Evidence:

Do quality improvements in primary care reduce secondary care costs?

Primary research into the impact of the Quality and Outcomes Framework on hospital costs and mortality

July 2010

Identify Innovate Demonstrate Encourage

Do quality improvements in primary care reduce secondary care costs?

July 2010

Authors

Stephen Martin (1), Peter C Smith (2), Mark Dusheiko (3), Hugh Gravelle (3) and Nigel Rice (3)

Institution

1. Department of Economics, University of York
2. Imperial College Business School, London
3. Centre for Health Economics, University of York

Contact

Peter Smith

Email: peter.smith@imperial.ac.uk

Abstract

The introduction in 2004 of the Quality and Outcomes Framework (QOF) in UK general practice represents one of the most ambitious efforts to measure and incentivise quality improvements in primary care. This report takes advantage of a large database of over 50 million English citizens to determine whether the levels of QOF attainment in general practices have led to improvements in two major outcomes: mortality and the costs of hospital inpatient and outpatient use. Our findings are that primary care performance improvements are associated with some modest but measurable improvements in subsequent outcomes and costs.

Acknowledgements

This research was jointly funded by the Department of Health and the Health Foundation. The report does not necessarily reflect the views of either funder. We would like to thank Catherine Fullwood of the National Primary Care Research and Development Centre at the University of Manchester for providing us with the QOF population achievement data. We are also grateful for comments from participants at a seminar organised by the Swiss School of Public Health in Crans-Montana.

Foreword

There can be no doubt that the NHS faces significant financial challenges over the next few years. Even if funding remains constant, or there is a small real terms increase, this will be fast outstripped by increasing demand and higher-than-inflation rises in costs of medicines and equipment. It is more important than ever that we understand how the service is using resources and where costs can be reduced by doing things differently.

Up until now, research on the Quality and Outcomes Framework (QOF), one part of the general practice contract that links pay to performance, has focused on how effective it has been in changing clinical practice. This new research, supported by funding from the Health Foundation, seeks to take our understanding of the impact of QOF to a new level, attempting to answer the crucial questions: does improved performance in the QOF clinical domain lead to reduced hospital costs and, does it lead to a reduction in mortality?

Peter Smith of Imperial College and a team at the University of York have made use of newly available data sources, and the ability to link data sources to relate achievement of QOF points by GP practices to data on costs of hospital care for patients registered with these practices. The size of the data set analysed (covering 50 million patients in England) and the rigorous methods used provide a novel and invaluable insight into the relationship between QOF attainment and hospital costs and health outcomes.

The headline finding from this research is that there is an association between achievement of QOF indicators and some measurable reduction in costs for hospital care and mortality outcomes. This association is stronger for some QOF indicators than others and particularly strong for stroke care.

The report also suggests that QOF attainment in one clinical area could have a positive impact on hospital costs in other clinical areas. This suggests that studies that examine the impact of improved quality by looking at the benefits for only one disease might seriously underestimate the total benefits of that quality improvement.

However, these findings should be interpreted cautiously. The higher achievement of QOF scores is associated with, not the cause of the reduction in hospital costs. In addition, the reduction in hospital costs needs to be considered alongside increased costs to primary care and other health services, though it is worth noting that the additional payment through QOF for a one point higher score is very small compared to the associated hospital cost savings.

This research makes an important contribution to a number of topical policy initiatives, including the merits of prevention and early intervention and shifting care from secondary settings to primary care. As we introduce new models of commissioning, such new evidence will help guide more effective commissioning processes and

decisions than we have seen in the past and will help determine resource allocation at a national level.

The Health Foundation intends to continue to support this work to increase our understanding of a complicated but fundamentally important issue for the health service and for policy makers.

We hope that this report will add to the debate and to the evidence-based decision making that will improve the quality of care for patients.

Martin Marshall
Clinical Director and Director of Research and Development
The Health Foundation

Contents

Tables

Executive summary

There is a widespread belief and hope among policy-makers that timely intervention, in the form of behavioural change, preventive medicine and disease management, can both reduce demands for healthcare expenditure and improve health outcomes in the form of length and quality of life. However, the current research evidence is equivocal: most such preventive interventions increase costs, and many are not even cost-effective when compared to more conventional clinical interventions. Research suggests that if resources are to be used wisely, there is a need to focus on preventive interventions that are carefully targeted at relevant at-risk groups.

The Quality and Outcomes Framework (QOF)

The Quality and Outcomes Framework (QOF), which was introduced into UK primary care in 2004, is one of the most ambitious efforts to embed preventive efforts into the health system. It seeks to reward general practitioners (GPs) for a wide range of care processes and outcomes, with about 20% of their income tied to QOF financial incentives. Considerable effort was made to ensure that the QOF was aligned with best contemporary clinical practice (to the extent that evidence permitted). However, until now, research has examined whether the QOF has succeeded in altering clinical practice. Little work to date has examined whether it has led to reduced health service costs or improved health outcomes.

The scope of this study

This study seeks to shed light on the following research questions: does improved performance in the QOF clinical domain lead to reduced future National Health Service (NHS) hospital costs; and does it lead to a reduction in mortality? The study takes advantage of a major new database that links the register of all citizens registered with an English GP over four years to the inpatient and outpatient Hospital Episode Statistics (HES) data for all English NHS patients during that time.

A focus on preventive effort

The theoretical model underlying our research is simple (see chapter 2). It suggests that health status is determined by individual characteristics, social circumstances, access to health services, preventive effort and a random element. This empirical research seeks to isolate the specific impact of the preventive effort on future hospital costs and mortality.

The QOF achievement scores measure preventive quality across eight clinical areas:
- asthma
- chronic obstructive pulmonary disease (COPD)
- coronary heart disease
- diabetes
- hypertension
- hypothyroidism
- mental health
- stroke.

We have also used an index of overall QOF attainment. The scores were calculated by aggregating scores on individual performance indicators, weighted by the total number of QOF points allocated for that indicator. Throughout, we have used attainment scores based on the total population 'at risk', and made no adjustment for patients reported by GPs as 'exceptions', who are excluded from the performance measure for the purposes of calculating GP reimbursement.

The study dataset

The study dataset includes about 50 million English citizens, and encompasses over 500 variables, grouped into three broad categories:

1. individual characteristics
2. local (small area) population needs characteristics
3. local (small area) supply characteristics.

The principal individual characteristics used are:

- age
- sex
- previous diagnoses (152 categories) based on previous hospital encounters
- intensity of previous hospital use.

Individual-level data are complemented by over 160 indicators of the socio-economic characteristics of the small area in which the patient lives, and QOF data on disease prevalence rates for various conditions for the GP with which the individual is registered. These are intended to reflect area influences on the individual's need for healthcare, and also act as a proxy for data (such as morbidity) that are not available at an individual level.

We have also used over 130 variables of local health services supply that may reflect supply-side influences on health expenditure and outcomes. These include factors such as waiting times, distance to services, general practice characteristics, and the QOF attainment scores – the principal focus of this study.

Isolating the impact of QOF attainment on costs

The technical challenge is to isolate the impact of QOF attainment on costs (and outcomes) after adjusting for all other possible determinants. To be done satisfactorily, this requires the development of a comprehensive but tractable statistical model of the determinants of costs (or mortality), and the use of advanced statistical methods.

First, we report the development of a 'base' model of the determinants of NHS hospital expenditure in 2007/08 on individuals (excluding mental health and maternity services). This builds on work for the Department of Health (Dixon, Bardsley et al 2009) that is being used as the basis for allocating general practices' indicative budgets for hospital use by their registered patients. In order to economise on computing time, the model is based on a 10% sample of the study population (about five million people). After exploring more advanced alternatives, we concluded that conventional ordinary least squares (OLS) regression methods could be used to identify the quantitative relationship between patient hospital costs and our extensive set of possible influences on cost.

Developing a 'parsimonious model'

Retaining over 500 variables in the model would be unhelpful. We therefore developed a 'parsimonious' model of hospital expenditure on individuals that is intended to be as manageable, as statistically valid and as informative as possible. It retains all the individual-level variables, but – using an explicit set of selection criteria – retains only the most statistically significant and plausible small area or general practice variables. It results in the selection of seven local needs variables and three local supply variables, including one QOF attainment score for 2005/06, for the quality of stroke care.

Impact on the stroke achievement rate

The results suggest that a one-point increase in the stroke QOF achievement rate will be associated with a fall of £0.44 per person in hospital costs. With a population of 50 million people, this implies that a one-point increase in the mean stroke population achievement rate from 79.64 to 80.64 would be associated with a reduction in annual total hospital costs of £22.15 million. Although this is a modest sum when compared with the total secondary care spend (in 2007/08 this was about £22 billion, excluding expenditure on maternity and mental health), it is consistent with the claim that improvements in the quality of primary care can be associated with reductions in the cost of secondary care. It may also be plausible to envisage an improvement of more than one point in the QOF achievement score: it increased by about 10 points over the three years up to 2007/08.

Impact on other clinical areas

We re-estimated this base model, replacing the QOF stroke achievement rate with the achievement

rate for each of the other seven clinical areas. None was statistically significant. However, when we substituted the overall clinical QOF population achievement rate, we achieved a very similar pattern to the stroke result, albeit with a lower level of statistical significance. The close association between results for the stroke QOF score and the overall QOF score may indicate that stroke achievement reflects general primary care quality rather than quality only in stroke care.

Variants of the parsimonious model

We then examined a series of variants of the base model. First, we explored the impact of including more recent QOF scores (for example, for 2006/07 and 2007/08) in explaining costs for 2007/08. As is to be expected, the association between QOF quality and cost reductions becomes stronger as the time period over which the measure of quality is recorded moves closer to the period in which the costs are incurred. Specifically, the strength of the association between the stroke achievement score and expenditure for 2007/08 increases by 60% compared to the results for the QOF stroke score for 2005/06.

In order to explore the association between QOF achievement and health outcomes, we used an indicator that records whether the patient died in 2007/08. Using the same model selection criteria as for costs, we could derive a parsimonious model for the probability of death. Very similar results to the hospital costs models were found, with the stroke attainment scores again dominating. The results for the 2007/08 stroke score suggest that a one-point improvement in QOF attainment is associated with 2,385 fewer annual deaths.

The QOF targets interventions in specific clinical domains. We therefore disaggregated 2007/08 hospital expenditure into 23 programme budgeting categories, based on broad international classification of diseases (ICD) chapter headings. A parsimonious model was then developed for each programme budgeting category. In general, we found little association of QOF attainment with these more detailed expenditure headings. The exceptions were:

- The stroke quality score has a significant negative association with circulatory disease costs.
- The diabetes quality score has a significant negative association with the 'other' costs model.
- The dementia quality score has a significant negative association with cancer costs.

The programme-specific savings implied by these results are quite modest, and are smaller in sum than the savings in total costs noted earlier.

Obtaining more accurate estimates

The findings above rely on 2007/08 expenditure data, and are cross-sectional in nature (they consider only a one-year snapshot). While we can report associations between QOF attainment and cost reductions, we urge caution in inferring causality, because there may be some unobserved variable that is correlated with both QOF attainment and costs that confounds the analysis. In order to obtain more accurate estimates, we therefore constructed the analogous expenditure and explanatory data for the two preceding years: 2005/06 and 2006/07. We could then re-estimate the parsimonious expenditure model for the three years using more advanced statistical techniques. This enabled us to control for time-invariant unobserved factors (such as practice characteristics) that are correlated with both quality and cost, but whose influence would otherwise be attributed to quality in the one-period, cross-section models.

We estimated a multi-year version of the favoured 2007/08 model, which includes the QOF stroke attainment score, and we report several variants of this multi-year model. We discuss the implications of the results found, which confirm qualitatively the results obtained using the one-period model.

Our favoured models suggest that the true estimate of the marginal impact of QOF attainment on costs is likely to be somewhat lower than that suggested by the cross-sectional models. Using the midpoint of our two favoured models, we find that a one-percentage-point increase in the stroke QOF score is associated with a £16.5 million annual reduction in total patient costs. Over the period studied, the mean practice QOF stroke score increased by 10 percentage points, and we therefore tentatively suggest that annual secondary care costs may have been about £165 million lower in 2007/08 than in 2004/05 as a result of the increase in primary care quality.

How QOF might affect hospital costs or mortality outcomes

We must emphasise that this study is not seeking to evaluate the QOF initiative, or to offer estimates of the cost-effectiveness of QOF interventions. Rather, it gives an indication of the extent to which the initiative may have affected hospital costs or mortality outcomes. We find in both respects that the QOF appears to be associated with material but limited gains. We are cautious about drawing inferences of causality from our work, but feel that the panel data results do offer solid grounds for believing that QOF improvements are contributing to the gains.

The stroke QOF score dominates our models. To some extent, this may be because it is an indicator of overall primary care quality. It is highly correlated with overall QOF attainment. However, its dominance, and the role it plays in the model of circulatory disease costs, suggests that the stroke quality metrics are capturing specific aspects of preventive care that do have a measurable impact on outcomes.

Chapter 1

Introduction

1.1 GP contracts

Virtually all GPs in England are paid under one of two contracts. About two-thirds of GPs are in practices that operate within the General Medical Services (GMS) contract, which is periodically re-negotiated by the British Medical Association (BMA) (acting as the doctors' representative) and central government. GMS practices receive a mixture of capitation, lump-sum allowances, items of service and target incentives. About one-third of practices operate within a personal medical services (PMS) contract, which is negotiated between the practice and its local primary healthcare organisation (primary care trust – PCT). These PMS practices receive a lump sum for the services that they would have provided under the GMS contract, plus further remuneration for the additional services that they provide for specific patient groups (National Audit Office 2008).

A new GMS contract for the delivery of primary care in England was introduced in April 2004.[1] This contract came with considerable additional funding for general practices, and expenditure on primary care increased from £5.8 billion in 2003/04 to £7.7 billion in 2005/06. The new contract and associated additional expenditure were designed to increase the number of GPs working in the NHS, particularly in deprived and under-doctored areas, and to improve the quality of primary care delivered to patients (National Audit Office 2008).

Previous attempts to introduce a pay-for-performance element into the GMS contract had either been rejected by the BMA or had been on a very small scale and had made little impact (Roland 2004). However, as part of the new contract, about 20% of GP income became tied to financial incentives for practices to improve the quality of care delivered to patients. Because of the difficulty of attributing an (improved) health outcome to the specific activity of a GP, the Quality and Outcomes Framework (QOF) element of the new contract tied payments mainly to process activities over which GPs have direct control, and for which there is evidence of subsequent benefits to the patient (Doran 2008; Roland 2004).

The new contract certainly benefited GPs. Average GP incomes increased by 34% in two years, rising from £84,795 in 2003/04 to £113,614 in 2005/06 (National Audit Office 2008). The new contract also reduced GPs' hours of work and removed the requirement for practices to provide an out-of-hours urgent care service (this responsibility passed to PCTs). However, in its early years, the new contract did not lead to a measurable improvement in moving services into either deprived or under-doctored areas (National Audit Office 2008).

1.2 Quality indicators

The new contract was also expected to benefit patients and the wider NHS. The initial 146 indicators were split between four 'domains':

- clinical (76 indicators)
- organisational (56 indicators)
- patient experience (4 indicators)
- additional services (10 indicators).

The 76 indicators in the clinical domain accounted for 550 of the available 1,050 points and, as a result, under the initial version of QOF, clinical quality determined about 10% of GP income. These 76 indicators related to various common chronic

diseases (such as diabetes) and typically referred to the regularity of monitoring (such as whether the patient's blood pressure has been recorded in the last 15 months). Practices were awarded points according to the proportion of eligible patients for whom each target was met. By stimulating an improvement in chronic disease management, the QOF was expected to lead to a reduction in avoidable hospital admissions (NHS Information Centre for Health and Social Care 2005).

1.3 Practice QOF scores and hospital admission rates

Although there is a growing literature on the impact of the QOF on GP activity and performance, there have been only a small number of studies of the association between practices' QOF scores and hospital admission rates. These have found only a weak association between quality scores and admissions (Downing, Rudge et al 2007; Shohet, Yelloly et al 2007; Bottle, Gnani et al 2008; Bottle, Millett et al 2008). Dusheiko, Doran et al (2009) noted that this might be due to the use of a relatively small sample of practices, to the focus on a single year of data, or to the characteristics of the pay-for-performance scheme (for example, the QOF quality indicators have upper achievement thresholds of between 50% and 90%, so that practices can score the maximum number of points without achieving the target for all patients).

Dusheiko, Doran et al (2009) undertook a more comprehensive study of the relationship between practice QOF scores and practice hospital admission rates. Their focus was on diabetes and, in particular, whether better diabetes management in primary care (as measured by the QOF indicators) was associated with fewer emergency hospital admissions for short-term complications of diabetes. Dusheiko and his team estimated a pooled cross-section regression model for 2004/05 to 2006/07 at the practice level, with the emergency diabetic admission rate as the dependent variable. They studied all English practices with a list size of at least 1,000 patients. Their estimated models included diabetes prevalence rates and baseline (pre-QOF) admission rates, together with several sets of covariates, including:

- practice and patient characteristics
- access measures to primary and secondary care
- local population characteristics (for example, indicators of deprivation)
- year and PCT dummies.

Dusheiko, Doran et al found that emergency admission rates for all short-term diabetic complications were significantly lower when practices had more patients with good and moderately well-controlled diabetes. They calculated that moving 10% of registered diabetic patients from poor to good control in an average practice was associated with a 14% decrease in the rate of emergency admissions for short-term complications, and a £1,928 reduction in hospital costs per practice in 2006/07. However, the authors noted some limitations to their study. Quality of care was measured at the practice level, and could not be adjusted for the age, sex, co-morbidity or type of diabetes for individual diabetic patients. Their study was also unable to determine which patients from each practice were admitted to hospital. Another limitation was that the study could not examine the impact of better diabetic care on admissions for other (non-diabetic) conditions. It might be that practices that provided better diabetic care (and so incurred fewer diabetic admissions) did so at the expense of care for patients with other diseases and that, as a result of spending more time on diabetic care, such practices devoted less time to other conditions. Although better diabetic care might be associated with fewer diabetic admissions, it might also be associated with more admissions for other diseases. To address this issue, the impact that better primary care had on all secondary care costs would need to be examined, and not just those costs associated with one particular disease.

1.4 Drawing on a new dataset

In this study, we take advantage of a major new dataset to examine whether higher practice QOF scores are associated with reduced hospital costs for each patient registered with an English practice on 1 April 2007. This dataset brings together practice-based patient registration data and patient-level hospital use data for all English citizens registered with a general practitioner (Dixon, Bardsley et al 2009). It enables us to study

whether patient hospital costs in 2007/08 are

associated with QOF scores in 2004/05, 2005/06, 2006/07 or 2007/08, conditional on, for example, the patient's age and gender and their recent use of hospital services. In addition, we split total hospital costs across 23 programme budget categories. This allows us to examine the impact of practice QOF scores on each patient's hospital costs in individual care programmes. Finally, the dataset also includes a binary variable that indicates whether the patient died during the 12-month period from 1 April 2007 to 31 March 2008. This enables us to examine the association between the quality of primary care and the probability of death, and how many fewer deaths might be expected if the quality of care were increased by a small amount.

1.5 Report structure

The structure of this report is as follows. Chapter 2 briefly reviews the policy and clinical context, and offers a rudimentary mathematical model of disease management. Chapter 3 describes the QOF data on which this study is based, while chapter 4 outlines the model to be estimated and relevant estimation issues. Chapter 5 presents the results and discusses several variants of the basic model. Chapter 6 contains some concluding remarks.

Chapter 2

Theoretical background

The QOF is probably the most advanced attempt to embed preventive medicine and 'disease management' into primary care. Considerable effort was made to ensure that it was aligned with best contemporary clinical practice (to the extent that evidence permitted), by engaging relevant professionals in detailed working groups. The intention of disease management is to ensure that at-risk groups, or those with established chronic conditions, are offered timely interventions and advice that increase their future health prospects and reduce expected future health services expenditure (Congressional Budget Office 2004).

2.1 Effectiveness or cost-effectiveness?

There is a rich literature on prevention efforts and disease management that is usually specific to the chosen clinical domain. In interpreting the literature, a crucial issue is whether the study is examining only the effectiveness of interventions (in terms of future health of the patient) or their cost-effectiveness. The majority of studies have demonstrated that prevention and treatment for chronic conditions, while usually improving health outcomes, tend also to increase health services costs. Fewer than 20% of studies have identified cost-saving interventions (Russell 2009). The cost-effectiveness of disease management is therefore a critical issue. As summarised by Cohen, Neumann and Weinstein (2008): 'careful analysis of the costs and benefits of specific interventions, rather than broad generalisations, is critical'. In other words, it is likely that the precise population groups targeted, and the frequency and mode of implementation, will be crucial determinants of an intervention's impact on health and health service costs.

2.2 Scope of this study

In this study, we are not seeking to undertake a comprehensive cost-effectiveness analysis of the various interventions embodied in the QOF. Rather, we are examining whether an improvement in a general practice's performance on specific QOF clinical areas is associated with reduced subsequent hospital costs, and with reduced subsequent mortality.

At a time of financial retrenchment in the NHS, such information is essential if limited general practice capacity is to be focused on interventions that will not only improve health but also reduce NHS expenditure.

2.3 A mathematical model

As shorthand, while recognising the limitations of the expression, we shall refer to all of the interventions covered by the QOF as 'prevention'. Then, the impact of prevention on costs and outcomes can be represented by a very simple dynamic mathematical model. We represent the health status of an individual in time t by h_t. This depends on health status in the previous period (the individual's 'stock' of health), the level of any preventive efforts in that period, and a random stochastic element. That is:

$$h_t = f(h_{t-1}, p_{t-1}, \boldsymbol{z}) + \varepsilon_t$$

where p_{t-1} is preventive effort in year $t-1$, $\boldsymbol{z}$ is a vector of personal characteristics unrelated to health (such as education level), and ε_t is a stochastic shock. So with multiple periods, ignoring the stochastic element:

$$h_t = g(h_0, p_{t-1}, p_{t-2}, \ldots, p_0, \boldsymbol{z})$$

We might assume that healthcare costs c_t in year t are straightforwardly related to h_t via (say) $c_t = c(h_t, \boldsymbol{z})$. This formulation would justify modelling c_t as a function of $(h_0, p_{t-1}, p_{t-2}, ..., p_0, \boldsymbol{z})$. In practice, actual costs in year t will give a signal of the magnitude of the stochastic shock in the year, so they have information content in addition to preventative effort. There is therefore also a case for entering $(c_{t-1}, c_{t-2}, ..., c_0)$ or some other indicators of previous health service use into any model explaining costs in year t.

It is straightforward to specify an analogous model for mortality. Suppose the probability of survival in year t is s_t. Assuming the patient is alive in year 0 ($s_0 = 1$), the probability of survival can then be modelled as:

$$s_t = s(h_{t-1}, p_{t-1}, \boldsymbol{z}).s_{t-1}$$

yielding

$$s_t = q(h_0, p_{t-1}, p_{t-2}, ..., p_0, \boldsymbol{z})$$

Survival depends on initial health status, personal characteristics $\boldsymbol{z}$, and the history of preventive effort. Again, stochastic elements can be integrated into this model, and might be captured empirically by measures of previous health service costs or use.

Chapter 3

The Quality and Outcomes Framework (QOF)

The Quality and Outcomes Framework (QOF) was introduced as part of the new General Medical Services (GMS) contract in April 2004. Participation by practices in the QOF was voluntary, although participation rates were (and remain) very high. The QOF component of the new contract measured practice achievement against 146 indicators.[2] Practices scored points on the basis of their achievement against each indicator, up to a maximum of 1,050 points, and an average-sized practice received £75 per point in 2004/05 and £125 per point in 2005/06. Some of the indicators included in the QOF and the points available for some QOF indicators changed in 2006/07, but the basic principles behind the scheme remained the same. The revised QOF allowed a possible maximum score of 1,000 points, according to the revised set of 135 indicators (NHS Information Centre for Health and Social Care 2007).

PMS practices could also take part in the QOF, but, because it was thought that they would already be being paid for some of the services counting towards the QOF, for the purposes of reimbursement, they had some points deducted from their QOF score.[3] The QOF achievement data are derived from the Quality Management Analysis System (QMAS), a national IT system that uses data from general practices to calculate individual practices' quality achievement scores and disease prevalence rates. QMAS is a live database to which practices can submit clinical and non-clinical data at any time. QOF scores for 2004/05 were based on practice submissions on 31 March 2005 for the complete financial year April 2004 to March 2005. These submissions might have been made late (all late submissions made by the end of June are included), or might have been adjusted by the PCT in the period April to June 2005 (NHS Information Centre for Health and Social Care 2005). Similar arrangements existed for the submission of data for the QOF in 2005/06, 2006/07 and 2007/08.

3.1 QOF 2004/05 and 2005/06

Quality indicators in the clinical domain

The QOF component of the new GMS contract rewarded practices according to quality indicators in four different 'domains':
clinical (76 indicators)
- organisational (56 indicators)
- patient experience (4 indicators)
- additional services (10 indicators).

As the focus of this report is on clinical quality and its impact on secondary care costs, table 1 lists the 11 disease areas within the clinical domain. By way of illustration, for two of the disease areas (diabetes and stroke), tables 2 and 3 list the indicators within these disease areas and also report the minimum and maximum achievement thresholds for each indicator.

Table 1: The 11 disease sub-domains within the clinical domain, QOF 2004/05 and 2005/06

Disease sub-domain	Number of indicators (including the existence of a disease register): in total	that refer to all patients with the disease	Total points available for all indicators
Asthma	7	2	72
Cancer	2	1	12
Chronic obstructive pulmonary disease (COPD)	8	5	45
Coronary heart disease (CHD)	12	7	101
Diabetes	18	12	99
Epilepsy	4	1	16
Hypertension	5	3	105
Hypothyroidism	2	2	8
Left ventricular dysfunction (LVD) (with CHD)	3	1	20
Mental health	5	2	41
Stroke	10	5	31
Total	**76**	**41**	**550**

Source: Department of Health (2004)

Table 2: Indicators present in the diabetes clinical sub-domain

Indicator number	Indicator description	Minimum threshold	Maximum threshold	Available points
DM 1	The practice can produce a register of all patients with diabetes mellitus			6
DM 2	% whose notes record their BMI in the previous 15 months	25	90	3
DM 3	% in whom there is a record of smoking status in the previous 15 months except those who have never smoked	25	90	3
DM 4	% who smoke and whose notes contain a record that smoking cessation advice has been offered in the last 15 months	25	90	5
DM 5	% who have a record of HbA1c or equivalent in the previous 15 months	25	90	3
DM 6	% in whom the last HbA1c is 7.4 or less in the last 15 months	25	50	16
DM 7	% in whom the last HbA1c is 10 or less in last the 15 months	25	85	11

continued

Table 2: Indicators present in the diabetes clinical sub-domain – *continued*

Indicator number	Indicator description	Minimum threshold	Maximum threshold	Available points
DM 8	% who have a record of retinal screening in the previous 15 months	25	90	5
DM 9	% with a record of presence or absence of peripheral pulses in the previous 15 months	25	90	3
DM 10	% with a record of neuropathy testing in the previous 15 months	25	90	3
DM 11	% who have a record of the blood pressure in the past 15 months	25	90	3
DM 12	% in whom the last blood pressure is 145/85 or less	25	55	17
DM 13	% who have a record of micro-albuminuria testing in the previous 15 months	25	90	3
DM 14	% who have a record of serum creatinine testing in the previous 15 months	25	90	3
DM 15	% with proteinuria or micro-albuminuria who are treated with ACE inhibitors (or A2 antagonists)	25	70	3
DM 16	% who have a record of total cholesterol in the previous 15 months	25	90	3
DM 17	% whose last measured total cholesterol within previous 15 months is five or less	25	60	6
DM 18	% who have had influenza immunisation in the preceding 1 September to 31 March	25	85	3

Note: for '%', read 'The percentage of patients with diabetes'.
Source: Department of Health (2004)

Table 3: Indicators present in the stroke clinical sub-domain

Indicator number	Indicator description	Minimum threshold	Maximum threshold	Available points
STROKE 1	The practice can produce a register of patients with stroke and transient ischaemic attack (TIA)			4
STROKE 2	The percentage of new patients with presumptive stroke who have been referred for confirmation of the diagnosis by CT or MRI scan	25	80	2
STROKE 3	% who have a record of smoking status in the last 15 months, except those who have never smoked	25	90	3
STROKE 4	% who smoke and whose notes contain a record that smoking cessation advice has been offered in the last 15 months	25	70	2

continued

Table 3: Indicators present in the stroke clinical sub-domain – *continued*

Indicator number	Indicator description	Minimum threshold	Maximum threshold	Available points
STROKE 5	% who have a record of blood pressure in the notes in the preceding 15 months	25	90	2
STROKE 6	% in whom the last blood pressure reading (measured in the last 15 months) is 150/90 or less	25	70	5
STROKE 7	% who have a record of total cholesterol in the last 15 months	25	90	2
STROKE 8	% whose last measured total cholesterol (measured in the last 15 months) is 5 mmol/l or less	25	60	5
STROKE 9	% who have a record that aspirin, an alternative anti-platelet therapy, or an anti-coagulant is being taken (unless a contraindication or sideeffect is recorded)	25	90	4
STROKE 10	% who have had influenza immunisation in the preceding 1 September to 31 March	25	85	2

Note: for '%', read 'The percentage of patients with stroke or transient ischaemic attack'.
Source: Department of Health (2004)

Key points about the clinical indicators:

- Some indicators refer to all patients with a disease (for example, DM2: the percentage of patients with diabetes whose notes record their BMI in the previous 15 months), while others relate to a subset of patients with the disease (for example, DM4: the percentage of patients with diabetes who smoke and whose notes contain a record that smoking cessation advice has been offered in the last 15 months).
- Most of the indicators within the clinical domain relate to the regularity of monitoring, and the number of points earned on an indicator increases linearly with the percentage of eligible patients for whom each target is met.
- However, this linear relationship between achievement and points earned only applies between two thresholds. The minimum achievement threshold for every target within the clinical domain in the initial version of QOF was 25%, and practices that achieved the target for less than 25% of eligible patients received no points for that indicator.
- The maximum achievement threshold was not constant across all clinical indicators but varied between 50 and 90%.
- Practices that recorded a score above the maximum threshold received no additional points beyond those available for meeting the upper threshold.
- Setting upper thresholds below 100% was designed to reduce the risk that GPs would inappropriately treat some patients (Roland 2004). However, this might also discourage practices from including the most hard-to-reach patients because no further points are received when a practice has achieved 90% coverage (National Audit Office 2008).

Exception reporting

Most of the clinical indicators are expressed as percentages – an approach designed to encourage practices to increase the number of treated patients from the set of patients eligible for treatment. However, practices can exclude some patients from the denominator by designating them as 'exceptions'. Patients can be exception-reported for several reasons (these are outlined in the Appendix). Exception reporting is intended to avoid penalising practices where, for example, patients do not attend for review, or where a medication cannot be prescribed due to a contraindication or sideeffect, and it is an important mechanism in the absence of any other adjustment for case-mix complexity.

However, there is always the possibility that practices might inappropriately exception report patients to increase their achievement rate on any particular indicator. Doran, Fullwood et al (2006) report that: 'a small number of practices appear to have achieved high scores by excluding large numbers of patients by exception reporting. More research is needed to determine whether these practices are excluding patients for sound clinical reasons or in order to increase income'. Gravelle, Sutton and Ma (2010) test for gaming of exception reporting by comparing the rates of exception reporting in 2005/06 for practices that were above the upper threshold in 2004/05 (which would have had no incentive to increase exception reporting) with practices that were below the threshold in 2004/05 (which would have had an incentive to increase exception reporting). They 'find evidence that practices which performed worse in 2004/05 were more likely to game exceptions in 2005/06'.

Given the possibility that reported achievement rates may involve some gaming, in this study, we use the population achievement rates (unadjusted for exceptions) as indicators of quality. Also, it is important to note that our study is interested in the impact of QOF achievement on future costs and health outcomes, and allows adjustment for any variation in achievement caused by local population characteristics. Therefore, the population achievement rate appears the most appropriate metric to use. Since 2005/06, the number of exception-reported patients by each practice has been publicly available and so the calculation of population achievement rates from this date is straightforward.

However, to derive population achievement rates for the first year of QOF (2004/05), exception reporting rates for each indicator have been imputed using Doran, Fullwood et al's method (2006).

Key points about his method:

- It uses practice disease-register counts on National Prevalence Day (14 February 2005) to estimate the number of patients for whom the indicator was relevant before the removal of exception-reported patients.
- These practice disease-register counts refer to all patients in the practice and are not available for subgroups of patients with a disease (for example, for those patients aged 8 years and over, or for those patients with a diagnosis since 1 April 2003).
- In this way, only those (30) clinical indicators that refer to all patients can be adjusted to include estimated exception reports for 2004/05.
- Consequently, the estimated population achievement rate for each disease area is based on a weighted average of the achievement score for those indicators that refer to all patients (with weights reflecting the maximum number of points available for each indicator).
- Indicators that refer to a subgroup of patients are excluded from our achievement rates.
- Although exception reports are available for 2005/06, the population weighted achievement rates for this year have been calculated on the same basis as those for 2004/05 (that is, they only use indicators that relate to all patients with a disease, and practice disease-register counts on National Prevalence Day have been used to estimate the number of patients for whom the indicator was relevant before the removal of exception-reported patients).
- This facilitates a comparison between achievements rates for 2004/05 and those for 2005/06.

Population achievement rates

Table 4 reports descriptive statistics for population achievement rates for eight of the 11 clinical sub-domains for 2004/05 and 2005/06.[4] There are no population achievement rates for the cancer, epilepsy and chronic heart disease (CHD)/left ventricular dysfunction (LVD) sub-domains because all of their constituent indicators refer to subgroups of patients with cancer, epilepsy and CHD/LVD, and so the reported achievement rates cannot be adjusted for patient exceptions.

As table 4 shows, practice population achievement rates varied considerably across the clinical sub-domains (for example, in 2004/05, population achievement rates ranged from about 69% for asthma to 94% for hypothyroidism). With the exception of the hypothyroidism domain (which had already recorded a very high achievement rate in 2004/05), the average achievement rate for each clinical sub-domain increased by about five percentage points in 2005/06.

Table 4: Descriptive statistics for population achievement rates in selected clinical sub-domains, QOF 2004/05 and 2005/06

QOF variable	Number of practices	Population achievement rate			
		Mean	Standard deviation	Min	Max
Asthma 2004/05	8,536	68.84	17.39	0.00	100
Asthma 2005/06	8,407	74.33	11.99	0.00	100
Chronic obstructive pulmonary disease (COPD) 2004/05	8,505	74.27	15.19	0.00	100
COPD 2005/06	8,385	83.10	10.13	2.95	100
Coronary heart disease 2004/05	8,526	77.68	7.08	4.18	100
Coronary heart disease 2005/06	8,399	81.37	5.15	29.38	100
Diabetes 2004/05	8,538	72.44	8.66	1.40	100
Diabetes 2005/06	8,404	76.70	6.35	26.54	100
Hypertension 2004/05	8,541	75.33	8.59	0.00	100
Hypertension 2005/06	8,407	79.38	6.63	24.18	100
Hypothyroidism 2004/05	8,525	93.53	7.92	0.00	100
Hypothyroidism 2005/06	8,402	95.12	5.01	0.00	100
Mental health 2004/05	8,230	77.82	22.96	0.00	100
Mental health 2005/06	8,360	81.90	17.63	0.00	100
Stroke 2004/05	8,508	74.54	9.34	4.54	100
Stroke 2005/06	8,391	79.54	7.05	0.00	100

Correlation coefficients

Table 5 shows that, although the practice population achievement rates for each clinical sub-domain are positively correlated with each other, the correlations are not as high as might have been expected. For example, the correlation between the asthma achievement rate and the other seven clinical achievements rates varies between 0.300 for mental health and 0.502 for hypertension. It is also noticeable that the mental health achievement rate is the least well correlated with the other sub-domains. The data for 2004/05 reveal a similar pattern of correlations.

3.2 QOF 2006/07 and 2007/08

A revised QOF was introduced in April 2006. This included some new clinical areas and changed some of the clinical indicators.

Key points about the revised QOF:

- It continued to measure achievement against a set of evidence-based indicators.
- The 146 indicators and three measures of the depth of care were replaced with 135 indicators and one measure of the depth of care (known as holistic care).
- The clinical domain was expanded from 76 to 80 indicators, and these covered not 11 but 19 clinical areas.
- The proportion of points available from the clinical domain increased from 52.4% (550 out of 1,050) of the total to 67.5% (675 out of 1,000).

Disease sub-domains

Table 6 lists the 19 disease sub-domains within the revised clinical domain together with the number of indicators and the total points available for all indicators within each sub-domain.

Table 5: Correlation coefficients for population achievement rates in 2005/06 for the clinical sub-domains

	Asthma	CHD	COPD	Diabetes	Hyper-tension	Hypo-thyroidism	Mental health	Stroke	Overall
Asthma	1.000								
CHD	0.493	1.000							
COPD	0.482	0.561	1.000						
Diabetes	0.469	0.728	0.533	1.000					
Hypertension	0.502	0.707	0.434	0.658	1.000				
Hypothyroidism	0.342	0.474	0.382	0.476	0.370	1.000			
Mental health	0.300	0.294	0.251	0.244	0.280	0.185	1.000		
Stroke	0.477	0.746	0.502	0.653	0.664	0.409	0.277	1.000	
Overall	**0.755**	**0.847**	**0.690**	**0.811**	**0.814**	**0.506**	**0.524**	**0.770**	**1.000**

Note: the number of practices is 8,335. The overall population achievement rate is a weighted average of clinical sub-domain achievement rates, with weights reflecting the number of points available in each sub-domain.

Table 6: The 19 disease sub-domains within the clinical domain, QOF 2006/07 and 2007/08

Disease sub-domain	Number of indicators (including the existence of a disease register): in total	that refer to all patients with the disease	Total points available for all indicators
Asthma	4	2	45
Atrial fibrillation	3	1	30
Cancer	2	1	11
Chronic kidney disease (CKD)	4	2	27
Chronic obstructive pulmonary disease (COPD)	5	4	33
Coronary heart disease (CHD)	10	6	89
Dementia	2	2	20
Depression	2	1	33
Diabetes	16	10	93
Epilepsy	4	1	15
Heart failure	3	1	20
Hypertension	3	2	83
Hypothyroidism	2	2	7
Learning disabilities	1	1	4
Mental health	6	3	39

continued

Table 6: The 19 disease sub-domains within the clinical domain, QOF 2006/07 and 2007/08 – *continued*

Disease sub-domain	Number of indicators (including the existence of a disease register): in total	that refer to all patients with the disease	Total points available for all indicators
Obesity	1	1	8
Palliative care	2	1	6
Smoking	2	1	68
Stroke	8	4	24
Sub-total	80	47	655
Holistic care[5]			20
Total			**675**

Source: British Medical Association/NHS Employers (2006)

Population achievement rates

Table 7 reports descriptive statistics for 10 of the 19 clinical sub-domains for 2006/07 and 2007/08. These are derived from practices' weighted mean population achievement rates for the indicators within each clinical sub-domain, where the weights reflect the maximum number of points available for each indicator.[6] As was the case for 2004/05 and 2005/06, only those clinical indicators that refer to all patients have been used to calculate the population achievement rate for each disease area. Indicators that refer to a subgroup of patients are excluded from our achievement rates and so there are no population achievement rates for nine of the clinical sub-domains because all of their constituent indicators refer to subgroups of patients. Although exception reports are available for 2006/07 and 2007/08, the population weighted achievement rates for these years have been calculated on the same basis as those for 2004/05 (that is, they only use indicators that relate to all patients with a disease, and practice disease-register counts on National Prevalence Day have been used to estimate the number of patients for whom the indicator was relevant before the removal of exception-reported patients). This facilitates the comparison of results from models that use achievements rates for different years.

As table 7 shows, practice population achievement rates varied considerably across the clinical sub-domains (for example, in 2006/07, they ranged from about 72% for mental health to 98% for chronic kidney disease). However, the average achievement rate across each clinical sub-domain remained largely unchanged between 2006/07 and 2007/08.

Table 7: Descriptive statistics for population achievement rates in selected clinical sub-domains, QOF 2006/07 and 2007/08

		Population achievement rate			
QOF variable	Number of practices	Mean	Standard deviation	Min	Max
Asthma 2006/07	8,368	75.73	9.98	0.00	100
Asthma 2007/08	8,289	76.38	9.11	0.00	100
Chronic kidney disease (CKD) 2006/07	8,295	97.87	3.22	0.00	100
CKD 2007/08	8,251	97.68	2.57	50.00	100
Chronic obstructive pulmonary disease (COPD) 2006/07	8,350	80.75	10.08	14.91	100
COPD 2007/08	8,279	81.97	8.61	0.00	100
Coronary heart disease (CHD) 2006/07	8,362	82.38	4.21	17.20	100
CHD 2007/08	8,284	82.74	3.85	35.00	100
Dementia 2006/07	8,272	76.98	17.44	0.00	100
Dementia 2007/08	8,200	75.44	17.48	0.00	100
Diabetes 2006/07	8,366	86.84	6.27	10.98	100
Diabetes 2007/08	8,290	87.42	5.57	9.38	100
Hypertension 2006/07	8,370	91.48	4.25	1.72	100
Hypertension 2007/08	8,292	91.28	4.11	14.87	100
Hypothyroidism 2006/07	8,362	95.41	4.04	0.00	100
Hypothyroidism 2007/08	8,281	95.37	3.78	9.09	100
Mental health 2006/07	8,362	71.60	15.27	0.00	100
Mental health 2007/08	8,284	74.73	13.66	0.00	100
Stroke 2006/07	8,352	85.94	6.08	20.00	100
Stroke 2007/08	8,276	86.17	5.77	6.67	100

Correlation coefficients

Table 8 reports correlation coefficients for practice population achievement rates for each clinical sub-domain. As was the case for 2005/06 rates, although the rates are positively correlated with each other, the correlations are not as high as might have been expected. Again, the mental health achievement rate is the least well correlated with the other sub-domains. The data for 2007/08 reveal a similar pattern of correlations.

Table 8: Correlation coefficients for population achievement rates in 2006/07 for the clinical sub-domains

	Asthma	Chronic kidney disease (CKD)	Chronic obstructive pulmonary disease (COPD)	Coronary heart disease (CHD)	Dementia	Diabetes	Hyper-tension	Hypo-thyroidism	Mental health	Stroke
Asthma	1.000									
Chronic kidney disease (CKD)	0.158	1.000								
Chronic obstructive pulmonary disease (COPD)	0.401	0.127	1.000							
Coronary heart disease (CHD)	0.372	0.200	0.506	1.000						
Dementia	0.271	0.132	0.231	0.216	1.000					
Diabetes	0.404	0.176	0.514	0.561	0.198	1.000				
Hypertension	0.407	0.237	0.375	0.511	0.208	0.503	1.000			
Hypothyroidism	0.274	0.136	0.303	0.381	0.159	0.404	0.430	1.000		
Mental health	0.328	0.086	0.242	0.230	0.314	0.223	0.211	0.105	1.000	
Stroke	0.366	0.221	0.484	0.645	0.245	0.544	0.493	0.358	0.239	1.000

Note: the number of practices is 8,217.

Chapter 4

The dataset and the development of a basic model of hospital costs

This study builds on work undertaken for the Department of Health to estimate the determinants of NHS hospital expenditure on individual patients (Dixon, Bardsley et al 2009). This chapter briefly summarises the dataset to be used, and outlines the development of a basic model that seeks to explain hospital expenditure on individuals in 2007/08. We briefly outline the scope of the model to be estimated and present details of both the estimation sample and the estimation method. In order to make the analysis tractable, we have derived a parsimonious model of expenditure. This base model includes a QOF-based indicator of the quality of primary care. The financial implications of this result for secondary care costs are outlined.

4.1 A model for patient expenditure

At the core of our analysis is a very large dataset, originally prepared for a project commissioned by the Department of Health to develop a Person-Based Resource Allocation (PBRA) formula. The intention was to develop a capitation formula with which to allocate 'fair shares' of funding to general practices, to serve as a basis for practice-based budgets for hospital care (Dixon, Bardsley et al 2009). The data were also used to explore the determinants of hospital expenditure in individual programmes of care (programme budgeting categories).

The dataset links the practice registration of individuals to Hospital Episode Statistics (HES) data over a five-year period.[7] In particular, after costing each HES inpatient spell and each HES outpatient attendance, Dixon's research team could calculate the total hospital cost for each person using NHS secondary care in 2007/08. This patient-level cost information was then merged with patient registration data for everyone registered with an English practice on 1 April 2007.

By linking these two data sources, the research team developed a patient-based model of NHS hospital expenditure based on three broad elements:

- individual characteristics, such as age, sex, and previous use of hospital services, which we denote as needs (individual)
- characteristics of the small area in which the individual lives, likely to be indicators of the need for medical care (such as measures of deprivation), which we denote as needs (attributed)
- characteristics of the small area in which the individual lives, likely to be indicators of the supply of medical care (such as proximity to hospitals), which we denote as supply (attributed).

Variables in the healthcare needs (individual) set

The healthcare needs (individual) set includes variables that reflect:

- The patient's age and gender on 1 April 2007 (38 dummies).
- HES diagnosis data from all of the patient's inpatient episodes in 2005/06 and 2006/07, based on the international classification of diseases 10 (ICD10) categories used by the NHS

Information Centre to summarise the volume of HES activity by diagnosis (152 dummies equal to 1 if the individual had one or more HES encounters, which were assigned the relevant ICD10 category; zero otherwise – see HES online). These variables are intended to capture citizens' previous health status.8

– Four further encounter variables:
  1. the number of inpatient episodes recorded by the patient in 2005/06 and 2006/07
  2. the number of outpatient attendances recorded by the patient in 2005/06 and 2006/07
  3. whether the patient had a priority referral to an outpatients department in 2005/06 or 2006/07
  4. whether the patient had received treatment in the course of an outpatient attendance during 2005/06 or 2006/07.

  The rationale for the use of these 'encounter' variables as indicators of need is that the number of times an individual has been admitted to hospital or attended an outpatients department in the past conveys something about the intensity of their morbidity experience, over and above the information contained in the binary ICD10 morbidity variables. The morbidity categories, while very powerful as explanatory variables, will not reflect repeated encounters falling in the same ICD10 category. Also, the use of morbidity categories on their own assumes that the presence of two different diagnoses is additive, and may not capture the possible non-additive effects of co-morbidity. The encounter variables may go some way to rectifying this.
– Two individual needs variables based on:
  1. whether the patient had a private inpatient spell in 2005/06 or 2006/07
  2. whether the patient had a private outpatient attendance in 2005/06 or 2006/07.

We include dummy variables for whether the patient had a private episode with an NHS provider and whether they had a private outpatient attendance with an NHS provider. The expectation is that individuals who have been private patients in the past are more likely to use private provision in the future if ill, and will therefore generate less NHS expenditure if ill. As the aim is to model NHS expenditure, we treat these private care variables as measures of need for NHS expenditure.

Variables in the healthcare needs (attributed) set

The healthcare needs (attributed) set includes over 160 variables that reflect:

- the socio-economic characteristics of the small area (Lower Layer Super Output Area – LSOA) in which the patient lives on 1 April 2007
- QOF data on disease prevalence rates for 2004/05 and 2005/06 for various conditions for the practice with which the patient is registered on 1 April 2007.

Variables in the healthcare supply (attributed) set

The healthcare supply (attributed) set includes over 130 variables that reflect:

– QOF population achievement scores for 2004/05 and 2005/06 for the practice with which the patient is registered. For each clinical area, we use a weighted average of the population achievement scores for those indicators that refer to all patients (with weights reflecting the maximum number of points available for each indicator). Indicators that refer to a subgroup of patients are excluded from our achievement rates. We use an overall QOF achievement score calculated on this basis, and a score for each of eight clinical areas (for example, for asthma, diabetes, stroke, etc)
– the distance from the LSOA in which the patient lives to local providers and various measures of waiting time
– practice characteristics (from General Medical Services [GMS] data) for the practice with which the patient is registered
– distance and population weighted measures of the accessibility of various healthcare facilities from the LSOA in which the patient lives.

PCT dummies

The model also includes 152 PCT dummies. These were introduced to pick up unobserved factors that vary across PCTs in order to prevent them from biasing the coefficients on other variables. PCT dummies will primarily reflect unobserved supply factors because we include a large number of need variables in our models. The PCT dummies are

likely to capture factors such as PCT variations in past spending levels, different provider diagnosis recording practices, and different referral or treatment thresholds at different locations.

The timing of the variables

The dependent variable to be used refers to total inpatient and outpatient costs (excluding maternity and mental health costs) for 2007/08. The needs and supply variables refer to previous years (for example, 2005/06 and 2006/07). Further details of the attributed needs and attributed supply variables can be found in Dixon, Bardsley et al's report (2009).

4.2 The estimation sample

In principle, our model could be estimated across all English citizens registered with an English practice on 1 April 2007. However, with over 50 million patients, such a model would require a prohibitively large amount of computer memory and would take an excessive amount of time to estimate (bearing in mind that hundreds of model variants were estimated). Therefore, to estimate this base model and the other models used in this study, we use a 10% random sample of all patients registered with an English practice with a list size of at least 1,000 patients in 2004, in 2005 and in 2006 on 1 April 2007.

4.3 The estimation method

Hospital cost data for individuals have a spike at zero and a long right-hand tail, and are heteroskedastic.[9] These characteristics can make ordinary least squares (OLS) regression methods inappropriate. Alternatives to OLS include generalised linear models (GLM) and transformed ordinary least squares where the dependent variable is transformed by, for example, taking its square root to help deal with skewness. Two-part estimation separately models the probability of a positive cost and the level of cost for individuals with positive cost. This latter stage can again be modelled using all the estimators available for one-part expenditure models, including OLS, GLM and transformations of the cost variable.

The performance of alternative estimators was investigated as part of the PBRA project and it was found that, because of the large sample size, a one-part OLS model of untransformed expenditure was preferable in terms of predictive power and ease of use. This is in line with findings from other researchers using very large samples. We have therefore used the one-part OLS estimator in this study.

4.4 Derivation of a parsimonious model for hospital expenditure

In developing the base model, all of the personal needs, attributed needs and attributed supply variables (over 500 in total) were initially included in the estimated model. However, most of these variables were statistically insignificant in this 'full' model, which is also difficult to interpret. We therefore embarked on a search for a much smaller subset of variables that were statistically significant, had intuitively plausible coefficients, and provided almost as much explanatory power (for example, in terms of the model R-squared) as all the variables. We called the resulting model the 'parsimonious' model. It retained all of the individual needs variables and the PCT dummy variables, but only a small subset of attributed (small area) variables.

Through the repeated process of dropping the least significant attributed variables and re-estimating the model, we could identify a parsimonious model that included just 10 of the 300 attributed variables.[10]

Attributed needs variables

The seven attributed needs variables retained in the parsimonious model are:

- people in social rented housing as a proportion of all people living in the LSOA
- all disability living allowance claimants as a proportion of all people living in the area (LSOA)
- people aged 16–74 with no qualifications – age standardised
- dummy for people living an area assigned to ONS subgroup 2.3b (mature city professionals)

– students as a proportion of all people living in the area
– whether the person had a private inpatient episode finishing in 2005/06 or 2006/07
– whether the person had a private outpatient attendance in 2005/06 or 2006/07.

Supply variables

The three supply variables retained in the parsimonious model are:
– the 2005/06 QOF stroke weighted population achievement for the practice with which the patient is registered
– a measure relating to distance and population weighted access to MRI scanners for the LSOA in which the patient lives
– a measure relating to distance and population weighted access to residential home beds for the LSOA in which the patient lives.

The coefficients on the age and gender dummies, on the encounter variables and on the 10 attributed needs and attributed supply variables are shown under Model 1 in table 9 (below – the coefficients on the 152 ICD10 morbidity markers and the 152 PCT dummies are not shown). The coefficients on the age and gender dummies exhibit the familiar pattern (increasing with each age group from age 25 onwards). The coefficients on the dummies for having private care in an NHS provider are negative. This is plausible since individuals who have been private patients in the past are more likely to use private facilities in the future if ill and therefore to generate less NHS expenditure. The other attributed needs variables reflect a positive association between deprivation (need) and expenditure.

Attributed supply variables

There are three attributed supply variables in the model. Access to beds in residential homes has a negative sign. It is likely that this partly reflects the fact that residential homes provide a substitute for NHS care, and can reduce the need for patients to use hospital services. The positive coefficient on the MRI scanner access variable reflects a supply effect (more access leads to higher use).

Comparing the stroke QOF score with other clinical QOF scores

Of particular interest for this study is the coefficient on the QOF stroke weighted population achievement for 2005/06 (–0.443), which was the only QOF variable retained in this general parsimonious model of hospital costs. The coefficient implies that a one-point increase in the stroke achievement rate will be associated with a fall of £0.44 per person in hospital costs.[11] With a population of 50 million people, this implies that a one-point increase in the mean stroke population achievement rate from 79.64 to 80.64 would be associated with a reduction in annual total hospital costs of £22.15 million. Although this is a modest sum when compared with the total secondary care spend (in 2007/08 this was about £22 billion, excluding expenditure on maternity and mental health), it is consistent with the claim that improvements in the quality of primary care are associated with reductions in the cost of secondary care. It may also be plausible to envisage an improvement of more than one point in the QOF achievement score: the increase was about 10 points over the three years up to 2007/08.

Because of the focus of this study on the QOF, we re-estimated this general model, replacing the QOF stroke achievement rate with the achievement rate for each of the other seven clinical sub-domains. Although these QOF scores had the anticipated negative coefficient in all seven models, none of them was statistically significant. We also re-estimated the parsimonious model with the overall clinical QOF population achievement rate replacing the stroke score.[12] The coefficient on this all clinical domains variable (–0.422) was slightly smaller than that on the stroke coefficient (–0.443). Also, the standard error was larger so that, although the stroke score was significant at the 1% level, the all clinical domain score was only significant at the 5% level. The close association between results for the stroke QOF score and the overall QOF score may indicate that stroke achievement reflects general primary care quality rather than quality only in stroke care.

Table 9: Ordinary least squares (OLS) models illustrating the impact of QOF scores on patient hospital costs, 2007/08

	Dependent variable				
	Model 1 all costs £, 2007/08	**Model 2 all costs £, 2007/08**	**Model 3 all costs £, 2007/08**	**Model 4 all costs £, 2007/08**	**Model 5 all costs £, 2007/08**
Regressors in model					
Males aged <1	197.405*** [17.145]	197.432*** [17.148]	197.159*** [17.172]	195.723*** [17.114]	197.419*** [17.287]
Males aged 1–4	–2.425 [5.096]	–2.414 [5.095]	–2.587 [5.090]	–2.15 [5.139]	–2.405 [5.080]
Males aged 5–9	–20.204*** [4.031]	–20.203*** [4.032]	–20.270*** [4.030]	–20.154*** [4.025]	–20.949*** [3.789]
Males aged 10–14	–2.010 [3.091]	–2.004 [3.091]	–1.858 [3.081]	–1.822 [3.104]	–2.121 [3.087]
Males aged 15–19	–9.382** [3.647]	–9.355** [3.650]	–9.454** [3.651]	–9.126** [3.677]	–9.816*** [3.654]
Males aged 20–24	–11.603*** [3.665]	–11.590*** [3.665]	–12.389*** [3.617]	–11.511*** [3.702]	–12.602*** [3.648]
Males aged 25–29	–18.092*** [3.396]	–18.125*** [3.395]	–18.077*** [3.381]	–17.797*** [3.383]	–18.198*** [3.380]
Males aged 30–34	–14.130*** [3.390]	–14.120*** [3.391]	–14.182*** [3.375]	–13.753*** [3.433]	–14.297*** [3.352]
Males aged 40–44	17.561*** [3.874]	17.559*** [3.874]	17.462*** [3.880]	17.675*** [3.894]	17.037*** [3.874]
Males aged 45–49	52.436*** [4.541]	52.444*** [4.543]	51.924*** [4.553]	52.753*** [4.585]	51.812*** [4.568]
Males aged 50–54	97.770*** [5.941]	97.794*** [5.942]	97.817*** [5.963]	97.340*** [5.918]	97.534*** [5.971]
Males aged 55–59	158.660*** [5.467]	158.681*** [5.470]	158.994*** [5.480]	158.162*** [5.479]	158.800*** [5.513]
Males aged 60–64	257.169*** [7.770]	257.182*** [7.772]	257.183*** [7.782]	257.239*** [7.782]	256.885*** [7.790]
Males aged 65–69	373.299*** [10.320]	373.327*** [10.322]	373.501*** [10.308]	373.300*** [10.373]	372.892*** [10.292]
Males aged 70–74	545.785*** [10.605]	545.760*** [10.606]	544.035*** [10.592]	545.521*** [10.546]	543.605*** [10.572]
Males aged 75–79	667.262*** [17.444]	667.277*** [17.444]	665.744*** [17.591]	664.473*** [17.283]	665.427*** [17.597]
Males aged 80–84	848.752*** [21.091]	848.772*** [21.094]	849.779*** [21.240]	850.407*** [21.249]	849.446*** [21.208]
Males aged >84	1,053.998*** [25.805]	1,054.029*** [25.805]	1,054.739*** [25.962]	1,053.985*** [25.864]	1,054.627*** [26.048]
Females aged <1	138.460*** [15.309]	138.526*** [15.313]	138.057*** [15.324]	139.045*** [15.385]	138.289*** [15.358]
Females aged 1–4	–25.499*** [4.491]	–25.483*** [4.491]	–25.710*** [4.496]	–25.648*** [4.527]	–25.905*** [4.524]
Females aged 5–9	–26.529*** [3.215]	–26.491*** [3.216]	–26.575*** [3.206]	–26.316*** [3.228]	–26.827*** [3.196]

continued

Table 9: Ordinary least squares (OLS) models illustrating the impact of QOF scores on patient hospital costs, 2007/08 – *continued*

	Dependent variable				
	Model 1 all costs £, 2007/08	**Model 2 all costs £, 2007/08**	**Model 3 all costs £, 2007/08**	**Model 4 all costs £, 2007/08**	**Model 5 all costs £, 2007/08**
Females aged 10–14	–3.746 [3.781]	–3.720 [3.783]	–3.812 [3.790]	–3.33 [3.780]	–3.985 [3.801]
Females aged 15–19	–5.850* [3.069]	–5.853* [3.070]	–5.888* [3.074]	–5.668* [3.114]	–6.266** [3.074]
Females aged 20–24	–1.949 [3.528]	–1.974 [3.530]	–2.097 [3.547]	–1.618 [3.611]	–2.239 [3.564]
Females aged 25–29	6.820* [3.558]	6.792* [3.557]	6.751* [3.567]	7.063* [3.625]	6.534* [3.594]
Females aged 30–34	27.721*** [3.761]	27.710*** [3.763]	27.692*** [3.774]	27.994*** [3.803]	27.344*** [3.796]
Females aged 35–39	42.508*** [3.716]	42.527*** [3.717]	42.395*** [3.722]	42.719*** [3.745]	42.123*** [3.740]
Females aged 40–44	55.601*** [4.116]	55.614*** [4.118]	55.531*** [4.119]	55.919*** [4.149]	55.333*** [4.127]
Females aged 45–49	83.970*** [5.247]	83.946*** [5.248]	83.746*** [5.254]	84.300*** [5.277]	83.608*** [5.256]
Females aged 50–54	116.634*** [5.478]	116.647*** [5.480]	116.837*** [5.487]	117.193*** [5.511]	116.499*** [5.475]
Females aged 55–59	146.756*** [5.595]	146.774*** [5.597]	146.370*** [5.611]	146.601*** [5.576]	145.882*** [5.613]
Females aged 60–64	211.030*** [7.273]	211.048*** [7.276]	210.937*** [7.337]	210.972*** [7.251]	210.316*** [7.340]
Females aged 65–69	305.452*** [8.906]	305.466*** [8.905]	305.400*** [8.885]	306.224*** [8.901]	305.235*** [8.924]
Females aged 70–74	442.360*** [11.034]	442.388*** [11.036]	442.814*** [11.040]	441.999*** [10.977]	442.255*** [11.018]
Females aged 75–79	577.492*** [12.487]	577.518*** [12.488]	577.676*** [12.395]	578.011*** [12.491]	577.341*** [12.413]
Females aged 80–84	730.881*** [13.831]	730.898*** [13.830]	730.020*** [13.816]	730.531*** [13.724]	730.523*** [13.866]
Females aged >84	985.362*** [19.523]	985.377*** [19.528]	985.262*** [19.485]	985.151*** [19.528]	984.768*** [19.499]
In social rented housing	0.279*** [0.093]	0.279*** [0.093]	0.279*** [0.093]	0.344*** [0.102]	0.342*** [0.101]
Disability living allowance	337.035*** [75.432]	336.458*** [75.363]	339.079*** [75.539]	295.695*** [79.302]	292.772*** [79.592]
No qualifications: age standardised	23.969*** [4.797]	23.942*** [4.783]	24.143*** [4.790]	22.507*** [4.839]	23.136*** [4.836]
ONS 15: mature city professionals	–23.822*** [6.993]	–23.836*** [6.944]	–23.794*** [7.024]	–23.211*** [6.957]	–22.845*** [7.027]
Students in population	–1,319.462*** [141.400]	–1,313.195*** [141.811]	–1,297.098*** [139.034]	–1,271.58*** [141.288]	–1,254.63*** [138.773]

continued

Table 9: Ordinary least squares (OLS) models illustrating the impact of QOF scores on patient hospital costs, 2007/08 – *continued*

	Dependent variable				
	Model 1 all costs £, 2007/08	**Model 2 all costs £, 2007/08**	**Model 3 all costs £, 2007/08**	**Model 4 all costs £, 2007/08**	**Model 5 all costs £, 2007/08**
Private episode 2005/06/07	–490.588*** [24.778]	–490.505*** [24.785]	–490.674*** [24.728]	–491.942*** [24.677]	–491.478*** [24.625]
Number of episodes 2005/06/07	299.068*** [9.894]	299.071*** [9.894]	298.948*** [9.900]	298.842*** [9.907]	298.628*** [9.938]
Number of attendances 2005/06/07	45.696*** [2.199]	45.697*** [2.199]	45.690*** [2.202]	45.721*** [2.209]	45.712*** [2.211]
Outpatient priority referral 2005/06/07	70.867*** [10.607]	70.889*** [10.606]	70.865*** [10.591]	70.717*** [10.625]	70.965*** [10.590]
Outpatient treatment 2005/06/07	55.527*** [12.393]	55.552*** [12.396]	55.710*** [12.398]	55.740*** [12.423]	55.845*** [12.415]
Private attendance 2005/06/07	–167.142*** [25.073]	–167.152*** [25.085]	–167.002*** [25.184]	–168.710*** [25.045]	–168.255*** [25.153]
QOF stroke score 2005/06	–0.443*** [0.141]				
QOF stroke score 2006/07	–0.555*** [0.164]				
QOF stroke score 2007/08		–0.712*** [0.182]		–0.664*** [0.190]	
QOF asthma score 2006/07			–0.364*** [0.096]		
MRI scanner access	5,436.256** [2,245.926]	5,456.198** [2,244.584]	5,470.898** [2,251.048]		
Residential home beds access	–7.850*** [2.404]	–7.940*** [2.399]	–8.047*** [2.383]		
Constant	128.858*** [18.027]	142.632*** [18.685]	157.214*** [18.368]	98.522*** [8.135]	130.476*** [17.584]
Observations	5,206,651	5,205,882	5,188,594	5,181,099	5,170,603
R-squared	0.266	0.266	0.266	0.266	0.265
Adj R-squared	0.266	0.266	0.266	0.266	0.265

Robust standard errors are in brackets with clustering by PCT.
* p<0.1, ** p<0.05, *** p<0.01
Notes:
1. Model 1 incorporates QOF scores for 2005/06 in the full model and is the original parsimonious model, in which stroke achievement is selected.
2. Model 2 re-estimates Model 1, replacing the QOF stroke score for 2005/06 with that for 2006/07.
3. Model 3 re-estimates Model 1, replacing the QOF stroke score for 2005/06 with that for 2007/08.
4. Model 4 incorporates QOF scores for 2006/07 in the full model and is the resulting parsimonious version.
5. Model 5 incorporates QOF scores for 2007/08 in the full model and is the resulting parsimonious version.
6. The dependent variable is the patient's hospital costs for 2007/08, excluding maternity and mental health costs.

Chapter 5

Variants of the base model, and further analysis

In this chapter, we build on the basic model described in chapter 4 to explore in more detail four issues associated with the findings:

1. We examine whether the use of more recent QOF scores (for example, for 2006/07 and 2007/08) might reveal a larger quantitative impact of quality on costs for 2007/08 (the PBRA study only had access to QOF quality scores for 2004/05 and 2005/06).
2. With the availability of a binary indicator that records whether the patient died in 2007/08, we can derive a parsimonious model for the probability of death, and we examine whether there is any association between the QOF quality scores and the probability of death.
3. We estimate parsimonious models for expenditure by programme budget (PB) category (where total hospital costs are split between 23 disease areas) in order to examine whether this reveals further evidence of the impact of quality in primary care on costs in secondary care.
4. Finally, to obtain more accurate parameter estimates, we re-estimate the parsimonious expenditure model using a panel dataset. This enables us to control for time-invariant unobserved factors that are correlated with both quality and cost, but whose influence would otherwise be attributed to quality in the one-period, cross-section models.

5.1 Updating the QOF stroke achievement score

We examine the implications of updating the QOF stroke score in three ways:

1. We retain the existing parsimonious model and re-estimate it with an updated QOF score.
2. We then derive new parsimonious models, with updated QOF scores (for 2006/07 and then for 2007/08) available in the full model.
3. Finally, we derive a further parsimonious model from a full model that includes QOF scores for all three years (for 2005/06, 2006/07 and 2007/08).

Replacing the QOF stroke score for 2005/06 with a more recent score

Although the derivation of the parsimonious model of hospital expenditure incorporated over 300 attributed variables in the full model (see Model 1 in table 9), the available QOF quality scores were at least two years behind the dependent variable. Costs were for 2007/08, and the available QOF quality scores were for 2004/05 and 2005/06. However, since the derivation of that model, QOF quality scores for 2006/07 and 2007/08 have become available, and we therefore re-estimated the parsimonious model (see Model 1 in table 9), replacing the practice QOF stroke population achievement score for 2005/06 with that for 2006/07 (see Model 2 in table 9), and then with that for 2007/08 (see Model 3 in table 9).

All three models have the same explanatory power in terms of the model R-squared. However, the coefficient on the QOF stroke population achievement score increases from –0.443 in the model with the QOF score for 2005/06 (Model 1), to –0.555 with the QOF score for 2006/07 (Model 2), and then to –0.712 in the model with the QOF score for 2007/08 (Model 3). As is to be expected, the association between QOF quality and cost reductions becomes stronger as the time period over which the measure of quality is recorded moves closer to the period in which the costs are incurred.[13]

The largest coefficient (in absolute terms) on the QOF stroke score is –0.712 for 2007/08. With a population of 50 million people, this implies that a one-percentage-point increase in the QOF score for 2007/08 is associated with a reduction in total secondary care costs for the same period of £35.6 million. This is a 60% increase on the saving implied by the coefficient (–0.443) on the QOF stroke score for 2005/06.

Deriving new parsimonious models with QOF quality scores for 2006/07 and then for 2007/08 in the full model

Because many of the attributed needs/supply variables are highly correlated with each other, it is possible that, if updated QOF scores for 2006/07 or 2007/08 had been available for the derivation of the parsimonious model, a different set of attributed needs and supply variables would have been selected to those chosen with only the availability of the QOF scores for 2005/06. Therefore, two new parsimonious models were derived: one with QOF quality scores for 2006/07 in the full model, and the other with QOF quality scores for 2007/08 in the full model. The coefficients on the age and gender variables, the four encounter variables, and the attributed needs/supply variables in the derived parsimonious models, are shown as Models 4 and 5 in table 9.

The derivation of a parsimonious model from a full model that includes the QOF quality scores for 2006/07 results in very few qualitative changes, but in the inclusion of a QOF quality score for asthma rather than stroke care (Model 4). The coefficient on this variable is –0.364. The derivation of a parsimonious model from a full model that includes the QOF quality scores for 2007/08 results in the inclusion of the QOF stroke score with a coefficient of –0.664 (Model 5). This coefficient on the QOF stroke score for 2007/08 implies that a one-percentage-point increase in this score is associated with an annual reduction in total secondary care costs of £33.2 million (=50 million × £0.664). This is a similar result to that obtained when the QOF stroke achievement rate for 2005/06 was replaced with the achievement rate for 2007/08 in the original parsimonious model (see Model 3 in table 9).

In the corresponding full model for Model 5 in table 9, the coefficient on the QOF stroke score for 2007/08 is –0.583. In the parsimonious version of this model – which excludes all other clinical QOF scores – this coefficient increases in absolute size to –0.664. One interpretation of the increase in this coefficient (as other QOF scores are removed from the model) is that, in the parsimonious model, the stroke QOF score reflects the quality of primary care provided in more than just the stroke domain.

Deriving a new parsimonious model with QOF quality scores for 2005/06, 2006/07 and 2007/08 in the full model

So far, we have examined the association between hospital costs for 2007/08 and QOF scores for a single year (that is, either for 2005/06, 2006/07 or 2007/08).

However, hospital costs in 2007/08 might be associated with the quality of primary care in more than one year. We therefore derived a parsimonious model for hospital costs with QOF scores for the clinical sub-domains for three years (2005/06, 2006/07 and 2007/08) included in the full model.

The usual estimation procedure – of repeated estimation, dropping the least significant regressors and re-estimation – generates the result shown as Model 6 in table 10 (the coefficients on the 152 international classification of diseases (ICD) 10 morbidity markers, the 152 PCT dummies and the age/sex dummies are not shown). Here, two QOF variables have a significant negative association with costs in 2007/08: the asthma quality score for 2006/07 and the stroke quality score for 2007/08. The asthma achievement rate was present in the parsimonious model when the full model included all quality scores but only for 2006/07 (see Model 4 in table 9), and the stroke achievement rate was present in the parsimonious model when the full model included all quality scores but only for 2007/08 (see Model 5 in table 9).

If Model 6 is re-estimated without the stroke achievement rate (which is not significant at the 1% level), the coefficient on the asthma score increases in absolute size, from –0.275 to –0.360 (see Model 7 in table 10). Similarly, if Model 6 is re-estimated

without the asthma achievement rate, the coefficient on the stroke score also increases in absolute size, from –0.506 to –0.680 (see Model 8 in table 10), and this variable is now significant at the 1% level. The implication of this is that the asthma and stroke scores are detecting similar effects, which may relate to the general quality of primary care rather than the quality of care for specific conditions.

Table 10: Ordinary least squares (OLS) models illustrating the impact of QOF scores for three years on patient hospital costs, 2007/08

	Dependent variable		
Regressors	**Model 6 all costs £, 2007/08**	**Model 7 all costs £, 2007/08**	**Model 8 all costs £, 2007/08**
Number of episodes 2005/06/07	298.459*** [9.952]	298.459*** [9.952]	298.459*** [9.952]
Number of attendances 2005/06/07	45.710*** [2.216]	45.711*** [2.216]	45.712*** [2.216]
Outpatient priority referral 2005/06/07	70.732*** [10.613]	70.723*** [10.615]	70.739*** [10.615]
Outpatient treatment 2005/06/07	56.164*** [12.440]	56.125*** [12.436]	56.171*** [12.441]
In social rented housing	0.327*** [0.099]	0.325*** [0.099]	0.328*** [0.099]
Disability living allowance	311.591***	314.407***	312.357***
	[76.641]	[76.475]	[76.684]
No qualifications: age standardised	22.578*** [4.814]	22.698*** [4.813]	22.550*** [4.831]
QOF asthma score 2006/07	–0.275*** [0.104]	–0.360*** [0.096]	
QOF stroke score 2007/08	–0.506** [0.207]		–0.680*** [0.191]
ONS 15: mature city professionals	–22.090*** [7.070]	–22.119*** [7.107]	–21.913*** [7.048]
Students in population	–1,328.004*** [140.068]	–1,331.920*** [140.039]	–1,328.288*** [140.145]
Private episode 2005/06/07	–491.493*** [24.658]	–491.467*** [24.657]	–491.490*** [24.657]
Private attendance 2005/06/07	–168.081*** [25.170]	–168.117*** [25.168]	–168.029*** [25.161]
Residential home beds access	–5.839*** [2.138]	–5.765*** [2.151]	–5.772*** [2.156]
Constant	169.988*** [18.978]	130.694*** [14.578]	165.164*** [19.068]
Observations	5,153,437	5,153,437	5,153,437
R-squared	0.266	0.266	0.266
Adj R-squared	0.266	0.266	0.266

Robust standard errors are in brackets with clustering by PCT.
* p<0.1, ** p<0.05, *** p<0.01
Note: The dependent variable is the patient's hospital costs for 2007/08, excluding maternity and mental health costs.

5.2 Deriving a new parsimonious model with a binary (patient died) dependent variable

The regression models in tables 9 and 10 have been estimated across a 10% sample of patients registered with each English practice on 1 April 2007. The dependent variable is the person's total hospital costs for 2007/08. This averaged about £427 across the estimation sample of five million patients.

In addition to the total hospital cost incurred by each patient on the practice list, we also know whether the patient died in 2007/08. We can use this information to construct a binary variable that takes a value of 1 if the patient died in 2007/08 or a value of 0 otherwise.[14] If this binary indicator replaces hospital expenditure as the dependent variable, the logit estimator can be used to derive a parsimonious model that reveals which variables significantly affect the probability of death.[15]

If QOF population achievement rates for 2005/06 are included in the full model, the derived parsimonious logit model is shown as Model 9 in table 11. This reveals the coefficients on the four encounter variables, and the 12 significant attributed needs and attributed supply variables. One of the 12 significant attributed supply variables is the QOF quality score for stroke care in 2005/06 and, although the coefficient on this variable is small, its negative coefficient implies that better quality primary care is associated with a reduced probability of death.

The derivation of the parsimonious model shown as Model 9 in table 11 included the QOF quality scores for 2005/06 in the full model. Model 10 in table 11 is the parsimonious model derived with QOF quality scores for 2006/07 in the full model, and Model 11 is the parsimonious model derived with QOF quality scores for 2007/08 in the full model. The coefficient on the QOF quality score increases (in absolute size) as the QOF achievement rate is updated: the coefficient increases from –0.003 for 2005/06 to –0.004 for 2006/07, and then to –0.007 for 2007/08. This implies that the strength of the negative association between primary care quality and the probability of death increases as the time period over which quality is measured moves closer to the period during which death is recorded. This finding is very similar to what we found in relation to the impact of quality on cost: the association of quality with costs strengthens as the time period over which the measure of quality is recorded moves closer to the period over which the costs are incurred.

The interpretation of the coefficient estimates in a logit model is more complex than in an ordinary linear regression. In particular, the coefficients in a logit model no longer reflect the impact on the dependent variable of a one-unit increase in a regressor. However, these marginal effects are routinely provided by most econometric software packages.[16] The average marginal effect of the QOF stroke score on the probability of death increases from –0.0000241 in Model 9 (with the score for 2005/06), to –0.0000425 in Model 10 (with the score for 2006/07), and increases further to –0.0000477 in Model 11 (with the score for 2007/08).[17] With a population of 50 million people, this implies that a one-unit increase in the stroke QOF score for 2007/08 would be associated with 2,385 (=50m × 0.0000477) fewer deaths in that year.

Table 11: Logit models illustrating the impact of QOF scores on the probability of patient death, 2007/08

	Dependent variable		
Regressors	**Model 9 lived/died, 2007/08**	**Model 10 lived/died, 2007/08**	**Model 11 lived/died, 2007/08**
Number of episodes 2005/06/07	0.006*** [0.001]	0.006*** [0.001]	0.006*** [0.001]
Number of attendances 2005/06/07	0.014*** [0.002]	0.014*** [0.002]	0.014*** [0.002]
Outpatient priority referral 2005/06/07	0.159*** [0.014]	0.159*** [0.015]	0.159*** [0.015]
Outpatient treatment 2005/06/07	–0.022 [0.026]	–0.021 [0.026]	–0.020 [0.026]
Indices of Deprivation (ID) 2007: difficulty of access to owner occupation	–0.006*** [0.001]	–0.006*** [0.001]	–0.007*** [0.001]
Indices of Deprivation (ID) 2007: years of potential life lost	0.003*** [0.001]	0.004*** [0.001]	0.003*** [0.001]
Proportion of residents in communal establishments	1.022*** [0.126]	1.004*** [0.126]	0.943*** [0.124]
Indices of Multiple Deprivation (IMD) 2004: income deprivation affecting older children	1.054*** [0.114]	1.241*** [0.117]	1.187*** [0.114]
Access to acute provider capacity (with deterrence function)		–0.037*** [0.012]	
Epilepsy prevalence rate 2006	0.181*** [0.045]	0.174*** [0.045]	0.184*** [0.044]
Dementia prevalence rate 2006	0.158*** [0.022]	0.171*** [0.022]	0.153*** [0.022]
Dispensing practice 2005	0.064*** [0.014]	0.048*** [0.015]	0.061*** [0.014]
ONS 47: resorts and retirement	–0.129*** [0.036]	–0.101*** [0.037]	–0.124*** [0.037]
ONS 51: urban terracing		0.150*** [0.053]	
Private attendance 2005/06/07	0.232*** [0.088]	0.233*** [0.087]	0.236*** [0.088]
Proportion of population widowed	1.349*** [0.190]		1.109*** [0.183]
Number of children <1yr per 1,000 women	0.004*** [0.001]		
QOF stroke score 2005/06	–0.003*** [0.001]		
QOF stroke score 2006/07		–0.006*** [0.001]	
QOF stroke score 2007/08			–0.007*** [0.001]

continued

Table 11: Logit models illustrating the impact of QOF scores on the probability of patient death, 2007/08 – *continued*

	Dependent variable		
Regressors	**Model 9 lived/died, 2007/08**	**Model 10 lived/died, 2007/08**	**Model 11 lived/died, 2007/08**
Constant	–7.243*** [0.152]	–6.915*** [0.161]	–6.835*** [0.165]
Observations	5,206,636	5,181,084	5,170,588

Robust standard errors are in brackets with clustering by PCT.
* p<0.1, ** p<0.05, *** p<0.01
Notes:
1. Model 9 incorporates QOF scores for 2005/06 in the full model and is the resulting parsimonious version.
2. Model 10 incorporates QOF scores for 2006/07 in the full model and is the resulting parsimonious version.
3. Model 11 incorporates QOF scores for 2007/08 in the full model and is the resulting parsimonious version.

5.3 Deriving parsimonious models for individual disease areas

The regression models in tables 9, 10 and 11 have been estimated across a 10% sample of patients registered with each English practice on 1 April 2007. The dependent variable has been either each person's total hospital costs for 2007/08 (see tables 9 and 10) or a binary variable indicating whether the person died in 2007/08 (see table 11). In all of these results, the QOF scores have been for particular disease areas (for example, for asthma, diabetes, stroke, etc). In other words, we have been estimating models that relate total costs across all disease areas (or death due to all causes) to QOF scores in specific disease areas.

Although we cannot disaggregate the death data by cause of death, it is possible to disaggregate the cost data into 23 care programmes (as shown in table 12). This enables us to derive a parsimonious model for costs in each individual care programme, and to test whether this division of total costs leads to the identification of stronger or weaker relationships between costs in secondary care and the quality of primary care.

Considering the total cost variable for each patient

The construction of the total cost variable for each patient involved the costing of each hospital inpatient episode and each hospital outpatient visit. This, in turn, involved attaching a unit cost to the most resource-intensive episode in each inpatient spell and attaching a cost to each outpatient attendance. Each inpatient spell includes a primary diagnosis code and, to facilitate the allocation of costs between the 23 care programmes, the Department of Health provided us with a mapping from this diagnosis code to programme budget (PB) category.[18]

There is little diagnosis information in the HES outpatient data. However, each outpatient attendance includes details of the consultant's specialty, and the Department of Health provided us with a mapping from outpatient specialty to PB category. We applied this mapping to the treatment specialty (or to the main specialty if the treatment specialty was missing), and this enabled us to assign each attendance and its cost to a PB category. By aggregating over spells and attendances in the same PB category for each individual, we obtained each individual's total hospital cost by care programme for 2007/08.

Table 12 shows the number of patients with a non-zero total (inpatient plus outpatient) cost by each of the 23 PB categories for the five million patients in our patient sample.[19] In none of the 23 categories do more than 10% of patients incur any expenditure. The PB category with the highest proportion of patients with some expenditure is PB category 23: Other. PB 23 acts as a residual category, particularly for outpatient costs, where the lack of diagnostic information makes it impossible to allocate some specialties (for

example, for general medicine) to a specific PB category. The remaining 22 PB categories can be divided into two groups:

1. nine categories:
   - cancer
   - neurology
   - vision
   - circulation
   - gastro-intestinal
   - skin
   - musculo-skeletal
   - trauma
   - and genito-urinary

   with between 4.7% and 6.7% of patients having non-zero costs
2. the remaining 13 categories, where fewer than 2% of patients record a non-zero cost.

These percentages are small relative to the percentage of patients (about 35%) that incur some inpatient or outpatient cost for at least one PB category. However, Martin, Dusheiko et al (2009) report that, despite the low proportions of non-zero expenditure, OLS remains the preferred estimation procedure, and we persevere with this estimation method here.

Table 12: Hospital costs in 2007/08 by programme budget (PB) category for the estimation sample

PBC	PBC description	Patients with non-zero costs		Total cost	Cost per patient with non-zero cost	Cost per patient in sample
		Number	%	£	£	£
1	Infectious diseases	10,542	0.20	15,553,682	1,475	2.99
2	Cancers and tumours	248,250	4.77	236,327,104	952	45.39
3	Disorders of blood	52,241	1.00	58,734,232	1,124	11.28
4	Endocrine, nutritional and metabolic	61,734	1.19	43,148,560	699	8.29
5	Mental health disorders	62,459	1.20	183,125,712	2,932	35.17
6	Problems of learning disability	2,404	0.05	6,030,665	2,509	1.16
7	Neurological	303,499	5.83	150,566,656	496	28.92
8	Problems of vision	245,493	4.71	78,373,040	319	15.05
9	Problems of hearing	25,084	0.48	12,764,938	509	2.45
10	Problems of circulation	295,215	5.67	272,416,928	923	52.32
11	Problems of the respiratory system	95,117	1.83	136,666,128	1,437	26.25
12	Dental problems	54,362	1.04	26,463,094	487	5.08
13	Problems of gastrointestinal system	312,569	6.00	226,994,080	726	43.60
14	Problems of the skin	291,564	5.60	76,618,104	263	14.72
15	Problems of musculoskeletal system	344,295	6.61	230,012,832	668	44.18
16	Problems due to trauma and injuries	308,233	5.92	160,722,784	521	30.87

continued

Table 12: Hospital costs in 2007/08 by programme budget (PB) category for the estimation sample – *continued*

PBC	PBC description	Patients with non-zero costs		Total cost	Cost per patient with non-zero cost	Cost per patient in sample
		Number	%	£	£	£
17	Problems of genito urinary system	268,344	5.15	238,556,320	889	45.82
18	Maternity and reproductive health	104,546	2.01	159,933,488	1,530	30.72
19	Conditions of neonates	574	0.01	762,148	1328	0.15
20	Adverse effects and poisoning	23,048	0.44	54,878,328	2,381	10.54
21	Healthy Individuals	8,521	0.16	8,001,166	939	1.54
22	Social care needs	776	0.01	3,515,645	4,530	0.68
23	Other	474,345	9.11	158,695,328	335	30.48
	PB category missing	9,146	0.18	204,881,536	22,401	39.35
	All categories			**2,743,742,498**		**526.97**

Note: the estimation sample size is 5,206,651.
Source: Martin, Dusheiko et al (2009)

Parsimonious models were derived for all 23 programme budget categories. However, significant QOF quality scores were present in the models for only three care programmes (cancer, circulatory disease and other).[20] The parsimonious model for each of these three programmes is shown in table 13 (the coefficients on the age/sex dummies, the ICD10 morbidity markers and the PCT dummies are not shown).

The dementia quality score

The dementia quality score has a significant negative coefficient (–0.056) in the cancer costs model. It implies that a one-percentage-point increase in the dementia quality score (with a mean value of 75.44 for 2007/08) is associated with a reduction in per capita cancer costs of £0.056. With a population of 50 million people, this implies an annual reduction in total cancer costs of £2.8 million. It is not immediately obvious why improved dementia care (monitoring the support needs of the patient and their carer) should reduce costs associated with cancer care. It might be that improved dementia care is associated with improved support for the patient and that the availability of this support reduces the cost of cancer care (for example, through shorter stays in hospital).

The stroke quality score

The stroke quality score has a significant negative coefficient (–0.207) in the circulatory disease costs model. It implies that a one-percentage point increase in the stroke quality score (with a mean value of 86.17 for 2007/08) is associated with a reduction in per capita circulatory disease costs of £0.207. With a population of 50 million people, this implies an annual reduction in total circulatory disease costs of £10.35 million.

The diabetes quality score

The diabetes quality score has a significant negative coefficient (–0.091) in the 'other' costs model. It implies that a one-percentage-point increase in the diabetes quality score (with a mean value of 87.42 for 2007/08) is associated with a reduction in per capita other costs of £0.091. With a population of 50 million people, this implies an annual reduction in total other costs of £4.55 million.

Together, these three models imply the possibility of annual cost savings totalling £17.7 million associated with a one-point increase in the QOF quality scores for three diseases areas (cancer, circulatory disease and other problems). This implies a rather modest saving of just under £6 million per percentage point increase in QOF quality.

This is a considerably smaller saving than that implied from the parsimonious model with all costs as the dependent variable: for example, the derivation of a parsimonious model with QOF scores for 2007/08 included in the full model generated the result that a one-percentage-point increase in the stroke QOF score is associated with an annual reduction in total secondary care costs of £33.2 million. However, if we look at the results for the PB models, a one-percentage-point increase in the stroke QOF score is associated with an annual reduction in circulatory disease costs of £10.35 million.

One explanation for this result is that improvements in quality (as measured by QOF scores) reduce costs in more than one programme. This is plausible and particularly likely where co-morbidities are often present (for example, with diabetes and circulatory disease). Consequently, we risk under-estimating the cost savings from quality improvements unless cost savings across all disease areas are monitored.

Table 13: Parsimonious models for individual care programmes with significant QOF quality variables for 2007/08

	Dependent variable is hospital cost in 2007/08 for:		
Regressors	**Cancer**	**Circulatory disease**	**Other**
Number of episodes 2005/06/07	7.569*** [0.945]	3.140*** [0.920]	1.788*** [0.460]
Number of attendances 2005/06/07	2.422*** [0.218]	3.867*** [0.258]	5.619*** [0.293]
Outpatient priority referral 2005/06/07	26.695*** [2.071]	12.996*** [2.029]	4.319*** [1.060]
Outpatient treatment 2005/06/07	9.639*** [2.834]	3.821* [2.284]	4.441** [2.005]
QOF dementia score 2007/08	–0.056*** [0.019]		
Proportion of students in population	–160.647*** [36.862]		
ONS rural/urban classification: urban >10k, sparse (urban settlement located in sparsely populated area)	–15.538*** [2.831]		
Private episode 2005/06/07	–68.595*** [10.497]	–58.282*** [8.998]	–9.671*** [2.353]
Private attendance 2005/06/07	–31.252*** [9.942]		–14.979*** [2.164]
Number of FTE medical staff at hospital practitioner/clinical assistant grade (distance and population weighted access measure)	433.816*** [84.838]		
Nursing home beds (distance and population weighted access)	–2.953*** [0.938]		

continued

Table 13: Parsimonious models for individual care programmes with significant QOF quality variables for 2007/08 – *continued*

	Dependent variable is hospital cost in 2007/08 for:		
Regressors	**Cancer**	**Circulatory disease**	**Other**
Indices of Multiple Deprivation (IMD) distance to nearest GP			–0.297*** [0.098]
Persons aged >75 living alone		7.228*** [2.081]	
ONS 15: mature city professionals		–6.586*** [1.633]	
No qualifications: age standardised		9.260*** [0.862]	
QOF stroke score 2007/08		–0.207*** [0.069]	
Average age of practice GP 2006		–0.138*** [0.050]	
Single GP practice 2006		3.737*** [1.329]	
ONS 36: multicultural suburbia		–4.510*** [1.582]	
ONS 53: small town communities		4.794*** [1.758]	
All pension credit claimants			9.367*** [1.597]
Proportion Bangladeshi			0.109*** [0.040]
Proportion separated			–47.809*** [17.408]
Proportion of inpatients waiting <12 months			35.859*** [9.821]
DLA claimants aged <16			44.522*** [8.631]
QOF diabetes score 2007/08			–0.091*** [0.029]
Proportion of female GPs 2006			1.670*** [0.516]
Constant	10.766** [4.727]	15.696** [6.742]	1.455 [10.600]
Observations	5,170,603	5,170,603	5,170,603
R-squared	0.0782	0.0411	0.0489
Adj R-squared	0.0781	0.0410	0.0488

Robust standard errors are in brackets with clustering by PCT.
* p<0.1, ** p<0.05, *** p<0.01

5.4 Panel data estimation

Panel data models

So far, we have estimated various cross-section models, using patient hospital costs in 2007/08 as the dependent variable. As regressors we have:

1. the patient's age and gender on 1 April 2007
2. various inpatient diagnosis variables, reflecting the patient's use of hospital services in 2005/06 and 2006/07 combined
3. various hospital encounter variables, based on the patient's use of hospital services in 2005/06 and 2006/07 combined
4. various attributed needs and supply variables, reflecting the characteristics of the area in which the patient lives or the practice with which they are registered on 1 April 2007.

We also have this set of variables for each of the two previous years. The availability of repeated observations for each patient can potentially increase the robustness of the parameter estimates. Providing we use an estimation method that takes account of this repetition, combined cross-sectional, time-series models like this one can generate more precise parameter estimates than a single cross-section. For example, there might be some practice characteristic that is correlated with a practice's QOF scores but which we have omitted from our model. Although we find an association between the QOF score and patient cost, there might be no causality involved, because it is an unobserved third factor that is causally associated with cost rather than the practice QOF score. The use of a panel model can control for the possible influence of such time-invariant unobserved factors.

More precisely, suppose we believe that any unobserved heterogeneity shifts the regression line up or down by a fixed amount for each patient so that:

$$y_{it} = a + bX_{it} + d_i + u_{it}$$

where y_{it} is the *i*th patient's total cost in period *t*, X comprises the time-varying regressors, d_i is the unobserved fixed effect for each patient (or practice), and u_{it} is the usual idiosyncratic error term. This fixed effect might reflect the patient's genetic predisposition to disease, and this representation is known as the fixed effects model.

Alternatively, if the individual differences can be considered to be random values, drawn from a distribution with zero mean and constant variance such that $d_i + u_{it} = e_{it}$, then $y_{it} = a + bX_{it} + e_{it}$ and then we have the random effects model where the unobserved effect (d_i) is assumed to be independent of the regressors (X_{it}). Note that a random effects model will lead to consistent (unbiased) estimates when the observed regressors are not correlated with the unobserved effects. However, if both fixed and random effects are consistent, the random effects estimator is more efficient (it produces smaller standard errors).

Panel data results

Our one-period, cross-section sample consists of just over five million patients living in England and registered with an English practice on 1 April 2007. To this database we added hospital cost and other covariate data including the practice with which the patient was registered on 1 April 2006 and on 1 April 2005, together with the LSOA in which the patient lived at these dates. Although some of our variables are, in principle, time invariant (such as the attributed needs and supply variables), some patients will have moved and/or changed their practice between 1 April 2005 and 1 April 2007. This means that, in principle, all of our regressors are time variant.[21] However, in practice, the within-individual variation in some variables may be too small to contribute to the analysis.[22]

The panel data results are shown in table 14. To aid comparison, Model 12 in table 14 is the cross-section parsimonious model with the stroke QOF score for 2007/08 (this is Model 3 in table 9). The coefficient on the stroke QOF score (–0.712) implies that annual patient costs in practices with a one-percentage-point higher QOF score are, on average, £0.712 lower. This result is likely to be an upper-point estimate for the impact of quality on costs for three reasons:

1. It is based on individual and therefore across-practice variation in QOF scores and, as there is plenty of across-practice variation, the model will be able to detect any association between cost and quality.
2. Any lag in the response of costs to a change in quality is less likely to reduce the (absolute) size of the estimated QOF coefficient in an across-individual than in a within-individual model.[23]

3. There might be some unobserved practice characteristic that is correlated with both quality and costs, but whose effect on costs is incorporated in the coefficient on practice quality.

The first panel data model estimated is a fixed effects (FE) version of the parsimonious model (with the current period stroke QOF score as a time-varying regressor). This FE model performs poorly and is shown as Model 13 in table 14. The coefficients on the age/gender variables and on the hospital encounter variables are very different to those obtained from the cross-section models, and they do not accord with our prior beliefs (for example, the coefficients on the age/gender variables do not exhibit the usual positive age gradient for adults). This result can be attributed to the fact that the fixed effects estimator uses only within-individual variation (and not between-individual variation) to estimate the regression coefficients, and that many patients will remain in the same five-year age group and record the same value for their encounter variables for all three data years. Also, apart from measurement error, the inclusion of a time trend will make it impossible to separate out the effects of age from the time trend.

The more plausible coefficients on the age and encounter variables revealed by the single cross-section model (Model 12) reflect the greater variation in these variables across individuals than within individuals. The coefficient on the stroke variable in the FE Model 13 is close to zero and is statistically insignificant. This is likely to reflect partly the lack of variation in the stroke QOF score within individuals, and partly the poor performance of the other regressors in the FE model. Also, because the FE estimator is using only within-individual variation, and there tends to be more noise (measurement error and random fluctuations, for example) in variables within than across individuals, the FE coefficient estimates will always tend to be closer to zero and have larger standard errors than their across-individual counterparts.

The FE model can be thought of as generating a dummy variable for each individual. With about five million people, this consumes many degrees of freedom. We therefore also estimated a practice FE model. This model controls for practice fixed effects but not individual fixed effects. This involves estimating the same model but with a dummy variable for each of the 8,083 practices. The results for this model are presented as Model 14 in table 14 and are much more in line with our prior beliefs. There is a plausible age gradient, and the coefficients on the encounter variables are similar to those in the one-period, cross-section model. The coefficient on the stroke QOF score remains insignificant at the 5% level, but it is now negative (–0.249). The practice fixed effects model generates more plausible results than the individual fixed effects model, because there is more variation across individuals (but within practices) than across repeated observations within individuals. However, the coefficient on the QOF score in the practice fixed effects model is smaller (in absolute terms) than in the cross-section model. This is partly because the FE model controls for any unobserved practice-level, time-invariant factors whose impact on costs might otherwise be attributed to the practice QOF score. It might also be partly because the impact on costs of a change in practice quality might take time (beyond the current period) to fully affect patient costs.

Both of the individual and practice fixed effects models ignore any across-individual/practice variation. In contrast, the between effects estimator uses only across-individual information and generates the result shown as Model 15. The coefficient on the QOF score is now significant and larger (in absolute terms) than in either of the fixed effect models (it is –0.431). This increase is to be expected as the between effects estimator does not control for unobserved heterogeneity (a disadvantage) and is less likely to be affected by the presence of a lagged response of costs to quality (an advantage). Also, with three years of data, the parameter estimates from this model should be more stable than those from its one-year counterpart (see Models 12 and 15).

Model 16 in table 14 is a random effects weighted average of the between and within estimates. The coefficient on the stroke QOF score (–0.377) is slightly smaller than in the between effects model, but it is still significant at the 1% level.[24] It is smaller than that obtained in the one-period, cross-section models, but it still implies that a small increase in the quality of primary care (as measured by the stroke QOF score) would be associated with an annual reduction in secondary care costs of about £19 million.[25]

A Hausman test could be used to examine the validity of the RE model (in particular, its assumption that any unobserved individual effect is uncorrelated with the regressors). This test draws on the result that the FE estimator generates estimates that are consistent, whether or not the unobserved individual effect is correlated with the regressors. However, in our models, the lack of variation in the regressors within individuals through time compromises the FE approach, and the results cannot be relied upon. Consequently, we do not use the Hausman test to assess the validity of the RE model.

Finally, we re-estimated Model 16 (using the same panel of 4.2 million patients) with an additional time-invariant variable: the stroke QOF score in 2004/05 (the first year of the new GMS contract) for the practice with which the patient was registered on 1 April 2005 (see Model 17 in table 14). The coefficient on this variable will reflect the association between the 2004/05 quality score of the practice with which the patient was registered on April 2005 and patient expenditure in the following years (2005/06, 2006/07 and 2007/08). It may therefore capture a lagged quality effect on costs, or it may capture the association between unobserved time-invariant practice characteristics that influence both quality and expenditure. The inclusion of this 'baseline' quality indicator is a way of allowing for unobservable time-invariant practice characteristics (that might be associated with both quality and patient costs) within an RE model.

Although the coefficient on the baseline quality variable in Model 17 is negative (it is –0.119 and has a standard error of 0.089), it is not statistically significant at the 5% level. This small negative coefficient suggests that patients who were registered with better-quality practices in 2004/05 were associated with lower expenditures in 2005/06 and 2007/08, although this association was not statistically significant. This suggests either that the lagged effect of quality on cost is small (given this reflects the pooled association across all three years) or that there is a positive correlation between quality and unobserved time-invariant, GP practice-level characteristics that are both negatively associated with the level of patient expenditure.

The addition of the baseline stroke QOF score to the RE model reduces the size of the coefficient on the time-varying stroke QOF score from –0.377 to –0.282, but this is still significant at the 5% level.[26,27] Also, the coefficient on the stroke QOF score in the practice fixed effects model (–0.249) is similar to the coefficient on the same variable in the RE model with the inclusion of a baseline quality score (–0.282). It appears that allowing for baseline quality in an RE model generates a similar point estimate on the time-varying quality variable as is obtained from a practice fixed effects model, but the RE model has the distinct advantage that it generates a much smaller standard error (0.133 rather than 0.195).

Conclusion from panel data estimates

From the cross-section and panel data models, we have a variety of estimates of the impact of primary care quality on secondary care costs. Given the nature of the data, the random effects (RE) model would seem to offer the best estimate of the impact of quality on costs. The RE model offers a compromise between the poor fixed effects model (which is hampered by the lack of variation at the individual patient level) and which is likely to underestimate the impact of quality on costs, and the one-period, cross-section model, which relies on a single cross-section and fails to incorporate variation over time.

We have two RE point estimates of the impact of quality on costs. The coefficient on the time-varying quality score in the RE model with baseline quality (–0.282) provides a lower-point estimate for the cost saving per patient from a one-percentage-point increase in the stroke QOF score. This estimate omits both the lagged impact of baseline quality on costs (a disadvantage) and the impact of any unobserved practice characteristics on costs (an advantage). The coefficient on the time-varying quality score in the RE model without baseline quality (–0.377) provides an upper-point estimate of the cost saving per patient for a one-percentage-point increase in the stroke QOF score. This estimate includes both the lagged impact of baseline quality on cost (an advantage) and also the impact of unobserved practice characteristics on costs (a disadvantage).

Ideally, we want a point estimate that includes the lagged impact of baseline quality on cost but which

excludes the impact of unobserved practice characteristics. With lower- and upper-point estimates of –0.282 and –0.377 respectively, and in the absence of any further evidence, we take the midpoint of these two estimates (–0.330) as our best-point estimate of the impact of a one-percentage-point increase in the stroke QOF score on each patient's annual secondary care costs. With a population of 50 million people, this implies that a one-percentage-point increase in the stroke QOF score is associated with a £16.5 million annual reduction in total patient costs.

Table 14: Panel models illustrating the impact of QOF scores on patient hospital costs, 2005/06 and 2007/08

	Model number and description					
	Model 12	Model 13	Model 14	Model 15	Model 16	Model 17
Regressors	One–period cross–section	Individual fixed effects	Practice fixed effects	Individual between effects	Individual random effects	Random effects with baseline quality 2004/05
Males aged <1	197.159*** [17.172]	734.630*** [19.093]	189.634*** [13.306]	42.039 [26.208]	196.931*** [14.393]	196.811*** [14.405]
Males aged 1–4	–2.587 [5.090]	647.882*** [13.794]	–10.625*** [3.372]	–43.660*** [5.194]	–6.041* [3.643]	–6.086* [3.647]
Males aged 5–9	–20.270*** [4.030]	526.911*** [12.341]	–20.809*** [2.474]	–12.928*** [4.311]	–21.183*** [2.702]	–21.331*** [2.704]
Males aged 10–14	–1.858 [3.081]	435.953*** [11.177]	–2.838 [2.340]	15.316*** [4.201]	–5.929** [2.512]	–6.028** [2.514]
Males aged 15–19	–9.454** [3.651]	342.728*** [10.008]	–11.344*** [2.482]	–2.052 [4.185]	–14.020*** [2.639]	–14.064*** [2.641]
Males aged 20–24	–12.389*** [3.617]	257.130*** [8.764]	–8.952*** [2.567]	–7.221* [4.259]	–11.581*** [2.765]	–11.613*** [2.768]
Males aged 25–29	–18.077*** [3.381]	156.234*** [6.583]	–17.097*** [2.477]	–9.020** [4.172]	–19.468*** [2.590]	–19.611*** [2.592]
Males aged 30–34	–14.182*** [3.375]	69.563*** [4.588]	–15.083*** [2.403]	–8.766** [4.286]	–17.163*** [2.528]	–17.235*** [2.530]
Males aged 40–44	17.462*** [3.880]	–66.376*** [5.336]	14.103*** [2.676]	6.217 [4.140]	15.457*** [2.890]	15.411*** [2.893]
Males aged 45–49	51.924*** [4.553]	–124.687*** [7.900]	41.175*** [3.026]	22.906*** [4.040]	45.009*** [3.185]	44.875*** [3.187]
Males aged 50–54	97.817*** [5.963]	–153.004*** [11.672]	84.533*** [3.502]	46.531*** [4.221]	93.046*** [3.809]	92.955*** [3.812]
Males aged 55–59	158.994*** [5.480]	–142.495*** [15.308]	144.182*** [3.776]	82.950*** [4.174]	157.359*** [4.141]	157.249*** [4.144]
Males aged 60–64	257.183*** [7.782]	–59.513*** [18.742]	236.485*** [4.563]	129.127*** [4.429]	254.322*** [5.023]	253.982*** [5.023]
Males aged 65–69	373.501*** [10.308]	60.018** [24.346]	344.198*** [5.962]	177.406*** [4.775]	372.603*** [6.566]	372.291*** [6.570]
Males aged 70–74	544.035*** [10.592]	275.851*** [30.586]	476.005*** [7.268]	254.325*** [5.136]	510.746*** [7.872]	510.836*** [7.878]
Males aged 75–79	665.744*** [17.591]	611.364*** [37.828]	605.382*** [9.491]	296.560*** [5.697]	662.575*** [10.518]	662.423*** [10.523]

continued

Table 14: Panel models illustrating the impact of QOF scores on patient hospital costs, 2005/06 and 2007/08 – *continued*

	Model number and description					
	Model 12	**Model 13**	**Model 14**	**Model 15**	**Model 16**	**Model 17**
Regressors	**One–period cross–section**	**Individual fixed effects**	**Practice fixed effects**	**Individual between effects**	**Individual random effects**	**Random effects with baseline quality 2004/05**
Males aged 80–84	849.779*** [21.240]	1,029.171*** [50.534]	712.218*** [12.373]	341.950*** [6.789]	778.410*** [13.793]	778.600*** [13.802]
Males aged >84	1,054.739*** [25.962]	1,652.916*** [63.794]	858.444*** [16.218]	375.701*** [8.166]	937.979*** [17.048]	938.306*** [17.057]
Females aged <1	138.057*** [15.324]	–112.725** [53.017]	134.804*** [10.818]	9.371 [26.499]	135.827*** [12.022]	134.724*** [11.991]
Females aged 1–4	–25.710*** [4.496]	–193.247*** [53.583]	–24.125*** [3.300]	–40.325*** [5.269]	–21.488*** [3.569]	–21.564*** [3.572]
Females aged 5–9	–26.575*** [3.206]	–301.954*** [53.855]	–30.065*** [2.276]	–11.620*** [4.365]	–32.067*** [2.447]	–32.294*** [2.447]
Females aged 10–14	–3.812 [3.790]	–383.973*** [54.039]	–7.189*** [2.307]	17.393*** [4.259]	–9.819*** [2.472]	–9.917*** [2.474]
Females aged 15–19	–5.888* [3.074]	–453.382*** [54.267]	–3.985 [2.440]	8.646** [4.241]	–5.519** [2.623]	–5.628** [2.625]
Females aged 20–24	–2.097 [3.547]	–520.644*** [54.637]	0.614 [2.392]	–3.526 [4.323]	0.252 [2.585]	0.184 [2.588]
Females aged 25–29	6.751* [3.567]	–582.069*** [54.915]	11.268*** [2.389]	7.762* [4.294]	13.199*** [2.649]	13.109*** [2.652]
Females aged 30–34	27.692*** [3.774]	–645.935*** [55.134]	32.471*** [2.616]	30.666*** [4.231]	33.839*** [2.783]	33.450*** [2.774]
Females aged 35–39	42.395*** [3.722]	–703.586*** [55.459]	40.568*** [2.552]	29.424*** [4.048]	43.920*** [2.769]	43.871*** [2.771]
Females aged 40–44	55.531*** [4.119]	–760.838*** [55.872]	51.392*** [2.570]	29.162*** [4.003]	57.342*** [2.831]	57.218*** [2.833]
Females aged 45–49	83.746*** [5.254]	–796.811*** [56.355]	70.283*** [3.023]	35.148*** [4.146]	76.829*** [3.155]	76.755*** [3.157]
Females aged 50–54	116.837*** [5.487]	–802.917*** [57.047]	99.653*** [3.502]	48.337*** [4.278]	108.325*** [3.701]	108.194*** [3.703]
Females aged 55–59	146.370*** [5.611]	–803.931*** [57.736]	130.719*** [3.467]	61.841*** [4.204]	141.877*** [3.694]	141.787*** [3.696]
Females aged 60–64	210.937*** [7.337]	–776.497*** [58.738]	190.837*** [4.366]	91.633*** [4.441]	206.272*** [4.650]	206.142*** [4.653]
Females aged 65–69	305.400*** [8.885]	–707.915*** [60.450]	276.969*** [5.368]	140.063*** [4.693]	302.894*** [5.792]	302.384*** [5.776]
Females aged 70–74	442.814*** [11.040]	–542.703*** [62.567]	393.192*** [6.285]	207.279*** [4.905]	426.904*** [6.727]	426.763*** [6.730]
Females aged 75–79	577.676*** [12.395]	–318.722*** [65.129]	514.935*** [7.129]	252.666*** [5.160]	557.301*** [7.639]	557.254*** [7.644]
Females aged 80–84	730.020*** [13.816]	–27.401 [68.539]	652.441*** [8.960]	295.141*** [5.642]	710.358*** [9.527]	710.452*** [9.532]

continued

Table 14: Panel models illustrating the impact of QOF scores on patient hospital costs, 2005/06 and 2007/08 – *continued*

	Model number and description					
	Model 12	**Model 13**	**Model 14**	**Model 15**	**Model 16**	**Model 17**
Regressors	**One–period cross–section**	**Individual fixed effects**	**Practice fixed effects**	**Individual between effects**	**Individual random effects**	**Random effects with baseline quality 2004/05**
Females aged >84	985.262*** [19.485]	311.897*** [74.489]	846.188*** [10.556]	370.145*** [5.759]	925.551*** [11.099]	925.516*** [11.106]
Year dummy for 2006/07		72.043*** [1.584]	17.427*** [1.755]	–21.779 [58.697]	19.783*** [1.452]	19.184*** [1.496]
Year dummy for 2007/08		135.557*** [1.668]	17.902*** [1.756]	–79.649** [34.544]	22.543*** [1.432]	21.869*** [1.480]
Number of episodes 2005/06/07	298.948*** [9.900]	–24.080*** [8.081]	306.067*** [5.047]	355.019*** [0.253]	293.793*** [5.307]	293.734*** [5.310]
Number of attendances 2005/06/07	45.690*** [2.202]	–30.126*** [0.875]	46.451*** [0.730]	46.256*** [0.160]	45.274*** [0.706]	45.255*** [0.706]
Outpatient priority referral 2005/06/07	70.865*** [10.591]	57.459*** [4.394]	72.855*** [3.728]	35.793*** [2.350]	75.433*** [3.771]	75.687*** [3.773]
Outpatient treatment 2005/06/07	55.710*** [12.398]	21.675*** [6.492]	39.481*** [5.107]	29.660*** [3.409]	37.723*** [5.170]	37.787*** [5.173]
Private attendance 2005/06/07	–167.002*** [25.184]	70.255*** [18.805]	–126.31*** [13.341]	–167.050*** [12.560]	–123.462*** [13.954]	–123.35*** [13.964]
Private episode 2005/06/07	–490.674*** [24.728]	422.954*** [23.463]	–451.02*** [16.283]	–859.607*** [14.107]	–381.645*** [17.512]	–381.67*** [17.519]
In social rented housing	0.279*** [0.093]	0.419** [0.191]	0.241*** [0.066]	0.065 [0.051]	0.336*** [0.060]	0.338*** [0.060]
Disability living allowance	339.079*** [75.539]	–612.749*** [161.541]	313.911*** [55.642]	13.826 [42.362]	347.448*** [53.707]	346.571*** [53.733]
No qualifications: age standardised	24.143*** [4.790]	36.860*** [10.490]	19.153*** [3.593]	5.320* [2.889]	19.859*** [3.190]	19.594*** [3.202]
Stroke QOF score 2005/06/07	–0.712*** [0.182]	0.061 [0.176]	–0.249 [0.195]	–0.431*** [0.123]	–0.377*** [0.117]	–0.282** [0.133]
ONS 15: mature city professionals	–23.794*** [7.024]	19.56 [15.468]	–12.715** [6.387]	–19.321*** [5.877]	–16.595*** [5.581]	–16.288*** [5.584]
Students in population	–1,297*** [139.034]	–810** [325.269]	–1,270*** [103.198]	–632*** [96.102]	–1,372*** [96.498]	–1,376*** [96.576]
Access to MRI scanners	5,470** [2,251]	10,735*** [3,876]	3,065 [3,322]	4,989*** [975]	5,717*** [1,167]	5,734*** [1,167]
Access to residential beds	–8.047*** [2.383]	–11.954** [5.319]	–6.656* [3.759]	–7.782*** [1.186]	–8.483*** [1.321]	–8.475*** [1.321]
Baseline stroke QOF 2004/05						–0.119 [0.089]
Constant	157.214*** [18.368]	679.615*** [68.259]	88.316*** [26.842]	116.745*** [32.454]	130.092*** [16.121]	131.998*** [16.256]
Observations	5,188,594	14,567,220	14,567,220	14,567,220	12,580,505	12,569,257

continued

Table 14: Panel models illustrating the impact of QOF scores on patient hospital costs, 2005/06 and 2007/08 – *continued*

	Model number and description					
	Model 12	**Model 13**	**Model 14**	**Model 15**	**Model 16**	**Model 17**
Regressors	**One–period cross–section**	**Individual fixed effects**	**Practice fixed effects**	**Individual between effects**	**Individual random effects**	**Random effects with baseline quality 2004/05**
Number of individuals	5,188,594	4,863,259	4,863,259	4,863,259	4,199,991	4,196,149
Within R-squared		0.0372		0.0228	0.0105	0.0105
Between R-squared		0.1459		0.5205	0.4815	0.4816
Overall R-squared	0.2662	0.0442	0.2079	0.1948	0.2037	0.2037

Robust standard errors are in brackets.
* p<0.1, ** p<0.05, *** p<0.01
Notes:
1. Model 12 is the same as Model 3 in table 9. It is a one-period, cross-section model with the QOF score for 2007/08. It reports robust standard errors with clustering by PCT.
2. Model 13 is an individual fixed effects model with robust standard errors clustered by individual.
3. Model 14 is a practice fixed effects model with robust standard errors clustered by general practice.
4. Model 15 is a between effects model with robust standard errors clustered by individual.
5. Model 16 is a random effects model with robust standard errors clustered by individual.
6. Model 17 is a random effects model with robust standard errors clustered by individual.
7. There are just under 4,000 patients whose practice on 1 April 2005 does not have a stroke QOF score for 2004/05. Therefore, the number of observations for Model 17 is slightly less than that for Model 16.
9. The dependent variable is the patient's hospital costs for 2007/08, excluding maternity and mental health cost (Model 12) or for 2005/06, 2006/07 and 2007/08 (Models 13–17).

Chapter 6

Conclusions

6.1 Base model findings

In this study, we have taken advantage of a major new dataset, which links practice-based patient registration data and patient-level hospital use data, to examine whether higher practice QOF scores are associated with reduced hospital costs. The base model is a one-period, cross-section study, which finds that the QOF stroke quality score for 2005/06 had a small but significant negative association with patients' hospital costs in 2007/08. We discovered that this negative association increases if the model is otherwise left unchanged but is re-estimated with the QOF stroke score updated to 2006/07. We also found that this negative association increases still further if the QOF stroke score is updated to 2007/08.

We also re-derived the parsimonious model for hospital costs but with the inclusion of the QOF quality scores for 2007/08 in the full model. The QOF stroke score still appears in the parsimonious model and has a coefficient of –0.664. This implies that a one-percentage-point increase in the QOF stroke score for 2007/08 is associated with a £33.2 million reduction in secondary care costs. The similarity between this result and that achieved by retaining the original parsimonious model and updating the QOF stroke variable suggests that these one-period, cross-section results are reasonably stable. However, they will tend to exaggerate any impact of quality on costs because they make no allowance for the presence of unobserved practice characteristics that are correlated with both quality and cost.

6.2 Panel data model findings

To overcome this shortcoming we also estimated several panel data models. The advantage of these over the one-period, cross-section model is that panel datasets can accommodate the impact of unobserved factors on patient costs. As anticipated, the panel data models generated smaller coefficients on the QOF stroke care variable than did the one-period, cross-section models. Our two preferred panel data models are random effects (RE) models. In our first RE model, the coefficient on the stroke QOF score is –0.377, and this is likely to be an upper-point estimate of the cost saving per patient (because this estimate includes both the lagged impact of baseline quality on cost and the impact of unobserved practice characteristics on costs). In our second RE model, the coefficient on the stroke QOF score is –0.282, and this is likely to be a lower-point estimate of the cost saving per patient (because this estimate excludes both the lagged impact of baseline quality on cost and the impact of unobserved practice characteristics on costs).

In the absence of any further evidence, we take the midpoint of these two estimates (–0.330) as our best-point estimate of the impact of a one-percentage-point increase in the stroke QOF score on each patient's annual secondary care costs. With a population of 50 million people, this implies that a one-percentage-point increase in the stroke QOF score is associated with a £16.5 million annual reduction in total patient costs. Over the period studied, the mean practice QOF stroke score increased by 10 percentage points, and this implies that annual secondary care costs were £165 million lower in 2007/08 than in 2004/05 as a result of the increase in primary care quality. It is not possible to estimate directly the costs to

primary care and other health services of securing this improvement. However, it is worth noting that the additional QOF incentive payments associated with a one-point improvement in stroke achievement is very small compared to the associated hospital cost savings.

It could also be argued that the reduction in hospital costs might be attributable to the re-location of some treatments out of hospitals and into primary care. However, careful scrutiny of the QOF indicates few opportunities for substituting primary care for treatments formerly delivered in hospitals. The QOF indicators are likely to reflect mainly improvements in clinical practice and outcomes in primary care.

We would like to emphasise that we are *not* claiming that improved primary care will reduce total lifetime healthcare costs. As is well known, the majority of disease prevention strategies generate additional costs (Cohen, Neumann et al 2008). Our focus is solely on the long-run annual cross-sectional impact on secondary care costs when there is an improvement in primary care. We do not consider, for example, the additional pharmaceutical and primary care costs associated with meeting QOF stroke targets, nor the impact on total lifetime healthcare costs if the patient lives longer as a result of better primary care.

6.3 Scope for future research

The dataset also includes a binary variable that indicates whether the patient died during the 12-month period from1 April 2007 to 31 March 2008. We estimated a one-period, cross-section logit model to identify those factors that are associated with the probability of death, and we found that the QOF stroke quality score was significantly negatively associated with the probability of death. We also found that the strength of this negative association – between primary care quality and the probability of death – increases as the time period over which quality is measured moves closer to the period over which death is recorded. These preliminary results suggest that a one-percentage-point increase in the QOF stroke score for 2007/08 would be associated with 2,385 fewer deaths in that year. Although this result is likely to exaggerate the impact of quality on mortality (not least because the cross-section model makes no allowance for unobserved practice characteristics), it does indicate that this may be a fruitful line of research, and we intend to estimate a panel model for mortality in the near future.

Finally, we split each patient's total hospital costs between 23 care programmes and derived parsimonious models for each programme. We found that QOF quality scores for 2007/08 (for dementia, stroke and diabetes care) had a significant negative effect on hospital costs in the cancer, circulatory disease and 'other' programmes respectively. However, even when aggregated, these one-programme effects were much smaller than those recorded when modelling costs across all programmes. This implies that quality improvements in one disease area may generate cost savings across more than one programme. Although this finding is based solely on a comparison of one-period, cross-section models, it does suggest that studies that examine the impact of improved quality by looking at the benefits in only one disease area might seriously underestimate the total benefits of that quality improvement. This is another topic that we intend to pursue in the near future.

6.4 QOF – material but limited gains?

We must emphasise that this study is not seeking to evaluate the QOF initiative, or to offer estimates of the cost-effectiveness of QOF interventions. Rather, it gives an indication of the extent to which the initiative may have affected hospital costs or mortality outcomes. We find in both respects that the QOF appears to be associated with material but limited gains. We are cautious about drawing inferences of causality from our work, but feel that the panel data results do offer solid grounds for believing that QOF improvements are contributing to the gains.

The stroke QOF score dominates our models. To some extent, this may be because it is an indicator of overall primary care quality. It is highly correlated with overall QOF attainment. However, its dominance, and the role it plays in the model of circulatory disease costs, suggests that the stroke quality metrics are capturing specific aspects of preventive care that do have a measurable impact on outcomes.

Endnotes

1. This contract was introduced throughout the UK but our focus is on its impact in England.
2. There were also three measures of 'depth' of care, which we do not consider further here.
3. The deduction was 168 points in 2004/05 and 109 points in 2005/06, from a maximum possible score of 1,050 points.
4. The binary indicator relating to the existence or absence of a disease register is excluded from the calculation of the weighted population achievement rate for each clinical sub-domain.
5. Holistic care payments to practices are designed to recognise the breadth of achievement across the clinical domain, and 20 QOF points are available. To calculate holistic care points, the practice's points totals in each of the clinical areas of the clinical domain are ranked on the basis of the proportion of available points achieved. The points relating to the highest proportion are ranked first. The proportion relating to the points total that is third-to-last is then taken as the proportion of 20 holistic care points to which the practice is entitled as the basis for its holistic care payment (NHS Information Centre for Health and Social Care 2007).
6. As was the case for 2004/05 and 2005/06, we do not include the binary indicators relating to the existence or absence of a disease register in the weighted population achievement rates for each clinical sub-domain.
7. Hospital Episode Statistics (HES) is a data warehouse containing details of all hospital admissions for NHS patients. HES also includes details of all NHS outpatient attendances in England. HES contains admitted patient care data from 1989/90 onwards, with more than 15 million new records added each year, and outpatient attendance data from 2003/04 onwards, with more than 60 million new records added annually. On admission to hospital, each patient is assigned to the care of a particular consultant, and HES opens a new consultant episode. When the patient is discharged from hospital or dies, the record is closed and becomes a finished consultant episode (FCE). The HES record also becomes a FCE if the patient is transferred from the care of one consultant to another while still in the same hospital. The period between admission to and discharge from the hospital is known as a spell of care, and a patient might record several FCEs within a single spell of care. Each HES record contains a wealth of information, including the patient's age, gender, length of stay, and diagnosis data.
8. Private care at private providers is excluded, but this is unlikely to be a problem because we are modelling NHS funded care.
9. For two-thirds of all patients registered with an English practice on 1 April 2007, their hospital cost for 2007/08 was zero (that is, they did not use any inpatient or outpatient services during the year).
10. More precisely, the derivation of the parsimonious model involved the following steps:
    - Step 1. Estimate the full model.
    - Step 2. Re-estimate the full model, retaining only those attributed need and supply variables whose absolute t-ratio was greater than 0.20.
    - Step 3. Re-estimate the model estimated in step 2, retaining only those

attributed need and supply variables whose absolute t-ratio was greater than 0.40.

Step 4. Re-estimate the model estimated in step 3, retaining only those attributed need and supply variables whose absolute t-ratio was greater than 0.60.

Step 5. Continue this process until only those variables with an absolute t-ratio greater than 2.00 remain.

Step 6. Inspect the coefficients on the remaining need and supply variables and drop those variables with 'incorrect/unexpected' signs.

Step 7. Re-estimate the model with the remaining attributed need and supply variables.

Step 8. Re-estimate the model estimated in step 7, retaining only those attributed need and supply variables whose absolute t-ratio is greater than 2.20.

Step 9. Continue the process outlined in steps 6 to 8 until only those variables with a significance level of 1% or lower remain.

11. The regression coefficients show the impact on the dependent variable of a one-unit increase in the regressors. In the case of the stroke QOF score, the regressor is a percentage achievement rate, and the dependent variable is the annual patient hospital cost (£) in 2007/08. Therefore, a coefficient of –0.443 implies that a one-unit increase in the stroke QOF rate is associated with a £0.443 reduction in each patient's annual hospital cost.

12. The overall clinical QOF population achievement rate is a weighted average of individual clinical sub-domain achievement rates, with weights reflecting the number of points available in each sub-domain.

13. Most of the indicators in the clinical domains reflect practice achievement in the 15 months to the financial year end (that is, the QOF indicators for 2007/08 reflect practice achievement from January 2007 to March 2008). This will be slightly (three months) behind the cost variable that relates to inpatient episodes and outpatient attendances finishing between April 2007 and March 2008.

14. Just under 45,000 people in our estimation sample died in 2007/08.

15. The ordinary least squares (OLS) estimator is inappropriate if the dependent variable is binary because, with OLS, predicted values can be greater than one and less than zero. Both types of value are theoretically inadmissible with a binary variable. Logistic regression analyses binomially distributed data of the form:

$Y_i \sim B(n_i, p_i)$, for $i = 1, ..., m$,

where the numbers of Bernoulli trials n_i are known and the probabilities of success p_i are unknown. The model proposes that, for each trial i, there is a set of explanatory variables that might inform the final probability. These explanatory variables can be thought of as being in a k vector X_i and the model then takes the form:

$$p_i = \mathrm{E}\left(\frac{Y_i}{n_i} \mid X_i\right).$$

The logits (natural logs of the odds) of the unknown binomial probabilities are modelled as a linear function of the X_i:

$$\mathrm{logit}(p_i) = \mathrm{In}\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_{1,}x_{1,i} + ... + \beta_k,x_{k,i}.$$

The unknown parameters β_j are usually estimated by maximum likelihood. The interpretation of the β_j parameter estimates is as the additive effect on the log odds ratio for a unit change in the jth explanatory variable. Because the relationship between the regressors and the probabilities is non-linear, the β_j parameters do not have a straightforward interpretation in this model as they do in ordinary linear regression. Nevertheless, each of the regression coefficients describes the size of the contribution of that risk factor. A positive regression coefficient means that that explanatory variable increases the probability of the outcome, while a negative regression coefficient means that the variable decreases the probability of that outcome; a large regression coefficient means that the risk factor strongly influences the probability of that outcome, while a near-zero regression coefficient means that that risk factor has little influence on the probability of that outcome.

16. In Stata v11 the relevant command to find the average marginal effect of the stroke QOF score on the probability of death is: margins, dydx(strokeQOFscore2007/08).

17. The average marginal effect calculates the marginal effect for each individual at the actual values taken for each explanatory variable. The average marginal effect can be contrasted with the marginal effect evaluated at the mean values of the regressors. The latter approach to the calculation of the marginal effect makes sense if the vector of means of the explanatory variables is a sensible point to examine the marginal effect. However, if the vector of mean values represents a unit that is not observed in the data, or cannot be observed due to the presence of many binary explanatory variables (as are present here), then the use of mean values is not a sensible point at which to examine the marginal effect. This explains our preference for the average marginal effect rather than the marginal effect at the mean. We also find that the marginal effect at the mean is much smaller than the average marginal effect: for example, the marginal effect at the mean of the 2007/08 QOF stroke score on the probability of death implies that a one-unit increase in the stroke QOF score for 2007/08 would be expected to lead to 350 (rather than 2,385) fewer deaths in that year.

18. The 23 care programmes are also known as programme budget (PB) categories. Since April 2003, each PCT has been required to allocate all of its expenditure – including expenditure on inpatient care, outpatient care, community care, primary care and pharmaceuticals – between the 23 programmes of care. The availability of corresponding PCT mortality rates for some of these programmes has facilitated a study that examines whether expenditure and health outcomes are positively associated with each other, given the need for healthcare in each local area (Martin, Rice and Smith 2008).

19. The costs included in table 12 incorporate mental health and maternity costs, but the costs for these two areas are excluded from the dependent variable for the models reported in tables 9 and 10.

20. The full model included QOF population achievement rates for 2007/08.

21. Although there are just over five million patients in our 2007/08 sample, it was only possible to update all variables for 2005/06 and 2006/07 for just over 4.8 million of these people. In addition, the estimation of a random effects (RE) model with 350 regressors (as we have here) requires a great deal of computing memory. The maximum amount of memory available to us was 48gb, and this enabled us to estimate the RE model across 4.2 million patients from our sample of 4.8 million (that is, 87.5%).

22. For example, for the attributed needs variables, the within-individual standard deviation is about one-fifth of the between-individual standard deviation.

23. In a within-individual model, and if we ignore patients who switch practice, any change in the QOF score will reflect quality changes in the previous 12 months. In an across-individual model, changes in the QOF score might reflect differences in quality that have existed across practices for several years.

24. The 95% confidence interval for this coefficient is that it lies between –0.143 and –0.611.

25. With a population of 50 million people, the coefficient on the QOF stroke score variable implies that a one-percentage-point increase in the value of this variable would be associated with a £18.85 million (=50 million × 0.377) reduction in annual secondary care costs.

26. The 95% confidence interval for this coefficient is that it lies between –0.016 and –0.548.

27. Note that, in Model 5, the coefficient on the time-varying QOF score reveals the effect on cost of a change in the quality of care, but in Model 6 the coefficient on the time-varying QOF score reveals the effect on cost of a change in the quality of care relative to the level of baseline quality in 2004/05.

References

1 Bottle A, Gnani S, Saxena S, Aylin P, Mainous A and Majeed A (2008). 'Association between quality of primary care and hospitalization for coronary heart disease in England: a national cross-sectional study'. *J Gen Intern Medicine,* vol 23, pp 135–141.

2 Bottle A, Millett C, Xie Y, Saxena S, Wachter R and Majeed A (2008). 'Quality of primary care and hospital admissions for diabetes mellitus in England'. *Journal of Ambulatory Care Management,* vol 31, pp 226–238.

3 British Medical Association/NHS Employers (2006). *Revisions to the GMS contract 2006/07. Delivering investment in general practice.* London: British Medical Association/NHS Employers.

4 Cohen J, Neumann P and Weinstein M (2008). 'Does preventive care save money? Health economics and the presidential candidates'. *New England Journal of Medicine,* vol 358, pp 661–664.

5 Congressional Budget Office (2004). *An analysis of the literature on disease management programs*. Washington DC: Congressional Budget Office.

6 Department of Health (2004). *Quality and outcomes framework. Guidance – updated August 2004.* London: Department of Health.

7 Dixon J, Bardsley M, Georghiou T, Steventon A, Smith P, Gravelle H, Martin S et al (2009). *Developing a person-based resource allocation formula for allocations to general practices in England*. Report to Department of Health. Centre for Health Economic/National Primary Care Research and Development Centre, University of York.

8 Doran T (2008). 'Lessons from early experience with pay for performance'. *Disease Management and Health Outcomes*, vol 16, pp 69–77.

9 Doran T, Fullwood C, Gravelle H, Reeves D, Kontopantelis E, Hiroeh U and Roland M (2006). 'Pay for performance programs in family practices in the United Kingdom'. *New England Journal of Medicine*, vol 355, pp 375–384.

10 Downing A, Rudge G, Cheng Y, Tu Y, Keen J and Gilthorpe M (2007). 'Do the UK government's new Quality and Outcomes Framework (QOF) scores adequately measure primary care performance? A cross-sectional survey of routine healthcare data'. *BMC Health Services Research*, vol 7, p 166.

11 Dusheiko M, Doran T, Gravelle H, Fullwood C and Roland M (2009). *Does higher quality of care in general practice reduce emergency hospital admissions? Evidence from diabetes management in the Quality and Outcomes Framework*.

12 Gravelle H, Sutton M and Ma A (2010). 'Doctor behaviour under a pay for performance contract: treating, cheating and case finding?' *Economic Journal 2010*, vol 120, F129–F156.

13 HES Online. 'Primary diagnosis summary'. Available at: www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=202.

14 Martin S, Dusheiko M, Gravelle H and Rice N (2009). *Estimating needs for programme budget areas.* Report to Department of Health. Centre for Health Economic/National Primary Care Research and Development Centre, University of York.

15 Martin S, Rice N and Smith PC (2008). 'Does health care spending improve health outcomes? Evidence from English programme budgeting data'. *Journal of Health Economics*, vol 27, pp 826–842.

16 National Audit Office (2008). *NHS pay modernisation: New contracts for general practice services in England.* London: The Stationery Office.

17 NHS Information Centre for Health and Social Care (2007). *National Quality and Outcomes Framework statistics for England 2006/07*. Leeds: NHS Health and Social Care Information Centre.

18 NHS Information Centre for Health and Social Care (2005). *National Quality and Outcomes Framework statistics for England 2004/05*. Leeds: NHS Health and Social Care Information Centre.

19 Roland M (2004). 'Linking physician pay to quality of care: a major experiment in the UK'. *New England Journal of Medicine*, vol 351, pp 1448–54.

20 Russell LB (2009). 'Preventing chronic disease: an important investment, but don't count on cost savings'. *Health Affairs*, vol 28, pp 42–45.

21 Shohet C, Yelloly J, Bingham P and Lyratzopoulos G (2007). 'The association between the quality of epilepsy management in primary care, general practice population deprivation status and epilepsy-related emergency admissions'. *Seizure,* vol 16, pp 351–355.

Appendix

Grounds for exception reporting patients

The Quality and Outcomes Framework includes the concept of exception reporting. This has been introduced to allow practices to pursue the quality improvement agenda and not be penalised, where, for example, patients do not attend for review, or where a medication cannot be prescribed due to a contraindication or sideeffect.

The following criteria have been agreed for exception reporting:

A) patients who have been recorded as refusing to attend review who have been invited on at least three occasions during the preceding twelve months

B) patients for whom it is not appropriate to review the chronic disease parameters due to particular circumstances (for example, terminal illness, extreme frailty)

C) patients newly diagnosed within the practice or who have recently registered with the practice, who should have measurements made within three months and delivery of clinical standards within nine months (for example, blood pressure or cholesterol measurements within target levels)

D) patients who are on maximum tolerated doses of medication whose levels remain suboptimal

E) patients for whom prescribing a medication is not clinically appropriate (for example, those who have an allergy, another contraindication or have experienced an adverse reaction)

F) where a patient has not tolerated medication

G) where a patient does not agree to investigation or treatment (informed dissent), and this has been recorded in their medical records

H) where the patient has a supervening condition which makes treatment of their condition inappropriate (for example, cholesterol reduction where the patient has liver disease)

I) where an investigative service or secondary care service is unavailable.

Source: Adapted from Department of Health (2004)

The Health Foundation is an independent charity working to continuously improve the quality of healthcare in the UK.

We want the UK to have a healthcare system of the highest possible quality – safe, effective, person-centred, timely, efficient and equitable. We believe that in order to achieve this, health services need to continually improve the way they work.

We are here to inspire and create the space for people, teams, organisations and systems to make lasting improvements to health services.

Working at every level of the healthcare system, we aim to develop the technical skills, leadership, capacity, knowledge, and the will for change, that are essential for real and lasting improvement.

Identify Innovate Demonstrate Encourage

The Health Foundation
90 Long Acre
London WC2E 9RA
T 020 7257 8000
F 020 7257 8001
E info@health.org.uk

Registered charity number: 286967
Registered company number: 1714937

www.health.org.uk