB① +B +L①	N/A
Mid deprived quintile 40-60%	+NB①② +B -C +L①②	+NB①② +B -C +L①②	N/A
Most deprived 20-40%	+NB①② +B -C +L①	+NB①② +B -C +L①②	N/A
Most deprived 20%	+NB①② +B -C +L①②③	+NB①② +B -C +L①②③	N/A
Outcome prior to study entry	+NB①②③ +B +C +L①②③	+NB①②③ +B +C +L①②③	N/A
Current smoker	+NB①② +B +C +L①②	+NB①② +B +C +L①②	+C

+/- Indicate direction of effect on outcome variable.
NB=Negative Binomial B=Binary hurdle model C=Count hurdle model L=Logit
①= Whole population ②=MI subgroup ③=Current smokers subgroup
All figures significant p≤0.05

7.5.7.1 Key findings

A number of significant and consistent findings interest are evident in these results. These are:

1. Regardless of the functional form specification of the outcome variable, whether fixed or random effects, or population averaged models were used; the QOF CHD target variables used in the model significantly reduced outcomes with a lag. This applied to the subgroup targets also, the only exceptions to this result being:
   i. For BP, all forms of the negative binomial and logit regression in the smoking subgroup
   ii. For Cholesterol, the fixed effects logit in the smoking subgroup
2. Meeting the cholesterol target was found to have an immediate and significant reductive impact on outcomes in all analyses aside from the fixed effects logit model.
3. The ACE inhibitor/A2 antagonist QOF target was additionally found to reduce outcomes in all analyses when applied to the whole study population.
4. Of all the co-morbidities only HF was a significant contributor to the number of outcomes in all analyses.
5. Where other co-morbidities were significant they led to increases in outcomes. The only exception to this being hypertension which increased outcomes in fixed effects analyses but reduced them in and non-fixed effects analyses. This result suggests that within patients, i.e. being diagnosed with hypertension increases outcomes, however comparing between patients, having a hypertension diagnosis reduced outcomes. The possibility that the result could be caused by covariance with the BP target was considered and explored but no evidence was found, correlation=0.0122.
6. Having a study outcome prior to entry significantly increased outcomes in all analyses.
7. Increased practice attainment on QOF CHD targets significantly reduced outcomes in all analyses, with each 1% increase in practice attainment generally increasing outcomes by a greater than equivalent amount in non-fixed effects analyses and by a lower than equivalent amount in fixed effects analyses.

8. More deprived quintiles, with the exception of the least deprived 20-40%, were associated with a significantly greater number of outcomes relative to the least deprived in all analyses outside of the smoking subgroup.
9. Increasing age predominantly increased outcomes in random effects analyses, with the exception of the smoking subgroup where increasing age decreased outcomes. In fixed effects analysis the variable was largely insignificant, but tended to reduce outcomes where it was in whole population models. In what could be considered more chronically ill subgroups, namely those who had a recent MI diagnosis or an outcome during their participation in the study; increasing age reduced outcomes within individuals. This could be indicative of a 'survivor effect.'
10. Males had a higher number of outcomes relative to females in all models except in MI subgroup analysis, and random effects in the hurdle count model.
11. Both Diabetes and Stroke TIA increase outcomes in all random effects analyses outside out of the count component of the hurdle model.
12. Smoking increases outcomes in all instances in non-fixed effects models but is only significant in fixed effects analysis in the hurdle count model. This suggests that at the individual level smoking does not increase outcomes but it does between non-smoking and smoking individuals. This could be the result of more chronically ill patients giving up smoking, hence in this context, the smoking subgroup would represent a healthier cohort of the study population.
13. Practice size and GP workload had no significant impact on the study outcomes in whole population analyses. In the MI subgroup practices in the lowest 20-40% quintile for GP workload had a significantly higher rates (between 12% and 20%) or odds (between 11% and 27%) of outcomes relative to those in the lowest GP workload quintile. In the smoking subgroup mid-sized practices had a significantly lower rate (approx 29%, $p<0.01$) and odds (approx 28%, $p<0.05$) of outcomes relative to the smallest practices

7.5.8 Discussion

Previous research on CHD patients in the QOF found improvement on QOF target measures and in the case of BP and Cholesterol the surrogate outcomes that mapped to those targets[119-122]. This research has confirmed those findings and taken a number of steps forward. Firstly in terms of the period covered, in comparison to previous

research which tended to examine incentives for short periods or limited time points, here individuals have been followed over an 11 year period. In comparison to the QOF, which does not have a control group, by using individual level panel data, patients have been able to act as form of comparator in this research. Finally attainment to QOF target standards has been linked at an individual level to linked clinical outcomes. No research to date has done this. This research suggests that when evidence based QOF CHD targets are met the odds ratios, or incidence rates of hospital admissions for CHD complications are reduced, within a one year lag. This is important as it suggests, in an uncontrolled real world environment outside of experimental conditions, there is significant patient benefits from meeting these targets. Furthermore, more generally, that other targets supported by high level evidence, may have benefits in the population. These benefits in the case of the targets examined, serious hospital admissions for complications arising from CHD avoided, could have service implications, although these went unexamined in this thesis. The targets examined represent a small subset of the total number of QOF targets, those which were strongly evidence based and had extractable outcomes in large datasets, and the findings cannot be therefore applied to the QOF more generally. It remains to be seen therefore, and is a challenge still remaining, to see if, or ensure that, this applies to the wider QOF P4P scheme.

This research, as with most economic analyses, has been undertaken in an uncontrolled setting and has consequently relied on observational data. As such, as with all work conducted outside of experimental conditions, it can only with certainty show causation and not causality. However it is often not possible to prove causation and correlation is sufficient (e.g. higher insurance for male drivers due to higher risk of accident); and a number of approaches have been used which suggest causation. Firstly through the use of multivariate analyses, the study has controlled for a number of potential confounders. Furthermore in fixed effects analyses, even if potential confounders have been omitted from the model, so long as these are fixed over time, these have been controlled for.

Correlation refers to variables moving together, in causality there is a cause and effect, with the change in the dependent variable preceding the change in the outcome. Fixed effects analyses allow us to analyse changes within patients, observe ‘shocks’, which are often used in econometrics which relies on observational data to

infer causation. Shocks refer to sudden unexpected changes, so for example if there is a sudden fall in a company's share prices following the unexpected resignation of its owner, causality is assumed. In the context of this thesis while not 'shocks'; changes within a patients have been examined, which act in a similar manner. Hence it has been demonstrated that a patient who goes from failing a QOF target, to meeting it, significantly reduces their odds or rates of having an outcome. In much the same manner, the use of lags, enables the examination of a causal effect. Here we see the impact of a variable in a previous period on an outcome in the present period, holding other variables constant. In this thesis the effects of QOF targets were all shown to be significant with a lag, suggesting causation rather than correlation. Finally all of the QOF targets and linked outcomes were selected on the basis of high level evidence using experimental study designs, suggesting a causal relationship.

Regression to the mean is also a possibility in any regression analysis. This refers to the phenomenon that a variable will regress towards the mean over time, so if a variable is extreme on one measure it will tend closer to the average on its next measure. In the context of this thesis this would manifest itself in an individual not having an outcome in the next period if they had one in this, due to regression to the mean rather than any treatment effects. While regression to the mean cannot be discounted there are a number of factors which mitigate against it. Firstly in panel data time is captured within the panel, and rather than looking at time trends where regression to the mean can be an issue; individual level panel data examines the effects of changes in state within an individual or differences in state between individuals. Hence the results show us that both within and between patients that meeting the QOF targets examined significantly reduced the odds or rates of an outcome. Finally the *hospre* variable, which recorded whether or not the individual had a study outcome prior to entry into the study, had a consistently large and highly statistically significant impact in all analyses. This suggests that for the outcomes examined in this study, rather than reverting to the mean, if the individual had an outcome, they had a higher 'risk' of a further outcome during the study period.

This research has demonstrated the benefit of panel data in terms of the ability to examine temporal effects and introduce lags into the data. The results, as discussed, show that evidence based CHD targets in the QOF are effective in reducing evidence linked CHD hospital admissions within a year. They also suggest that in the current

period, when those same targets are met, with the exception of cholesterol control and smoking cessation, those hospital admissions increase. This reversal of effect could be argued is evidence for regression to the mean, as present period values suggest the targets are ineffective and indeed do harm. This thesis does not believe this to be the case for the reasons discussed previously. Rather it argues that the present period values are more illustrative of correlation and the lags demonstrate causation. This is due to a number of reasons. Firstly most treatments once initialised require a period of repeated compliance before the treatments effects are evident. Secondly due to the way the panel data was organised, in the present period, treatment did not necessarily precede the outcome. Lags on the other hand ensured that treatment preceded any recorded outcome. The fact that the cholesterol control target was also shown to be largely effective without a lag, may have been due to the study population receiving statins prior to entry. These were recommended for use in patients with a 20% or greater 10 year risk of CVD, so it is possible that cholesterol was being monitored and controlled prior to entry. The thesis did not explore this and it should be noted that the recommendation pre-dates much of the study period. For smoking cessation, the target is effective in the present period but not when lagged. This target was examined in the current smokers only, hence if they quit, they would not feature in the subsequent period. In this population the lagged values could give a distorted picture for a persistent and hard to treat smoking population.

The results show that rates or odds of outcomes did not differ significantly between SHA regions in the main analyses outside of the hurdle model. However the hurdle models produced some interesting results. The ‘at risk’ model suggests that relative to the North East, individuals in the Yorkshire and the Humber and the East of England, who had CHD, had lower odds of having a study outcome. In those individuals who had a study outcome, the count component, relative to the North East, individuals from the West Midlands, South West, London and South East Coast, all had significantly lower rates of outcomes. In the USA much work has been undertaken by the Dartmouth Institute looking at disparities in Medicare spending across the USA and the impact of location on the delivery of, and access to, services, and health outcomes[140]. This shows big discrepancies in per capita Medicaid expenditure across the states, not explained by poverty. Furthermore that greater spending does not explain better health outcomes, and that there is enormous scope

for improving efficiency and quality in US health care, by rewarding providers for providing evidence based care and reducing unnecessary care. The results in this thesis would suggest that the NHS performs better in this respect, and on the whole people with similar study characteristics can expect to have similar outcomes, for CHD at least, regardless of SHA in England. This is perhaps not surprising for a condition such as CHD where care protocols are standardised, and in a universally provided, free at the point of use, healthcare system. At the same time incentive schemes like the QOF and other initiatives like National Service Frameworks may have played a role. The differences in the hurdle model suggest two things: Firstly that some SHAs are more successful at preventing first admissions for complications arising from CHD following first diagnosis. Secondly that other SHAs are better at preventing subsequent admissions once a first admission has occurred. No SHAs feature in both of these results. Strong support could not be found for this result from national hospital statistics which showed that with the exception of the South East Coast and the South West all these SHAs had above national average rates for emergency hospitalisations for CHD, and MI, and for deaths from CHD[141]. This suggests that the individuals from the SHAs outside of the exceptions mentioned, used in this study are not representative of the population in those areas as a whole. The deprivation profile of the population would support this. Or, that the individuals in those SHAs receive care which is not representative of the region as a whole. This would be in keeping with the work by the Dartmouth Institute, but requires further research.

7.5.9 Limitations

This research has assumed that the explanatory variables are strictly exogenous, namely that they are uncorrelated with the error term, see (3). Some attempts have been made to adjust for potential endogeneity of the explanatory variables, such as by using fixed effects, but others such as the use of instrumental variables has not.

No commands existed within Stata to analyse panel count models. The study therefore used existing Stata commands to create an approximate hurdle model. A possible alternative however was to use facilities within Stata to write commands specifically for this purpose, which may have produced a better specified model and more efficient estimators.

This study did not explore the effect of interactions between treatments and/or conditions. QOF treatment is condition specific hence this is in keeping with an analysis of QOF targets. Nonetheless this may have highlighted different responses to treatments in subgroups with different co-morbidities, which could have informed policy.

The statistical significance of the *Hospre* variable in all models is suggestive of dynamics in the outcome measure, namely whether or not the individual had an outcome in the previous period or periods, plays a significant role in whether they have one in the present period. Lagged dependent variables create problems with respect to the error term and exogeneity assumption, and require complex econometric solutions, which were beyond the scope of the study

This study looked at specific outcome measures using IHD ICD 10 codes in the primary diagnosis by hospitalisation file. Other outcomes were considered or examined briefly before rejection such as hospital length of stay. Other important outcomes were not, such as number of relevant surgical procedures, time to an outcome and death within a specified duration of diagnosis.

7.6 Implications for practice

This work has shown that meeting the selected evidence based CHD targets used in the QOF significantly reduces evidence linked hospital admissions. This is the case whether the outcome is specified as a count or binary measure; fixed or random effects are used; when adjustments are made for excess zeros, and whether whole population or subgroup targets are considered; in the vast majority of cases. As such the effectiveness of the chosen QOF targets has been conclusively analysed and demonstrated. However the improvement is lagged meaning that compliance with targets reduces outcomes in the subsequent period. The cholesterol target is alone in having an immediate reductive impact on outcomes in addition to a lagged one.

ACE inhibitors were shown to be effective in all patients. This would support extending the existing QOF target to all patients with CHD, not just those with MI. Deprivation was generally a strong predictor of outcomes. This shows that more deprived CHD patients had worse clinical outcomes, and future design changes in the QOF may wish to focus care more intensively on those patients. Descriptive analysis showed that the study population became increasingly comorbid over the study period. Panel data models that most comorbidities played a significant role in

increasing outcomes. There are very few multi-morbidity targets in the QOF, despite this situation becoming increasingly common and having significant health implications, particularly with respect to polypharmacy[142]. There are a number of targets that could be merged across conditions in the QOF to avoid the potential harm and costs brought about by the proliferation of drugs in those patients.

7.7 Implications for future research

The study has highlighted the possibility for research utilising individual level linked data and panel data techniques and hence demonstrated the opportunity for future research in this area. This could explore alternative hurdle and zero inflated model specification to correct for excess zeros. An excess of zeros is a problem common to many measures in health care, whether they be hospital outcomes or visits to a GP, which require zero inflated model solutions. Hence there is scope for much greater research and refinement in this area, with large administrative datasets offering the opportunity to undertake it on a range of outcome measures and to conduct it relatively quickly. There is also opportunity to study dynamics in the outcome measure and interactions between included variables. Finally this research looked at hospitalisations where selected IHD ICD 10 codes were the primary diagnosis for that hospitalisation. Future research may wish to consider alternative outcomes such as hospital length of stay and surgical procedures.

7.8 Conclusion

This chapter has explored the key explanatory and outcome variables in the study dataset and examined the impact of co-morbidities. The results from the descriptive analysis show that the data extracted is largely representative of the wider English disease population, and that trends in QOF target attainment and their mapped surrogate measures found in the study population, are largely mirrored in the primary literature and QOF statistics produced by the HSCIC.

Those variables were then included in a number of econometric models. The results showed that meeting the selected evidence based targets in the QOF consistently reduced evidence linked hospital admissions at an individual patient level. In addition to this co-morbidities, higher deprivation and having an outcome prior to entry, in the main resulted in a greater number of outcomes. Improvements in QOF CHD practice attainment consistently reduced outcomes.

Chapter 8 Discussion and synthesis

8.1 Objective

To summarise and synthesise the information gathered in this thesis, discuss its relevance to the research aim, identify policy implications, and areas for future research.

8.2 Background

The thesis set out to explore the effectiveness of physician incentives on individual level clinical outcomes. This was undertaken from an NHS perspective, where service reorganisation on the back of significant new investments in primary care, saw the introduction of the QOF in 2004, recognised as one of the biggest investments in P4P in the world[1 3]. To examine the effectiveness of evidence based targets in the QOF this thesis used individual level data extracted from large administrative medical databases. Rather than examining the whole QOF clinical domain, CHD target were chosen because they had good evidence of clinical effectiveness and extractable data from administrative datasets. The effectiveness of evidence based targets in the QOF were analysed using panel data econometric models.

This chapter details how the separate chapters informed this thesis, and together addressed the research aim. It further identifies the specific strengths and limitations of each chapter and any overall implications for practice and future research.

8.3 Executive summary

The literature reviews identified a number of weaknesses in research to date. There was little high level study design evidence and incentives were examined over short durations. The outcome measures used were mostly process and surrogate outcomes, no study on P4P looked at clinical outcomes, and none outside of P4P considered condition specific clinical outcomes. The research on the QOF and CHD usually involved only two time points, had no control, and was largely reliant on data from one PCT. This work has sought to correct the problems identified in the present literature. It has used administrative datasets to extract data on a large sample of CHD patients from across England and tracked them, at an individual level, in primary and secondary care. QOF targets and clinical outcome measures were

selected on the basis of high level clinical evidence. The impact of evidence based standards in the QOF has been examined over an 11 year period four of which were prior to the QOF. By utilising panel data model techniques, individual patients have been able to act as their own comparator. Adjustments were made to model specifications for over dispersion, excess zeros, heterogeneity in, and covariance with, the error term. In all whole population and most subgroup models evidence based CHD targets in the QOF were found to significantly reduce evidence linked hospital admissions after a one year delay. This shows that evidence based targets can have a significant and positive impact on patient health in real world uncontrolled settings outside of clinical trials.

8.4 The effectiveness of physician incentives and evidence for QOF clinical targets

8.4.1 Context and relevance to the thesis

The QOF was introduced in April 2004 and near universally adopted; hence there is no control group by which to judge its effectiveness. Furthermore there is no published research conducted prior to the QOFs introduction to justify the choice of targets, the design of the QOF, or the specific choice of P4P as the means of incentivising GPs. This thesis attempted to establish a clear methodological approach to assess the impact of QOF targets in this context. It did so over the course of three literature reviews and a review of the clinical guideline evidence for QOF clinical targets.

Chapter 2 undertook a review of systematic reviews of incentives in primary care, limited to high level study designs. This sought to determine the evidence for the effectiveness of physician incentives in primary care to provide a context for the QOF and P4P. This revealed that:

1. There is a paucity of high level research evidence on the effect of physician incentives.
2. A lack of high level research, and heterogeneity of outcome measures, prevented meta-analyses of findings.
3. There was no evidence found that would justify the use of P4P as the chosen payment method for the QOF.
4. There is an urgent need for research which meets the following requirements:
    i. Examines the effectiveness of incentives over a longer duration

ii. Considers the effectiveness of physician incentives on clinical outcomes

The first literature review in Chapter 2 clearly showed that there is great uncertainty over the effectiveness of financial incentives and a number of methodological weaknesses in the high level research evidence, common among which were selection bias, a lack of blinding for both physicians and patients, and poor external validity.

To remove bias in the selection of targets and outcomes, it was important to base those decisions on high level clinical evidence[11 18]. Chapter 3 therefore selected from the QOF, exemplar targets, with demonstrable, and relevant, clinical outcomes, supported by high level clinical evidence. The process surrounding the introduction and retirement of new targets has undoubtedly become more robust and formalised with the introduction of the QOF committee in 2009. Nonetheless there is a lack of critical appraisal of the clinical effectiveness of QOF targets. Specifically there are no studies which incorporate the evidence base in their analysis of the QOF or incentives generally. This chapter identified that many of the QOF targets, 2009/10, were not supported by high level clinical evidence. Hence studies which do not consider the clinical evidence, ignore clinical outcomes, and simply report process measures in their work, often wrongly assume that improvements in recording and activity are, or have, a clinical benefit.

Following the selection of CHD as the exemplar condition Chapter 4 undertook a literature review to look at the effects of physician incentives on CHD clinical outcomes. This was in order to determine if research similar to that about to be attempted by this thesis had been undertaken, and how previous research could inform this study. Chapter 2 included only high level design and primary care incentives; it was felt based on the evidence of that review that replicating that approach would draw few if any studies. It was therefore decided to include any study design with a within or between group comparator, so long as some attempt was made to match patients on key socio-demographic variables at baseline or the results stage. Key findings from the review that informed this thesis were:

1. There was a paucity of research on the impact of physician incentives on clinical outcomes
2. No studies were found linking primary care physician incentives to clinical outcomes

3. No strong evidence was found for the effectiveness of P4P.

Lastly a review of the research literature on CHD and the QOF was undertaken. Since the previous reviews had required some form of comparator arm, QOF studies had not been returned in those reviews. As the QOF formed the policy and financial intervention in this thesis, this was an important omission. This review therefore was a broad literature review on the CHD QOF literature with few limitations on study design or outcome measures. Key findings from the review that informed this thesis were:

1. There was a very limited research evidence base on the effects of the QOF on individuals with CHD
2. The research evidence was predominantly focused on one administrative database in Wandsworth PCT, an inner London borough, and was therefore unlikely to be reflective of the UK as a whole.
3. There was no assessment of the clinical effectiveness of QOF targets beyond the surrogate and process measures which formed those targets.
4. The CHD QOF literature generally used uncontrolled before and after cross sectional study designs. Only one study followed or observed patients over a significant period and considered pre-existing trends when looking at the impact of the QOF[119]. This study found that the QOF did not lead to a significant improvement on existing trends[119].

The literature reviews highlighted the difficulty of conducting research in this area, hence a limited role for what are considered high level study designs, and a number of limitations in existing research. Nonetheless a number of important findings were revealed with respect to incentivisation. The evidence suggests that physicians do respond according to the incentives on offer and these can be used to direct physician behaviour. Hence in FFS physicians tended to provide more of the services that drew incentives. Another way of interpreting this is that while incentives may influence physician behaviour, physicians remain the principal and not the agent. In this regard physicians ensure that regardless of the incentives on offer they provide a certain amount and type of activity in order to maximise their income or utility, in line with the target income hypothesis examined in the paper by Krasnik[52]. Perhaps unsurprisingly therefore was little evidence for one form of incentivisation being 'superior', and limitations in all forms of physician incentives were uncovered. With respect to P4P no significant body of evidence was found to justify its use in the

QOF. Overall, and of specific interest to this thesis there was little examination of the effects of incentives on evidence based outcomes.

8.4.2 Strengths

There have been a number of reviews on the effect of incentivisation in health care, P4P specifically. Hence there are recent examples by Eijkenaar et al who undertook a review of systematic reviews on P4P and Van Herck et al who undertook a broad based systematic review of primary studies on P4P[43 143]. Likewise for the QOF P4P incentive scheme specifically, Steel et al have undertaken a review of the existing primary literature[144]. Within the Cochrane collaboration there have been a number of systematic reviews looking at the impact of P4P and physician incentives more generally[36 40 45].

The literature reviews in this thesis do not attempt to replicate any of those studies. Rather in conjunction with the clinical guidelines they have sought to triangulate the evidence around physician incentives, and their impact on clinical outcomes, focusing on the exemplar condition, CHD. None of the previous reviews have analysed incentives from this perspective, linking into the clinical evidence, to see if research has focused on clinically significant and relevant outcomes. The approach adopted by this thesis provided a formal process for the analysis of evidence based QOF targets in this research and could do so in further work. This is important for research on the QOF due to the lack of a control and questions over its design and implementation discussed previously.

8.4.3 Limitations of the reviews

There are a number of weaknesses in the reviews with regard to their inclusion criteria specifically and generally with regard to the studies found beyond those previously mentioned. These are:

1. Due to heterogeneity in the conditions examined, study designs used, and outcomes measured, there was no opportunity to undertake a meta-analysis of their outcomes
2. As with any literature review there is the potential for publication bias. A number of other sources of bias were mentioned within the systematic reviews. Common among the P4P literature was a lack of blinding in the intervention, bias in the selection of physicians and patients; and poor external validity of the outcome

measures[45]. These limitations equally applied to the literature reviewed in Chapters 4 and 5.

3. The review of the evidence for QOF targets focused on NICE and SIGN guidelines evidence to recommendations and publications by large disease representative bodies. The primary literature was consulted only if the evidence from those sources was unclear.
4. Two of the reviews focused on CHD care only, which has been the focus of this thesis. Extending these reviews to other conditions may have found more substantive evidence on the effectiveness of incentives.
5. The reviews only considered English language publications and for the most part search engines available through the OvidSP search platform. Consequently the literature was predominantly from the USA and UK. English language journals not available through OvidSP and PubMed were missed from the reviews; as was literature from the majority of non-English speaking countries and their respective health services.

8.5 Analysis of the QOF using large administrative datasets

Following the literature and guidelines reviews the thesis moved onto data analysis. This was conducted and described in Chapters 6 and 7. These two chapters are described below.

8.5.1 Using CPRD and HES data

8.5.1.1 Context and relevance to the thesis

Chapter 6 explained how the CPRD and HES administrative datasets were used to extract patient linked data and construct variables for inclusion in panel data models. This was primarily a descriptive chapter outlining the file structures present within the data sources and describing in detail the construction of key variables. Any concerns with respect to data quality, and the protocols used to overcome them were discussed. Strengths and limitations of the data outside of clear anomalies, or missing data, were examined during data descriptive analysis in Chapter 7.

8.5.2 Data description and analysis

8.5.2.1 Context and relevance to the thesis

In this chapter the study variables were descriptively analysed to find out how representative the study population was of the wider English national population, and whether variables followed trends or agreed with levels published in national level statistics and previous research. Econometrics, panel data models specifically, were described before formal analyses was undertaken using those techniques. Panel data models were then used to determine if there was a statistically significant relationship between evidence based CHD targets in the QOF and evidence linked hospital admissions, at an individual patient level.

8.5.2.2 Principal finding

From the review of the evidence on the QOF in this thesis, the literature identified simply showed a link between the introduction of the QOF and attainment on QOF targets, or their mapped surrogate measures. This was usually done for two time periods and for an inner London borough. The data analysis in this research has advanced that research significantly. It has shown at an individual level, that there is a statistically significant relationship between meeting an evidence based standard in the QOF and the rates or odds of evidence linked hospital admissions. This relationship has been verified in a number of model specifications, where the variable coefficients have been largely consistent in magnitude and significance, showing that this is a robust finding. The effectiveness of QOF targets has been examined over an 11 year period and from a large sample population drawn from across England, suggesting the results are generalisable and have high validity. Importantly the results have been demonstrated in a 'real world' uncontrolled setting.

8.5.3 Critical appraisal

8.5.3.1 Strengths

There are a number of strengths of the approaches taken in this chapter. These are:

1. Descriptive analysis found the study population to be representative of trends evidenced in previous research and national statistics in terms of the direction of QOF target attainment and levels on mapped surrogate outcomes over the study period. The same was true of the study outcome measure as well as the age and gender mix of the study population.

2. Individuals were followed for up to eleven years, seven following the introduction of the QOF. This enabled the study to look at the impact of repeated compliance with evidence based standards in the QOF, and state changes.
3. Panel data analyses were used to model individual level outcomes. The benefits of this approach were:
   i. Given the absence of a control or comparator group in the QOF, individual level panel data enables patients to act as their own comparator. In this context, within patients, the previous period in the panel provides the comparator/control for the present period. In much the same way as they do in a controlled before and after study.
   ii. Variables could be lagged to assess temporal effects. This is important for the QOF as treatment can be initiated at any point in the financial year, including post a study outcome. Also with prescriptive targets in particular full clinical benefits are often realised some time after treatment initiation.
4. Analyses allowed for fixed and random effects, included robust standard errors and adjusted for over dispersion and excess zeros in the outcome variable.

8.5.3.2 Limitations

There were a number of limitations to the approaches taken in the chapter that may limit its findings. These were:

1. This research used hurdle models to account for excess zeros in the count outcome variable. As there were no existing commands to do this in Stata, it relied on adapting pre-existing commands to create a hurdle model.
2. This research did not include dynamics in the outcome variable due to the complications to econometric modelling these create, which were beyond the scope of this research. However the significance of the *Hospre* variable in all models is highly suggestive of dynamics in the outcome variable
3. The effect of workload in relation to GP's only was examined. This ignored other staff grades that play an important role in the delivery of the QOF such as nurses, and hence may not have been a representative figure. Nor was the effect of staff case mix on outcomes examined.
4. The focus in this research has been on the effects of QOF targets on clinical outcomes at the individual level. Effects have not been examined at the practice level and due to this nor has the impact of exception reporting.

5. The outcome measures used have all been primary diagnoses by hospitalisation pertinent to CHD Secondary prevention. Other potential outcomes not considered are surgical procedures; death within a set time frame of diagnosis or hospitalisation; and other measures of resource usage such as time spent in intensive care or hospital.
6. To act as their own comparator in a panel model, individuals are required to transition between states, in this study specifically, move from meeting to not meeting a QOF target and vice versa. A significant number of individuals remained in the same state throughout their time in the study, meaning that they could not perform that role; the variables became time invariant and were lost in fixed effects analysis.

8.6 Implications for practice

This research suggests strongly that the selected evidence based targets in the QOF targets work, namely that meeting the QOF standard leads to significant improvements in evidence linked hospital admissions, at the individual patient level. However there are aspects of the QOF that could be improved. This research shows that there are significant clinical benefits to be gained from extending ACE inhibitors to the entire CHD population, and not limiting it to those with recently diagnosed MI.

The CHD study population aged, and became increasingly co-morbid and multi-morbid over the course of the study. Despite this situation becoming increasingly common and having health and cost implications, particularly in the area of polypharmacy, the QOF still largely focuses on single conditions[142]. There remains scope within the QOF to merge conditions such as the cardiovascular diseases, and targets such as blood pressure which may not only have clinical benefits but could free up QOF points for use in other areas.

This research demonstrated the importance of linking targets to evidence, and was surprised at how little of the QOF was evidence based and how few targets had process measures or surrogate outcomes that mapped to clinical outcomes. There have been significant improvements made since the end of the study period in this area. These have been brought about by innovations such as the NICE QOF committee which has led to more robust use of evidence to guide the introduction of

new targets and retirement of existing ones. The most recent QOF guidance, 2013/14, is now much slimmed down and most of the targets based on simple record keeping have been stripped out. Nonetheless there are still a number of incumbent targets which lack a high level evidence base and have clinical benefits that are difficult to measure and extract from administrative datasets.

This work found in most models, that there still remains a significant inequality gap, which was statistically significant throughout in the most deprived quintile. The most deprived quintile were found to have higher odds or rates for the outcome measures of around 30%-50%, compared to the least deprived quintile. Given the significant sums of money spent on the QOF, this may have been a missed opportunity to make more progress in eliminating these inequalities. Future changes therefore may want to address this by for instance increasing incentives to those who treat more deprived patients or incorporating socio-economic measures into QOF targets.

There was also evidence found to support the target income hypothesis which argues that individuals have a target income and so will work a requisite amount to get to that income. Hence incentives can lead to an initial spurt in activity until the physician determines how much activity is required to meet their target income, at which point it settles. While this was not applied in a QOF context, it is possible to see supporting evidence for it in the QOF, where greater incentives and contractual changes saw significant numbers of GPs decide to opt out of out of hours surgery, and a fall in the hours worked by GPs as incomes rose[145]. This finding would support a reduction in QOF payments or more demanding payment thresholds, the latter of which has been shown to increase attainment, most significantly among poorer performing practices[146].

8.7 Comparison to previous research

The extension of ACE inhibitors to those with stable CHD has been shown to be both clinically and cost effective[147]. This was based on data from the EUROPA trial which involved over 12,000 randomised patients from 24 countries with a mean follow up of over 4 years[147]. The findings do however relate only to one specific ACE inhibitor, perindopril. The American Heart Association recommends that long term ACE therapy should be considered in all patients with coronary or other

vascular diseases unless contraindicated[148]. It has also been recognised by the Joint British Societies, which notes that ACE inhibitors have wider benefits outside of those who have just had an MI, in particular in individuals with elevated BP [87]. These findings however have yet to be incorporated in guideline to evidence recommendations in the UK, by bodies such as NICE and SIGN[81].

The investigation of lagged effects is common in economics, particularly in macroeconomic models when studying the implications of changes in interest rates for example. Within the medical literature and health economics they are examined far less frequently. Hence it is difficult to find evidence which looks at lagged patient level effects in clinical treatment, and none was found which looked at them in the context of the QOF targets used in this study. This highlights the fact that this research is novel. Nonetheless there is general evidence to support a lag in clinical benefits, for the QOF targets examined, and indeed for immediate benefits from the use of statins. However these benefits have been mostly demonstrated using general mortality measures and not for the outcomes measured in this study.

The clinical benefits of statins, specifically atorvostatin, have been shown to be in evident within 30 days of treatment initiation in the treatment of artherosclerotic vascular disease, in terms of significant reductions in all-cause mortality[149]. This would support the immediate benefits evidenced in the majority of models in the study from meeting the cholesterol QOF target, which related to statin use. Studies show clinical benefits of ACE inhibitors within patients post MI, occurring within 6 weeks, when measured in terms of mortality. However these are often not significant, and build with long term initiation, which would support a lagged effect certainly in the MI subgroup, the group covered by the QOF target[150] [151]. The benefits of beta blockers have been demonstrated in terms of reduced mortality over a one year to four year follow up period. This does provide some evidence for a delay in treatment effect, however no studies could be found that measured its effects over a shorter period to fully validate this, and the results referred specifically to post MI patients[152]. Anti-platelets have been shown to significantly reduce risks of serious vascular event in those with stable angina and in post MI patients at 2 years[153]. This outcome measure links closely to the study outcome measure.

The finding that the co-morbidities were largely significant was expected as they are all atherosclerotic conditions or risk factors. However while risk models have been

developed for the primary prevention of CHD which look at co-morbidities, no similar research could be found for secondary prevention[154]. This would have provided some assistance in interpreting the results in this study. Heart Failure was the only condition that significantly increased outcomes in all models. It is known that an MI is a significant risk factor for heart failure, and it is unsurprising therefore that those who had damage to the heart and may have suffered a previous MI were a high risk group[155]. What is not clear however is why this condition should figure more prominently than Stroke, AF, and Diabetes. Rheumatoid Arthritis is a known risk factor in the primary prevention of CHD, and was included in the study for that reason[154]. Research from a Danish wide cohort has shown that rheumatoid arthritis significantly increases the risk of having an MI (IRR =1.7, 95% CI 1.5-1.9)[156]. This was of a similar magnitude to the increased risk of MI witnessed in patients with diabetes in the same study. Given the significance of diabetes in this study, albeit not consistently so, this makes the results found for RA in this study potentially questionable. However it should be noted that the results were for the general population of Denmark, and not those with CHD, hence results do not translate across into this study. Since RA is an autoimmune disease and not atherosclerotic, the causal pathway in relation to the outcome measures is not as clear as it is for the other co-morbidities. As the least common of the co-morbidities it is also possible that there were insufficient number of patients with the condition to pick up a significant effect, assuming one exists.

The evidence on the impact of the QOF on inequalities, shows that the QOF had a very limited impact on inequalities, which largely persisted[157 158]. This is supportive of the finding in this thesis. However whereas previous research has measured inequalities largely in relation to QOF attainment and at a practice level, this work has done so for clinical outcomes at the individual patient level.

8.8 Implications for future research

This research has opened up a wide range of future research opportunities, particularly in relation to individual level analysis using linked primary and secondary care data, which may improve upon or expand this work. This study used specific ICD 10 codes extracted from the primary diagnoses over a hospitalisation file to create its outcome measure. Future research may wish to look at alternative outcomes available in linked HES. Examples include clinical procedures, ICU or

hospital length stay of stay. There are a number of additional individual linked data sources available to users of CPRD among them Office for National Statistics (ONS) mortality data and MINAP, that could also be used to create alternative outcome measures, or in the case of the latter alternative population specifications. Examples include time from diagnosis to death, or MI's confirmed using MINAP.

This research looked at the impact of dynamics in the study explanatory variables but not in the outcome measures, which could be analysed in future econometric models. Existing Stata commands were used to generate the hurdle model in this thesis. Future work could look at alternative specifications of the hurdle model and explore the possibility of using Stata commands to generate a panel hurdle model.

In terms of the literature this research has only looked at the English language evidence on the effectiveness of incentives. Future research could look at non English peer reviewed and grey literature to see if this can provide more evidence on the effects of incentivisation.

8.9 Summary and conclusion

To ensure there was a strong evidence link and the outcome was extractable using large administrative datasets this research used clinical guideline evidence to direct the choice of targets. This ensured the process was transparent, methodical, and importantly evidence driven. The literature reviews identified a clear need for research which linked primary care incentives to clinical outcomes which this thesis then addressed.

This thesis has shown that meeting evidence based CHD targets in the QOF significantly reduces evidence based hospital admissions after a one year lag. It has done so using individual linked data from large administrative datasets over an 11 year period. The data extracted has been analysed using panel data econometric models. These models adjusted for over dispersion and excess zeros in the outcome term, used random and fixed effects to deal with potential correlation with included explanatory variables, and reported robust standard errors to account for heterogeneity in the error term. The analyses undertaken in this thesis are innovative in a QOF context.

Chapter 9 Appendices

9.1 Appendix 1 Incentives in primary care search terms

9.1.1 Medline

1. Physician Incentive Plans/ [ML]

2. ((physician? or practitioner? or doctor?) adj4 (bonus$ or incentive? or financial or monetar$ or payment? or profit shar$ or reward? or salar$)).ti,ab.

3. exp Physicians/ and (incentiv$ adj (economic or financial or monetar$ or payment? or reimburs$)).ti,ab,hw.

4. or/1-3

5. Fees, Medical/ [ML]

6. exp Income/ [ML]

7. insurance, health, reimbursement/ or reimbursement mechanisms/ or fee-for-service plans/ or physician payment review commission/ or medicare payment advisory commission/ or reimbursement, disproportionate share/ or reimbursement, incentive/ or relative value scales/ [ML]

8. ((fees.mp. and charges/) or capitation fee/ or fee-for-service plans/ or fees, medical/ or fees, pharmaceutical/ or prescription fees/ or rate setting.mp.) and review/ [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]

9. (capitation or capitated or capitating or fundhold$ or fund-hold$).ti,ab.

10. (rate setting or rate review).ti,ab.

11. (gainshar$ or payer-provider? or payer-patient?).ti,ab.

12. ((incentiv$ or bonus$ or reward?) adj (economic or employee? or financ$ or insurer? or insurance or market$ or monetar$ or payment? or physician? or practitioner? or program$ or provider? or reimburs$ or salary or staff or team$ or value-based)).ti,ab.

13. (reimburs$ adj (disproportion$ or health$ or insur$ or mechanism? or plan$ or physician? or practitioner? or program$ or proportion$ or provider? or relative or

scale? or share? or sharing or value-based or performance-base? or QI or quality or scheme?)).ti,ab.

14. (pay for compliance or pay for participation or pay for performance or performance pay$ or P4P or pay for quality improvement? or P4QI or fee-for service?).ti,ab.

15. (payment? adj (blend$ or blue cross or bonus$ or capped or episode of care or fixed or government$ or insurance or insurer? or level? or linear or medicaid or medicare or non-linear or per-patient or per-episode or per-visit or performance or prospectiv$ or retroactiv$ or retrospectiv$ or reward$ or schedule? or system? or target$ or third-part$ or threshold? or uncap$ or shared or variable or per-visit?)).ti,ab.

16. ((compensation or compensatory) adj (doctor? or physician? or plan? or practitioner? or system?)).ti,ab.

17. (copay$ or co-pay$ or cost-shar$ or prepaid or pre-paid or prepay$ or pre-pay$).ti,ab.

18. or/5-17

19. exp Primary Health Care/ [ML]

20. (primary adj2 care).ti,ab.

21. Physicians, Family/ [ML]

22. exp physicians/ [ML]

23. Family Practice/ [ML]

24. exp Group practice/ [ML]

25. ((community or family or general or group) adj2 (doctor? or physician? or practice? or practitioner?)).ti,ab.

26. Partnership practice/ or Private practice/ [ML]

27. ((partner$ or private) adj (practice? or practitioner?)).ti,ab.

28. or/19-27

29. 4 or 18

30. 28 and 29

31. limit 30 to (yr="1990 - 2013" and systematic reviews)

9.1.2 Embase

1. (Physician? adj2 Incentiv$ adj (plan? or program or policy)).ti,ab.

2. ((physician? or practitioner? or doctor?) adj4 (bonus$ or incentive? or financial or monetar$ or payment? or profit shar$ or reward? or salar$)).ti,ab.

3. exp Physicians/ and (incentiv$ adj (economic or financial or monetar$ or payment? or reimburs$)).ti,ab,hw.

4. or/1-3

5. *Medical fee/ [EM]

6. *Income/ or *Salary/ or Physician Income/ [EM]

7. Reimbursement/ or reimburs$.ti. [EM]

8. *Income/ or *Salary/ or Physician Income/ [EM]

9. (pay for performance or pay for compliance or pay for participation or performance pay$ or P4P or pay for qualityimprovement? or P4QI or fee-for service?).ti,ab.

10. ((doctor? or physician? or general practitioner? or pa?ediatrician?) adj2 (fee? or income or salary or salaries)).ti,ab.

11. ((incentiv$ or bonus$ or reward?) adj (economic or employee? or financ$ or insurer? or insurance or managed care or HMO or market$ or monetar$ or payment? or performance based or physician? or practitioner? or program$ or provider? or reimburs$ or salary or staff or team$ or value-based)).ti,ab.

12. (reimburs$ adj (disproportion$ or health$ or insurer? or mechanism? or plan$ or physician? or practitioner? or program$ or proportion$ or provider? or relative or scale? or share? or sharing or value-based or performance-base? or QI or quality or scheme?)).ab.

13. ((compensation or compensatory) adj2 (doctor? or physician? or plan? or practitioner? or system?)).ti,ab.

14. (capitation or capitated or capitating or fundhold$ or fund-hold$).ti,ab.

15. (rate setting or rate review).ti,ab.

16. (gainshar$ or payer-provider? or payer-patient?).ti,ab.

17. (pay for compliance or pay for participation or pay for performance or performance pay$ or P4P or pay for quality improvement? or P4QI or fee-for service?).ti,ab.

18. (payment? adj (blend$ or blue cross or bonus$ or capped or episode of care or fixed or government$ or insurance or insurer? or level? or linear or medicaid or medicare or non-linear or per-patient or per-episode or per-visit or performance or prospectiv$ orretroactiv$ or retrospectiv$ or reward$ or schedule? or system? or target$ or third-part$ or threshold? or uncap$ or shared or variableor per-visit?)).ti,ab.

19. ((copay$ or co-pay$ or cost-shar$ or prepaid or pre-paid or prepay$ or pre-pay$) adj4 (physician? or practitioner? or performance)).ti,ab.

20. or/5-19 [Incentives]

21. physician/ or female physician/ or general practitioner/ or pediatrician/ [EM]

22. (physician? or doctor?).ti.

23. primary health care/ or primary medical care/ [EM]

24. (primary adj2 care).ti,ab.

25. General Practitioner/ [EM]

26. General Practice/ [EM]

27. Group practice/ [EM]

28. ((community or family or general or group) adj2 (doctor? or physician? or practice? or practitioner?)).ti,ab.

29. Private practice/ [EM]

30. ((partner$ or private) adj (practice? or practitioner?)).ti,ab.

31. or/21-30 [Practitioners or Primary Care]

32. 20 and 31

33. 4 or 32

34. limit 33 to ("systematic review" and yr="1990 - 2013" and journal)

9.1.3 PubMed

systematic[sb] AND ((financial incentives) AND primary care)

systematic[sb] AND ((payment mechanisms) AND primary care)

9.2 Appendix 2 PRISMA checklist

Table 9-1 PRISMA checklist details

Section/Topic	Number	Checklist item
TITLE		
Title	1	Identify the report as a systematic review, meta-analysis, or both.
ABSTRACT		
Structured summary	*2* ***(1)***	*Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.*
INTRODUCTION		
Rationale	*3* **(2)**	*Describe the rationale for the review in the context of what is already known.*
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).

METHODS		
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.
Eligibility criteria	*6* **(3)**	*Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.*
Information sources	*7* **(4)**	*Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.*
Search	*8* **(5)**	*Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.*
Study selection	*9* **(6)**	*State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).*
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and

		simplifications made.
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.
Risk of bias across studies	*15* **(7)**	*Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).*
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.
RESULTS		
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.

Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).
DISCUSSION		
Summary of	*24* **(8)**	*Summarize the main findings including the strength of evidence for each main outcome; consider their relevance*

evidence		*to key groups (e.g., healthcare providers, users, and policy makers).*
Limitations	*25* **(9)**	*Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).*
Conclusions	*26* **(10)**	*Provide a general interpretation of the results in the context of other evidence, and implications for future research.*
FUNDING		
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.
Source: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097[47] *Checklist items shown in italics represent those selected for the review of reviews in Chapter 2.* **(Question number used in Chapter 2 shown in bold and parentheses)**		

9.3 Appendix 3 Population calculations

The population percentage figures represent the prevalence of the condition in 2009/10 based on England population level data published on the HSCIC QOF website[139 159]. If the target applied to a subset of, rather than the whole condition population, that percentage figure has been calculated by finding the target denominator for that target and dividing it by the denominator of a target that applied to the whole condition population. In each instance the blood pressure record target has been used for the latter as this is present in all the conditions examined and is routinely collected. As before the information to populate these figures was retrieved from the HSCIC QOF website[139]. SHA level data has been used to make these calculations as the national level data is not broken down by targets, only by condition. This calculation produced the percentage of the condition population treated by the QOF targets. The figures in the tables represent the percentage of the UK population treated by the QOF target. To produce this figure the result from the calculation described above was multiplied by the condition prevalence. This will very closely match the actual disease population but not exactly as a small number of exception reported patients will have been removed from it. For example the population figure for target CHD11 has been calculated as follows:

$$\frac{Sum\ of\ CHD11\ demominators}{Sum\ \ of\ CHD5\ denominators} \times Condition\ prevalence$$

Where the numerator represents all patients who qualified for the CHD11 target, namely all patients with recently diagnosed MI minus those exception reported. While the denominator is all those who qualified for the BP record target, which is a very close approximation of the condition population, the difference being the number of patients exception reported.

As an illustrative example to show how the figures shown in the table would relate to events of interest in HES based on a conservative figure for the HES linked population of two million. Assuming an outcome event rate of 5% a year, and that the study wanted to be assured of finding at least 3,000 events a year; would require a population percentage figure of 3% or more.

9.4 Appendix 4 CHD Physician Incentives Search Terms

9.4.1 Medline

1. Physician Incentive Plans/ [ML]

2. ((physician? or practitioner? or doctor?) adj4 (bonus$ or incentive? or financial or monetar$ or payment? or profit shar$ or reward? or salar$)).ti,ab.

3. exp Physicians/ and (incentiv$ adj (economic or financial or monetar$ or payment? or reimburs$)).ti,ab,hw.

4. or/1-3

5. Fees, Medical/ [ML]

6. exp Income/ [ML]

7. insurance, health, reimbursement/ or reimbursement mechanisms/ or fee-for-service plans/ or physician payment review commission/ or medicare payment advisory commission/ or reimbursement, disproportionate share/ or reimbursement, incentive/ or relative value scales/ [ML]

8. ((fees.mp. and charges/) or capitation fee/ or fee-for-service plans/ or fees, medical/ or fees, pharmaceutical/ or prescription fees/ or rate setting.mp.) and review/ [mp=title, abstract, original title, name of substance word, subject heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]

9. (capitation or capitated or capitating or fundhold$ or fund-hold$).ti,ab.

10. (rate setting or rate review).ti,ab.

11. (gainshar$ or payer-provider? or payer-patient?).ti,ab.

12. ((incentiv$ or bonus$ or reward?) adj (economic or employee? or financ$ or insurer? or insurance or market$ or monetar$ or payment? or physician? or practitioner? or program$ or provider? or reimburs$ or salary or staff or team$ or value-based)).ti,ab.

13. (reimburs$ adj (disproportion$ or health$ or insur$ or mechanism? or plan$ or physician? or practitioner? or program$ or proportion$ or provider? or relative or scale? or share? or sharing or value-based or performance-base? or QI or quality or scheme?)).ti,ab.

14. (pay for compliance or pay for participation or pay for performance or performance pay$ or P4P or pay for quality improvement? or P4QI or fee-for service?).ti,ab.

15. (payment? adj (blend$ or blue cross or bonus$ or capped or episode of care or fixed or government$ or insurance or insurer? or level? or linear or medicaid or medicare or non-linear or per-patient or per-episode or per-visit or performance or prospectiv$ or retroactiv$ or retrospectiv$ or reward$ or schedule? or system? or target$ or third-part$ or threshold? or uncap$ or shared or variable or per-visit?)).ti,ab.

16. ((compensation or compensatory) adj (doctor? or physician? or plan? or practitioner? or system?)).ti,ab.

17. (copay$ or co-pay$ or cost-shar$ or prepaid or pre-paid or prepay$ or pre-pay$).ti,ab.

18. or/5-17

19. myocardial ischemia/ or angina pectoris/ or coronary disease/ or myocardial infarction/

20. 4 or 18

21. 19 and 20

22. limit 20 to yr="1990 -Current"

9.4.2 Embase

1. *Medical fee/ [EM]

2. *Income/ or *Salary/ or Physician Income/ [EM]

3. Reimbursement/ or reimburs$.ti. [EM]

4. *Income/ or *Salary/ or Physician Income/ [EM]

5. (pay for performance or pay for compliance or pay for participation or performance pay$ or P4P or pay for quality improvement? or P4QI or fee-for service?).ti,ab.

6. ((doctor? or physician? or general practitioner? or pa?ediatrician?) adj2 (fee? or income or salary or salaries)).ti,ab.

7. ((incentiv$ or bonus$ or reward?) adj (economic or employee? or financ$ or insurer? or insurance or managed care or HMO or market$ or monetar$ or payment? or performance based or physician? or practitioner? or program$ or provider? or reimburs$ or salary or staff or team$ or value-based)).ti,ab.

8. (reimburs$ adj (disproportion$ or health$ or insurer? or mechanism? or plan$ or physician? or practitioner? or program$ or proportion$ or provider? or relative or scale? or share? or sharing or value-based or performance-base? or QI or quality or scheme?)).ab.

9. ((compensation or compensatory) adj2 (doctor? or physician? or plan? or practitioner? or system?)).ti,ab.

10. (capitation or capitated or capitating or fundhold$ or fund-hold$).ti,ab.

11. (rate setting or rate review).ti,ab.

12. (gainshar$ or payer-provider? or payer-patient?).ti,ab.

13. (pay for compliance or pay for participation or pay for performance or performance pay$ or P4P or pay for quality improvement? or P4QI or fee-for service?).ti,ab.

14. (payment? adj (blend$ or blue cross or bonus$ or capped or episode of care or fixed or government$ or insurance or insurer? or level? or linear or medicaid or medicare or non-linear or per-patient or per-episode or per-visit or performance or prospectiv$ orretroactiv$ or retrospectiv$ or reward$ or schedule? or system? or target$ or third-part$ or threshold? or uncap$ or shared or variableor per-visit?)).ti,ab.

15. ((copay$ or co-pay$ or cost-shar$ or prepaid or pre-paid or prepay$ or pre-pay$) adj4 (physician? or practitioner? or performance)).ti,ab.

16. or/1-16 [Incentives]

17. ischemic heart disease/

18. heart infarction/

19. 17 or 18

20. 16 and 19

38. limit 20 to yr="1990 -Current"

9.4.3 PyscINFO

1. monetary incentives/ or monetary rewards/

2. incentives/

3. salaries/ or bonuses/ or income/ or economic/ or income level/ or professional fees/

4. ((incentiv$ or bonus$ or reward?) adj (economic or employee? or financ$ or insurer? or insurance or market$ or monetar$ or payment? or physician? or practitioner? or program$ or provider? or reimburs$ or salary or staff or team$ or value-based)).ti,ab.

5. (reimburs$ adj (disproportion$ or health$ or insur$ or mechanism? or plan$ or physician? or practitioner? or program$ or proportion$ or provider? or relative or scale? or share? or sharing or value-based or performance-base? orQI or quality or scheme?)).ti,ab.

6. (pay for compliance or pay for participation or pay for performance or performance pay$ or P4P or pay for quality improvement? or P4QI or fee-for service?).ti,ab.

7. (payment? adj (blend$ or blue cross or bonus$ or capped or episode of care or fixed or government$ or insurance or insurer? or level? or linear or medicaid or medicare or non-linear or per-patient or per-episode or per-visit or performance or prospectiv$ or retroactiv$ or retrospectiv$ or reward$ or schedule? or system? or target$ or third-part$ or threshold? or uncap$ or shared or variable or per-visit?)).ti,ab.

8. ((compensation or compensatory) adj (doctor? or physician? or plan? or practitioner? or system?)).ti,ab.

9. (copay$ or co-pay$ or cost-shar$ or prepaid or pre-paid or prepay$ or pre-pay$).ti,ab.

10. (gainshar$ or payer-provider? or payer-patient?).ti,ab.

11. (rate setting or rate review).ti,ab.

12. (capitation or capitated or capitating or fundhold$ or fund-hold$).ti,ab.

13. fee for service/

14. reimburs$.ti.

15. medical fee?.ti,ab.

16. or/1-15 [Financial Incentives]

17. exp Heart Disorders/

18. 16 and 17

9.4.5 Health Management Information Consortium (HMIC)

1. exp Coronary diseases/

2. limit 1 to (abstracts and yr="1990 -Current")

3. exp Performance indicators/ or exp Performance related pay/

4. limit 3 to (abstracts and yr="1990 -Current")

5. exp payment by results/

6. limit 5 to yr="1990 -Current"

7. 4 or 6

8. 2 and 7

9.5 Appendix 5 Down and Black ranking scale checklist

Reporting

1. Is the hypothesis/aim/objective of the study clearly described?

Yes	1
No	0

2. Are the main outcomes to be measured clearly described in the Introduction or Methods section? **(1)**

If the main outcomes are first mentioned in the Results section, the question should be answered no.

Yes	*1*
No	*0*

3. Are the characteristics of the patients included in the study clearly described?

In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and the source for controls should be given. **(2)**

Yes	*1*
No	*0*

4. Are the interventions of interest clearly described?

Treatments and placebo (where relevant) that are to be compared should be clearly described.

Yes	1
No	0

5. Are the distributions of principal confounders in each group of subjects to be compared clearly described?

A list of principal confounders is provided.

Yes	2
Partially	1
No	0

6. Are the main findings of the study clearly described?

Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical tests which are considered below). **(3)**

Yes	*1*
No	*0*

7. Does the study provide estimates of the random variability in the data for the main outcomes?

In non normally distributed data the inter-quartile range of results should be reported. In normally distributed data the standard error, standard deviation or confidence intervals should be reported. If the distribution of the data is not described, it must be assumed that the estimates used were appropriate and the question should be answered yes.

Yes	1
No	0

8. Have all important adverse events that may be a consequence of the intervention been reported?

This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).

Yes	1
No	0

9. Have the characteristics of patients lost to follow-up been described?

This should be answered yes where there were no losses to follow-up or where losses to follow-up were so small that findings would be unaffected by their inclusion. This should be answered no where a study does not report the number of patients lost to follow-up.

Yes	1
No	0

10. Have actual probability values been reported(e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001? **(4)**

Yes	*1*
No	*0*

External validity

All the following criteria attempt to address the representativeness of the findings of the study and whether they may be generalised to the population from which the study subjects were derived.

11. Were the subjects asked to participate in the study representative of the entire population from which they were recruited?

The study must identify the source population for patients and describe how the patients were selected. Patients would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random

sample. Random sampling is only feasible where a list of all members of the relevant population exists. Where a study does not report the proportion of the source population from which the patients are derived, the question should be answered as unable to determine.

Yes	1
No	0
Unable to determine	0

12. Were those subjects who were prepared to participate representative of the entire population from which they were recruited?

The proportion of those asked who agreed should be stated. Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.

Yes	1
No	0
Unable to determine	0

13. Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?

For the question to be answered yes the study should demonstrate that the intervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.

Yes	1
No	0
Unable to determine	0

Internal Validity-bias

14. Was an attempt made to blind study subjects to the intervention they have received?

For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.

Yes	1
No	0
Unable to determine	0

15. Was an attempt made to blind those measuring the main outcomes of the intervention?

Yes	1
No	0
Unable to determine	0

16. If any of the results of the study were based on "data dredging", was this made clear?

Any analyses that had not been planned at the outset of the study should be clearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.

Yes	1
No	0
Unable to determine	0

17. In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls ?

Where follow-up was the same for all study patients the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow-up are ignored should be answered no. **(5)**

Yes	*1*
No	*0*
Unable to answer	*0*

18. Were the statistical tests used to assess the main outcomes appropriate?

The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes. **(6)**

Yes	*1*
No	*0*
Unable to answer	*0*

19. Was compliance with the intervention/s reliable?

Where there was non compliance with the allocated treatment or where there was contamination of one group, the question should be answered no. For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.

Yes	1
No	0
Unable to answer	0

20. Were the main outcome measures used accurate (valid and reliable)?

For studies where the outcome measures are clearly described, the question should be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes. **(7)**

Yes	*1*
No	*0*
Unable to answer	*0*

Internal validity - confounding (selection bias)

21. Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population? **(8)**

For example, patients for all comparison groups should be selected from the same hospital. The question should be answered unable to determine for cohort and case control studies where there is no information concerning the source of patients included in the study.

Yes	*1*
No	*0*
Unable to answer	*0*

22. Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time? **(9)**

For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.

Yes	*1*
No	*0*
Unable to answer	*0*

23. Were study subjects randomised to intervention groups?

Studies which state that subjects were randomised should be answered yes except where method of randomisation would not ensure random allocation. For example alternate allocation would score no because it is predictable.

Yes	1
No	0
Unable to answer	0

24. Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?

All non-randomised studies should be answered no. If assignment was concealed from patients but not from staff, it should be answered no.

Yes	*1*
No	*0*
Unable to answer	*0*

25. Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?

This question should be answered no for trials if: the main conclusions of the study were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomised studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no. **(10)**

Yes	*1*
No	*0*
Unable to answer	*0*

26. Were losses of patients to follow-up taken into account?

If the numbers of patients lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small

to affect the main findings, the question should be answered yes.

Yes	1
No	0
Unable to answer	0

Power

27. Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?

Sample sizes have been calculated to detect a difference of x% and y%.

	Size of smallest intervention group	
A	$< n_1$	0
B	$n_1 - n_2$	1
C	$n_3 - n_4$	2
D	$n_5 - n_6$	3
E	$n_7 - n_8$	4
F	$n_8 +$	5

Source[110]

Questions selected for ranking scale used in Chapter 4 are shown in italics

(Question number used in Chapter 4 shown in bold and parentheses)

9.6 Appendix 6 CHD QOF Search Terms

9.6.1 PsycInfo

1. exp Heart Disorders/

2. Quality Outcome* Framework.mp.

3. QOF.mp.

4. P4P.mp.

5. Pay for performance.mp.

6. 2 or 3 or 4 or 5

7. 1 and 6

9.6.2 Medline

1. Coronary Disease/

2. Quality Outcome* Framework.mp.

3. QOF.mp.

4. P4P.mp.

5. Pay for performance.mp.

7. 2 or 3 or 4 or 5

8. 1 and 7

9. limit 8 to yr="2004 -Current"

9.6.3 HMIC

1. exp Coronary diseases/

2. Quality Outcome* Framework.mp.

3. QOF.mp.

4. P4P.mp.

5. Pay for performance.mp.

6. 2 or 3 or 4 or 5

7. 1 and 6

8. limit 7 to yr="2004 -Current"

9.6.4 Embase

1. ischemic heart disease/

2. Quality Outcome* Framework.mp.

3. Pay for performance.mp.

4. P4P.mp.

5. QOF.mp.

6. 2 or 3 or 4 or 5

7. 1 and 6

8. limit 7 to yr="2004 -Current"

9.6.5 ISI Web of Knowledge

Topic=(Quality ADJ outcome* ADJ framework) AND Topic=(Coronary Heart Diseases)

9.7 Appendix 7 Codes used in the study

9.7.1 Codes used to define the CHD population

Table 9-2 lists all the qualifying Read and medical codes for the CHD study population a description of what each Read terms relate to.

Table 9-2 CPRD CHD Read and Medcodes

Medcode	Readcode	Readterm
241	G30..00	Acute myocardial infarction
1204	G30..14	Heart attack
1430	G33..00	Angina pectoris
1431	G311.13	Unstable angina
1677	G30..15	MI - acute myocardial infarction
1678	G308.00	Inferior myocardial infarction NOS
2491	G30..12	Coronary thrombosis
3704	G307.00	Acute subendocardial infarction
4656	G311.11	Crescendo angina
5387	G301.00	Other specified anterior myocardial infarction
7320	G343.00	Ischaemic cardiomyopathy
8935	G302.00	Acute inferolateral infarction
9507	G307000	Acute non-Q wave infarction
10562	G307100	Acute non-ST segment elevation myocardial infarction
12139	G300.00	Acute anterolateral infarction
13566	G30..11	Attack - heart

13571	G30..16	Thrombosis - coronary
14658	G30z.00	Acute myocardial infarction NOS
14897	G301z00	Anterior myocardial infarction NOS
14898	G305.00	Lateral myocardial infarction NOS
17689	G30..17	Silent myocardial infarction
17872	G301100	Acute anteroseptal infarction
18842	G35..00	Subsequent myocardial infarction
23892	G304.00	Posterior myocardial infarction NOS
24126	G360.00	Haemopericardium/current comp folow acut myocard infarct
25842	G33z.00	Angina pectoris NOS
28736	G30y000	Acute atrial infarction
29421	G344.00	Silent myocardial ischaemia
29553	G366.00	Thrombosis atrium,auric append&vent/curr comp foll acute MI
29643	G303.00	Acute inferoposterior infarction
29758	G30X.00	Acute transmural myocardial infarction of unspecif site
29902	G330z00	Angina decubitus NOS
30330	G309.00	Acute Q-wave infarct
30421	G30..13	Cardiac rupture following myocardial infarction (MI)
32272	G38..00	Postoperative myocardial infarction

32854	G30B.00	Acute posterolateral myocardial infarction
34803	G30y.00	Other acute myocardial infarction
36523	G311.00	Preinfarction syndrome
36609	G342.00	Atherosclerotic cardiovascular disease
38609	G351.00	Subsequent myocardial infarction of inferior wall
39546	Gyu3000	[X]Other forms of angina pectoris
40429	G301000	Acute anteroapical infarction
41221	G30y200	Acute septal infarction
41835	G384.00	Postoperative subendocardial myocardial infarction
45809	G350.00	Subsequent myocardial infarction of anterior wall
46017	G30yz00	Other acute myocardial infarction NOS
46112	G380.00	Postoperative transmural myocardial infarction anterior wall
46166	G35X.00	Subsequent myocardial infarction of unspecified site
46276	G381.00	Postoperative transmural myocardial infarction inferior wall
52517	Gyu3.00	[X]Ischaemic heart diseases
62626	G30y100	Acute papillary muscle infarction
63467	G306.00	True posterior myocardial infarction
68748	G38z.00	Postoperative myocardial infarction, unspecified
72562	G353.00	Subsequent myocardial infarction of other sites

9.7.2 Codes for MI, co-morbidities and smoking cessation

A number of other searches were done in CPRD on co-morbidities of interest, to determine if and when the CHD population was coded for that condition. The details of these codes are given in Table 9-3.

Table 9-3 CPRD Read and Medcodes for MI, co-morbidities and smoking cessation

Condition	Read codes v2	Medcodes
Diabetes	C10E.% C10F.% (excluding C10F8)	758,1549,1407,18390,10692,26054,22884,12640, 18496,25627,47954,51261,32627,30323,10418, 18387,30294,34450,35288,12455,40837,25591, 18777,53392,18425,34268,35385,44982,55239, 49074,64668,39070,47315,60796,47321,63690, 62674,22871,46917,18683,54008,69676,95636, 12736,43921,47582,49655,40682,47650,37806, 5267,51756,42831,59253,62209,46301,47649, 49554,45276,18642,43227,69993,96235,97849, 49949,54600,57278,68105,85991,91646,93875, 93878,95343,95351,102112,50527,62613,72702, 91942,93468,95539,99719,100964,66145,91943, 93727,97894,98071,98616,98704,98723,99311, 99716,100770,101311,101735,102163,102201, 102620
Atrial Fibrillation	3272. 8H41. 8H44. 8HR1. 8H4R.	1664, 2212, 1268, 23437, 96076, 96277, 35127

	8HVJ.	
Hypertension	G2... G20..% G24.. - G2z.. (Excluding G24z1)	799,204,351,10818,3712,7057,1894,8732, 4372,16292,15377,7329,4668,8857,18765, 29310,16059,16173,17434,83473,15106, 34744,31387,42229,39649,31464,32423, 25371,57288,31755,28684,61166,51635, 43935,62718,21837,63466,52427,50157, 52127,59383,61660,57987,68659,73293, 63000,67232,95334,72668
Stroke	G61..% (excluding G617.) G63y0 - G63y1 G64..% G66..% G6760 G6W.. G6X.. G65..- G654. G656.- G65zz F4236	1469,504,1433,1298,5363,3149,8837,6116,1195, 5051,6960,2417,5268,15788,3132,5602,569,7780, 12833,16517,6253,6155,10794,3535,18604,15019 ,9985,10504,17322,13564,23942,33543,8443, 23671,25615,36717,26424,23465,15252,18689, 5185,33499,24446,19260,7912,19280,53745, 40758,39344,19354,40338,28314,27975,19201, 51767,31595,46316,34758,31060,44765,91627, 47642,57315,30045,53810,92036,94482,62342, 90572,50594,96630
Heart Failure	G58..% G1yz1 662f. – 662i.	398,884,2062,2906,4024,1223,5942,13189, 18853,5255,19066,32671,10079,9524,17278, 10154,27964,23481,23707,27884,51214,11424, 22262,43618,12590,94870,101138,101137

Rheumatoid Arthritis	N040.00 N04..00 N040S00 N040N00 N040Q00	844,27603,31054,30548,18155
Smoking cessation	6791. 67A3. 67H6. 8B2B. 8CAL. 8HkQ. 8HTK. 13p..% 9OO..% 9N4M. 9N2k. 8H7i. 67H1. 8B3Y. 8B3f. 745H.% 8IAj.	2111,7130,7622,9833,10184,10211,10742,10898, 11356,11527,12953,18573,18926,19485,21637, 28834,28886,32083,32572,34126,34127,34374, 38112,40417,40418,41405,42722,53101,58597, 60720,63901,66387,74907,81440,85247,85975, 89464,90522,91708,94958,98137,98154,100099, 101764
CHD Myocardial Infarction	G30..% (Excluding G30A.)	241,14658,1677,1678,1204,2491,3704, 14897,5387,12139,8935,17872,9507,23892, 14898,1357,29643,17689,46017,18842,

codes	G35..% G38..% Gyu34	13566,32272,34803,30421,41221,29758, 30330,40429,28736,38609,45809,63467, 32854,41835,46276,68748,46166,46112, 72562,62626,96838,10562

9.7.3 Codes for QOF prescription targets

In table 9-4 are the BNF codes used for each of the prescribing targets along with the product codes which qualified under the BNF code.

Table 9-4 CPRD BNF and Product codes: Prescriptions

BNF Header	BNF code	Product codes
Beta-blockers	18	5;24;26;197;220;297;472;581;594;599;707;739;751;753;769;786;817;822;940;1006;1048;1050;1124;1288;1290;1295;1333;1334;1448;1572;1597;1684;1788;2361;2414;2432;2499;2587;2590;2629;2775;2780;3005;3087;3167;3344;3474;3516;3526;3588;3691;3748;3827;4004;4025;4265;4410;4429;4542;4588;4605;4725;4771;4796;4983;5284;5330;5478;5713;5721;5858;5968;6066;6751;7049;7066;7091;7429;7474;7528;7543;7553;7620;7852;7853;7974;8023;8061;8068;8071;8113;8147;8172;8189;8262;8290;8331;8369;8555;8623;8642;8673;8707;8807;8935;8978;8987;9016;9143;9178;9185;9273;9292;9783;10191;10294;10429;10627;10716;10777;10892;11338;11380;11711;11793;12037;12054;12141;12296;12456;12495;12517;12519;12651;13051;13394;13415;13487;13499;13526;13871;14030;14057;14058;14117;14126;14146;14438;14502;14552;14673;14808;15042;15117;15176;

		15488;15619;15730;16645;16776;16786;17082;17149; 17322;17462;17615;17679;17783;18185;18287;18414; 18743;18950;19055;19068;19142;19172;19178;19182; 19191;19200;19202;19437;19853;19858;19998;20012; 20082;20093;20169;20468;20502;20728;21025;21133; 21182;21838;21839;21866;21873;21885;21905;21966; 22208;22793;22912;23131;23134;23326;23587;24083; 24094;24191;24195;24218;24280;24461;24635;24832; 25359;25363;25367;25462;25644;25730;26211;26228; 26229;26248;26255;26529;26741;26895;26922;27357; 27486;27700;27719;27727;27946;27964;28048;28128; 28177;28700;28788;28996;29180;29230;29368;29398; 29427;29610;29762;29763;29827;29998;30400;30519; 30541;30636;30770;31214;31470;31536;31708;31776; 31833;31934;32094;32114;32135;32162;32552;32630; 32787;32836;33079;33085;33092;33184;33374;33376; 33569;33578;33602;33644;33650;33657;33659;33836; 33839;33850;33909;34012;34034;34092;34094;34125; 34171;34177;34185;34188;34208;34214;34265;34365; 34371;34378;34407;34430;34443;34449;34492;34501; 34509;34520;34575;34584;34585;34600;34640;34690; 34695;34740;34741;34754;34783;34804;34821;34825; 34854;34867;34868;34882;34884;34890;34899;34925; 34945;34949;34963;34976;35054;35062;35695;35710; 35778;35938;35940;36261;36576;36603;37118;37725; 37837;38370;38433;38498;38991;39233;39423;39646;

		39819;39846;40167;40240;40241;40761;41555;41572;41591;41740;41827;42152;42795;43251;43525;43549;43564;44000;44083;44808;44858;45250;45289;45297;45309;45343;45494;45765.
Aspirin; Anti-platelets; Anti-coagulants; Dipyridamole; Clopidogrel	15 196 912 918 919	3;16;34;45;61;254;377;393;434;489;572;657;689;714;771;833;836;1049;1137;1781;1814;1902;2105;2106;2607;2628;3832;4446;4679;5305;5882;6006;6007;6262;6666;6696;7516;8185;8186;8466;8467;8645;9144;9301;9939;10031;10305;11977;13348;13501;13502;13503;13504;13505;13644;13882;15006;15376;16597;17130;17449;17704;17920;17965;18030;18217;18329;19189;21019;21380;21382;21921;21989;22138;22232;22618;23078;23488;23593;23878;23932;24025;24960;25232;25284;25335;25718;28810;29759;29848;30202;30203;30554;30920;30975;30976;31192;31210;31211;31511;31858;31870;31937;31938;31953;31954;31956;32036;32210;32992;33293;33320;33656;33662;33668;33676;33711;33877;34019;34086;34087;34088;34095;34299;34309;34385;34386;34416;34417;34418;34434;34485;34517;34526;34576;34611;34666;34691;34709;34758;34762;34796;34797;34864;34918;34942;35108;35809;36099;36521;36543;37541;38041;38044;38349;38998;39119;39444;39503;39639;39738;39755;39866;39932;40114;40143;40144;40381;40591;40913;41229;41512;41569;41594;41766;42474;42750;43060;43407;43408;43409;43434;43530;43655;43679;43709;43806;44639;44866;45576;

		45643.
ACE inhibitors; Angiotensin II antagonist	5 142	65;69;78;80;82;97;147;196;277;448;520;529;531;575;593;624;633;654;709;756;761;764;828;1021;1121;1143;1144;1293;1299;1520;1780;1807;1904;2971;2982;3069;3203;3222;3310;3720;3839;3929;4103;4155;4226;4540;4571;4645;4685;4741;4818;5013;5047;5117;5159;5189;5275;5612;5723;5735;5800;5861;5988;6078;6200;6217;6243;6261;6285;6288;6314;6351;6359;6362;6364;6408;6437;6468;6518;6765;6786;6794;6806;6807;6877;6939;7043;7314;7338;7419;8025;8026;8105;8106;8268;8800;8830;9196;9646;9693;9731;9745;9764;9915;9948;10316;10323;10882;10902;11133;11197;11251;11252;11348;11351;11448;11469;11526;11561;11567;11641;11864;11937;11965;11983;11987;12313;12411;12412;12574;12815;12836;12858;12874;13026;13123;13589;13755;13821;14228;14283;14387;14477;14478;14738;14870;14943;14960;14965;14983;15031;15085;15096;15108;15121;15135;15605;15958;16060;16161;16196;16197;16212;16285;16371;16701;16708;16710;16924;17006;17120;17474;17545;17624;17633;17655;17686;17689;18200;18202;18219;18223;18263;18269;18325;18903;18910;19198;19204;19208;19223;19690;20117;20188;20579;20849;20975;21053;21162;21231;21423;21943;22439;22708;23252;23456;23478;23642;24041;24268;24359;24482;24484;24632;25382;25998;26995;27520;27871;28127;28438;28486;28586;28724;28725;28820;28902;

		29130;29530;29627;29634;30039;30921;31072;31307; 31587;31716;31810;32048;32166;32241;32514;32560; 32597;32857;32934;33057;33078;33095;33336;33353; 33646;33811;33894;33977;34357;34382;34390;34400; 34412;34429;34431;34432;34453;34471;34490;34505; 34528;34539;34540;34544;34562;34567;34583;34589; 34651;34652;34657;34696;34698;34710;34712;34719; 34732;34768;34798;34799;34877;34893;34936;34937; 34943;34952;34953;35007;35096;35173;35174;35189; 35196;35302;35304;35317;35329;35343;35380;35481; 35697;35731;35794;36742;36753;36939;37080;37087; 37573;37650;37655;37710;37747;37778;37908;37930; 37964;37965;37971;37978;38026;38034;38285;38308; 38367;38395;38459;38510;38854;38889;38899;38995; 39021;39137;39147;39199;39227;39242;39355;39421; 39512;39786;39944;39984;40316;40355;40384;40571; 40639;40668;40711;41203;41205;41232;41417;41522; 41532;41538;41573;41617;41633;41694;41743;41746; 42081;42285;42723;42894;42901;42902;42908;43012; 43322;43411;43412;43413;43416;43418;43432;43507; 43563;43566;43649;43813;43915;44527;44657;44778; 45217;45228;45264;45300;45319;45324;45337;45340; 45554;45600.

9.7.4 ICD 10 codes used in HES

9.7.4.1 Codes for outcome variable name miadmis

I21.0 Acute transmural myocardial infarction of anterior wall

I21.1 Acute transmural myocardial infarction of inferior wall

I21.2 Acute transmural myocardial infarction of other sites

I21.3 Acute transmural myocardial infarction of unspecified site

I21.4 Acute subendocardial myocardial infarction

I21.9 Acute myocardial infarction, unspecified

I22.0 Subsequent myocardial infarction of anterior wall

I22.1 Subsequent myocardial infarction of inferior wall

I22.8 Subsequent myocardial infarction of other sites

I22.9 Subsequent myocardial infarction of unspecified site

I23.0 Haemopericardium as current complication following acute myocardial infarction

I23.1 Atrial septal defect as current complication following acute myocardial infarction

I23.2 Ventricular septal defect as current complication following acute myocardial infarction

I23.3 Rupture of cardiac wall without haemopericardium as current complication following acute myocardial infarction

I23.4 Rupture of chordae tendineae as current complication following acute myocardial infarction

I23.5 Rupture of papillary muscle as current complication following acute myocardial infarction

I23.6 Thrombosis of atrium, auricular appendage, and ventricle as current complications following acute myocardial infarction

I23.8 Other current complications following acute myocardial infarction

9.7.4.2 Codes for outcome variable emang

I20.0 Unstable angina

I20.1 Angina pectoris with documented spasm

I20.8 Other forms of angina pectoris

I20.9 Angina pectoris, unspecified

I24.0 Coronary thrombosis not resulting in myocardial infarction

I24.1 Dressler's syndrome

I24.8 Other forms of acute ischaemic heart disease

I24.9 Acute ischaemic heart disease, unspecified

I25.0 Atherosclerotic cardiovascular disease, so described

I25.1 Atherosclerotic heart disease

I25.6 Silent myocardial ischaemia

I25.8 Other forms of chronic ischaemic heart disease

Variable name, *alladmis*, combines these two variables, forms the main outcome measure and is a count of all poor outcomes.

9.8 Appendix 8 Study variables ruleset

Details on how QOF target and outcomes variables were created are provided in Chapter 6. Details on how the databases were used to create the remaining variables are detailed in the Table 9-5. These follow the same protocols set out in Chapter 6 namely:

- The absence of a record of an event is interpreted as meaning the event did not take place
- The targets applied to the whole study population except where specified
- Outcomes measured in HES had to take place prior to the patient's death or transfer out date and when their practice was up to standard.

Table 9-5 Remaining study variables ruleset

Variable	CPRD/HES Ruleset
age	Created using Birth year, yob, variable in CPRD. This figure was subtracted from 1999 to create an age prior to the study start. To this figure the study year (1 to 11) was added for each year the patient was present on the database to give their age at that point
gender	This is recorded in the patient file and mapped to the SEX lookup file
Years	With the exception of Year 1 which spanned 15 months, 01/01/2000 to 31/03/2001, it refers to a 12 month 01 April to 31 March the following year period. The year figure relates to the date the event measured took place or the period the patient had CHD.
Denominator	All patients in the study population
Exclusions	An event recorded prior to the patient having a Read code for CHD The patients' practice was not up to standard It did not occur in the patients current registration period The event occurred after the patients recorded transfer out or death date

Missing data	Where the date of the event was missing the system date was used in its place providing it was in the patient's current registration and the practice up to standard period.
ACE	At least one prescription for an ACE inhibitor or Angiotension II antagonist (bnfcodes 5 or 142) recorded in the therapy file in the last six months of a QOF year.
HF, AF, RA, DM, Hyper, CKD, STIA	These are separate variables which show whether or not the CHD population had Heart Failure (HF), Atrial Fibrillation (AF), Rheumatoid Arthritis (RA), Diabetes (DM), Hypertension (Hyper), Chronic Kidney Disease (CKD), and Stroke and Transient Ischaemic Attack (STIA). They were generated by a search on the define tool for disease specific medcodes which related to QOF Read codes for the conditions. In the case of RA, which is not in the QOF, the Read codes were determined by searching for Rheumatoid in the Read term in the Medical browser
Exclusions	None matched study years, namely, periods prior to first CHD coding
newMI	This relates to a newly diagnosed MI. In the pre-QOF period this relates to a diagnosis after 01/04/1999 in the post QOF period to a diagnosis after 01/04/2003
Denominator	All patients in the study population
Exclusions	Any subsequent MI code after the first code First code prior to 01/04/1999 in the pre QOF period First code prior to 01/04/2003 in the QOF period
Missing data	Assumed no diagnosis had been made
Region	This shows which of the ten strategic health authorities (SHA) in England, created by the reorganisation of SHA's in July 2006, the patient's practice resides in. It is recorded in the practice file and

	mapped to the PRG lookup file
pracatt	This is a mean percentage figure for the practice's QOF performance. It has been created by adding the percentage of their patients who met the seven QOF targets that are examined, and dividing it by the total number of targets, 7. This variable was also separated into quintiles based on all practice years figures, variable name *pracattq*, to see if effects were more evident between categories
Denominator	All patients present on the database that belong to that practice
size	This is the total number of patients within the practice for each study year. It has also been broken down into quintiles to create the variable, sizeq.
Denominator	All patients registered to that practice in CPRD based on Feb 2012 upload
workld	This is a measure of GP workload in a practice. It was calculated by dividing the number of patients present in the practice for each year (size variable) by the number of staff given a GP grade. These are staff given the following codes: 1, Senior Partner; 2, Partner; 8, GP registrar; 10, Sole Practitioner; and 47, Salaried Partner, in the role field of the staff file. Staffs with these codes were then matched across the clinical, test, therapy and referral files to determine the number of GP's active over each of the study years. The workload variable generated was also split into quintiles based on aggregate workload data for all practice years to produce the variable *workldq*
Denominator	All staff present in the database for the specific year given the GP codes identified above
IMD	This is a measure of deprivation using the IMD 2007 algorithm using the most up to date data on all my study population at the point of request, March 2012

Missing data	Given a missing value code
Smoker	A current smoker QOF Read code recorded in the clinical file
MIpre	This records whether or not an MI primary diagnosis by hospitalisation, occurred on or before the same day as the patient's first CHD Read code. These were created following the same rules as those for the *miadmis* outcome variable. As these occurred prior to or on the first QOF Read code date, before the patient's study period, they were not recorded in the *miadmis* variable.
Angpre	This records whether or not a non MI primary diagnosis for an emergency hospitalisation, occurred on or before the same day as the patient's first CHD Read code. These were created following the same rules as those for the *emang* outcome variable. As these occurred prior to or on the first QOF Read code date, before the patient's study period, they were not recorded in the *emang* variable.
Hospre	This is formed from the variables *MIpre* and *Angpre* and shows whether or not the patient had one of these events prior to their entry into the study

Table 9-6 details how the study variables are recorded in the study dataset.

Table 9-6 Variable labels and formats

Name	**Label**	**Format**
DEPENDENT VARIABLE		
alladmis	All hospital admissions	Count (0-13)
This consists of the sum of the following two measures		
miadmis	MI admission post QOF CHD coding	Count (0-8)
emang	Emergency angina admission post QOF CHD coding	Count (0-13)

EXPLANATORY VARIABLES		
gender	Gender	Binary (Female/Male)
age	Age in Years	Continuous
BP	Met QOF BP target	Binary (No/Yes)
Chol	Met QOF Cholesterol target	Binary (No/Yes)
BB	Met QOF BB prescription target	Binary (No/Yes)
Anti	Met QOF Aspirin, anticoagulant or anti-platelet prescription target	Binary (No/Yes)
Acea2q	Met QOF ACEI/A2 prescription target	Binary (No/Yes)
Smcess	Met QOF smoking cessation target	Binary (No/Yes)
Smoker	Current Smoker	Binary (No/Yes)
Smokstat	Smoking status	Nominal categorical (1 Never smoked, 2 Ex, 3 Current)
Years	Year in study	Ordered categorical (1, 2000/01 to 11, 2010/11)
ACE	Received ACEI or A2 in the previous 6 months	Binary (No/Yes)
MI	QOF Read code for MI	Binary (No/Yes)
HF	QOF Read code for HF	Binary (No/Yes)
AF	QOF Read code for AF	Binary (No/Yes)
Diab	QOF Read code for Diabetes	Binary (No/Yes)

Hyper	QOF Read code for Hypertension	Binary (No/Yes)
CKD	QOF Read code for CKD	Binary (No/Yes)
RA	Read code for Rheumatoid Arthritis	Binary (No/Yes)
size	Practice size (Patient numbers)	Continuous
sizeq	Practice size in quintiles	Ordered categorical (1-5)
workld	Practice GP workload	Continuous
workldq	Practice GP workload in quintiles	Ordered categorical (1-5)
pracatt	Practice QOF attainment	Continuous (%)
pracattq	Practice QOF attainment in quintiles	Ordered categorical (1-5)
IMD	Index of Multiple Deprivation 2007	Ordered categorical (0-4)
Region	SHA Region	Nominal categorical (1-10)
MIpre	Had an MI hospital admission prior to CHD coding	Binary (No/Yes)
Angpre	Had an emergency angina hospital admission prior to CHD coding	Binary (No/Yes)
Hospre	Had one of the outcomes of interest (MIpre or Angpre) prior to CHD QOF coding	Binary (No/Yes)

All of the explanatory variables are expected to explain some of the variance in the outcome variable, hence their inclusion in the study dataset. The expected direction of that relationship and the reason behind it is explained in Table 9-7. The +/- column shows prior expectations of the variables regression coefficient, namely its expected relationship with the dependent variable. A positive means that the dependent variable should increase as the variable increases or moves from zero to one. A negative signifies the reverse. In the case of Region both are shown as while

it is expected that certain regions will be associated with higher counts it is impossible to be certain for all regions, and will depend on which region is used as the baseline.

Table 9-7 Explanatory variables-expected relationship with dependent variables with rationale

Name	**+/-**	**Explanation**
gender	+	Prior to menopause CHD is two to five times less common in females than males[160]. Although the risk of CHD increases markedly in post menopausal women, the risk of CHD increases with age in males also so there remains a noticeable difference in risk of CHD between the genders[160 161]. Once diagnosed with CHD men have a high risk of mortality from CHD relative to females with CHD[162]
age	+	As the body ages, fatty deposits accumulate on the lining of the coronary arteries increasing the risk of CHD and its progression[160]. Therefore the risk of adverse CHD events should increase with age.
BP	-	As all the targets have been selected on the basis of high level clinical evidence, it is expected that patients who meet these targets will have fewer of the study outcomes.
Chol	-	
BB	-	
Anti	-	
Acea2q	-	
Smcess	-	
Smoker	+	Smoking is a risk factor for cardiovascular disease, but one that mitigates when someone ceases to smoke, even in those who have done so for long periods[88]. Hence smokers are expected to be at higher risk of adverse outcomes relative to ex smokers
Smokstat	+	

		whom have a higher risk relative to non smokers.
Years	-	It is expected, particularly following the QOF, that as better care has become incentivised and embedded, that the number of outcomes has fallen.
ACE	-	ACE inhibitors/A2 antagonists have wider benefits in preventing disease progression in CHD patients, outside of those who have had an MI: In particular in those with elevated blood pressure. Therefore there is expected to be a reduction in adverse outcomes in patients receiving these products[73 87 163].
HF	+	The presence of these co-morbidities means the patient is likely to be in poorer health increasing odds of disease specific complications, CHD among them. Furthermore they are all cardiovascular diseases, or in the case of rheumatoid arthritis and diabetes risk factors for cardiovascular disease[154]; and hence are themselves stand alone risk factors for the study outcomes.
AF	+	
Diab	+	
Hyper	+	
CKD	+	
Arth	+	
size	-	The balance of evidence tends to show that larger practices have the ability to provide a greater package of care and therefore have better patient outcomes[164]
sizeq	-	
MIpre	+	Having an MI prior to the study is expected to increase the odds of poor outcomes because the patient will have suffered irreparable damage to heart muscle tissue, thereby increasing the likelihood of subsequent coronary events[89].
Angpre	+	Although this is a less serious event than an MI, it is still a reflection of disease progression and will therefore increase the chances of another event or progression onto MI[89].
Hospre	+	See Angpre and MIpre above for reasons

workld	+	While not totally reflective of care as nurses are commonly used in the delivery of the QOF, it expected that there will be some trade off between the quality of a consultation and the workload of a GP. In the sense that a GP with a higher workload will have less time to devote to a consultation, and getting to know their patient needs, and is therefore more likely to offer worse care and have poorer outcomes, than a GP who has a lower workload.
workldq	+	
pracatt	-	A higher overall score on the QOF targets is taken to be an indication that the practice provides good care and therefore it is expected that patients in these practices will have better outcomes as a result
pracattq	-	
IMD	+	Patients in more deprived quintiles will have poorer outcomes as unemployment, morbidity, poor housing, crime, income and the living environment all impact on an individual's quality of life, and their health[165 166]. There is evidence that the QOF has reduced inequalities in outcomes but that having a lower socio-economic status is still associated with poorer health outcomes[13 157 167]
Region	+/-	This variable uses the 10 English Strategic Health Authorities (SHA's) created in July 2006, which replaced the previous 28. As these cover large areas with a heterogeneous population it is difficult to be specific on expected effects. Nonetheless it is expected to show differences across the regions, with areas like the North West and North East having comparatively poor outcomes compared to the generally more affluent South East and South West.

9.9 Appendix 9 Regression model specifications

9.9.1 Negative binomial

The negative binomial is used to model count data, and is employed in preference to the Poisson when the equal dispersion restriction imposed by that model needs to be relaxed. It is specified below[168]

$$E[y_i\,|x_i, \varepsilon_i] = \exp(\alpha + x_i'\beta + \varepsilon_i) = h_i\lambda_i$$

Where $h_i = \exp(\varepsilon_i)$ is assumed to be a one parameter gamma distribution, $\Gamma(\theta,\theta)$ with mean 1 and variance $1/\theta = \kappa$

It has a conditional mean $E[y_i\,|x_i] = \lambda_i$ and

Variance, $Var[y_i\,|x_i] = \lambda_i[1 + \kappa\lambda_i]$

Where $\kappa = \mathrm{Var}[h_i]$, is referred to as the index parameter, which is the inverse of the dispersion parameter, θ. The inclusion of this enables the negative binomial to fit count data where the mean ≠ variance. If $\kappa = 0$ the negative binomial is the equivalent of the Poisson distribution.

9.9.2 Logit

The logit is the inverse of the logistic function and is used where the observed outcome is binary. The logit of an outcome, p, which takes a value between 0 and 1 is given by the formula:

$$\operatorname{logit}(p) = \log\left(\frac{p}{1-p}\right) = \log(p) - \log(1-p) = -\log\left(\frac{1}{p} - 1\right).$$

The linear predictor for the logit as a function of several regressors is given is represented as[169]:

$$\pi_i = \frac{1}{1+e^{-x_i'\beta}}$$

Where β represents the coefficient parameters for the included explanatory variables $(\beta_1 \ldots, \beta_k)$ and x the included explanatory variables $(x_1 \ldots, x_k)$

9.9.3 Hurdle model specification

The hurdle model removes excess zeros from a count outcome variable by breaking it down into two separate components which are then modelled separately. An initial binary model includes the whole study population. The coefficients in this model are interpreted as the effect of that variable on the individual having an event. Those who do have an outcome cross the hurdle and move into a truncated at zero count model; which shows the relationship between the explanatory variable and the number of outcomes an individual has.

This approach removes the excess zeros problem from the count model by removing them entirely in the case of cross sectional data. However this becomes somewhat more complicated in a count panel data model. In such circumstances the hurdle could reset and apply separately to each period in the panel. This is a logical approach if it can be assumed that all patients are the same at the beginning of each period regardless of whether or not they crossed the hurdle (had an event) in a previous period. There are examples where this may be the case but for this study the outcomes were serious CHD complications, and hence this was unlikely to be the case. Consequently it was decided that those who breached the hurdle should be treated as a distinct group, with different risks of having an event to those who did not, for the remainder of their time in the study period. In this study the hurdle model was implemented as follows:

The whole study population participated in the binary component of the hurdle model until the point of their first outcome event. This was modelled using the logit described above. If the patient had an outcome they crossed over into the count component of the hurdle model, where they remained for the remainder of their time in the study. As it was possible for those patients to have no events in subsequent time periods in the panel, namely a zero count, that count model was not truncated at zero. The negative binomial was used to model these patients, which is described above.

The specific setup of the hurdle in this thesis is best described in the fictional example presented in Table 9-8 for patient 2260. *Patid* represents the unique patient identifier, *alladmis*, the study outcome count measure, which for purposes of the hurdle is split into a binary model (*risk*), and a count model (*event*). Patient 2260 remained in the binary component until 2006/07 when they had two admissions. Prior to that period they did not participate in the count component, where they were represented by the missing value (.). When the patient first had an outcome this was recorded in the binary model as a 1 value, and the number of events was counted in the count component. This fits the standard application of the hurdle model in single observation data, as at this point the count model is truncated at zero. Had the hurdle reset for the following period, 2007/08 the count component would have remained truncated at zero. However in this study patients remained in the count model (*event*) following an outcome, and were missing from the binary model: Shown by counts in the event column and missing values, (.), in the risk column. This approach fitted the circumstances of the study population and also meant patients had continuous observations for their periods in either model, rather than potentially intermittent periods in both. This was important because analysis of the effect of the QOF variables on the outcome measures included lags of those explanatory variables

Table 9-8 Application of the hurdle model: An example

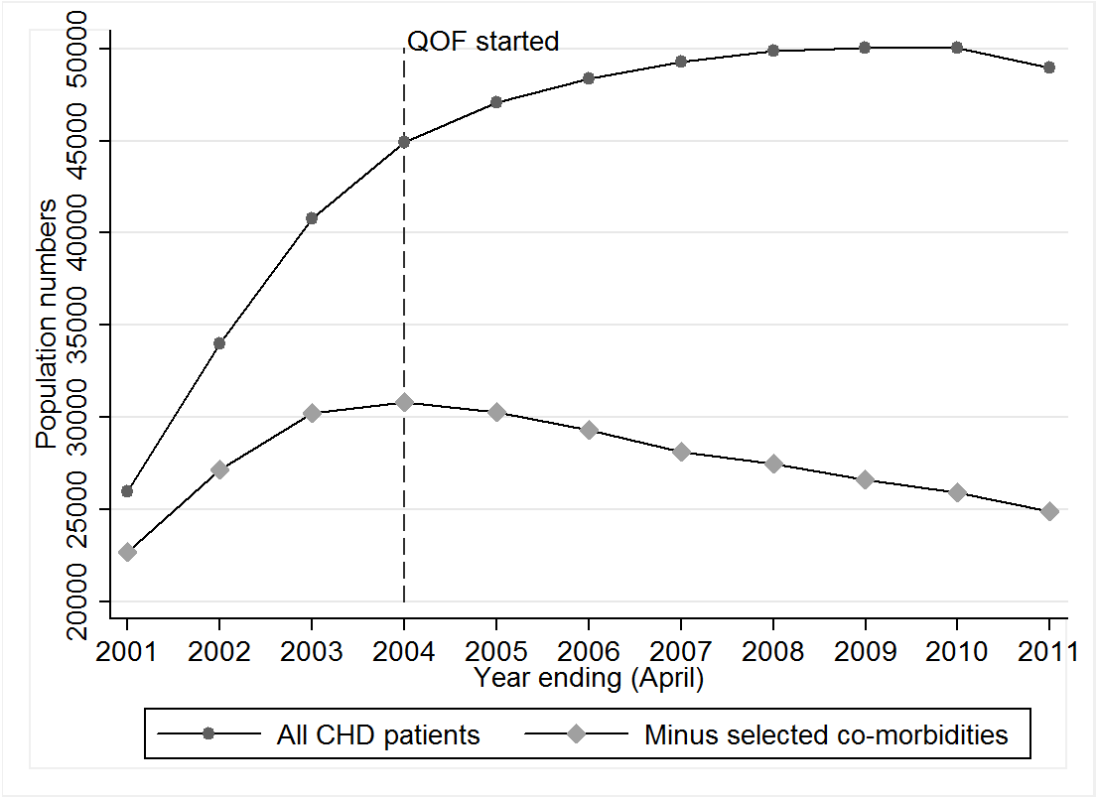
patid	*years*	*alladmis*	*risk*	*event*
2260	2004/05	0	0	.
2260	2005/06	0	0	.
2260	2006/07	2	1	2
2260	2007/08	0	.	0
2260	2008/09	2	.	2
2260	2009/10	0	.	0
2260	2010/11	1	.	1

9.10 Appendix 9 The effect of co-morbidities on key study variables

9.10.1 Introduction and specification

To explore the effects of co-morbidities in relation to the study population, a sub group of the study population, those without selected co-morbidities of interest, are presented. In this subgroup study population, patients who had a QOF coding for any of the following: Diabetes, Hypertension, Atrial Fibrillation, Heart Failure and Stroke were excluded, from the point where their first co-morbidity was recorded onwards. Chronic Kidney Disease (CKD) was not excluded due to uncertainty over the conditions prevalence prior to entry into the QOF. Rheumatoid Arthritis as the only condition not in the QOF, and not derived using QOF Read codes, was also not excluded. This study population has been termed a CHD only population as important co-morbidities have been removed though it is possible that the patient had other serious clinical conditions not considered as well as CKD and RA.

Figure 9-1 Total and CHD only study population, 2000-2011



The impact of removing co-morbidities from the study population is shown in Figure 9-1, as compared against the total study population. Population numbers show the number of patients who were present for any period of that year. Year's run in 01

April to 31st March the following year cycles, with the exception of the first year which ran from 1st January 2000 to 31st March 2001, and the figures given are for the year end totals.

Removing patients with chosen co-morbidities (CHD only) reduces the total patient population considerably. Both populations increase initially as more patients join the study. However the differential between the two increases over time as more of those patients acquire the selected co-morbidities. The CHD only portion of the total population falls from the beginning of the QOF onwards, which may be the result of better recording of co-morbidities, and greater efforts to diagnose, in response to incentives. This proposition concurs with the continued increase in the total population until 2007 which is consistent with the continued increase in the diagnosis and recording of CHD. A static total population from 2008, and its subsequent fall, could be the result of better primary prevention resulting in fewer cases of CHD. This could explain the relationship seen as improved primary prevention could lead to the number of new cases falling below the number exiting the study to die or transfer out.

9.10.2 Comparisons between the study and CHD only population on key explanatory variables

Comparisons are made between the study population and the CHD only component of that population on a number of key explanatory variables in a series of tables below.

Starting with differences in age; table 9-9, looks at the mean age of the study population specifications and their standard deviations over the course of the study. As the study progresses, patients acquired co-morbidities and the two populations separated. Unsurprisingly given that health deteriorates and chronic conditions are acquired as people grow older the whole population, which includes individuals with co-morbidities, is older than the part of it which has none of the selected co-morbidities, CHD only.

Table 9-9 Age by population specification, 2000-2011

	All Patients		CHD only	
Study year	Mean Age (Years)	SD (Years)	Mean Age (Years)	SD (Years)
2000/01	69.30	11.09	69.06	11.08
2001/02	69.60	11.23	69.19	11.24
2002/03	69.93	11.29	69.40	11.33
2003/04	70.25	11.34	69.59	11.43
2004/05	70.53	11.38	69.74	11.48
2005/06	70.83	11.37	69.92	11.52
2006/07	71.09	11.44	70.08	11.65
2007/08	71.37	11.43	70.27	11.67
2008/09	71.59	11.48	70.36	11.75
2009/10	71.84	11.49	70.53	11.80
2010/11	72.06	11.55	70.65	11.87

Table 9-10 looks at the gender split in the two populations, in terms of their percentage numbers who are male. The percentage share of the population that are male, increases year on year, in both specifications. Throughout the study period males have a greater representation than female in both population specifications as expected, however that percentage is lower in the CHD only part of the total population. A possible explanation is that individuals are more likely to develop co-morbidities as they age and health deteriorates. Females live longer than males and there is also evidence to show they live longer in poorer health states which would account for the figures seen[170 171].

Table 9-10 Percentage of population male, by population specification, 2000-2011

Male	All Patients		CHD only	
Year	N	%	N	%
2000/01	15,199	58.5%	13,291	58.7%
2001/02	19,921	58.6%	15,993	58.9%
2002/03	23,910	58.6%	17,863	59.1%
2003/04	26,463	58.9%	18,306	59.4%
2004/05	27,897	59.3%	18,126	59.9%
2005/06	28,780	59.5%	17,666	60.2%
2006/07	29,525	59.9%	17,036	60.6%
2007/08	30,051	60.2%	16,754	61.0%
2008/09	30,374	60.7%	16,420	61.7%
2009/10	30,570	61.1%	16,036	61.9%
2010/11	30,087	61.4%	15,492	62.3%

Tables 9-11 and 9-12 show mean diastolic and systolic blood pressure respectively over a QOF accounting period, for the total population and its CHD only component.

Table 9-11 Mean diastolic blood pressure by population specification, 2000-2011

BP	All Patients		CHD Only	
Diastolic	Mean (mmHg)	SD (mmHg)	Mean (mmHg)	SD (mmHg)
2000/01	80.39	8.56	80.25	8.45
2001/02	78.79	9.18	78.63	9.05
2002/03	78.05	9.22	77.91	9.14
2003/04	77.41	9.13	77.25	8.99
2004/05	76.63	8.87	76.48	8.72
2005/06	76.05	8.80	75.98	8.67
2006/07	75.42	8.83	75.42	8.68
2007/08	75.12	8.85	75.12	8.72
2008/09	74.84	8.76	74.88	8.58
2009/10	74.65	8.80	74.63	8.64
2010/11	74.51	8.56	74.55	8.42

Table 9-12 Mean systolic blood pressure by study population specification, 2000-2011

BP	All Patients		CHD Only	
Systolic	Mean (mmHg)	SD (mmHg)	Mean (mmHg)	SD (mmHg)
2000/01	143.19	18.19	142.71	18.00
2001/02	141.03	18.76	140.40	18.48
2002/03	140.09	18.41	139.53	18.25
2003/04	138.96	17.96	138.13	17.71
2004/05	137.36	16.94	136.41	16.66
2005/06	136.21	16.35	135.38	16.13
2006/07	135.12	16.01	134.24	15.85
2007/08	134.61	15.87	133.74	15.75
2008/09	134.11	15.65	133.29	15.54
2009/10	133.83	15.63	132.91	15.42
2010/11	133.67	15.14	132.79	14.99

Mean diastolic and systolic blood pressure are very similar between the two population specifications, and fall in most years. The CHD only component of the total population has a slightly higher mean and standard deviation throughout for systolic blood pressure, and in most years for diastolic blood pressure.

Table 9-13 Mean cholesterol by study population specification, 2000-2011

	All Patients Cholesterol level		CHD only Cholesterol level	
Year	Mean (mmol/l)	SD	Mean (mmol/l)	SD
2000/01	5.40	1.06	5.40	1.06
2001/02	5.04	1.05	5.05	1.05
2002/03	4.86	1.04	4.88	1.04
2003/04	4.69	1.01	4.71	1.00
2004/05	4.57	0.99	4.61	0.99
2005/06	4.42	0.96	4.46	0.95
2006/07	4.32	0.95	4.38	0.94
2007/08	4.28	0.95	4.33	0.94
2008/09	4.26	0.97	4.32	0.96
2009/10	4.26	0.98	4.33	0.97
2010/11	4.25	0.99	4.32	0.99

Figures for total cholesterol are presented in Table 9-13 and represent the mean of all reading taken over a 15 month period, for all patients in the study population in the

respective years, as well as their standard deviations. Again the results are similar for the two population specifications. However there appears to have been bigger improvements in the co-morbid population following the QOF, leading to that group having noticeably lower level in the latter years of the study. For both groups mean cholesterol levels have been falling, throughout the majority of the study period, though we seem to have reached a plateau for both.

9.11 Appendix 10 Training, Presentation and Publications

9.11.1 Graduate School Courses

26/10/2010, Essential skills for new researchers in the social sciences, 1 credit

12/11/2010, Using posters to communicate research, 1 credit

15/11/2010, Creating and managing long documents in Microsoft word, 2 credit

18/11/2010, Referencing and citing using Endnote, 1 credit

18/11/2010, Nature of the doctorate and supervision process, 1 credit

25/11/2010, Planning your research, 1 credit

01/12/2010, Systematic review, 1 credit

14/12/2010, An introduction to Health Economics, 2 credits

19/04/2011, Presentation skills for researchers (Arts and Social Sciences), 2 credits

08/01/2013, Finishing your thesis, 1 credit

9.11.2 University of Nottingham Postgraduate MPH Courses

All undertaken in the Academic Year 2010/2011

CHS-A34591 Research Methods in Epidemiology with Basic Statistics, 15 credits

CHS-A34582 Advanced Statistical Methods, 10 credits

CHS-A34583 Advanced Epidemiology, 10 credits

CHS-A34585 Practical Uses of Routine Data for Epidemiological Research, 10 credits

CHS-A34589 Health Economics, 10 credits

9.11.3 Other training courses

14th & 15th June 2011, Centre for Health Economics, University of York, Analysing Patient Data using Hospital Episode Statistics

14th-16th December 2011, Applied Methods of Cost-effectiveness Analysis, University of Oxford

Jan/Feb 2012, Stata NetCourse 101: Introduction to Stata

March/April 2012, Stata NetCourse 151: Introduction to Stata programming,

2nd & 3rd April 2012, Writing a Dissertation, Englishwise, Chesterfield

3rd-6th July 2012, Nottingham Systematic Review Course, Cochrane Collaboration, University of Nottingham

24th-26th September 2012, Microeconometrics using Stata, Cambridge University, Timberlake

10th-12th October 2012, Introductory Econometrics, Cemmap, University College London

1st & 2nd November 2012, Panel / longitudinal data analysis, Cemmap, University College London

3rd-7th December 2012, Economic Evaluation in Health Care, University of Birmingham

16th-18th January 2013, Discrete Choice Modelling, Cemmap, University College London

24th & 25th January 2013, Microsimulation, Cemmap, University College London

30th January – 1st February 2013, Statistics for Health Economics, University of Birmingham

4th -8th March 2013, Modelling for Health Economics, University of Birmingham

14th & 15th March 2013, Utility Data for Health Technology Assessment, University of Sheffield

25th & 26th March 2013, Survival Analysis, Cemmap, University College London

8th- 10th April 2013, Applied Econometrics with Stata, CASS Business School, Timberlake

22nd April 2013, Stata Graphics, Cemmap, University College London

23rd- 26th April 2013, Policy Evaluation Methods, Pepa, University College London

9th & 10th May 2013, Data Management, Regression, Panel Data Analysis & Research Output using Stata, CASS Business School, Timberlake

20th & 21st June 2013, Panel Data Analysis using Stata, CASS Business School, Timberlake

10th – 12th July 2013, Systematic Reviews in Health Care: Meta-Analysis in Context, University of Bristol

13th & 14th July 2013, Advanced Meta Analysis, University of Bristol

9.11.4 Presentations

14th September 2011, National School of Primary Care Research (NSPCR) Trainee Conference, *Evaluation of the evidence for QOF clinical targets*, Oxford University, Poster presentation

20th September 2012, NSPCR Trainee Conference, *The effect of primary care interventions on secondary care outcomes: The Quality Outcomes Framework and Secondary prevention of CHD*, Oxford University, Poster presentation

19 October 2012, NSPCR Research Showcase, *An analysis of the effect of CHD QOF targets on secondary care outcomes using large patient databases*, Royal College of Physicians, London, Oral Presentation

5 July 2013, Society of Academic Primary Care conference, *An econometric analysis of the effect of CHD QOF targets on adverse cardiac complications*, University of Nottingham, Oral Presentation (Nominated for travel prize)

References

1. Millett C, Majeed A, Huckvale C, et al. Going local: devolving national pay for performance programmes. Bmj 2011;**342**:475-77.
2. Department of Health. Delivering Investment in General Practice: Implementing the new GMS contract, 2003.
3. Martin S, Smith PC, Dusheiko M, et al. Do quality improvements in primary care reduce secondary care costs?: The Health Foundation, 2010.
4. British Medical Association. Focus on how your practice is funded. Secondary Focus on how your practice is funded 2009. bma.org.uk/-/media/Files/PDFs/.../Contracts/gpfocuspracticefund.pdf.
5. Robinson S. QOF set for 'biggest ever change' in DH overhaul. Secondary QOF set for 'biggest ever change' in DH overhaul 2012. http://www.gponline.com/News/article/1156218/QOF-set-biggest-ever-change-DH-overhaul/.
6. Simon C, Morton A. Getting the most out of the QOF. In: Gillam S, Niroshan Siriwardena A, eds. The Quality and Outcomes Framework: QOF - Transforming General Practice. Abingdon: Radcliffe Publishing, 2011:111-28.
7. Lester H, Campbell S. Developing Quality and Outcomes Framework (QOF) indicators and the concept of 'QOFability'. Qual Prim Care 2010;**18**(2):103-9.
8. Burr T, et al. NHS Pay Modernisation: New Contracts for General Practice Services in England. In: Office NA, ed. London: The Stationery Office, 2008.
9. McDonald R. Paying for performance in primary medical care : learning about and learning from 'success' and 'failure' in England and California. Journal of Health Politics, Policy and Law 2009;**34**(5):747-67.
10. McDonald R, Checkland K, Harrison S. Incentives and control in primary health care : findings from English pay-for-performance case studies. Journal of Health Organization and Management 2008;**22**(1):48-62.
11. Lilley R. *The new GP contract: How to make the most of it*. Oxford: Radcliffe Medical Press, 2003.
12. Campbell SM, Reeves D, Kontopantelis E, et al. Effects of pay for performance on the quality of primary care in England. N Engl J Med 2009;**361**(4):368-78.
13. Purdy S, Griffin T, Salisbury C, et al. Emergency admissions for coronary heart disease: A cross-sectional study of general practice, population and hospital factors in England. Public Health 2011;**125**(1):46-54.
14. Gravelle H, Sutton M, Ma A. DOCTOR BEHAVIOUR UNDER A PAY FOR PERFORMANCE CONTRACT: TREATING, CHEATING AND CASE FINDING? Economic Journal 2010;**120**(542):F129-F56.
15. Sutton M, Elder R, Guthrie B, et al. Record rewards: the effects of targeted quality incentives on the recording of risk factors by primary care providers. Health Econ 2010;**19**(1):1-13.
16. Steel N, Maisey S, Clark A, et al. Quality of clinical primary care and targeted incentive payments: an observational study. Br J Gen Pract 2007;**57**(539):449-54.
17. Fleetcroft R, Steel N, Cookson R, et al. "Mind the gap!" Evaluation of the performance gap attributable to exception reporting and target thresholds in the new GMS contract: National database analysis. BMC Health Serv Res 2008;**8**:8.
18. Ashworth M, Kordowicz M. Quality and Outcomes Framework: smoke and mirrors? Qual Prim Care 2010;**18**(2):127-31.
19. National Health Service. Coronary heart disease Secondary Coronary heart disease 2013. http://www.nhs.uk/Conditions/Coronary-heart-disease/Pages/Introduction.aspx.

20. Townsend N, Wickramasinghe K, Bhatnagar P, et al. Coronary heart disease statistics: A compendium of health statistics, 2012 edition: British Heart Foundation Health Promotion Research Group, 2012.
21. McCartney D, Scarborough P, Webster P, et al. Trends in social inequalities for premature coronary heart disease mortality in Great Britain, 1994–2008: a time trend ecological study. BMJ Open 2012;**2**(3).
22. Smolina K, Wright FL, Rayner M, et al. *Determinants of the decline in mortality from acute myocardial infarction in England between 2002 and 2010: linked national database study*, 2012.
23. CPRD. Release Notes – GOLD On-Line Data October 2012: Clinical Practice Research Datalink, 2012:2.
24. GPRD. The GPRD patient and patient populations October 2011 release: General Practice Research Database, 2011.
25. Wijlaars L. The Health Improvement Network (THIN) research team: THIN Database. Secondary The Health Improvement Network (THIN) research team: THIN Database 2013. http://www.ucl.ac.uk/pcph/research-groups-themes/thin-pub/database.
26. University of Nottingham. What is QResearch. Secondary What is QResearch 2013. http://www.qresearch.org/SitePages/What%20Is%20QResearch.aspx.
27. Health and Social Care Information Centre. What is HES? Secondary What is HES? 2012. http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=456.
28. Hayes J, Prah P, Nazareth I, et al. Prescribing Trends in Bipolar Disorder: Cohort Study in the United Kingdom THIN Primary Care Database 1995–2009. PLoS ONE 2011;**6**(12):e28725.
29. Scowcroft ACE, Lee S, Mant J. Thromboprophylaxis of elderly patients with AF in the UK: an analysis using the General Practice Research Database (GPRD) 2000–2009. Heart 2013;**99**(2):127-32.
30. Scottish Intercollegiate Guidelines Network. SIGN 50 A guideline developer's handbook. In: NHS Quality Improvement Scotland, ed. Edinburgh: Scottish Intercollegiate Guidelines Network,, 2012.
31. Department of Health. Equity and Excellence: Liberating the NHS. In: Department of Health, ed. Norwich: The Stationery Office, 2010:61.
32. J. Sørensen R, Grytten J. Competition and supplier-induced demand in a health care system with fixed fees. Health Econ 1999;**8**(6):497-508.
33. Manning WG, Newhouse JP, Duan N, et al. Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment. The American Economic Review 1987;**77**(3):251-77.
34. Coast J, Smith RD, Millar MR. Superbugs: Should antimicrobial resistance be included as a cost in economic evaluation? Health Econ 1996;**5**(3):217-26.
35. Nyman JA. The economics of moral hazard revisited. J Health Econ 1999;**18**(6):811-24.
36. Giuffrida A, Gosden T, Forland F, et al. Target payments in primary care: effects on professional practice and health care outcomes. In: Reviews CDoS, ed., 1999.
37. The Cochrane Library: The Cochrane Collaboration, 2013.
38. Gosden T, Pedersen L, Torgerson D. How should we pay doctors? A systematic review of salary payments and their effect on doctor behaviour. Q J Med 1999;**92**(1):47-55.
39. Gosden T, Forland F, Kristiansen IS, et al. Impact of payment method on behaviour of primary care physicians: a systematic review. J Health Serv Res Policy 2001;**6**(1):44-55.
40. Gosden T, Forland F, Kristiansen IS, et al. Capitation, salary, fee-for-service and mixed systems of payment: effects on the behaviour of primary care physicians. Cochrane Database of Systematic Reviews 2000(3):CD002215.

41. Peckham S, Wallace A. Pay for performance schemes in primary care: what have we learnt? Quality in Primary Care 2010;**18**(2):111-6.
42. Vahidi RG, Mojahed F, Jafarabadi MA, et al. A systematic review of the effect of payment mechanisms on family physicians service provision and referral rate behavior. Journal of Pakistan Medical Students 2013;**3** (1):54-60.
43. Van Herck P, De Smedt D, Annemans L, et al. Systematic review: Effects, design choices, and context of pay-for-performance in health care. BMC Health Serv Res 2010;**10**:247-59.
44. Petersen LA, Woodard LD, Urech T, et al. Does pay-for-performance improve the quality of health care? Ann Intern Med 2006;**145**(4):265-72.
45. Scott A, Sivey P, Ait Ouakrim D, et al. The effect of financial incentives on the quality of health care provided by primary care physicians. Cochrane Database of Systematic Reviews 2011(9):CD008451.
46. Town R, Kane R, Johnson P, et al. Economic incentives and physicians' delivery of preventive care: a systematic review. Am J Prev Med 2005;**28**(2):234-40.
47. Moher D, Liberati A, Tetzlaff J, et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 2009;**6**(7):e1000097.
48. Davidson SM, Manheim LM, Werner SM, et al. Prepayment with office-based physicians in publicly funded programs: results from the Children's Medicaid Program. Pediatrics 1992;**89**(4 Pt 2):761-67.
49. Gosden T, Sibbald B, Williams J, et al. Paying doctors by salary: a controlled study of general practitioner behaviour in England. Health Policy 2003;**64**(3):415-23.
50. Hickson G, Altemeier W, Perrin J. Physician reimbursement by salary or fee-for-service: effect on physician practice behavior in a randomized prospective study. Pediatrics 1987;**30**(3):344-50.
51. Hutchison B, Birch S, Hurley J, et al. Do physician-payment mechanisms affect hospital utilization? A study of Health Service Organizations in Ontario. Canadian Medical Association Journal 1996;**154**(5):653-61.
52. Krasnik A, Groenewegen PP, Pedersen PA, et al. Changing remuneration systems: effects on activity in general practice. British Medical Journal 1990;**300**(6741):1698-703.
53. An L, Bluhm JH, Foldes SS, et al. A randomized trial of a pay-for-performance program targeting clinician referral to a state tobacco quitline. Arch Intern Med 2008;**168**(18):1993-99.
54. Fairbrother G, Hanson K, Friedman S, et al. The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates. Am J Public Health 1999;**89**(2):171 - 75.
55. Fairbrother G, Siegel MJ, Friedman S, et al. Impact of Financial Incentives on Documented Immunization Rates in the Inner City: Results of a Randomized Controlled Trial. Ambulatory pediatrics : the official journal of the Ambulatory Pediatric Association 2001;**1**(4):206-12.
56. Grady K, Lemkau J, Lee N, et al. Enhancing mammography referral in primary care. Prev Med 1997;**26**:791 - 800.
57. Hillman A, Ripley K, Goldfarb N, et al. Physician financial incentives and feedback: Failure to increase cancer screening in Medicaid managed care. Am J Public Health 1998;**88**(11):1699 - 701.
58. Hillman A, Ripley K, Goldfarb N, et al. The use of physician financial incentives and feedback to improve pediatric preventive care in Medicaid managed care. Pediatrics 1999;**104**(4):931 - 35.

59. Kouides RW, Bennett NM, Lewis B, et al. Performance-based physician reimbursement and influenza immunization rates in the elderly. American Journal of Preventive Medicine 1998;**14**(2):89-95.
60. Mullen KJ, Frank RG, Rosenthal MB. Can you get what you pay for? Pay-for-performance and the quality of healthcare providers. RAND Journal of Economics (Blackwell Publishing Limited) 2010;**41**(1):64-91.
61. Ritchie LD, Bisset AF, Russell D, et al. Primary and preschool immunisation in Grampian: progress and the 1990 contract. Bmj 1992;**304**(6830):816-19.
62. Rosenthal M, Frank RG, Li Z, et al. Early experience with pay-for-performance: From concept to practice. Jama 2005;**294**(14):1788-93.
63. Roski J, Jeddeloh R, An L, et al. The impact of financial incentives and a patient registry on preventive care quality: increasing provider adherence to evidence-based smoking cessation practice guidelines. Prev Med 2003;**36**(3):291-99.
64. Twardella D, Brenner H. Effects of practitioner education, practitioner payment and reimbursement of patients' drug costs on smoking cessation in primary care: a cluster randomised trial. Tobacco Control 2007;**16**(1):15-21.
65. Young G, Meterko M, Beckman H, et al. Effects of Paying Physicians Based on their Relative Performance for Quality. J Gen Intern Med 2007;**22**(6):872-76.
66. NHS Confederation. Investing in General Practice: The New General Medical Services Contract: NHS Confederation, 2003.
67. The National Collaborating Centre for Chronic Conditions. Atrial Fibrillation: National clinical guideline for management in primary and secondary care: National Health Service, 2006.
68. Williams B, Poulter NR, Brown MJ, et al. Guidelines for management of hypertension: report of the fourth working party of the British Hypertension Society, 2004-BHS IV. Journal of Human Hypertension 2004;**18**(3):139-85.
69. Andrews J, Guyatt G, Oxman AD, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. J Clin Epidemiol 2013;**66**(7):719-25.
70. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;**64**(4):383-94.
71. Scottish Intercollegiate Guidelines Network. Risk estimation and the prevention of cardiovascular disease: A national clinical guideline: NHS Quality Improvement Scotland, 2007.
72. National Institute for Health and Clinical Excellence. Hypertension: Clinical management of primary hypertension in adults. London: National Institute for Health and Clinical Evidence, 2011.
73. National Clinical Guideline Centre. Stable Angina: Methods, Evidence and Guidance: Royal College of Physicians, 2011.
74. Scottish Intercollegiate Guidelines Network. Diagnosis and management of chronic kidney disease: A national clinical guideline: NHS Quality Improvement Scotland, 2008.
75. The National Collaborating Centre for Chronic Conditions. Chronic Kidney Disease: National clinical guideline for early identification and management in adults in primary and secondary care: National Health Service, 2008.
76. Scottish Intercollegiate Guidelines Network. Management of diabetes: A national clinical guideline: NHS Quality Improvement Scotland, 2010.
77. The National Collaborating Centre for Chronic Conditions. Stroke: Diagnosis and initial management of acute stroke and transient ischaemic attack: National Health Service, 2008.

78. National Institute for Health and Clinical Excellence. Smoking cessation services in primary care, pharmacies, local authorities and workplaces, particularly for manual working groups, pregnant women and hard to reach communities: National Institute for Health and Clinical Excellence, 2008.
79. National Collaborating Centre for Mental Health. Bipolar disorder: The management of bipolar disorder in adults, children and adolescents, in primary and secondary care: National Institute for Health and Clinical Excellence, 2006.
80. National Collaborating Centre for Mental Health. Schizophrenia: The NICE guidline on core interventions in the treatment and management of Schizophrenia in adults in primary and secondary care: National Institute for Health and Clinical Excellence, 2010.
81. The National Collaborating Centre for Primary Care. Post Myocardial Infarction: Secondary prevention in primary and secondary care for patients following a myocardial infarction. In: Care NCCfP, ed.: Royal College of General Practitioners, 2007.
82. Scottish Intercollegiate Guidelines Network. Management of stable angina: A national clinical guideline: NHS Quality Improvement Scotland, 2007.
83. Scottish Intercollegiate Guidelines Network. Management of chronic heart failure: A national clinical guideline: NHS Quality Improvement Scotland, 2007.
84. National Clinical Guideline Centre. Chronic heart failure: the management of chronic heart failure in adults in primary and secondary care. London: National Institute for Health and Clinical Excellence, 2010.
85. National Clinical Guideline Centre. The Epilepsies: The diagnosis and management of the epilepsies in adults and children in primary and secondary care: National Institute for Health and Clinical Excellence, 2012.
86. The National Collaborating Centre for Primary Care. Lipid modification: Cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease: The Royal College of General Practitioners, 2010.
87. Prepared by: British Cardiac Society BHS, Diabetes UK, HEART UK, Primary Care Cardiovascular Society, The Stroke Association. JBS 2: Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice. Heart 2005;**91**(suppl 5):v1-v52.
88. Wannamethee S, Shaper A, Whincup P, et al. Smoking cessation and the risk of stroke in middle-aged men. JAMA: The Journal of the American Medical Association 1995;**274**(2):155-60.
89. Shaper AG, Pocock SJ, Walker M, et al. Risk factors for ischaemic heart disease: the prospective phase of the British Regional Heart Study. Journal of Epidemiology and Community Health 1985;**39**(3):197-209.
90. Ockene IS, Miller NH, Reduction FtAHATFoR. Cigarette Smoking, Cardiovascular Disease, and Stroke: A Statement for Healthcare Professionals From the American Heart Association. Circulation 1997;**96**(9):3243-47.
91. Manning WG. The Use of Pathology Services: A Comparison of Fee-for-Service and A Prepaid Group Practice. Santa Monica, CA: RAND Corporation, 1983.
92. Assaf AR, Lapane KL, McKenney JL, et al. Possible influence of the prospective payment system on the assignment of discharge diagnoses for coronary heart disease. N Engl J Med 1993;**329**(13):931-34.
93. Brown SL. The impact of fee-for-service and managed care insurance on access to cardiac care in Maryland. J Health Care Finance 2004;**30**(4):68-77.
94. Choudhry NK, Avorn J, Glynn RJ, et al. Full coverage for preventive medications after myocardial infarction. N Engl J Med 2011;**365**(22):2088-97.

95. Heidenreich PA, McClellan M, Frances C, et al. The relation between managed care market share and the treatment of elderly fee-for-service patients with myocardial infarction. Am J Med 2002;**112**(3):176-82.
96. Hilleman DE, Faulkner MA, Monaghan MS. Cost of a pharmacist-directed intervention to increase treatment of hypercholesterolemia. Pharmacotherapy 2004;**24**(8):1077-83.
97. Jackevicius CA, Tu JV, Demers V, et al. Cardiovascular outcomes after a change in prescription policy for clopidogrel. N Engl J Med 2008;**359**(17):1802-10.
98. Jha AK, Joynt KE, Orav EJ, et al. The long-term effect of premier pay for performance on patient outcomes. N Engl J Med 2012;**366**(17):1606-15.
99. Nelson EC, McHorney CA, Manning Jr WG, et al. A longitudinal study of hospitalization rates for patients with chronic disease: Results from the medical outcomes study. Health Serv Res 1998; **32**(6):759-74.
100. Ryan AM, Burgess Jr JF, Tompkins CP, et al. The relationship between Medicare's process of care quality measures and mortality. Inquiry 2009;**46**(3):274-89.
101. Shen Y-C. The effect of financial pressure on the quality of care in hospitals. J Health Econ 2003;**22**(2):243-69.
102. Carlisle DM, Siu AL, Keeler EB, et al. HMO vs fee-for-service care of older persons with acute myocardial infarction. Am J Public Health 1992;**82**(12):1626-30.
103. Glickman SW, Ou F-S, DeLong ER, et al. Pay for performance, quality of care, and outcomes in acute myocardial infarction. Journal of the American Medical Association 2007;**297**(21):2373-80.
104. Luft HS. Variations in patterns of care and outcomes after acute myocardial infarction for Medicare beneficiaries in fee-for-service and HMO settings. Health Serv Res 2003;**38**(4):1065-79.
105. Paone G, Higgins RS, Spencer T, et al. Enrollment in the Health Alliance Plan HMO is not an independent risk factor for coronary artery bypass graft surgery. Circulation 1995;**92**(9 Suppl):II69-72.
106. Petersen LA, Normand S-LT, Leape LL, et al. Regionalization and the underuse of angiography in the Veterans Affairs Health Care System as compared with a fee-for-service system. N Engl J Med 2003;**348**(22):2209-17.
107. Sada MJ, French WJ, Carlisle DM, et al. Influence of payor on use of invasive cardiac procedures and patient outcome after myocardial infarction in the United States. Participants in the National Registry of Myocardial Infarction. J Am Coll Cardiol 1998;**31**(7):1474-80.
108. Starr A, Furnary AP, Grunkemeier GL, et al. Is referral source a risk factor for coronary surgery? Health maintenance organization versus fee-for-service system. J Thorac Cardiovasc Surg 1996;**111**(4):708-16; discussion 16-7.
109. Young GJ, Cohen BB. The process and outcome of hospital care for Medicaid versus privately insured hospital patients. Inquiry 1992;**29**(3):366-71.
110. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. Journal of Epidemiology and Community Health 1998;**52**(6):377-84.
111. Hootman JM, Driban JB, Sitler MR, et al. Reliability and validity of three quality rating instruments for systematic reviews of observational studies. Research Synthesis Methods 2011;**2**(2):110-18.
112. Bottle A, Gnani S, Saxena S, et al. Association between quality of primary care and hospitalization for coronary heart disease in England: National cross-sectional study. 2008.

113. Campbell S, Reeves D, Kontopantelis E, et al. Quality of primary care in England with the introduction of pay for performance. N Engl J Med 2007;**357**(2):10.
114. Crawley D, Ng A, Mainous AG, 3rd, et al. Impact of pay for performance on quality of chronic disease management by social class group in England. J R Soc Med 2009;**102**(3):103-7.
115. Kiran T, Hutchings A, Dhalla IA, et al. The association between quality of primary care, deprivation and cardiovascular outcomes: a cross-sectional study using data from the UK Quality and Outcomes Framework. Journal of Epidemiology and Community Health 2010;**64**(10):927-34.
116. Levene LS, Baker R, Bankart MJG, et al. Association of features of primary health care with coronary heart disease mortality. 2010.
117. McLean G, Guthrie B, Sutton M. Differences in the quality of primary medical care services by remoteness from urban settlements. Quality and Safety in Health Care 2007;**16**(6):446-49.
118. Strong M, Maheswaran R, Radford J. Socioeconomic deprivation, coronary heart disease prevalence and quality of care: A practice-level analysis in Rotherham using data from the new UK general practitioner Quality and Outcomes Framework. Journal of Public Health Medicine 2006;**28**(1):4.
119. Lee JT, Netuveli G, Majeed A, et al. The effects of pay for performance on disparities in stroke, hypertension, and coronary heart disease management: Interrupted time series study. PLoS ONE 2011;**6**(12).
120. Murray J, Saxena S, Millett C, et al. Reductions in risk factors for secondary prevention of coronary heart disease by ethnic group in south-west London: 10-year longitudinal study (1998-2007). Fam Pract 2010;**27**(4):430-38.
121. Millett C, Gray J, Wall M, et al. Ethnic disparities in coronary heart disease management and pay for performance in the UK. J Gen Intern Med 2009;**24**(1):8-13.
122. McGovern MP, Boroujerdi MA, Taylor MW, et al. The effect of the UK incentive-based contract on the management of patients with coronary heart disease in primary care. Fam Pract 2008;**25**(1):33-39.
123. Herrett E, Shah AD, Boggon R, et al. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. Bmj 2013;**346**.
124. Hsiao C. Panel data analysis—advantages and challenges. TEST 2007;**16**(1):1-22.
125. Gujarati D. *Basic Econometrics, 4th Edition*. New York, USA: McGraw−Hill, 2002.
126. Farbmacher H. EXTENSIONS OF HURDLE MODELS FOR OVERDISPERSED COUNT DATA. Health Econ 2013;**22**(11):1398-404.
127. Professor Simon Capewell, Dr Steve Allender, Dr Julia Critchley, et al. Modelling the UK burden of Cardiovascular Disease to 2020. London: British Heart Foundation, 2008.
128. Maas A, Appelman Y. Gender differences in Coronary Heart Disease. Netherlands Heart Journal 2010;**18**(12):598-602.
129. Barnett K, Mercer SW, Norbury M, et al. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet 2012;**380**(9836):37-43.
130. Zhang Q-L, Rothenbacher D. Prevalence of chronic kidney disease in population-based studies: Systematic review. BMC Public Health 2008;**8**(1):117-29.
131. Fraser SDS, Roderick P, Bailey L, et al. Novel use of a combined primary and secondary care data resource: investigation of inequalities in chronic kidney disease identification in the UK Hampshire Health Record. The Lancet 2012;**380, Supplement 3**(0):S40.

132. NHS Employers British Medical Association. QOF frequently asked questions. In: Department for Health, ed., 2011.
133. Health and Social Care Information Centre. GP List Populations of Primary Care Organisations (aggregated to Strategic Health Authorities)-2011. Secondary GP List Populations of Primary Care Organisations (aggregated to Strategic Health Authorities)-2011 2012. http://www.hscic.gov.uk/searchcatalogue?productid=4710&q=title%3a%22Attribution+data+set+GP-registered+populations%22&sort=Relevance&size=10&page=1#top.
134. Sundquist K, Malmström M, Johansson S-E. Neighbourhood deprivation and incidence of coronary heart disease: a multilevel study of 2.6 million women and men in Sweden. Journal of Epidemiology and Community Health 2004;**58**(1):71-77.
135. Sundquist K, Winkleby M, Ahlen H, et al. Neighborhood socioeconomic environment and incidence of coronary heart disease: a follow-up study of 25,319 women and men in Sweden. Am J Epidemiol 2004;**159**(7):655-62.
136. Woodward M. Small area statistics as markers for personal social status in the Scottish heart health study. Journal of Epidemiology and Community Health 1996;**50**(5):570-76.
137. Lawlor DA, Davey Smith G, Patel R, et al. Life-course socioeconomic position, area deprivation, and coronary heart disease: findings from the British Women's Heart and Health Study. Am J Public Health 2005;**95**(1):91-97.
138. Office for National Statistics. General Lifestyle Survey Overview: A report on the 2010 General Lifestyle Survey. In: Dunstan S, ed. London: Office for National Statistics, 2012.
139. Health and Social Care Information Centre. Quality and Outcomes Framework - 2009/10. Secondary Quality and Outcomes Framework - 2009/10 2012. http://www.ic.nhs.uk/catalogue/PUB04431.
140. The Dartmouth Institute for Health Policy and Clinical Practice. The Dartmouth Atlas of Health Care. Secondary The Dartmouth Atlas of Health Care 2014. http://www.dartmouthatlas.org/.
141. Public Health England. Local Health. Secondary Local Health 2013. http://www.localhealth.org.uk/.
142. Duerden M, Avery AJ, Payne R. Polypharmacy and medicines optimization: making it safe and sound. In: Anna Brown, ed. London: The King's Fund, 2013.
143. Eijkenaar F, Emmert M, Scheppach M, et al. Effects of pay for performance in health care: A systematic review of systematic reviews. Health Policy 2013;**110**(2-3):115-30.
144. Steel N, Willems S. Research learning from the UK Quality and Outcomes Framework: a review of existing research. Qual Prim Care 2010;**18**(2):117-25.
145. Whalley D, Gravelle H, Sibbald B. Effect of the new contract on GPs' working lives and perceptions of quality of care : a longitudinal survey. Br J Gen Pract 2008;**58**(546):8-14.
146. Feng Y, Ma A, Farrar S, et al. THE TOUGHER THE BETTER: AN ECONOMIC ANALYSIS OF INCREASED PAYMENT THRESHOLDS ON THE PERFORMANCE OF GENERAL PRACTICES. Health Econ 2013.
147. Briggs A, Mihaylova B, Sculpher M, et al. Cost effectiveness of perindopril in reducing cardiovascular events in patients with stable coronary artery disease using data from the EUROPA study. Heart 2007;**93**(9):1081-86.
148. Williams MA, Fleg JL, Ades PA, et al. Secondary Prevention of Coronary Heart Disease in the Elderly (With Emphasis on Patients ≥75 Years of Age): An American Heart Association Scientific Statement From the Council on Clinical Cardiology

Subcommittee on Exercise, Cardiac Rehabilitation, and Prevention. Circulation 2002;**105**(14):1735-43.
149. Topol EJ. Intensive Statin Therapy — A Sea Change in Cardiovascular Prevention. N Engl J Med 2004;**350**(15):1562-64.
150. Dagenais GR, Pogue J, Fox K, et al. Angiotensin-converting-enzyme inhibitors in stable vascular disease without left ventricular systolic dysfunction or heart failure: a combined analysis of three trials. The Lancet 2006;**368**(9535):581-88.
151. Abdulla J, Barlera S, Latini R, et al. A systematic review: Effect of angiotensin converting enzyme inhibition on left ventricular volumes and ejection fraction in patients with a myocardial infarction and in patients with left ventricular dysfunction. European Journal of Heart Failure 2007;**9**(2):129-35.
152. Gheorghiade M, Goldstein S. β-Blockers in the Post–Myocardial Infarction Patient. Circulation 2002;**106**(4):394-98.
153. Antithrombotic Trialists' Collaboration. Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. Bmj 2002;**324**(7329):71-86.
154. Hippisley-Cox J, Coupland C, Vinogradova Y, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. Bmj 2008;**336**(7659):1475-82.
155. Kenchaiah S, Narula J, Vasan RS. Risk factors for heart failure. Med Clin North Am 2004;**88**(5):1145-72.
156. Lindhardsen J, Ahlehoff O, Gislason GH, et al. The risk of myocardial infarction in rheumatoid arthritis and diabetes mellitus: a Danish nationwide cohort study. Annals of the Rheumatic Diseases 2011;**70**(6):929-34.
157. Dixon A, Khachatryan A, Gilmour S. Does general practice reduce health inequalities? Analysis of quality and outcomes framework data. Eur J Public Health 2012;**22**(1):9-13.
158. Alshamsan R, Majeed A, Ashworth M, et al. Impact of pay for performance on inequalities in health care: Systematic review. Journal of Health Services Research and Policy 2010;**15**(3):178-84.
159. Health and Social Care Information Centre. QOF Prevalence data tables 2009/10. Secondary QOF Prevalence data tables 2009/10 2012. http://www.ic.nhs.uk/statistics-and-data-collections/supporting-information/audits-and-performance/the-quality-and-outcomes-framework/qof-2009-10/data-tables/prevalence-data-tables.
160. Jousilahti P, Vartiainen E, Tuomilehto J, et al. Sex, Age, Cardiovascular Risk Factors, and Coronary Heart Disease: A Prospective Follow-Up Study of 14 786 Middle-Aged Men and Women in Finland. Circulation 1999;**99**(9):1165-72.
161. National Heart LaBI. Coronary Heart Disease Risk Factors. Secondary Coronary Heart Disease Risk Factors 2011. http://www.nhlbi.nih.gov/health/health-topics/topics/hd/atrisk.html.
162. Natarajan S, Liao Y, Cao G, et al. Sex differences in risk for coronary heart disease mortality associated with diabetes and established coronary heart disease. Arch Intern Med 2003;**163**(14):1735-40.
163. Gibbons G. Cardioprotective Mechanisms of ACE Inhibition. Drugs 1997;**54**(5):1-11.
164. Doran T, Campbell S, Fullwood C, et al. Performance of small general practices under the UK's Quality and Outcomes Framework. Br J Gen Pract 2010;**60**(578):e335-44.
165. Stafford M, Marmot M. Neighbourhood deprivation and health: does it affect us all equally? Int J Epidemiol 2003;**32**(3):357-66.
166. Sloggett A, Joshi H. Higher mortality in deprived areas: community or personal disadvantage? Bmj 1994;**309**(6967):1470-74.

167. Ashworth M. The relationship between social deprivation and the quality of primary care. Br J Gen Pract 2007;**57**(539):441-48.
168. Greene W. Functional forms for the negative binomial model for count data. Economics Letters 2008;**99**(3):585-90.
169. Fox J. Logit and Probit Models. Toronto: York University, Ontario, 2010.
170. Drumond Andrade FC, Guevara PE, Lebrão ML, et al. Gender Differences in Life Expectancy and Disability-Free Life Expectancy Among Older Adults in São Paulo, Brazil. Women's Health Issues 2011;**21**(1):64-70.
171. VAN DEN BOS GAM. The burden of chronic diseases in terms of disability, use of health care and healthy life expectancies. The European Journal of Public Health 1995;**5**(1):29-34.