QQUIP Quest for Quality and Improved Performance

The link between healthcare spending and health outcomes

Evidence from English programme budgeting data

Stephen Martin,
Department of Economics,
University of York

Nigel Rice and Peter C Smith,
Centre for Health Economics,
University of York

June 2007

The Health Foundation

QQUIP and the Value for Money project

QQUIP (Quest for Quality and Improved Performance) is a five-year research initiative of The Health Foundation. QQUIP provides independent reports on a wide range of data about the quality of healthcare in the UK. It draws on the international evidence base to produce information on where healthcare resources are currently being spent, whether they provide value for money and how interventions in the UK and around the world have been used to improve healthcare quality.

The Value for Money component of the QQUIP initiative provides a series of reports that enable comparisons to be made between the scale of benefits and costs across a number of different disease groups. It also provides a methodological framework for examining the costs and benefits of national policies for treatment of conditions such as coronary heart disease and mental health.

For more information visit www.health.org.uk/QQUIP

Acknowledgements

This study was produced as part of the Quest for Quality and Improved Performance (QQUIP), an initiative of The Health Foundation. We are grateful to Hugh Gravelle, Andrew Jackson, Chris Young, Peter Brambleby and anonymous referees for helpful comments and advice. The material has also benefited greatly from comments from Chris Murray (Harvard Initiative for Global Health), Dean Jamison (University of California, San Francisco) and participants at a seminar at the Harvard School of Public Health.

Published by:

The Health Foundation
90 Long Acre
London WC2E 9RA
Telephone: 020 7257 8000
Facsimile: 020 7257 8001

www.health.org.uk

Registered charity number 286967
Registered company number 1714937

First published 2007

ISBN 978-0-9548968-8-1

Contents

List of tables and figures

Executive summary

This report presents preliminary results from research undertaken as part of the Quest for Quality and Improved Performance (QQUIP), an initiative of The Health Foundation.

For three years, each primary care trust (PCT) in the English National Health Service (NHS) has prepared data on expenditure on healthcare across 23 'programmes' of care, based on the International Classification of Diseases (ICD) Version 10 disease categories. These programme budgeting data seek to allocate exhaustively to disease categories all items of NHS expenditure, including expenditure on inpatient care, outpatient care, community care, primary care and pharmaceuticals. In 2004/05 the average size of the programmes varied considerably, with the three largest being mental health (£145 per head per year), circulatory disease (£122) and cancer (£75).

The programme budgeting data offer immense opportunities for examining the link between healthcare expenditure and health outcomes across PCTs. There is extensive international literature on this topic, but very little solid empirical evidence on the magnitude of the link. Indeed, many authors claim that – at the margin – extra healthcare spending has little impact on health.

The main reason for the lack of evidence is the difficulty of disentangling cause and effect. Areas with high health needs and poor outcomes tend to attract high levels of healthcare spending. This phenomenon is confirmed by examining the link between programme budgeting expenditure and health outcomes (standardised mortality rates or SMRs) among the 303 PCTs (the figure relevant for the study period 2004/05). For example, there is a strong positive correlation between expenditure and under-75 SMRs in cancer and circulatory disease.

However, the question for policy-makers is whether – after adjusting for need – extra spending gives rise to better health outcomes. Addressing this question requires substantial additional data (in order to model needs) and advanced statistical methods. This report examines the link between expenditure and outcomes in two programmes: cancer and circulatory disease. It models both programme expenditure as a function of needs, and then outcomes as a function of expenditure.

These preliminary results are encouraging. For both cancer and circulatory disease programmes, it proved possible to develop robust and well-specified statistical models in line with expectations. They demonstrate a strong positive link between expenditure and better health outcomes (lower SMR) in the two disease categories, and that the link is stronger in circulatory disease than in cancer.

Using a measure of 'years of life lost' instead of SMR as the measure of health outcome, it is also possible to estimate the expenditure required to 'save' a year of life in each disease category. Our estimate is that, for a PCT with average needs and expenditure, the marginal cost of a life year saved in cancer is (on average) about £13,100, and in circulatory disease about £8,000. It must be emphasised that these results have quite large confidence intervals, will vary between PCTs and should be treated with caution. Very importantly, they are not adjusted for quality of life. However, it is noteworthy that they do appear to compare favourably with the threshold of £30,000 per quality adjusted life year (QALY) often attributed to the National Institute for Health and Clinical Excellence (NICE).

These results are useful from a number of perspectives. Scientifically, they challenge the widely held view that healthcare has little marginal impact on health. From a policy perspective, they can help set priorities by informing resource allocation across programmes. They can also help NICE decide whether its current QALY threshold is at the right level.

1. Introduction

One of the most fundamental yet unresolved issues in health policy is the extent to which additional healthcare expenditure yields benefits for patients, in the form of improved health outcomes. The work of health technology agencies such as the English National Institute for Health and Clinical Excellence (NICE) has greatly improved our understanding at the micro-level of the costs and benefits of individual technologies. However, there remains a dearth of evidence at the macro-level on the benefits of increased health system expenditure.

The empirical problems of estimating the link between spending and health outcomes are manifest. If one relies on a time series of health outcome data for an individual health system it is difficult to disentangle the impact of expenditure from a wide range of other temporal influences on health, such as technological advances, epidemiological changes and variations in broader economic circumstances. Similar methodological difficulties arise if one attempts a cross-sectional comparison of different health systems. In particular, when seeking to draw inferences from international comparisons, researchers have found it hard to adjust for all the potential external influences on health outcomes.

Furthermore, there is the possibility that indicators of health system inputs, such as expenditure, are endogenous, in the sense that they have been influenced to some extent by the levels of health outcome achieved in the past. The difficulty of satisfactorily estimating the impact of health system inputs on outcomes may also be in part the result of the great heterogeneity of healthcare, and the rather general nature of the outcome mortality measure traditionally used.

This report takes advantage of a major new dataset developed in English healthcare, in the form of programme budgets, which enables us to address some of the difficulties associated with estimating the impact of healthcare expenditure on health outcomes. The data present expenditure on 23 broad programmes of care at the level of geographically defined local health authorities, known as primary care trusts (PCTs), and embrace most items of publicly funded expenditure, including inpatient, outpatient and community care and pharmaceutical prescriptions. It therefore becomes possible to examine the link between aggregate expenditure in a programme of care and the health outcomes achieved, notably in the form of disease-specific mortality rates.

The report models the link between spending and outcomes in two of the largest programmes of healthcare: circulatory disease and cancer. We start with a brief review of relevant empirical studies, which have often yielded conflicting results. We then describe the programme budgeting data, and present some descriptive statistics. Next, we explain a simple theoretical model of the budgetary problem faced by a PCT manager seeking to allocate limited funds to competing programmes of care. We then develop well-specified econometric models that estimate first, the budgetary expenditure choices and second, the health outcomes achieved by PCTs in the two selected programmes of care. In contrast to many previous studies, the model results show a strong positive impact of expenditure on health outcomes. Finally, from the model results we offer a quantitative estimate of the current marginal cost of a life year saved in the two programmes of care. We conclude by considering the important policy implications of these findings.

2. Previous studies

There is an extensive literature on the determinants of international variations in healthcare spending in which income levels often play a central role (Gerdtham and Jonsson, 2000). However, whether more expenditure generates better outcomes – such as reduced mortality – remains a matter of debate. For example, Fisher and Welch (1999) noted various ways in which more healthcare might harm patients and cited various studies supporting their arguments. In a comprehensive review, Nolte and McKee (2004) discussed many studies that had examined the impact of healthcare and other explanatory variables on some measure of healthcare outcome. Usually, this production function approach employs regression analysis: for example, in an early cross-sectional study of 18 developed countries, Cochrane et al (1978) applied regression analysis to examine the statistical relationship between mortality rates and per capita gross national product and per capita consumption of inputs such as healthcare provision. They found that the indicators of healthcare were generally not associated with outcomes in the form of mortality rates. Thereafter, the failure to identify strong and consistent relationships between healthcare expenditure and health outcomes (after controlling for other factors) has become a consistent theme in the literature, while – in contrast – socio-economic factors are often found to be good determinants of health outcomes (Nolte and McKee, 2004, p 58; Young, 2001; St Leger, 2001).

However, Gravelle and Backhouse (1987) have examined some of the fundamental methodological difficulties associated with empirical investigation of the determinants of mortality rates. These include simultaneous equation bias and the associated endogeneity problem, and the lag between expenditure and outcomes that may occur. To avoid the difficulties imposed by data heterogeneity inherent in international analyses, the study by Cremieux et al (1999) examined the relationship between expenditure and outcomes across ten Canadian provinces over the period between 1978 and 1992. They found that lower healthcare spending was associated with a significant increase in infant mortality and a decrease in life expectancy.

Although challenging the received empirical wisdom, a difficulty with the Cremieux et al (1999) study is that the estimated regression equation consists of a mixture of potentially endogenous variables (such as the number of physicians, health spending, alcohol and tobacco consumption and expenditure on meat and fat) and exogenous variables (such as income and population density). The authors' chosen estimation technique (generalized least squares or GLS) does not allow for this endogeneity and consequently the coefficients on the endogenous variables may be biased (Gravelle and Backhouse, 1987, p 428). Or's (2001) study of the determinants of variations in mortality rates across 21 Organisation for Economic Co-operation and Development (OECD) countries between 1970 and 1995 may suffer from the same weakness. She finds that the contribution of the number of doctors to reducing mortality in OECD countries is substantial but her estimation technique assumes that the number of doctors is exogenous to the health system.

Nixon and Ulmann (2006) have provided a detailed review of 16 studies that have examined the relationship between healthcare inputs and health outcomes, using macro-level data. They also undertook their own study using data for 15 European Union (EU) countries over the period from 1980 to 1995. They employed three health outcomes measures – life expectancy at birth for males and females, and the infant mortality rate – and a dozen or more explanatory variables including: per capita health expenditure, number of physicians (per 10,000 head of population), number of hospital beds (per 1,000 head of population), the average length of stay in hospital, the inpatient admission rate, alcohol and tobacco consumption, nutritional characteristics and environmental pollution indicators. Nixon and Ulmann conclude that, although health expenditure and the number of physicians have

made a significant contribution to improvements in infant mortality, ‘health care expenditure has made a relatively marginal contribution to the improvements in life expectancy in the EU countries over the period of the analysis’ (Nixon and Ulmann, 2006, p14). Again, however, the study does not allow for the possibility that some of the explanatory variables may be endogenous.

Although loosely based on the notion of a health production function, the traditional empirical study described above rarely has been informed by an explicit theoretical model. This is understandable, as the processes giving rise to observed health outcomes are likely to be very complex, and any theoretical model will become unwieldy. However, it leads to an atheoretical search for measures demonstrating a statistically ‘significant’ association with health outcomes. In contrast, in this study we seek to inform our empirical modelling with a theoretical model. We believe that this may lead to a more convincing and better specified model of health outcomes than that used in many previous studies.

3. Programme budgeting in England

The English NHS is an archetypal centrally planned and publicly funded health system. Its revenue derives almost entirely from national taxation and access to the system is generally free to the patient. Primary care is an important element of the system and general practitioners (GPs) act as gatekeepers to secondary care and pharmaceuticals. The system is organised geographically with the responsibility for the local administration of the NHS devolved to local health authorities known as primary care trusts (PCTs). In 2004/05 – the year relevant to this study – there were 303 PCTs with average populations of 160,000. PCTs are allocated fixed annual budgets by the national Department of Health, within which they are expected to meet expenditure on most aspects of healthcare, including inpatient, outpatient and community care, primary care and prescriptions.

3.1 The rationale behind programme budget data

Traditionally, PCTs have reported expenditure on the basis of inputs (for example, total expenditure on pay and non-pay items). However, for some time NHS policy-makers have realised that this approach does not create clinically meaningful financial data or help in the design and evaluation of programmes of patient care. It therefore initiated a 'programme budgeting' project, which has sought to create an accounting system that is more aligned with the distinct outputs and health outcomes of the health system. Since April 2003, in addition to its conventional accounting data, each PCT has prepared expenditure data disaggregated according to 23 programmes of healthcare. These programmes are defined by reference to the International Classification of Diseases Version 10 (ICD-10) codes at the four digit level, and most programme budget categories reflect ICD-10 chapter headings (for example, cancer and tumours, circulation problems, renal problems, neonates, problems associated with the skin, vision, hearing and so on). In some cases, the 23 categories are broken down into further sub-areas to achieve a closer match with the various national service frameworks: for example, the large mental health category is broken down into 'substance misuse', 'dementia' and 'other'.

Programme budgeting seeks to allocate all types of PCT expenditure to the various programme budget categories, including secondary care, community care and prescribing. However, the system acknowledges that a medical model of care may not always be appropriate, and two specific non-clinical groups – 'healthy individuals' and 'social care needs' – have been created. These are intended to capture the costs of disease prevention programmes and the costs of services that support individuals with social rather than healthcare needs. In addition, in some cases it is not possible to assign activity by medical condition, preventative activity or social care need, in which case expenditure is assigned to a category entitled 'other'. The most important element of this programme is expenditure on GP services. The use of this category ensures all expenditure can be assigned to a programme of care (Department of Health, 2005a, p 7).

The aim of the programme budget classifications is to identify the entire volume of healthcare resources assigned to broad areas of illness according to the primary diagnosis associated with an intervention. It serves a number of purposes, most notably to assist in the local planning of healthcare. But its crucial merit for this study is that it opens up the possibility of examining the statistical relationship between local programme spending and associated disease-specific outcomes.

3.2 The collection of programme budget data

Programme budgeting information is collected centrally by the Department of Health as part of the annual accounts process. Each PCT is required to submit an annual programme budgeting return to the department that shows how its total expenditure is allocated across the 23 programme budgeting categories.

Various forms of data collection and analysis are required to map PCT expenditure on acute, community and other services to the 23 programme budget categories. From the PCT perspective, however, the construction of each PCT's return largely involves collating information provided by other bodies and drawing on other information already in the PCT's own annual accounts. Thus General/Personal Medical Service expenditure, which is already reported in PCT accounts, relates to direct primary care and is mapped in its entirety to programme budget category 23a ('other'); General Ophthalmic Service expenditure (again from PCT accounts) maps directly to programme budget category 8 ('eye/vision problems'); and General Dental Service expenditure maps directly to programme budget category 12 ('dental problems'). Prescribing and pharmaceutical services expenditure is allocated to programme budget categories on the basis of an annual apportionment report provided by the Prescription Pricing Authority for each PCT as part of the annual accounts process. This apportionment report allocates each PCT's annual Family Health Services (FHS) prescribing expenditure across the 23 programme budget categories. The balance of any primary healthcare purchased by the PCT is apportioned across the 23 programme budget categories on the basis of local records, with any remaining expenditure allocated in line with the distributions already made across the budget categories.

It is the responsibility of all NHS providers – which includes PCTs, NHS trusts and foundation hospitals – to allocate admitted patient care expenditure across the programme budgeting categories, specific to each PCT that utilises its services. These allocations are constructed using 'finished consultant episodes' (FCEs) from the mandatory administrative Hospital Episode Statistics data set returned by all providers, each of which is assigned to a Healthcare Resource Group (HRG), an English version of diagnosis related groups (or DRGs). National grouping software automatically assigns each HRG to one of the 23 programme budgeting categories and attaches the provider's average reference cost for the relevant HRG to each record. For each PCT this information generates a split of inpatient care expenditure by programme budget category for each of its secondary healthcare providers.

There are numerous difficulties faced when attempting to allocate non-admitted patient care activity (that is, outpatients, community services, direct access, A&E and so on) to programme budget categories. The difficulties are primarily due to the absence of clear diagnostic codes. The 'primary reason for care' (equivalent to a diagnosis code) is information that is not routinely collected for community patients. Because of this, the approach prescribed is for service providers to produce a generic allocation analysis/report, for all PCTs making use of their services, for all non-admitted patient care costs across the 23 programme budget categories. Once derived, this generic allocation analysis/report is made available to PCTs at the same time as the unique (PCT-specific) inpatient care information described above. Unlike the first apportionment report relating to admitted patient care, the non-admitted patient care apportionment report will not be unique to the PCT, but will represent the provider's overall experience. PCTs are expected to use this data to inform the apportionment of their own spend on non-admitted patient care across the 23 programme budget categories.

The Department of Health recognises that this approach – the provision of a PCT-specific breakdown of admitted patient care costs and a generic allocation of all PCTs non-admitted

patient care spend by providers – is likely to generate a crude method for apportioning non-admitted patient care costs. Therefore PCTs and their providers are encouraged to put in place other arrangements that allow a more sophisticated analysis of non-admitted patient care spend. Such arrangements may well rely on an activity sampling approach (Department of Health, 2005a, p 33).

Mental health providers may not need to complete and forward detailed admitted and non-admitted patient care apportionment reports to PCTs. The nature of the services they provide may be such that the entire spend with them relates exclusively to the 'mental health' programme budget category. Ambulance trusts are required to provide non-admitted patient care information to those PCTs for whom they provide services. Where it is not possible to split the activity by PCT, a generic non-admitted patient care report is produced for all purchasers (Department of Health, 2005a, p 24).

The Department of Health has been criticised for the rather simplistic way in which it has apportioned certain costs among categories, and there are obvious issues with the allocation of costs associated with patients who have multiple disorders. However, the programme budgeting project is very much work-in-progress and the department is investigating ways to improve the accuracy with which costs are allocated across programmes (for example, the department is investigating the possibility of allocating training expenditures to specific programmes rather than to a generic medical training programme).

3.3 Programme budget data for 2004/05: all England

Programme budgeting information was first collected in the financial year 2003/04, and here we report information for the second year of implementation, 2004/05 (information for 2005/06 was released on 29 November 2006). The first column of Table 1 shows the national average NHS expenditure per person by programme budget category. Across England as a whole, NHS expenditure per person is £1,183. The single largest category is the 'other' category (category 23) with expenditure per person of almost £155 in 2004/05. This category includes primary care expenditure, workforce training expenditure and a range of other miscellaneous expenditure items. Primary care expenditure is by far the largest element at £127 per head.

There are two other categories with expenditure of over £100 per head: mental health problems (category 5) attract an annual expenditure of £145 per person and circulation problems (category 10) receive £122 per person. Next come four programme budget categories – cancers and tumours, gastro-intestinal problems, musculoskeletal problems, and trauma, burns and injuries – with an annual expenditure of between £71 and £75 per person. Respiratory and genito-urinary problems both record an expenditure of £62 per person, with maternity and reproductive conditions being allocated £55 per head. There is no evidence of major shifts in the data from 2003/04, suggesting they are reasonably stable.

Table 1: Expenditure by programme budget category, per person, all England, 2004/05 and descriptive statistics for cost adjusted expenditure by PCT, 2004/05

	Programme budget category	National net spend per head, £, 2004/05 (see below)	3PCT spend per head £, 2004/5, cost adjusted			
			Mean	Minimum	Maximum	CV
1	Infectious diseases	20.1	18.6	8.9	137.6	0.68
2	Cancers/tumours	75.1	75.8	39.1	133.4	0.21
3	Blood disorders	16.9	16.4	3.8	58.1	0.46
4	Endocrine/metabolic	31.7	31.7	12.4	51.5	0.18
4a	Diabetes	13.5	13.4	0.0	33.3	0.34
4x	Other	18.2	18.2	0.0	40.9	0.30
5	Mental health	145.3	142.9	51.2	323.3	0.28
5a	Substance abuse	11.9	12.2	-2.0	146.8	1.37
5b	Dementia	16.1	16.3	0.0	158.3	1.28
5x	Other	117.3	114.3	0.0	247.8	0.34
6	Learning disability	42.0	42.5	4.7	163.3	0.46
7	Neurological system	34.9	35.5	18.6	70.6	0.24
8	Eye and vision	27.5	28.2	4.5	65.7	0.30
9	Hearing	6.3	6.3	1.7	32.7	0.47
10	Circulation (CHD)	122.0	124.1	64.0	186.3	0.19
11	Respiratory	62.5	63.7	30.3	147.6	0.25
12	Dental	13.3	13.4	0.0	96.4	0.80
13	Gastro-intestinal	73.0	74.4	34.4	132.3	0.22
14	Skin	24.8	24.9	13.2	49.7	0.27
15	Musculoskeletal	71.2	72.3	19.1	157.6	0.23
16	Trauma/injuries	71.9	72.7	35.2	209.1	0.26
17	Genito/urinary	62.1	61.6	30.8	151.3	0.27
18	Maternity/reproductive	54.7	53.8	25.1	151.3	0.31
19	Neonate conditions	13.9	13.8	0.3	53.2	0.53
20	Poisoning	12.3	12.5	4.2	24.5	0.28
21	Healthy individuals	21.7	21.5	4.2	90.1	0.51
22	Social care needs	25.1	24.5	-80.4	140.1	0.85
23	Other areas	154.7	156.8	98.2	574.2	0.29
23a	GMS/PMS*	126.9	128.8	90.8	237.4	0.14
	All categories	1183.1	1188.1	820.2	1705.9	0.13

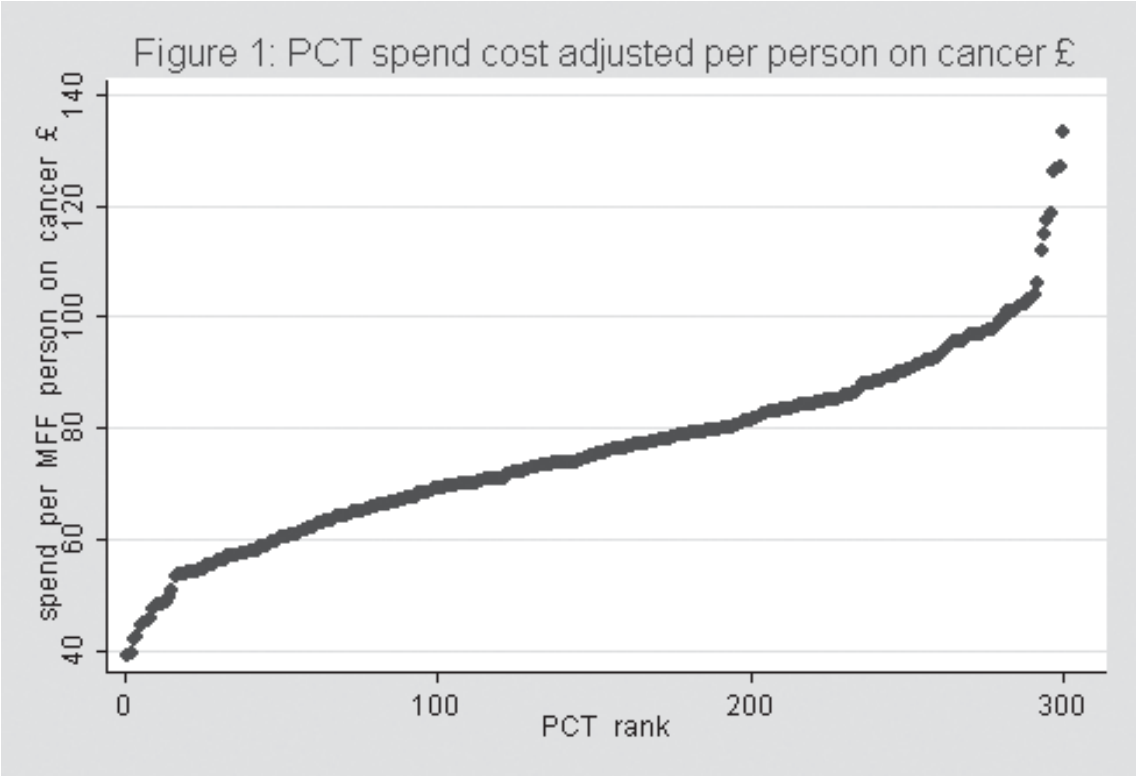
NB Descriptive statistics across PCTs are unweighted for population size and, for any given PCT, its expenditure per head figures reflect its raw population adjusted for unavoidable cost variations. The coefficient of variation (CV) is a measure of dispersion and is calculated as the standard deviation divided by the mean.

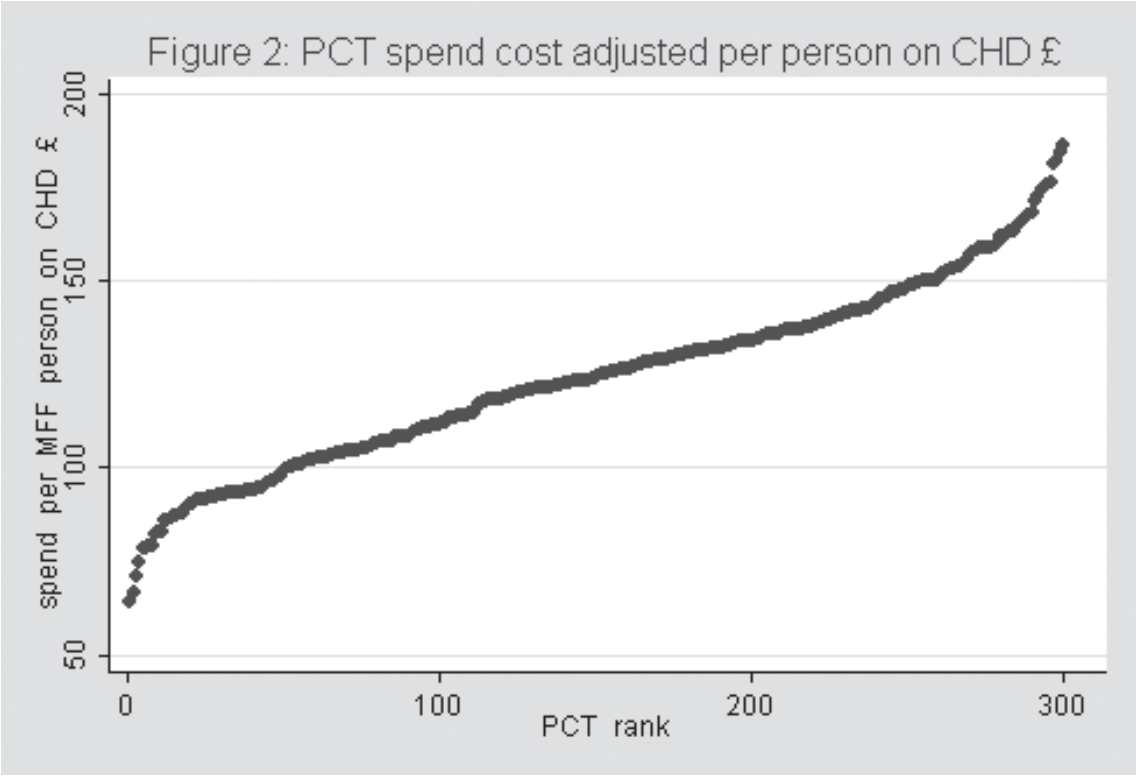
*The general medical services/personal medical services (GMS/PMS) figures exclude three PCTs where the reported expenditure figures are either zero or implausibly low.

The remaining columns of Table 1 indicate the variation in expenditure among PCTs. For each programme budget category, the PCT's per capita expenditure is adjusted for unavoidable geographical variation in costs. This is necessary because input prices in London and the south east of England are up to 30 per cent higher than elsewhere. The cost

adjustment is achieved by adjusting raw figures according to a price index reflecting input costs in the local health economy (the Hospital and Community Health Services Market Forces Factor, in Department of Health, 2005b). The unweighted average of these PCT expenditure per capita figures – adjusted for unavoidable geographical variation in costs – is reported for each programme budget category in the second column of Table 1, followed by the observed minimum and maximum. The final column shows the coefficient of variation.

The variation in expenditure levels across PCTs is considerable. For example, expenditure per head on cancers and tumours averages £76 across all PCTs but this varies between £39 and over £133 per head (see also Figure 1 which provides a plot of per capita PCT expenditure on cancer). Similarly, expenditure per head on circulation problems averages £124 across all PCTs but this varies between £64 and over £186 per head (see also Figure 2 which provides a plot of per capita PCT expenditure on circulation problems). Although there is some variation within these two particular programme budget categories, it is small relative to other programmes of care. Categories such as infectious diseases and blood disorders have much larger coefficients of variation, indicating substantially more variation than in the cancer and circulation categories.

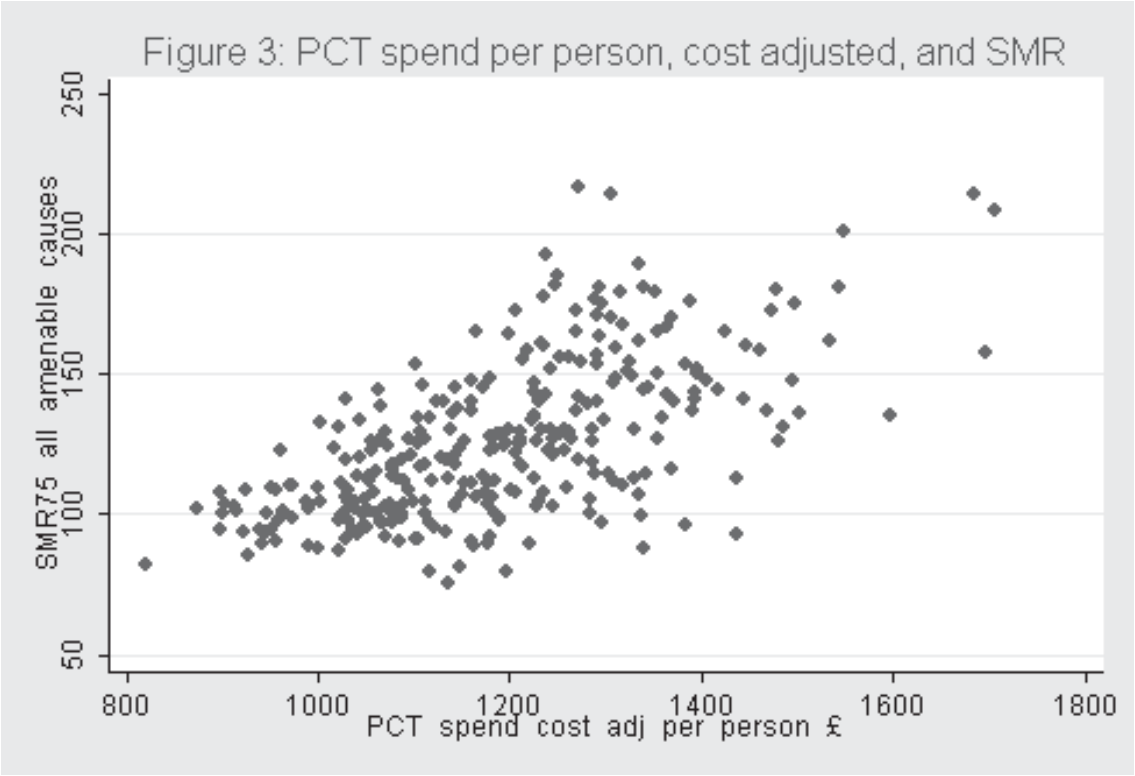




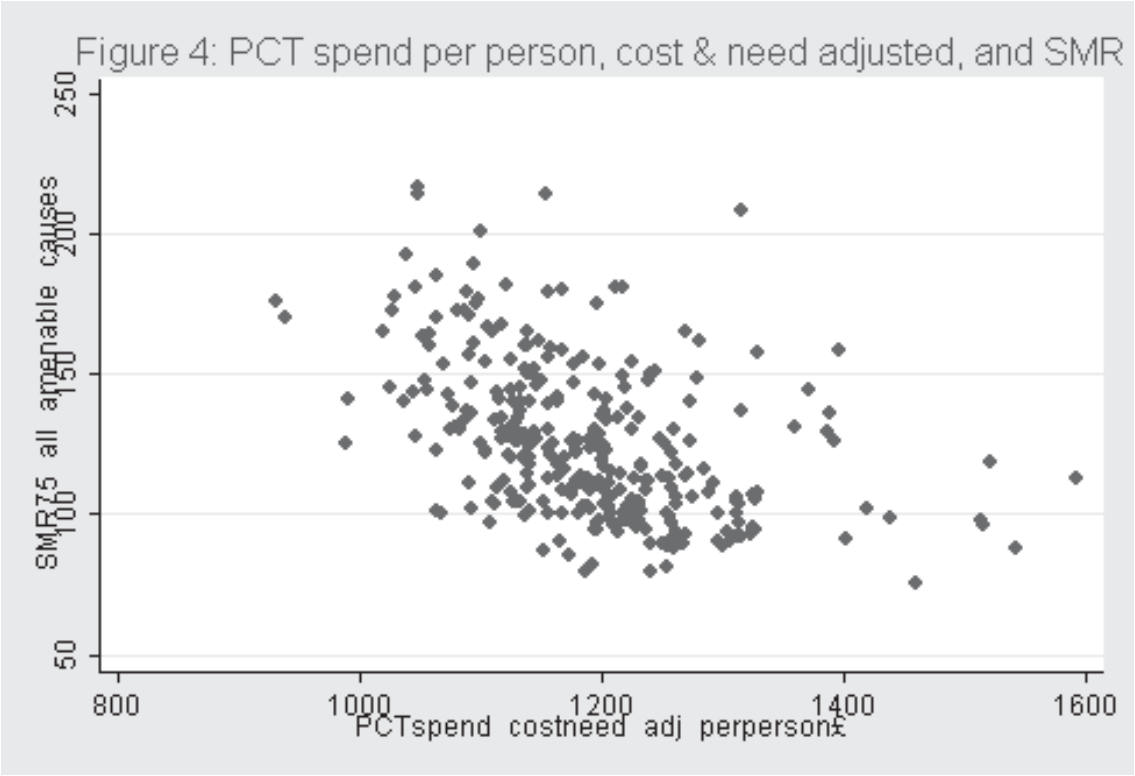
3.4 The association between expenditure and mortality rates

Figure 3 plots total PCT per capita expenditure (adjusted for local input price conditions) against the PCT mortality rate for those aged under 75 for deaths from all causes amenable to healthcare from 2002 to 2004 (see Table A1 in Appendix A for details of the causes defined as 'amenable to healthcare'). Figure 3 reveals a clear positive relationship between these two variables and the correlation coefficient is 0.624. A similar but slightly weaker positive relationship (not shown here) also exists between PCT per capita expenditure on cancer services and the cancer SMR for those aged under 75 (the correlation coefficient is 0.213). The relationship between PCT per capita expenditure on circulatory problems and the circulatory standardised mortality rates (SMR) for those aged under 75 is also positive and slightly stronger than that for cancer (the correlation coefficient between expenditure and deaths is 0.304).

Thus, as is frequently the case, the programme budgeting data indicate a strong positive relationship between healthcare spending and adverse outcomes, apparently contradicting the hypothesis that PCTs that spend more on healthcare will achieve better health outcomes. However, the interpretation of this finding is not straightforward, as much of the variation in expenditure across PCTs will reflect different levels of the need for healthcare. Areas with a large proportion of older residents, or operating in deprived locations, can be expected to experience relatively high levels of spending. Adjusting for the relative healthcare needs of different populations is therefore a central requirement of any analytic effort in this domain. Fortunately the Department of Health has a well-developed methodology for estimating the relative healthcare needs of PCTs, in the form of the weighted capitation formula it uses as the basis for allocating healthcare funds to PCTs (Smith et al, 2001). The current 'needs' formula is derived from an adjustment for the demographic profile of the PCT and a series of econometric analyses of the link between healthcare expenditure and other socio-economic factors at a small area level within England (Department of Health, 2005b).



We therefore use the current Department of Health needs and demographic adjustments as the basis for adjusting raw PCT expenditure for the healthcare 'needs' of its population. The plot in Figure 4 is similar to that in Figure 3, but holds constant the local need for healthcare by dividing expenditure by the index of needs used by the Department of Health. It therefore plots total expenditure per capita (adjusted for local cost and need conditions) against the SMR for those aged under 75 from all causes of death amenable to healthcare. The positive association between expenditure and deaths (with a correlation coefficient of 0.624), as shown in Figure 3, is now dramatically reversed (with a correlation coefficient of -0.451). That is, Figure 4 offers evidence to suggest that – once the need for healthcare is held constant – more expenditure is associated with a better outcome (a lower death rate). Similar results are obtained in specific programmes of care when allowance is made for local cost and need conditions: for example, the correlation coefficient between expenditure and the mortality rate becomes -0.323 for cancers and tumours, and it becomes -0.358 for circulation problems.



Adjusting for cost and need dramatically affects the bivariate relationship between expenditure and outcomes. It also reduces the variation in per capita expenditure across PCTs, although this reduction, as measured by the coefficient of variation, is modest and there remain substantial differences in per capita expenditure levels across the country. For example, for cancer and tumours the minimum and maximum spend per head is £39 and £139 respectively using cost adjusted expenditure data, but £34 and £136 using cost and need adjusted data. Similarly, expenditure per head in the circulation problems category varied between £66 and £183 with expenditure adjusted for local cost conditions, but falls between £66 and £171 using cost and need adjusted data.

This section suggests that, after adjusting for need, healthcare expenditure may have a distinctly beneficial impact on health outcomes. However, our adjustment is crude and without more detailed modelling it is impossible to say whether we have correctly captured the influence of need on spending and outcomes. It is for this reason that the statistical analysis reported in later sections must be undertaken.

3.5 Correlates with cost adjusted per capita expenditure levels

Table 2 reports the degree of correlation between various socio-economic indicators and per capita PCT expenditure (adjusted for unavoidable geographical variations in cost) for the three largest programme budget categories (cancer services, mental health problems and circulation problems) using PCT expenditure data for 2004/05. With about 300 observations, a correlation coefficient of 0.113 is significant at the 5 per cent level and a coefficient of 0.148 is significant at the 1 per cent level. These data become very important in the detailed modelling described in later sections.

Although many of the indicators of deprivation are strongly associated with expenditure per head, cancer services generally exhibit a lower level of association than mental health problems or circulatory disease. Both cancer and circulatory disease expenditure are highly

correlated with the proportion of households that are one pensioner households, possibly reflecting the fact that patients from single pensioner households have longer stays in hospital and require more medical and social support once they leave hospital.

In contrast, there is virtually no correlation between the proportion of households that are one pensioner households and expenditure per head on mental health problems. For this programme budget category, expenditure is most highly correlated with the proportion of households without a car – a standard measure of deprivation – and expenditure is also strongly correlated with several other low income and high deprivation measures including the proportion of prescriptions that are exempt from charges (positively) and the proportion of all households that are owner occupied (negatively).

Table 2: Correlation between expenditure (cost adjusted) per head (2004/05) and various socio-economic indicators for three programme budgeting categories, across all PCTs

	Expenditure (cost adjusted) per head on:		
	cancers/ tumours	mental health	circulation problems
Proportion of residents born outside EU	-0.171	0.374	-0.290
Proportion of residents in white ethnic group	0.170	-0.362	0.218
Proportion of working age population with long-term illness	0.233	0.382	0.558
Proportion of population providing some unpaid care	0.234	-0.139	0.558
Proportion of population providing <20 hours week unpaid care	0.069	-0.437	0.157
Proportion of population providing 20-49 hours week unpaid care	0.193	0.274	0.558
Proportion of population providing >50 hours week unpaid care	0.281	0.204	0.634
Proportion of population aged 16-74 no qualifications	0.212	0.170	0.487
Proportion of population aged 16-74 full-time students	-0.006	0.415	-0.164
Proportion of households without a car	0.200	0.669	0.273
Proportion of households that are owner occupied	-0.121	-0.663	-0.079
Proportion of households that rented from LA or HA	0.096	0.536	0.123
Proportion of households rented from private landlords	0.091	0.453	-0.068
Proportion of households that are one pensioner households	0.417	0.072	0.610
Proportion of households that are one parent households	0.033	0.528	0.251
Proportion of population aged 16-74 that are permanently sick	0.229	0.389	0.561
Proportion of population aged 16-74 that are long-term unemployed	0.160	0.538	0.316
Proportion in employment working in agriculture	0.111	-0.231	0.101
Proportion in employment working in management/ professions	-0.263	-0.165	-0.565
Population weighted ward-based Index Multiple Deprivation 2000	0.200	0.525	0.408
Exemptions from prescription charges (LISI 2002)	0.114	0.610	0.241

Sources: Data are from Population Census 2001 and Prescribing Support Unit. For further details about these indicators see Table A2.

3.6 Conclusion

The Department of Health's programme budgeting project has allocated all PCT expenditure to one of 23 mutually exclusive categories of illness according to the primary diagnosis associated with an intervention. This data set opens up the possibility of examining the statistical relationship between local programme spending and associated disease-specific outcomes. We have found evidence of a strong positive bivariate association between per capita expenditure (adjusted for unavoidable geographical variations in costs) and the mortality rate for three programmes of care. However, the interpretation of this result is not straightforward because some of the variation in expenditure across PCTs will reflect different levels of the need for healthcare. When per capita expenditure is adjusted for variations in local cost and need, the association between expenditure and mortality becomes negative implying that more expenditure may be associated with a better outcome (less mortality).

The remainder of this report seeks to integrate the rudimentary findings illustrated in this section into a coherent model of expenditure and outcomes, and to estimate the strength of the relationships suggested in the findings. The next section therefore presents a theoretical model of PCT expenditure allocation across the 23 programme budgeting categories. Section 5 presents our empirical results.

4. Theoretical model

We assume that each PCT i receives an annual financial lump sum budget yi from the national Department of Health, and that total expenditure cannot exceed this amount. The PCT must then decide how to allocate the budget across the J programmes of care (J=23 in this case). For each programme of care there is a 'health production function' fi(.) that indicates the link between local spending xij on programme j and health outcomes in that programme hij. Health outcomes might be measured in a variety of ways, but the most obvious is to consider some measure of improvement in life expectancy, possibly adjusted for quality of life, in the form of a quality adjusted life year (QALY).

The nature of the specific health production function confronted by a PCT will depend on two types of local factors: the clinical needs of the local population relevant to the programme of care (which we denote nij) and broader local environmental factors zij relevant to delivering the programme of care (such as input prices, geographical factors or other uncontrollable influences on outcomes). Both clinical and environmental factors may be multidimensional in nature. Increased expenditure then yields improvements in health outcomes, as expressed for example in improved local mortality rates, but at a diminishing rate. That is:

$$h_{ij} = f_j(x_{ij}, n_{ij}, z_{ij}); \partial f_j / \partial x > 0; \partial^2 f_j / \partial x^2 < 0 \quad (1)$$

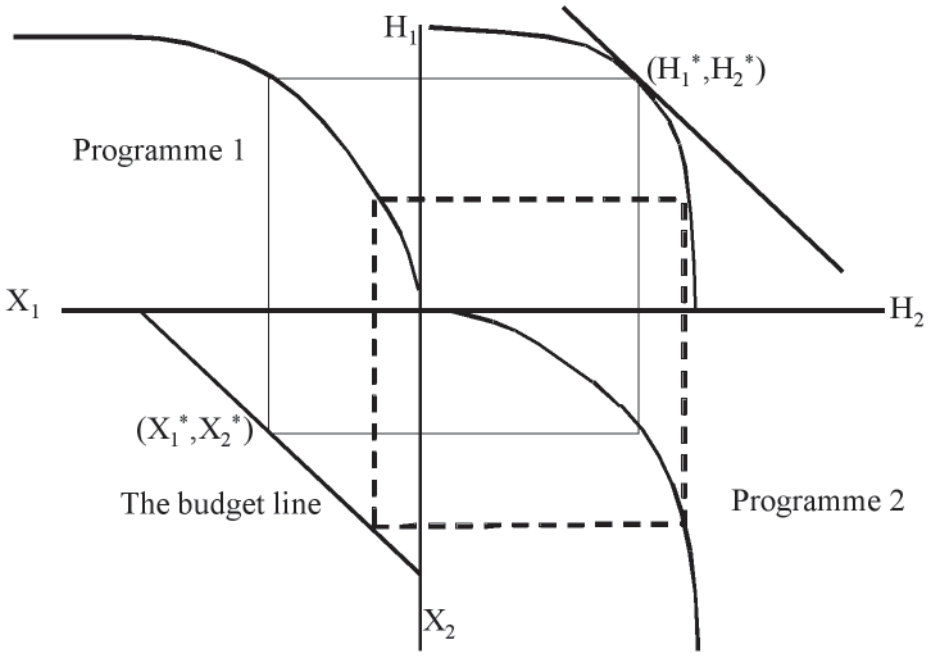
We assume there is a PCT social welfare function W(.) that embodies health outcomes across the J programmes of care. Assuming no interaction between programmes of care, each PCT allocates its budget so as to maximise total welfare subject to local budget constraint and the health production functions for each programme of care:

$$\begin{aligned} &\max && W(h_{i1}, h_{i2}, \ldots, h_{iJ}) \\ &\text{subject to} && \sum_j x_{ij} \le y_i \\ & && h_{ij} = f_j(x_{ij}, n_{ij}, z_{ij}); \quad \text{j} = 1, \ldots \text{J} \end{aligned} \qquad (2)$$

Of course it can be argued quite plausibly that decision-makers do not discriminate between health outcomes in different programmes of care, and that W(.) is merely the sum of such outcomes. However, there is no need for that assumption in our formulation.

Each PCT allocates expenditure across the 23 programmes of care so that the marginal benefit of the last pound spent in each programme of care is the same. This can be represented diagrammatically: Figure 5 shows the trade-off between just two programmes of care. The top left hand quadrant indicates the health production function for programme 1, while the bottom right hand quadrant indicates the health production function for programme 2. The bottom left hand quadrant indicates the budget constraint: the expenditure choice must lie on the budget line. This means that for each feasible pair of expenditure choices (points on the budget constraint line) a pair of health outcomes in the two programmes emerges, which is traced out as the production possibility frontier in the top right quadrant. The PCT will choose the point on this frontier that maximises welfare. In this example, we have indicated a simple health maximising approach (the maximum health summing across the two programmes), leading to optimal health outcomes (H1*, H2*) and expenditure (X1*, X2*).

Figure 5: Optimal trade-off between two programmes of care



Solving the constrained maximisation problem yields the result that the optimal level of expenditure in each category, x_{ij*}, is a function of the need for healthcare in each category (n_{i1}, n_{i2},..., n_{iJ}), environmental variables affecting the production of health outcomes in each category (z_{i1}, z_{i2},..., z_{iJ}), and PCT income (y_i). Thus

$$x_{ij}^{*} = g_i(n_{i1},\ldots n_{iJ}, z_{i1},\ldots z_{iJ}, y_i); \quad j = 1,\ldots,J \qquad (3)$$

Thus, for each programme of care there exists an expenditure equation (3) explaining expenditure choice of PCTs and a health outcome equation (1) that models the associated health outcomes achieved. We estimate these equations empirically for two programmes of care: cancer and circulatory disease. This requires the use of a suite of advanced econometric tools, as described in Appendix B. These are required to ensure that the models are correctly specified and that none of the fundamental modelling assumptions is violated.

Our model is static, in the sense that the health production function (1) assumes that all health benefits occur contemporaneously with expenditure. We acknowledge that, for some programmes of care, benefits might occur one or more years after expenditure has occurred. This is particularly likely to be the case for those programmes aimed at encouraging healthy lifestyles, where some benefits may occur decades after the actual programme expenditure. For other programmes, such as maternity/reproductive conditions and neonate conditions, benefits may be largely contemporaneous with expenditure. Furthermore, we do not model the decision-maker's time preferences. However, for our empirical modelling we are constrained by the data we have available, which are largely cross-sectional in nature. Implicitly we have to assume that the data represent a quasi long-run equilibrium position, and that relative expenditure levels and health outcomes within each PCT have been reasonably stable over a period of time. In the English context this appears to be a reasonable assumption, although as panel data become available it will become possible to develop a more dynamic model.

5. Empirical results

We first present results for the cancer programme of care. These are provided in Table 3. Columns under (1) present ordinary least squares (OLS) results, columns under (2) present two-stage least squares (2SLS) using SMRs as the measure of health outcomes, while columns under (3) present 2SLS estimates using years of life lost (YLL) as the outcome measure. All variables have been log transformed and accordingly parameter estimates can be interpreted as elasticities.

Conventional OLS results are presented as a starting point, but they are unlikely to be well-specified because they ignore the possibility that some of the explanatory variables may be 'endogenous' to the system of equations, as discussed in Section 1. They suggest that both cancer death rates and cancer expenditure are positively correlated with need. However, although the results indicate that expenditure on cancer services is negatively related to cancer deaths, the effect is very small and fails to achieve significance. Expenditure on cancer is also negatively related to other calls on expenditure – the non-cancer death rate – as proxied here by the circulatory disease death rate. The estimated coefficient suggest that a 10 per cent increase in calls on other expenditures, provided through a corresponding increase in circulatory death rates, results in a 3.8 per cent reduction in cancer expenditure. We further observe a large and positive relationship between expenditure on cancer services and total PCT budget, indicating that a 10 per cent increase in budget leads to a 9.3 per cent increase in cancer expenditure. This suggests that increases in income are distributed across programme budgets approximately in equal proportion to existing allocations. This would appear rational and provides additional face validity to the model specifications. It should be noted that the need variable employed here is a generic, all condition, measure of the need for healthcare. Ideally we should use a condition (cancer) specific measure of need. However, at the time of this study no such indicator was readily available and, as we shall see, the 2SLS models for cancer deaths show no evidence of omitted variable bias.

Table 3: Results for cancer programme of care

N = 295	OLS (1)		2SLS (2)		2SLS (3)	
	Cancer deaths	Cancer expenditure	Cancer deaths	Cancer expenditure	Cancer YLL	Cancer expenditure
Constant	4.966 (.103)	-.546 (1.171)	6.919 (.419)	.751 (1.267)	6.712 (.364)	.725 (1.302)
Need	.684 (.034)	.305 (.167)	.916 (.068)	.588 (.197)	.845 (.059)	.654 (.212)
Cancer expenditure	-.038 (.024)		-.491 (.097)		-.378 (.085)	
Total income		.933 (.153)		.874 (.155)		.877 (.159)
Non-cancer deaths		-.383 (.071)		-.576 (.099)		
Non-cancer YLL						-.556 (.100)
Test statistics						
Sargan (χ^2_1)			1.575 (.210)	.314 (.575)	2.750 (.097)	1.357 (.244)
Anderson (χ^2_1)			42..23 (.000)	214.2 (.000)	42.23 (.000)	177.4 (.000)
Cragg-Donald			22.39 (<.05)	154.7 (<.05)	22.39 (<.05)	119.6 (<.05)
Partial R^2			.133	.516	.133	.452
Reset:						
F(3,289)	11.29 (.000)					
F(3,288)		1.68 (.171)				
Pesaran-Taylor (χ^2_1)			.15 (.985)	.33 (.954)	.03 (.998)	.03 (.999)
Endogeneity (χ^2_1):						
Cancer expenditure			55.88 (.000)		32.63 (.000)	
Non-cancer deaths				7.92 (.005)		
Non-cancer YLL						12.18 (.000)

Note: Parentheses show standard errors for parameter estimates and p-values for the statistics. The instrument set for cancer expenditure and non-cancer deaths includes the proportion of households that are one pensioner households and the proportion of the population providing unpaid care.

The second set of results in Table 3 present 2SLS estimates for SMRs, which allow for the possibility that some of the explanatory variables may be endogenous. These results suggest that both cancer deaths and expenditure are more elastic with respect to health needs compared to the OLS. However, expenditure becomes less responsive to non-cancer deaths (.993 for OLS compared to 0.874 for 2SLS). The main difference between OLS and 2SLS is, however, the increased negative coefficient on cancer expenditure and its relationship with cancer deaths. This change is to be expected as 2SLS treats expenditure as endogenous to health outcomes. This means that the possibility that expenditure may be influenced by measured health outcomes is modelled, as well as the more obvious influence of expenditure on outcome. The 2SLS results indicate that a 10 per cent increase in cancer programme expenditure results in approximately a 5 per cent reduction in adverse health outcomes, observed through cancer deaths.

Substituting years of life lost for SMRs (equations (3)) results in substantively similar results. Moreover, they allow us to calculate the implied marginal 'cost' of saving a life year in the cancer disease category given the national average levels of expenditure and mortality. They suggest that a 1 per cent increase in cancer expenditure per head – which was £75.10 in 2004/05 – gives rise, other things being equal, to a 0.378 per cent reduction in years of life lost. From 2002 to 2004, total life years lost to cancer deaths in those aged under 75 was 2,268,541. This averaged 756,180 life years per annum which, across the English population of roughly 50 million, averages out at 0.015 life years (5.52 days) per person. Thus a 1 per cent increase in expenditure per head (£0.751) is associated with a 0.378 per cent reduction in life years lost (0.021 days) and implies that an extra life year would cost £13,137 in a PCT exhibiting national average levels of need and expenditure. Using the estimated standard errors suggests that the 95 per cent confidence interval surrounding this estimate is quite large (£9,118 to £23,490).

There is clear evidence that the OLS model is misspecified ($F(3,289) = 11.29;\ p = .000$) and it should therefore be rejected in favour of the 2SLS models. Further support for the 2SLS models is provided through the Sargan test of overidentifying restrictions, the Anderson and Cragg-Donald tests of instrument relevance and the partial R-squared values from the first stage regressions of the set of exogenous variables on the relevant endogenous variable (see Appendix B for explanations of these tests). These tests indicate that the instrument set is both valid and relevant. Further, the assumption of exogeneity of deaths or expenditure can be rejected in all models.

The analogous results for circulatory diseases are shown in Table 4. In general, the set of estimated coefficients exhibit the same qualitative characteristics as for cancer, but are more elastic than the corresponding cancer results. For example, a 10 per cent increase in health need results in an increase in circulatory expenditure of between 6 per cent (OLS) and 12.8 per cent (2SLS for years of life lost). Again as we move from OLS to 2SLS we observe an increase in the absolute value of the estimated coefficients attached to the endogenous regressors. For the 2SLS circulatory deaths model we observe a near three and a half fold increase over OLS in the estimated coefficient on circulatory expenditure. Further, the coefficient of -1.387 implies that circulatory deaths are more responsive to increases in expenditure compared to cancer. A 10 per cent increase in expenditure is associated with a 13.9 per cent reduction in the death rate. Similarly, the coefficient on non-circulatory deaths in the circulatory expenditure model increases from -.400 for OLS to -1.052 for 2SLS.

Table 4: Results for circulatory diseases programme of care

N = 295	OLS (1)		2SLS (2)		2SLS (3)	
	Circulatory deaths	Circulatory expenditure	Circulatory deaths	Circulatory expenditure	Circulatory YLL	Circulatory expenditure
Constant	6.492 (.245)	1.072 (.911)	11.23 (.728)	4.49 (1.242)	11.57 (.766)	6.63 (1.633)
Need	1.595 (.068)	.606 (.115)	2.450 (.153)	1.069 (.161)	2.652 (.161)	1.283 (.203)
Circulatory expenditure	-.402 (.051)		-1.387 (.151)		-1.427 (.159)	
Total income		.804 (.109)		.764 (.118)		.716 (.130)
Non-circulatory deaths		-.400 (.092)		-1.052 (.176)		
Non-circulatory YLL						-1.349 (.239)
% White ethnic group		.372 (.050)		.369 (.054)		.365 (.059)
Test statistics						
Sargan (χ^2_1)			4.113 (.128)	4.273 (.118)	5.034 (.081)	1.699 (.428)
Anderson (χ^2_2)			86.92 (.000)	111.4 (.000)	86.92 (.000)	68.93 (.000)
Cragg-Donald			33.12 (<.05)	44.04 (<.05)	33.12 (<.05)	25.27 (<.05)
Partial R^2			.255	.315	.255	.208
Reset:						
F(3,289)	2.09 (.035)					
F(3,288)		2.17 (.092)				
Pesaran-Taylor (χ^2_3)			6.78 (.08)†	0.42 (.936)	4.45 (.217)	1.33 (.723)
Endogeneity (χ^2_1):						
Circulatory expenditure			129.5 (.000)		112.4 (.000)	
Non-circulatory deaths				23.48 (.000)		
Non-circulatory YLL						29.75 (.000)

Note: Parentheses show standard errors for parameter estimates and p-values for the statistics. The instrument set for circulatory expenditure and non-circulatory deaths includes the proportion of households that are one pensioner household, the proportion of the population providing unpaid care, and the population weighted index of multiple deprivation based on ward level Index of Multiple Deprivation 2000 scores.

An important additional consideration in the modelling of circulatory disease was the failure of initial 2SLS models to pass the specification tests. Careful scrutiny of these results indicated that they arose from the failure to model expenditure satisfactorily in a small number of PCTs with high levels of non-white residents. The expenditure models were therefore re-estimated with an additional 'needs' variable, in the form of the percentage of the population in a 'white' ethnic group. This variable exhibited strong positive association with expenditure, other things being equal, and its introduction led to a well-specified model. In common with models for cancer, all circulatory models now appear well specified with valid and relevant instruments.

The results of the circulatory expenditure model for years of life lost can be used in a similar manner to those for cancer to calculate the marginal cost of a life year lost given the national average levels of expenditure and mortality associated with circulatory disease. The circulatory expenditure coefficient of -1.427 implies that a 1 per cent increase in expenditure gives rise to a 1.4 per cent reduction in life years lost. From 2002 to 2004, total life years lost to all circulatory disease deaths in those aged under 75 was 1,607,171. This averaged 535,724 life years per annum which, across an English population of 50 million, averages out at 0.0107144 life years (3.91 days) per person. Thus a 1 per cent increase in expenditure per head (£1.22) is associated with a 1.4 per cent reduction in life years lost (0.056 days) and implies that for national average levels of need and expenditure one life year would cost an extra £7,979. Using the estimated standard errors suggests that the 95 per cent confidence interval surrounding this estimate (£6,549 to £10,208) is considerably smaller than for cancer.

We have presented our results in terms of unadjusted life years. In order to give a very rough indication of how they might be adjusted to yield QALYs, we have applied the utility scores made available by the HODaR project (Health Outcome Data Repository, University Hospital of Wales) using the UK EQ-5D scoring algorithm. Quality of life scores are available for ICD-10 codes and can be assigned to the programme budget categories used here. We have therefore simply assigned scores to each of the ICD-10 categories with the programme budgeting areas of cancer and circulatory diseases where these match with the HODaR categories, and averaged the scores across the categories. Using this method, for cancer expenditure cost of a QALY is £19,070, while for circulatory diseases the corresponding figure is £11,960. We emphasise that these results are at best indicative and cannot offer an accurate calculation of a QALY saved, but they do suggest that the cost of a QALY from these programmes of care may be lower than many commentators have assumed.

6. Conclusions

This study has shown that healthcare expenditure has a strong positive effect on outcomes in the two programmes of care investigated. Our estimates suggest that, relative to received wisdom, the marginal cost of a 'life year' saved is quite low, at approximately £8,000 for circulatory disease and £13,100 for cancer. Although these estimates are not adjusted for quality of life, and they are associated with quite large confidence intervals (especially for cancer), they appear to compare quite favourably with a sum of £30,000 for a QALY commonly attributed to NICE as the threshold for accepting new technologies.

There is clear evidence that expenditure on circulatory disease yields greater benefits in terms of life years than expenditure on cancer. This is to be expected. Recent developments in circulatory drug therapies (especially statins) are acknowledged to be highly cost effective. Furthermore, a substantial element of cancer care is in the form of palliative care, the benefits of which are unlikely to be measured to any great extent in increased life expectancy.

The models offer evidence of a strong substitution effect between expenditure on programmes of care. Other things being equal, expenditure on a specific programme is depressed in the face of higher need in other programme areas. This suggests PCTs may be acting appropriately by directing their budget rationally to the programme areas that will yield greatest health benefit for their locality.

A by-product of the modelling has been the discovery of strong evidence of either lower levels of need or 'unmet' need among the non-white population in circulatory disease. This is suggested by the need to incorporate the 'white' ethnic group variable into the circulatory expenditure models. Further analysis of this finding is beyond the remit of this study, given the limited data available to us. However, the strength of this effect leads us to recommend that the Department of Health should examine with some urgency whether circulatory expenditure on 'non-white' ethnic groups is below that on their 'white' counterparts, after adjusting for clinical need.

The dramatic change in inference that arises from moving from the misspecified OLS models to the well-specified 2SLS models illustrates why proper econometric modelling is needed if the nature of the relationship between expenditure and outcome is to be investigated correctly. The models and methods described here are of necessity rather complex and unfamiliar to many commentators, but they are essential if incorrect inferences are to be avoided. In particular, they suggest a far more marked influence of healthcare spending on health outcomes than is often indicated by more conventional analyses.

Nevertheless, we recognise that this study has limitations. It uses limited health outcomes data (in the form of mortality rates for just two programmes of care). For the purposes of this study we were able to use only data made publicly available by the Department of Health, and we would hope that in time a greater range of outcome and epidemiological data will become available.

Furthermore, we have modelled just a single year's data. In practice, health outcomes are the results of years of expenditure by local PCTs, and conversely current expenditure is expected to yield outcome benefits beyond the current year. Implicitly, our analysis assumes that PCTs have reached some sort of equilibrium in the expenditure choices they make and the outcomes they secure. This is probably not an unreasonable assumption given the relatively slow pace at which both types of variable change. But a longer time series of data may enable us to model the effects with more confidence.

The English programme budgeting project is a major new data development. However, it is still under development, and there remain unresolved issues. Some health system expenditure is difficult to assign to programmes, most notably in primary care. Furthermore, accounting practice is variable and we would recommend that programme budgeting accounts be properly audited.

We nevertheless believe that programme budgeting is a major initiative that should be actively and vigorously promoted by the Department of Health. Most importantly, it brings together for the first time clinical data (in the form of health outcomes) and expenditure data. It therefore has the potential for engaging clinicians in value-for-money issues where more conventional budgetary approaches fail, thereby offering the potential for better clinical engagement in budgetary choices and better informed purchasing decisions by PCTs.

Furthermore, programme budgeting permits researchers to model links between healthcare expenditure and health outcomes in a much more secure manner than has previously been possible. This report has offered a glimpse of its potential in this respect. The results can help the Treasury and national politicians make more informed decisions on whether healthcare expenditure offers value for money. They can help the Department of Health and local purchasers make better informed decisions about where their limited budgets are best spent. And they can also inform the decisions of NICE on whether their current threshold for accepting new technologies is set at an appropriate level.

References

Anderson TW (1984). Introduction to Multivariate Statistical Analysis. 2nd ed. John Wiley & Sons.

Cochrane A, St Leger AS and Moore F (1978). 'Health service "input" and mortality "output" in developed countries'. Journal of Epidemiology and Community Health, vol 32, pp 200–05.

Cragg JG and Donald SG (1993). 'Testing identifiability and specification in instrumental variable models'. Econometric Theory, vol 9, pp 222–40.

Cremieux P, Ouellette P and Pilon C (1999). 'Health care spending as determinants of health outcomes'. Health Economics, vol 8, pp 627–39.

Department of Health (2005a). NHS Finance Manual. December 2005 edition. See www.dh.gov.uk/assetRoot/04/13/18/26/04131826.pdf

Department of Health (2005b). Unified Exposition Book: 2003/04, 2004/05 and 2005/06 PCT revenue resource limits. Department of Health.

Durbin J (1954). 'Errors in variables'. Review of the International Statistical Institute, vol 22; pp 23–32.

Fisher ES and Welch HG (1999). 'Avoiding the unintended consequences of growth in medical care: how more might be worse?'. Journal of the American Medical Association, vol 281, pp 446–53.

Gerdtham U and Jonsson B (2000). 'International comparisons of health expenditure' in Culyer A and Newhouse J (eds), Handbook of Health Economics. Elsevier.

Gravelle H and Backhouse M (1987). 'International cross-section analysis of the determination of mortality'. Social Science Medicine, vol 25, pp 427–41.

Nixon J and Ulmann P (2006). 'The relationship between health care expenditure and health outcomes'. European Journal of Health Economics, vol 7, pp 7–18.

Nolte E and McKee M (2004). Does Health Care Save Lives?. The Nuffield Trust.

Or Z (2001). Exploring the Effects of Health Care on Mortality Across OECD Countries. OECD Labour Market and Social Policy Occasional Paper No 46. OECD.

Pesaran MH and Taylor LW (1999). 'Diagnostics for IV regressions'. Oxford Bulletin of Economics and Statistics, vol 61, pp 255–81.

Ramsey JB (1969). 'Tests for specification errors in a classical linear least squares regression analysis'. Journal of the Royal Statistical Society, Series B, vol 31, pp 350–71.

Sargan JD (1958). 'The estimation of economic relationships using instrumental variables'. Econometrica, vol 26, pp 393–415.

Smith PC, Rice N and Carr-Hill R (2001). 'Capitation funding in the public sector'. Journal of the Royal Statistical Society, Series A, 164, pp 217–57.

St Leger S (2001). 'The anomaly that finally went away'. Journal of Epidemiology and Community Health, vol 55, pp 79.

Stock JH and Yogo M (2002). 'Testing for weak instruments in linear IV regression'. National Bureau of Economic Research Technical Working Paper 284.

Young FW (2001). 'An explanation of the persistent doctor–mortality association'. Journal of Epidemiology and Community Health, vol 55, pp 80–84.

Appendix A: Data considerations

Table A1: Deaths considered amenable to healthcare

Deaths considered amenable to healthcare are defined as those from the following causes for the specific age groups stated. See www.nchod.nhs.uk/ for further details.

Intestinal infections (ICD-10 A00-A09, ICD-9 001-009), ages 0–14 years;

Tuberculosis (ICD-10 A15-A19, B90; ICD-9 010-018, 137), ages 0–74 years;

Other infectious diseases (diptheria, tetanus, poliomyelitis) (ICD-10 A36, A35, A80; ICD-9 032, 037, 045), ages 0–74 years;

Whooping cough (ICD-10 A37, ICD-9 033), ages 0–14 years;

Septicaemia (ICD-10 A40-A41, ICD-9 038), ages 0–74 years;

Measles (ICD-10 B05, ICD-9 055), ages 1–14 years;

Malignant neoplasm of colon and rectum (ICD-10 C18-C21, ICD-9 153-154), ages 0–74 years;

Malignant neoplasm of skin (ICD-10 C44, ICD-9 173), ages 0–74 years;

Malignant neoplasm of female breast (ICD-10 C50, ICD-9 174), ages 0–74 years;

Malignant neoplasm of cervix uteri (ICD-10 C53, ICD-9 180), ages 0–74 years;

Malignant neoplasm of unspecified part of the uterus (ICD-10 C54-C55, ICD-9 179, 182), ages 0–44 years;

Malignant neoplasm of testis (ICD-10 C62, ICD-9 186), 0–74 years;

Hodgkin's disease (ICD-10 C81, ICD-9 201), ages 0–74 years;

Leukaemia (ICD-10 C91-C95, ICD-9 204-208), ages 0–44 years;

Diseases of the thyroid (ICD-10 E00-E07, ICD-9 240-246), ages 0–74 years;

Diabetes mellitus (ICD-10 E10-E14, ICD-9 250), ages 0–49 years;

Epilepsy (ICD-10 G40-G41, ICD-9 345), 0–74 years;

Chronic rheumatic heart disease (ICD-10 I05-I09, ICD-9 393-398), ages 0–74 years;

Hypertensive disease (ICD-10 I10-I13, I15, ICD-9 401-405), ages 0–74 years;

Ischaemic heart disease (ICD-10 I20-I25, ICD-9 410-414), ages 0–74 years;

Cerebrovascular disease (ICD-10 I60-I69, ICD-9 430-438), ages 0–74 years;

All respiratory diseases (excl. pneumonia, influenza and asthma) (ICD-10 J00-J09, J20-J44, J47-J99; ICD-9 460-479, 488-492, 494-519), ages 1–14 years;

Influenza (ICD-10 J10-J11, ICD-9 487), ages 0–74 years;

Pneumonia (ICD-10 J12-J18, ICD-9 480-486), ages 0–74 years;

Asthma (ICD-10 J45-J46, ICD-9 493), ages 0–44 years;

Peptic ulcer (ICD-10 K25-K27, ICD-9 531-533), ages 0–74 years;

Appendicitis (ICD-10 K35-K38, ICD-9 540-543), ages 0–74 years;

Abdominal hernia (ICD-10 K40-K46, ICD-9 550-553), ages 0–74 years;

Cholelithiasis & cholecystitis (ICD-10 K80-K81, ICD-9 574-575.1), ages 0–74 years;

Nephritis and nephrosis (ICD-10 N00-N07, N17-N19, N25-N27; ICD-9 580-589), ages 0–74 years;

Benign prostatic hyperplasia (ICD-10 N40, ICD-9 600), ages 0–74 years;

Maternal deaths (ICD-10 O00-O99, ICD-9 630-676), ages 0–74 years;

Congenital cardiovascular anomalies (ICD-10 Q20-Q28, ICD-9 745-747), ages 0–74 years;

Perinatal deaths (all causes excl. stillbirths), ages 0–6 days;

Misadventures to patients during surgical and medical care (ICD-10 Y60-Y69, Y83-Y84; ICD-9 E870-E876, E878-E879), ages 0–74 years.

Table A2: Socio-economic indicators employed in correlation analysis and available as potential instruments in the 2SLS estimation

Indicator name	Short description	Long description
BORNEXEU	Residents born outside the European Union	Residents born outside the European Union divided by all residents (census cell definition: KS005008/KS005001)
WHITEEG	Residents in white ethnic group	Population in white ethnic group divided by total population (KS006002+KS006003+KS006004)/KS006001
PCWALLTI	Population of working age with illness	Proportion of population of working age with limiting long-term illness divided by population aged 16–74 (KS008003/KS09A001)
POPPUCAR	Unpaid care providers in population	Proportion of population providing unpaid care (KS008007/KS008001)
POPPUCA1	Unpaid care (<20 hrs week) in population	Proportion of population providing unpaid care of 1–19 hours a week (KS008008/KS008001)
POPPUCA2	Unpaid care (20–49 hrs) in population	Proportion of population providing unpaid care for 20–49 hours per week (KS008009/KS008001)
POPPUCA3	Unpaid care (>50 hrs week) in population	Proportion of population providing unpaid care for over 50 hours week (KS008007/KS008001)
NQUAL174	Proportion aged 16–74 with no qualifications	Proportion of population aged 16–74 with no qualifications (KS013002/KS013001
FTSTUDEN	Proportion aged 16–74 full-time students	Proportion of population aged 16–74 that are full-time students ((KS013008+KS013009)/KS013001)
HHNOCAR	Households without a car	Proportion of households without a car (KS017002/KS017001)
OWNOCC	Owner occupied households	Proportion of households that are owner occupied (KS018002+KS018003+KS018004)/KS018001)
LAHARENT	Rented social housing	Proportion of households that are rented from local authority or housing association ((KS018005+KS018006)/KS018001)
PRIVRENT	Rented private housing	Proportion of households that are rented from private landlords (KS018007/KS018001)
LONEPENH	Lone pensioner households	Proportion of households that are one pensioner households (KS020002/KS020001)
LONEPARH	Lone parent households	Proportion of households that are lone parent households with dependent children (KS020011/KS020001)
PERMSICK	Permanently sick of those aged 16–74	Proportion of population aged 16–74 that are permanently sick (KS09A010/KS09A001)
PC74LTUN	Long-term unemployed of those 16–74	Proportion of those aged 16–74 that are long-term unemployed (KS09A015/KS09A001)
WORKAGRI	Employed in agriculture	Proportion of those aged 16–74 in employment that are working agriculture (KS11A002/KS11A001)
PROFOCCU	People in professional occupations	Proportion of those aged 16–74 in managerial and professional occupations ((KS14A002+KS14A003+KS14A004)/KS14A001)
POPWIMD	Index of multiple deprivation	Population weighted index of multiple deprivation based on ward level IMD 2000 scores
LISI2002	Exemptions from prescription charges	Low income supplement index (LISI). A measure of deprivation based on claims for exemption from prescription charges on grounds of low income. December 2001 to November 2002.
UNIFIED NEED	All NHS services needs index	This incorporates age, Hospital and Community Health Services, prescribing, General Medical Services, and HIV/AIDS adjustments. See Department of Health (2005b) for details.

Table A3: Outcomes data

The National Centre for Health Outcomes Development reports mortality rates by PCT for all causes and selected individual causes (see http://www.nchod.nhs.uk/) averaged over the three-year period from 2002 to 2004. Age and sex standardised mortality rates are available for those aged under 75. The correlation between the directly and indirectly age–sex standardised mortality rates is very high (for example, the correlation coefficient for the two standardised rates is 0.9986 for all causes of death amenable to healthcare) so here we concentrate on just one rate, the directly standardised rate.

Of the various outcome measures available, the ICD-10 coverage of the following correspond most closely to the ICD-10 coverage of the relevant programme budgeting category:

- all deaths from causes amenable to healthcare (a cause of death is included if there is evidence that it is amenable to healthcare intervention and – given timely, appropriate, and high quality care – the death rate should be low among the age groups specified). For details for the causes of death included in this categorisation see Table A1
- deaths from all cancers (this comprises mortality from all malignant neoplasms ICD-10 C00-C97 equivalent to ICD-9 140-208)
- deaths from all circulatory diseases (this comprises mortality from ICD-10 I00-I99 equivalent to ICD-9 390-459).

These three death rates can be compared with NHS expenditure levels for 2004/05 in the three broadly comparable programme budgeting categories:

- all expenditure
- expenditure associated with cancers and tumours (category 2) (but, unlike the deaths data, the expenditure data also includes expenditure on ICD-10 codes D00-D50 (carcinomas and melanomas in situ, benign and uncertain neoplasms))
- expenditure associated with circulatory problems (category 10) (but, unlike the deaths data, the expenditure data also includes expenditure on ICD-10 codes Q20-Q28 (problems with heart)).

Appendix B: Model estimation

The theoretical models (1) and (3) in Section 4 suggest the specification and estimation of a system of equations, with an expenditure and health outcome model for each of the 23 programmes of care. In the absence of endogenous regressors the system would reduce to the estimation of seemingly related regressions. However, this approach would be data intensive requiring variables to identify expenditure, need, environmental factors and health outcomes in each of the 23 programmes of care. In the presence of endogenous expenditure and outcome data, the approach would further require a set of exogenous variables to act as instruments to identify the system. This is beyond the scope of current data availability and at the time of writing we only have reliable health outcome indicators (standardised mortality rates) for two disease categories: cancer and circulatory problems. Further, we do not have convincing data on environmental factors likely to affect the production of healthcare. Accordingly, we concentrate on the two programmes of care

– cancer and circulatory diseases – where data availability is sufficient for our purposes. For each programme we specify the following reduced forms for models of expenditure (4) and health outcome (5):

$$x_{il} = \alpha_1 + \beta_1 n_{il} + \lambda y_i + \varepsilon_{1il} \qquad i = 1,K\ ,m; \quad l = 1,K\ ,2. \qquad (4)$$

$$h_{il} = \alpha_2 + \beta_2 n_{il} + \delta x_{il} + \varepsilon_{2il} \qquad (5)$$

In the absence of prevalence data relevant to each of the 23 programmes of care, we proxy healthcare need in each of the cancer and circulatory disease models using a combination of own programme need – proxied by the 'needs' component of the resource allocation formulae – and need for competing programmes by standardised mortality data. The needs element of the formula was specifically designed to adjust PCT allocations for local healthcare needs and accordingly, all things being equal, we would expect a positive relationship between expenditure x_{il} and need n_{il} for the two programmes of care. Total expenditure y_i represents expenditure across all categories of care.

Further we would expect a positive relationship between need and adverse health outcomes h_{il}. Health need for programme categories outwith the category of interest $n_{ir}\ (r \neq l)$ are proxied by death rates. For circulatory disease expenditure we use the all age standardised mortality rate for cancer; likewise, for cancer expenditure we use the all age standardised mortality rate for circulatory diseases. In addition, we further consider an alternative specification of the model where health outcomes are measured by years of life lost (YLL). As explained in Section 5, for expenditure in the circulatory disease programme we also include the proportion of white ethnic groups as an additional needs regressor.

Our estimation strategy is as follows. First we estimate the reduced form models using OLS. Assuming exogeneity of health outcomes in the expenditure model (4), and of expenditure in the health outcome model (5), OLS is a consistent estimator of the model parameters. However, should these variables be endogenous then we violate one of the assumptions of least squares as the endogenous variables will be correlated with the disturbance term in their respective model. We can test for endogeneity using the test proposed by Durbin (1954). Under the null hypothesis of exogeneity, OLS will yield consistent parameter estimates. The test consists of comparing OLS estimates to those produced by instrumental variables estimators such as two-stage least squares (2SLS). A large discrepancy between the estimates indicates a rejection of exogeneity. Under the null, the test statistic is distributed as chi-squared with degrees of freedom equivalent to the number of regressors deemed endogenous.

Assuming endogeneity of expenditure and health outcomes we implement 2SLS. Should the instrument set be relevant and valid, 2SLS will produce consistent estimates of the parameters of the reduced form models. We subject the instrument sets to tests for validity using the Sargan (1958) test of overidentifying restrictions. Under the null hypothesis that the instruments are uncorrelated with the disturbance and are correctly excluded from the equation of interest, the test statistic is distributed as chi-squared in the number of overidentifying restrictions.

In addition to the Sargan test, we test for instrument relevance using the Anderson (1984) canonical correlations likelihood-ratio test. The relevance of an instrument set refers to its ability to predict the endogenous variable of concern. If the instrument set is considered weak (marginally relevant) this may lead to biased 2SLS estimates of our equation of

interest. The likelihood ratio test of Anderson (1984) is a test of whether the equation is identified and under the null that the equation of interest is underidentified, the Anderson statistic is distributed as chi-squared with degrees of freedom equal to $(l-k+1)$, where l is the number of instruments (included and excluded exogenous variables) and k is the total number of regressors. Rejection of the null indicates that the model is identified.

While the Anderson statistic provides a test of the null hypothesis of underidentification, Stock and Yogo (2002) suggest a test for the null that the instruments are weak and provide appropriate critical values. The test is an extension of the Cragg-Donald (1993) test for identification. In the presence of a single endogenous regressor the statistic is based on the F statistic for testing the null hypothesis that the instruments do not enter the first stage regression of 2SLS. A general test of model specification is provided through the use of Ramsey's (1969) regression error specification test for OLS and an adapted version of the test for instrumental variables (Pesaran and Taylor, 1999). This test operates under the null hypothesis that there are no neglected nonlinearities in the functional form of the model specified. The standard Reset test implemented using OLS estimation follows an F distribution while the 2SLS equivalent follows a chi-square distribution. Both have degrees of freedom equal to the number of polynomial terms chosen for the fitted values. We implement the test using $\hat{y}^2$, $\hat{y}^3$ and $\hat{y}^4$, with three degrees of freedom.