VARIATIONS IN HEALTH CARE

The good, the bad and the inexplicable

John Appleby
Veena Raleigh
Francesca Frosini
Gwyn Bevan
Haiyan Gao
Tom Lyscom

The King's Fund

The King's Fund seeks to understand how the health system in England can be improved. Using that insight, we help to shape policy, transform services and bring about behaviour change. Our work includes research, analysis, leadership development and service improvement. We also offer a wide range of resources to help everyone working in health to share knowledge, learning and ideas.

Published by
The King's Fund
11–13 Cavendish Square
London W1G 0AN
Tel: 020 7307 2591
Fax: 020 7307 2801
www.kingsfund.org.uk

© The King's Fund 2011

First published 2011 by The King's Fund

Charity registration number: 1126980

All rights reserved, including the right of reproduction in whole or in part in any form

ISBN: 978 1 85717 614 8

A catalogue record for this publication is available from the British Library

Available from:
The King's Fund
11–13 Cavendish Square
London W1G 0AN
Tel: 020 7307 2568
Fax: 020 7307 2801
Email: publications@kingsfund.org.uk
www.kingsfund.org.uk/publications

Edited by Edwina Rowling

Typeset by Soapbox, www.soapbox.co.uk

Printed in the UK by The King's Fund

Contents

About the authors v

Executive summary vii

Introduction 1

1 Variations: what's good, what's bad? 2

Measuring variation and adjusting for need 5

2 Variations in elective hospital admissions 11

Ubiquity: variation in common, clinically effective procedures 12
Persistence: trends in variation 14
Efficiency: variation in day case rates 16
Effectiveness: variation in low effectiveness procedures 18
Equity: variation in pre-operative health 21
Choice: preference-sensitive variations 22

3 Conclusions and recommendations 25

References 27

Appendix 30

© The King's Fund 2011

About the authors

John Appleby is Chief Economist at The King's Fund. He is an honorary professor at the Department of Economics, City University.

The focus of his research and commentary work at The King's Fund is on current health policy matters, in particular the economic issues associated with the government's reform agenda for health care such as the introduction of competitive forces into the NHS, the use of targets to reduce waiting times, patient choice, Payment by Results and patient-reported outcome measures (PROMs). He has published widely on a range of health care finance and economic issues in books, academic journals, reports, magazines and newspapers. He co-authored a major study of NHS performance since 2002 with Sir Derek Wanless and colleagues, published by The King's Fund (2007), an analysis of the public's attitudes to the NHS, an analysis of future prospects for NHS funding written in partnership with the Institute for Fiscal Studies (The King's Fund 2009) and *Getting the Most Out of PROMs* (The King's Fund 2010).

John has also acted as an adviser to the UK government and parliament in various capacities, including carrying out a review for ministers of the future funding needs of Northern Ireland's health service, and, currently, serving as a task force member for the Marmot Commission on Health Inequalities. He is also a special adviser to the House of Commons Health Select Committee, a member of the National Quality Board's priorities sub-committee and a health economics expert for the Prime Minister's Commission on the Future of Nursing and Midwifery.

Veena Raleigh is an epidemiologist with extensive research experience and publications in public health, health inequalities, and quality, safety and patient experience in health care. In 2009 she joined The King's Fund as a Senior Fellow in Policy, where she is working on quality information and measurement issues. Prior to that, she spent eight years at the Healthcare Commission (and its predecessor the Commission for Health Improvement) as a Fellow in Information Policy, and was a Reader at the Postgraduate Medical School, University of Surrey. Veena was awarded a Fellowship of the Faculty of Public Health in 2005, and a Fellowship of the Royal Society of Medicine in 2007. Her recent publications for The King's Fund include *Getting the Measure of Quality, How do Quality Accounts Measure Up*, and a contribution to the Fund's Inquiry into the Quality of General Practice in England.Veena has also worked on health and population issues in Third World countries for international agencies (Department for International Development, the World Bank, UNFPA and The Ford Foundation), and on international health care quality measurement projects with the Commonwealth Fund and the OECD.

Francesca Frosini is a Senior Researcher in Policy at The King's Fund. Francesca joined The King's Fund in 2009 to work on various areas of health policy requiring the use of large datasets such as Hospital Episode Statistics. Her key areas of interest include the impact of patient choice and competition, organisational change, and variation in health care.

Francesca previously worked as a research fellow at Royal Holloway University of London, conducting research on the impact of decentralisation and incentives in the

© The King's Fund 2011

NHS. Francesca holds a PhD in Public Policy from New York University, focused on the impact of decentralisation of public hospitals.

Gwyn Bevan is Professor of Management Science and Director of the MSc in Public Management and Governance in the Department of Management at the London School of Economics and Political Science.

His current research includes comparisons of performance of health care and schools across the UK and developing methods for reducing expenditure on health care for least harm. He is a member of the Department of Health's two advisory groups on resource allocation, and his report for the Secretary of State for Health on the developments of the formulas used in the allocation of resources was published in 2008. He was seconded (2001 to 2003) to the Commission for Health Improvement (CHI), where he was Director of the Office for Information on Health Care Performance. This was responsible, for the NHS in England, for reporting annual performance 'star' ratings; national surveys of staff and patients; developing national clinical audits; and undertaking analyses for CHI's reviews, investigations, and national studies. He has worked for the National Coal Board, HM Treasury, and an economic consultancy; and, as an academic, at Warwick Business School, and the Medical Schools of St Thomas's Hospital and Bristol University.

Haiyan Gao is a Senior Researcher/Statistician at The King's Fund. She has more than 10 years' research experience and numerous publications in statistical methodology, mathematical and statistical modelling, and evaluation of the effects of policy changes and new services. Since joining The King's Fund in 2010, she has provided statistical analysis and advice for several projects including the Fund's Inquiry into the Quality of General Practice in England.

Haiyan worked as a Statistical Research Fellow in Intensive Care National Audit & Research Centre, where she led three national sub-studies in an SDO-funded project. She is the leading author of two publications that were used in NICE national guideline 50. She also worked in University College London and Imperial College as an Analyst/ Statistician and Research Associate, providing research design and analysis on a number of projects.

Tom Lyscom worked at The King's Fund through 2010 as a researcher analyst in the Policy directorate. He worked on many topics including referral management, patient choice, quality accounts, waiting times and health care variations. Tom now works at Imperial College conducting research and analysis in the Railway and Transport Strategy Centre.

© The King's Fund 2011

Executive summary

As this report illustrates with reference to differences in admission rates for many routine surgical procedures, variations in health care are not limited to relatively rare or new interventions. Nor are such variations a new phenomenon or simply concerned with the efficient use of scarce health service resources. The existence of persistent unwarranted variations in health care directly impacts on equity of access to services, the health outcomes of populations and efficient use of resources. But the eradication of all variation is not the task. As Al Mulley has noted:

> *If all variation were bad, solutions would be easy. The difficulty is in reducing the bad variation, which reflects the limits of professional knowledge and failures in its application, while preserving the good variation that makes care patient centred. When we fail, we provide services to patients who don't need or wouldn't choose them while we withhold the same services from people who do or would, generally making far more costly errors of overuse than of underuse.*
>
> (Mulley 2010)

The task, then, is more complicated. However, the NHS is well placed in terms of the data it collects to begin to map out in a systematic way where variations exist and then to move forward with strategies to tackle them. A key focus will need to be to tackle clinical decisions through greater emphasis on shared decision-making with patients as a way of driving out unwarranted, and promoting warranted, variation.

A critical issue in tackling 'bad' variations in future will be the impact of the current reforms of the NHS on the ability of organisations and individuals to make headway with this most persistent of problems. To what extent, for example, will GPs and GP commissioning consortia (GPCC) perhaps be better placed to tackle inequalities in access and implement shared decision-making as a way of establishing a more appropriate level of warranted variation? How will or could new roles proposed for the National Institute for Health and Clinical Excellence (NICE) help in generating the right kind of clinical evidence and guidance necessary for clinicians and patients on how to weigh up the trade-offs such evidence inevitably reveals?

Recommendations

- The first step in addressing unwarranted variations in health care is the **systematic and routine collation and publication of data on such variations**. Although there have been many examples of reporting of particular aspects of health care variations, these have tended to be sporadic and piecemeal. The recent launch of the Department of Health Atlas of Variations (Department of Health 2010a) will hopefully act as a real trigger for the analysis and reporting of variations in future.

© The King's Fund 2011

- While collating and presenting evidence of health care variations is a key first step, and while the broad and in some cases specific causes for variations are known or at least postulated, there is a subsequent need for **a programme of work not only to identify causes of variation at specific local level, but also to prioritise those variations and causes that have the most important impact on equity, effectiveness, efficiency and patient health outcomes**.
- Knowledge does not, unfortunately, always lead to action. Publicising the existence of unwarranted variations and their causes does not guarantee that they will be tackled. There is a further need for **local health organisations – both providers of care and commissioners – to be required to publicly justify and explain in a consistent way their relative position on key aspects of health care variation**. Further, it may also be necessary to **explore the development of harder-edged, locally focused incentives to encourage action to deal with unwarranted variation**.
- Most importantly, while publication of variations, the use of incentives, and the development and promulgation of clinical guidelines and other strategies have their place, what is also needed is a much **greater encouragement of shared decision-making to establish the *right* level of variation based on patients' own assessments of needs and risk aversion**.

© The King's Fund 2011

Introduction

It is now more than 20 years since The King's Fund published a review of variations in health care (Ham 1988). That report set out some of the history of variations in health care and policies since the 1970s designed to address differences in resource allocation, variations in local health organisations' performance and geographical variations in local populations' access to, and utilisation of, health services.

The issues the 1988 report raised about efficiency, equity and patient safety, and the causes of variations – the influence of demand, supply and professional decision-making – remain the same today. Unfortunately, what also remains the same is the prevalence of large variations.

For example, in 2008/9, 680 patients who lived within the Wiltshire Primary Care Trust (PCT) were admitted for a primary hip replacement operation. Taking account of Wiltshire's population – its age and gender structure – this meant that the age–gender standardised admission rate for hip operations was around 141 for every 100,000 residents of the PCT area. Meanwhile, across the country in Leicester, 174 hip operations were carried out in the same year – equivalent to an admission rate of 72 – just half the rate for Wiltshire. This may seem an extreme variation, but in fact the difference in admission rates for hip operations between the highest (Shropshire) and lowest (Kensington and Chelsea) PCTs in 2008/9 was nearly four-fold.

Such variations are not unusual nor, as is clear from The King's Fund's 1988 report and decades of research since the 1930s, are they new. While their ubiquity and persistence might suggest variations in medical care simply reflect actual and warranted variations in, for example, the need for care in different populations, equally ubiquitous and persistent research over decades both in the UK and internationally suggests otherwise. While the persistence of health care variations suggests a high degree of intractability, as NHS funding growth all but stalls over the next four years, the need to once again examine where the variation occurs, its principal causes and how unwarranted variation can be addressed will become more urgent.

In the first section of this report we first set out some of the explanations for the causes of variations in health care, what might be considered 'good' or warranted variation and what might be viewed as 'bad' or unwarranted. We also highlight the complex interaction of the multiple causes of variation – an indication of the possible difficulties in designing and implementing effective policy and action to deal with unwarranted variation. Section 1 also covers alternative ways of quantifying variation and the pros and cons of adjusting variations data to standardise for legitimate causes of variation, notably differences in need arising from demographic structure and the socio-economic conditions of different populations.

Section 2 then sets out some evidence for the existence of variations. This is a new analysis focusing on differences in elective admission ratios across PCTs in England for selected interventions. Different examples illustrate various aspects of variations: their ubiquity and persistence; their impact on efficiency, effectiveness and equity; and their reflection of variations in patient/clinician choice. Finally, we conclude with some recommendations.

© The King's Fund 2011

1 Variations: what's good, what's bad?

As Section 2 will go on to show, the fact that variations exist in, for example, hospital admission rates, is unquestionable. The complex question, however, is which variations – or what proportions of variation – are 'good', or warranted, and which are 'bad', or unwarranted. As Bob Evans (1990) has noted: 'If variations represent evidence of inappropriate care, *which* care is inappropriate? Are the regions, or institutions, or practitioners with high rates over-providing, or are the low ones under-providing, or does the 'best' rate lie somewhere in the middle (or beyond either end)?'

One approach to thinking about this issue is to start with mapping the many possible causes of variation, from possible spurious variations caused by data quality problems or differences in the geographic pattern of illness through to substitution effects arising from differences in the use of private health care and differences in clinician behaviour arising from the interaction of payment systems and the characteristics of clinicians.

So, a map of the causes of variation could look like Figure 1 opposite, which is deliberately drawn to convey the complexities and interactions of possible causes. Identifying which are important causes of variation becomes an empirical issue involving the construction and testing of a statistical model. Such statistical and qualitative research has suggested that a particularly important factor in health care variations arises from variations in the practice of medicine.

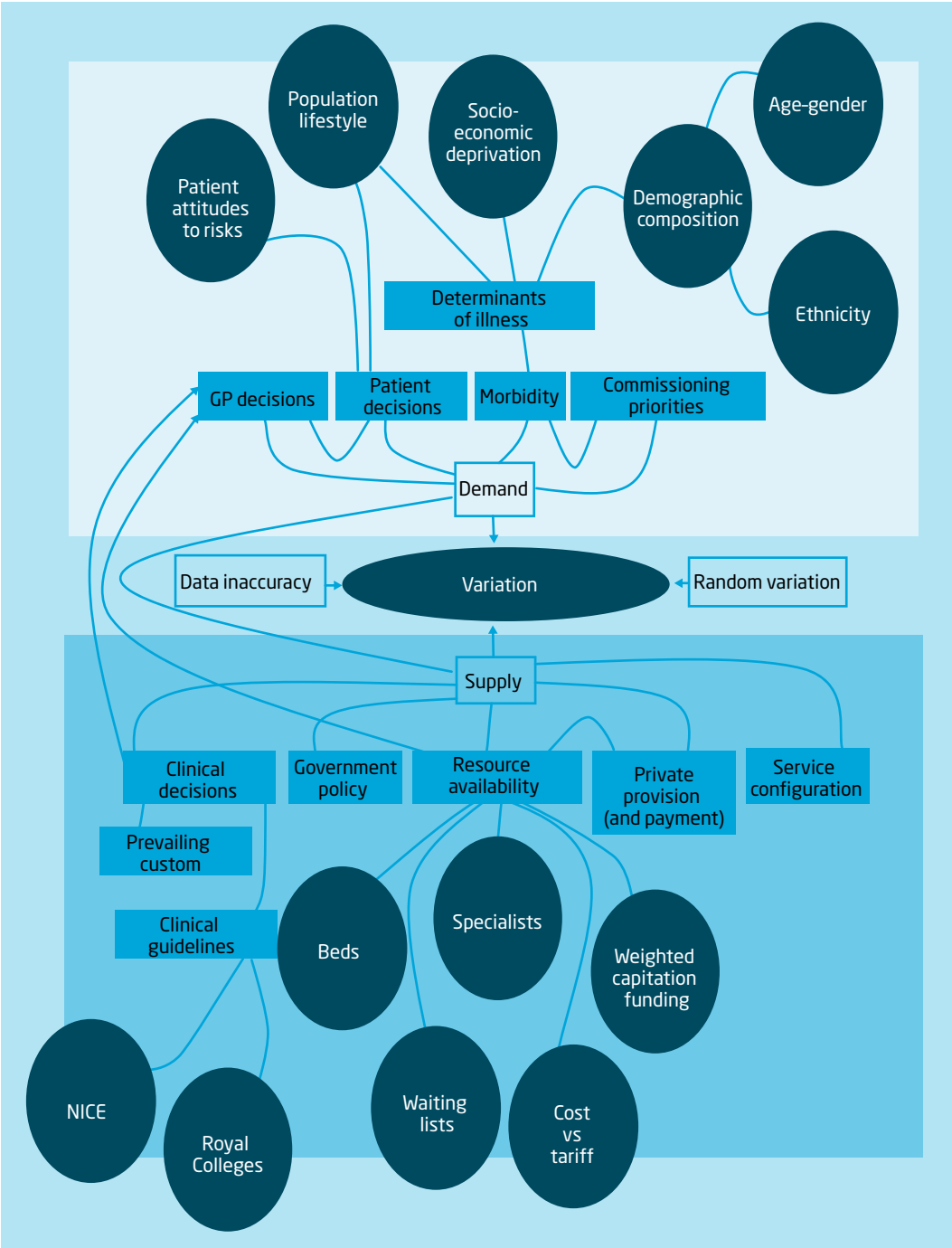
For example, in the 1930s Glover (1938) undertook pioneering analysis of small-area variations in clinical practice by examining rates of tonsillectomy across local authorities. He found 20-fold differences across London boroughs. At that time a report from the Medical Research Council pointed out that there was no evidence that wholesale tonsillectomy reduced the incidence of tonsillitis, but there was a 'tendency for the operation to be performed as a routine prophylactic ritual for no particular reason and with no particular result' (Burkinshaw 1956).

One tragic outcome that so troubled Glover was that children in poverty who returned home after the operation were vulnerable to infections and some died as a consequence: about seven children died every month in England in the 1930s as a result of tonsillectomy.

At around the same time as Glover's research, a study of 1,000 New York schoolchildren with acute and recurrent tonsillitis found that 61 per cent had already had their tonsils removed (Bakwin 1958, Bloor 1976). Such a finding – an indication of the poor efficacy of the procedure – was not, however, the most surprising finding of the study. The research went on to present the remaining 39 per cent of the children for assessment by a group of school doctors, who then recommended that 45 per cent of these children should undergo tonsillectomy/adenoidectomy. The rejected children were then sent to a second group of doctors who recommended surgery for 46 per cent of them. Those children twice rejected were finally sent to a third group of doctors who recommended surgery for 44 per cent of them. If all those recommended to have their tonsils removed had done so, then, together with those who had already undergone the operation, around 935 out of 1,000 children would have had their tonsils removed.

© The King's Fund 2011

Figure 1 Mapping causes of variation



The most recent evidence for tonsillectomy from a Cochrane review (Burton and Glasziou 2009) is that this may be beneficial for children with severe and recurrent tonsillitis (ie, those in the New York study) but there are risks associated with surgery, and children may 'grow out' of the problem. For less severely affected children the potential benefits of surgery are even more modest. Nevertheless, a recent study of paediatric tonsillectomy rates still found a seven-fold difference in rates between English regions, which could not simply be explained by a small number of high or low 'outliers' (Suleman *et al* 2010).

Research led by John Wennberg (2010) over a long and distinguished career has shown that when there is strong evidence and a professional consensus that an intervention is effective, there tends to be little or no variation in clinical practice (as, for example, surgery following a hip fracture); admission rates for these conditions can be predicted

© The King's Fund 2011

from knowledge of population statistics. However, clinical practice variations are manifested for admissions – like tonsillectomy – where there is weak evidence and professional uncertainty that hospital admission is effective. (However, as Bob Evans (1990) has pointed out, while there may be uncertainty at a group level, this does not necessarily mean that individual practitioners are uncertain: individual doctors may feel sure of the correctness of their decisions – it's just that each makes different decisions based on their experience, knowledge and interpretation of the evidence on effectiveness.)

Tonsillectomy is not uncomon, and evidence suggests that clinical practice variations are not exceptional; by implication, such variations are unlikely to have unusual or exceptional causes. The assumption that for most patients, for most of the time, the medical care that doctors decide they need depends on their illnesses rather than medical discretion has been shown to be wrong. US studies by Wennberg (2010) have shown that practice variations are, if anything, endemic. In the 1980s, for example, 90 per cent of hospital admissions in Maine were in a high variation category (similar to, but not as extreme as, those found for tonsillectomy) (Wennberg *et al* 1984).

But as Figure 1 on the previous page makes clear, many possible factors could explain health care variations, including, for example, the nature of the incentives inherent in the way health care is funded and financed. In the USA, for example, total health care spending emerges typically from fees paid to doctors and charges to hospitals for the services they supply, but in England the government aims to allocate total NHS resources equitably to populations (with reference to a formula that takes account of their estimated relative needs and differences in 'unavoidable' variations in provider costs) and hospital doctors are salaried (Bevan 2009). For Medicare spend in the USA, the ratio of states with the highest to lowest spend is nearly 300 per cent. (In 2006 spend per capita for the elderly varied from almost $16,000 in Florida to just $5,000 in Honolulu (Gottlieb *et al* 2010)). The ratio of PCTs with the highest to lowest allocations was 113 per cent. (In England in 2009 the allocations to PCTs showed that these ranged from being 10 per cent below (Bassetlaw) to 24 per cent above (Richmond and Twickenham) their estimated fair target allocation adjusted for need: this implies a ratio between these extremes of about 13 per cent. (*See* Table 20 in Department of Health (2010b)). Nevertheless, studies have found the scale of 'high' variation in hospital admission rates in England varies from 40 to 90 per cent. Newton *et al* (1994) in their study of six districts in the former Oxford Regional Health Authority reported 40 per cent of hospital admissions to be high variation; McPherson *et al* (1996) in their study of four English regions reported this to be over 90 per cent; Bevan *et al* (2004) reported this to be between 50 and 75 per cent in their study across England.

While there is evidence that there is high variation in areas with high overall rates of admission (Wennberg *et al* 1987; Price *et al* 1992), it should not be assumed that doctors in areas with *low* admission rates necessarily make more appropriate clinical decisions. In fact, studies by RAND (Chassin *et al* 1987; Leape *et al* 1990) of discretionary admissions in the USA in the 1980s found no systematic relationship between rates of appropriateness and overall admission rates: high proportions of admissions were classed as inappropriate or equivocal for areas with both high *and* low admission rates. Studies in the Trent region of England found that, despite its low rates of admission for coronary angiography and coronary artery bypass operations (when compared with the USA and England as a whole), British doctors, using their own criteria, deemed only about half of these to have been appropriate (Gray *et al* 1990).

There is also an important economic or efficiency consequence arising from variations in health care. On the assumption that all care is supplied only if this has good prospects of benefiting patients, and is hence appropriate, then reductions in volume in areas of high admission rates as a response to variation and as a tactic to reduce costs will result in harm

© The King's Fund 2011

to people who would have been admitted. But the now formidable literature on clinical practice variation shows that this is not necessarily so. Studies have suggested that the principal driver of variation in per capita spending in the US is not from variations in *costs per admission* but in *rates of admission* (Gottlieb *et al* 2010) and that there was scope by tackling admission rate variations to reduce spend on Medicare (for the elderly) by nearly 30 per cent ($40 billion) (Wennberg *et a*l 2002). Indeed, there is scope to make efficiency savings by reducing discretionary admissions that can harm patients.

More generally, variations research has prompted definitions of categories of care in the way proposed by the Dartmouth Atlas Project (Wennberg *et al* 2002). On this view there are three distinct categories:

- Effective care: this includes evidence-based services (such as haemoglobin A1c testing for diabetic patients) where variations will reflect failure to deliver needed care.
- Preference-sensitive care: this includes patient decisions where options have different risks and benefits and patients' attitudes toward these risks may vary. Such care would, for example, include coronary artery bypass surgery for heart disease; this can relieve chest pain but carries a small risk of causing memory loss. Another dimension is the choice or clinical discretion exercised by clinicians, a hypothesis widely advanced as a cause of variation in rates for common procedures such as mastectomy.
- Supply-sensitive care: this includes services where the supply of a resource (such as hospital beds, GPs, diagnostic equipment, or indeed skills and experience of specialists) has an influence on utilisation rates. How often patients consult their GP and the intensity of use of diagnostic scanning technology are examples of supply-sensitive care. In crude terms, if the resource is available it is clearly more liable to be used than if it is not.

Such a framework can be useful, not only in helping to categorise good and bad variation, but also in identifying where to direct efforts to deal with unwarranted variation. Moreover, such categorisation can turn the problem around by highlighting a solution to actively promote *warranted* variation through a greater focus on informed patient preferences, better information on effectiveness and so on.

Measuring variation and adjusting for need

Variation can be measured in a number of ways, from simple comparisons of extreme values through to more complex measures that consider the entire distribution of values. Some of the most commonly used statistics are the range and the related extremal quotient (EQ), the standard deviation (SD), the coefficient of variation (CV), and the systematic component of variation (SCV). Figure 2 overleaf shows some of the characteristics of these measures and the stylised relationship between complexity and ease of understanding.

The SCV is derived from a model that recognises two sources of total variation in area admission rates: (a) across areas – a difference in their rates, which is called systematic variation, and (b) within areas – random variation of observed rates around each area's true rate. Thus the SCV is an estimate of the true or non-random part of total variation. It is generally considered a robust measure of variation and is the most commonly used measure in such studies (McPherson *et al* 1996; Bevan *et al* 2004; Aylin *et al* 2005a; Dartmouth Atlas 2007).

The methodology used in this report for calculating the SCV was taken from McPherson *et al* (1982). It has been suggested that variations giving SCVs greater than 3 are likely to be due largely to differences in practice style or medical discretion, and that high

© The King's Fund 2011

Figure 2 Measures of variation

© The King's Fund 2011

variation admissions are those with an SCV of between 5.4 and 10.0, with SCVs greater than 10 being very high variation (McPherson *et al* 1996). Bevan *et al* (2004) identified high variation as healthcare resource groups (HRGs) with an SCV greater than 6.6, the SCV for hip replacement.

Adjusting for need

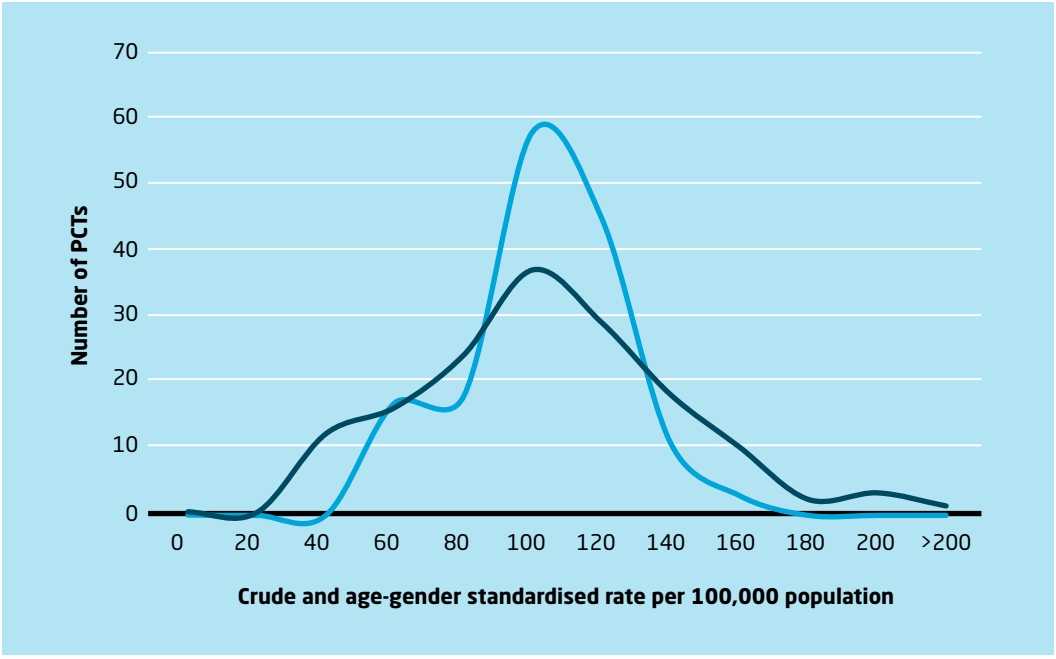
A problematic aspect of measuring variation is the extent to which it is desirable (and indeed possible) to adjust for variations in populations' need for care. Given the epidemiology of diseases, it is to be expected, that health care utilisation rates will vary from area to area on the basis of differences in, for example, the demographic pattern of populations, socio-economic conditions, and other determinants of the prevalence of disease.

Age and gender

Some differences across PCTs in utilisation rates would be expected as a result of differences in the age and gender structure of populations. For example, the epidemiology of osteoarthritis suggests a higher prevalence among women and older people (Dixon *et al* 2004). PCTs with a higher proportion of women and older people would be expected to have higher rates of hip and knee replacements. Age is generally the strongest single predictor of health care need.

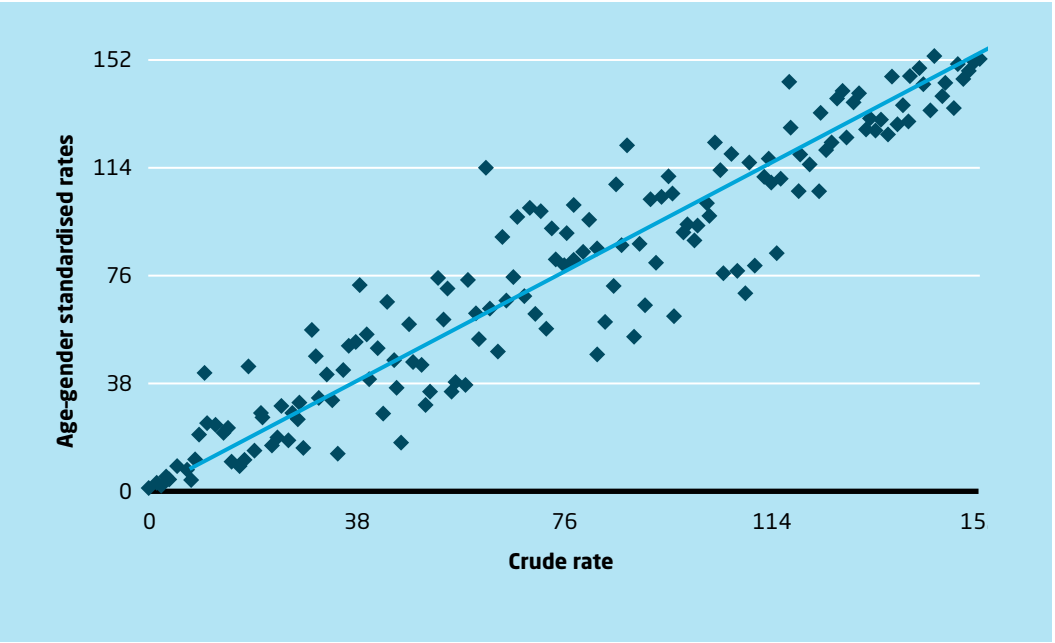
To adjust for these demographic differences, standardisation for age and gender is routinely undertaken in epidemiological analyses. As Figure 3 below shows, adjusting for age and gender does have an impact on the variation of, in this example, hip replacement admission rates, tending to pull in the extremes of variation. However, while the adjustment can also have an impact on the rankings of PCTs (*see* Figure 4 overleaf), overall the post-adjustment distribution still shows considerable variation.

Figure 3 Distribution of crude rates and age-gender standardised rates for primary hip replacement (English PCTs, 2009/10)



© The King's Fund 2011

Figure 4 PCT rankings: crude rates vs age-gender standardised rates for primary hip replacement (2009/10)



Economic and social characteristics

Apart from the demographic structure of populations, their socio-economic characteristics can also be a marker of health care need, with, generally speaking, more deprived or poorer areas suffering higher rates of illness than more affluent areas. By extension, areas of higher deprivation are likely to have higher rates of health care use. Adjustments to take account of the socio-economic characteristics of populations are often carried out as a way to eliminate legitimate or warranted variations (for example, Aylin *et al* 2005a and the Better Care Better Value indicators). However, such adjustments may not always be warranted and, where they might be, can pose technical difficulties in establishing the appropriate level of adjustment.

Table 1 Correlation between Index of Multiple Deprivation and procedure ratios for PCTs, 2009/10

Procedure	Spearman's rs
Hip replacement (primary)	-0.45*
Hip replacement (revision)	-0.30*
Knee replacement (primary)	0.02
Knee replacement (revision)	0.01
Cataract	0.42*
Coronary artery bypass graft	0.23*
Percutaneous coronary intervention	0.05
Cholecystectomy	0.01

Note: * Significance at $p < 0.05$

Spearman's rank correlation coefficient ranges from −1 to +1, and is calculated on the rank position of PCTs, not the actual values of the admission ratios and IMD. It is used when the variables being measured are not normally distributed. A positive sign means that as one variable increases the other also increases; a negative sign means that as one variable increases the other decreases. -1 indicates perfect negative correlation, zero indicates no correlation, and +1 indicates a perfect positive correlation.

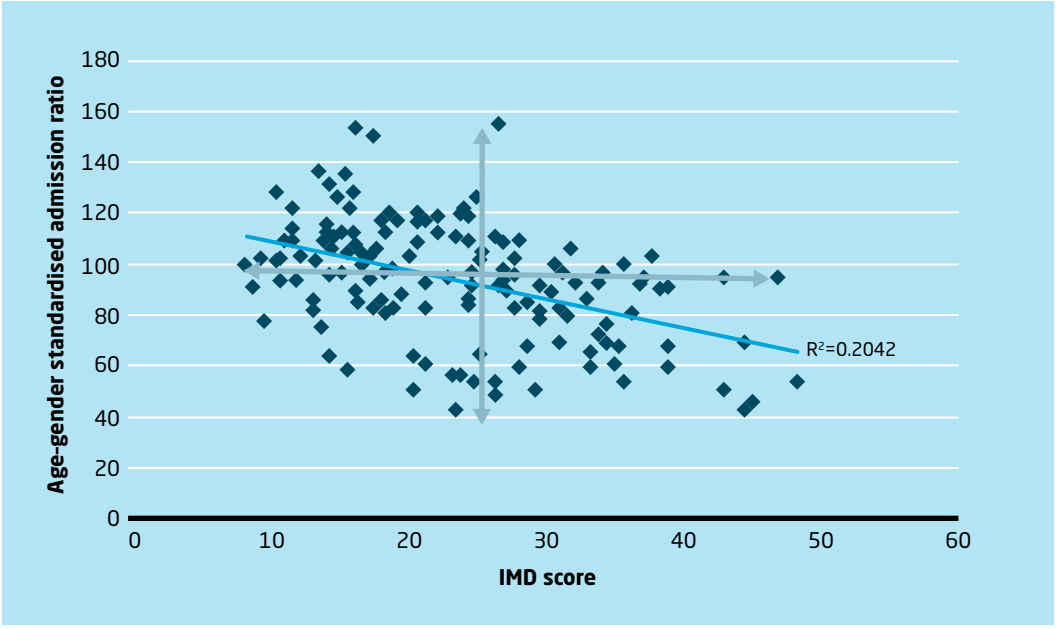
© The King's Fund 2011

First, as Table 1 on the previous page shows, the correlation between a generalised measure of deprivation – the Index of Multiple Deprivation (IMD[1]) – and admission ratios for a number of interventions are not consistent. Ratios for hip replacement (primary and revision) show a statistically significant negative association with deprivation – that is, more deprived areas tend to have *lower* admission ratios than those with lower levels of deprivation (*see* Figure 5 below). In contrast, ratios for cataract and coronary artery bypass graft surgery (CABG) show a positive association with deprivation. The other procedures show no association with deprivation. Furthermore, where it exists, the strength of any association with deprivation varies between procedures.

These associations (or lack of) with deprivation do not provide any indication of whether or not the variation in access is commensurate with actual need. The absence of, or a negative, association between admission ratios and deprivation for some procedures might be a marker of the inverse care law, reflecting inequality in access rather than differences in need. And higher rates in deprived areas can conceal an element of unmet need, if the ratios are not as high as they should because of inequitable access.

Second, there is significant variation between PCTs with similar levels of deprivation. For example, hip replacement admission ratios vary up to four-fold between PCTs with the same deprivation score, and conversely IMD scores can vary up to 4.5-fold between areas with the same admission ratio (blue arrows, Figure 5).

Figure 5 Age-gender standardised admission ratios for primary hip replacement vs IMD score, 2009/10



Third, in the absence of absolute markers for what constitutes appropriate levels of population provision for some procedures, some of the variation could reflect over- rather than under-provision of services.

Finally, even where there appears to be a significant statistical correlation between deprivation and admission ratios in a simple bivariate model, more sophisticated analyses of variations in admission ratios involving a number of possible explanatory factors

[1] The IMD is a composite measure of deprivation combining, with different weights, income, employment, health, education, crime, housing and environmental measures into one index where high scores indicate more deprived areas.

© The King's Fund 2011

are likely to alter the contribution of deprivation as an explanation. Without testing more comprehensive models to explain variations it is therefore difficult to establish the appropriate degree of adjustment to apply to the crude admission rates.

For these reasons, across the board adjustment of data on admission ratios, using a measure of deprivation or the PCT allocation formula need index, is not always warranted and can in fact hide issues of interest, for example, in the case of hip replacement, adjusting for deprivation might obscure unwarranted variation that is not need-related.

© The King's Fund 2011

2 Variations in elective hospital admissions

To illustrate the types of geographical variation in health care provision that exist in the NHS across England, this section provides an analysis of variations by primary care trust (PCT) in rates of elective surgery for selected procedures and modes of treatment.

Thirty-six procedures were selected on the basis that they are either generally recognised to be clinically effective, or there is a degree of clinical uncertainty concerning intervention, and/or there is evidence of cost-effective alternatives for conducting surgery, in particular, as a day case rather than as an inpatient. The box below details the procedures analysed and the calculation of admission ratios.

Procedures analysed

Clinically effective procedures

Hip and knee replacements (primary and revision), cataract surgery, coronary artery bypass graft (CABG), percutaneous coronary intervention (PCI) and cholecystectomy.

Clinically uncertain or low effectiveness procedures

Abdominal excision of uterus, vaginal excision of uterus, myringotomy with/without grommets, tonsillectomy, dilation and curettage/hysteroscopy and lumbar spine procedures.

Day case surgery

This indicator includes a basket of 25 procedures identified by the Audit Commission as being amenable to safe and cost-effective day case surgery (Audit Commission 2001): cataract removal, correction of squint, inguinal hernia repair, excision of breast lump, orchidopexy, varicose vein stripping/ligation, carpal tunnel decompression, dilation and curettage/hysteroscopy, excision of Dupuytren's contracture, operation for bat ears, reduction of nasal fracture, haemorrhoidectomy, removal of metalware, termination of pregnancy, circumcision, bunion operations, myringotomy, anal fissure dilation or excision, laparoscopy, laparoscopic cholecystectomy, sub mucous resection, tonsillectomy, excision of ganglion, arthroscopy, transurethral resection of bladder.

Note that three of these procedures (myringotomy, tonsillectomy and dilation and curettage/hysteroscopy) overlap with the clinically uncertain procedures group.

Admission ratios and data

Apart from the examination of day case rates, all data are presented as indirectly age and gender standardised admission ratios, standardised to the national average for England (= 100). All admission ratio data use provisional Hospital Episodes Statistics data for 2009/10 unless otherwise stated.

© The King's Fund 2011

Using these procedures as examples, this section shows that variations are **ubiquitous** and **persistent**, to be found even among common interventions of known effectiveness. Moreover, even where there is extensive evidence of more **efficient** modes of treatment (day case surgery, for example) and evidence of a lack or uncertainty of clinical **effectiveness**, wide variations in admission ratios are evident. Further, the existence of variations across PCTs for clinically effective operations that are associated negatively with socio-economic deprivation is indicative of **inequity** of access. Finally, we examine examples of admission variations, which in part reflect variation in **choices** by clinicians about treatment options and by patients as to the care they receive that reflect, for example, attitudes to risk concerning side-effects and outcomes of competing interventions.

Ubiquity: variation in common, clinically effective procedures

One of the most striking features of variations in health care is their ubiquity. In fact, it is hard to find examples where there is little or negligible variation. As Figures 6 and 7 below show, even for common, clinically effective procedures, admission ratios vary significantly between PCTs.

Figure 6 Age-gender standardised ratios for selected elective procedures (2009/10), England=100

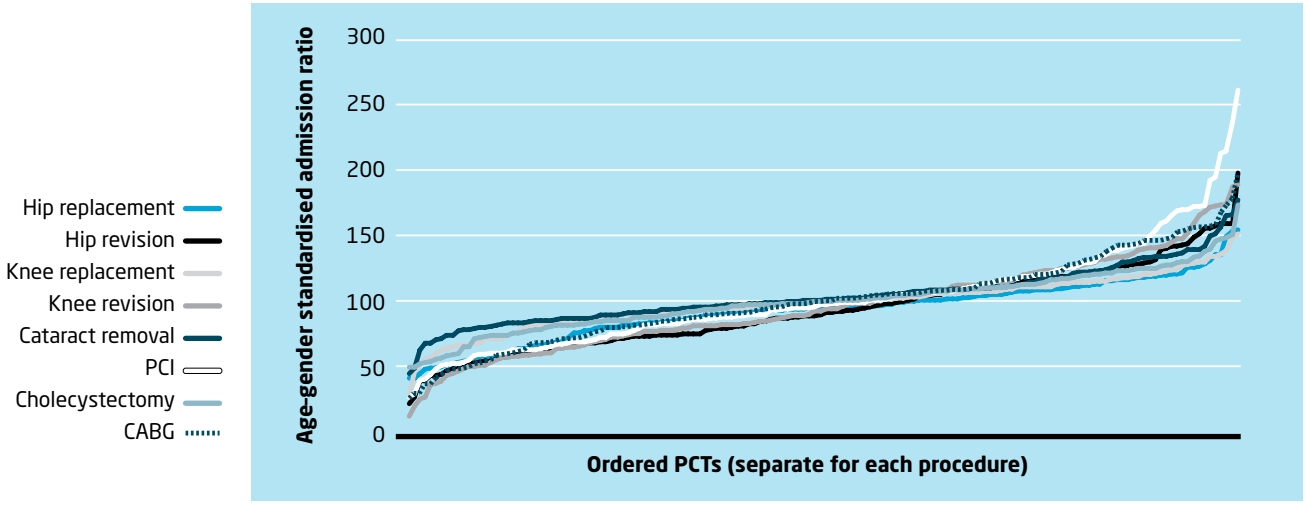
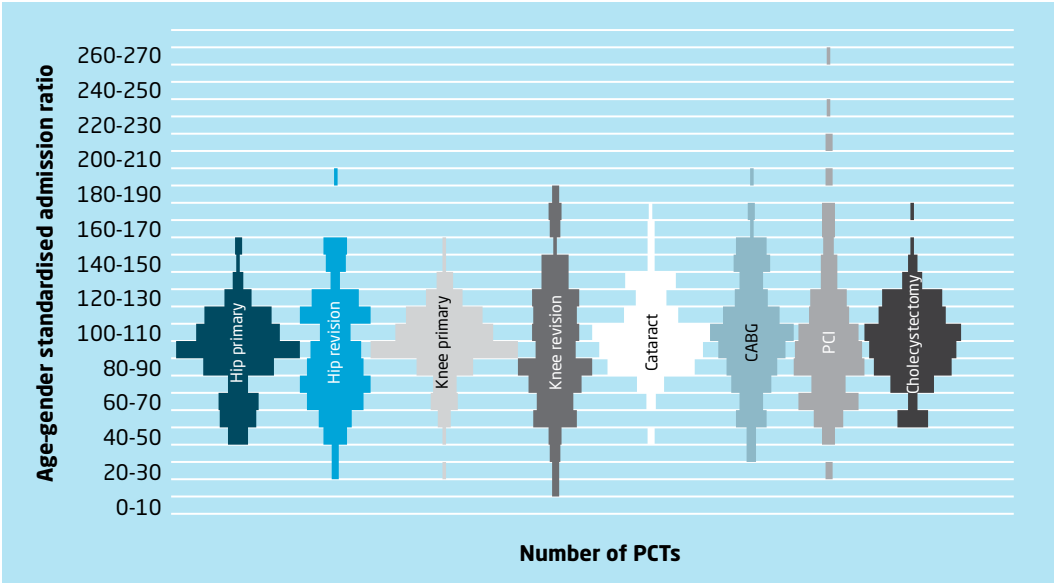


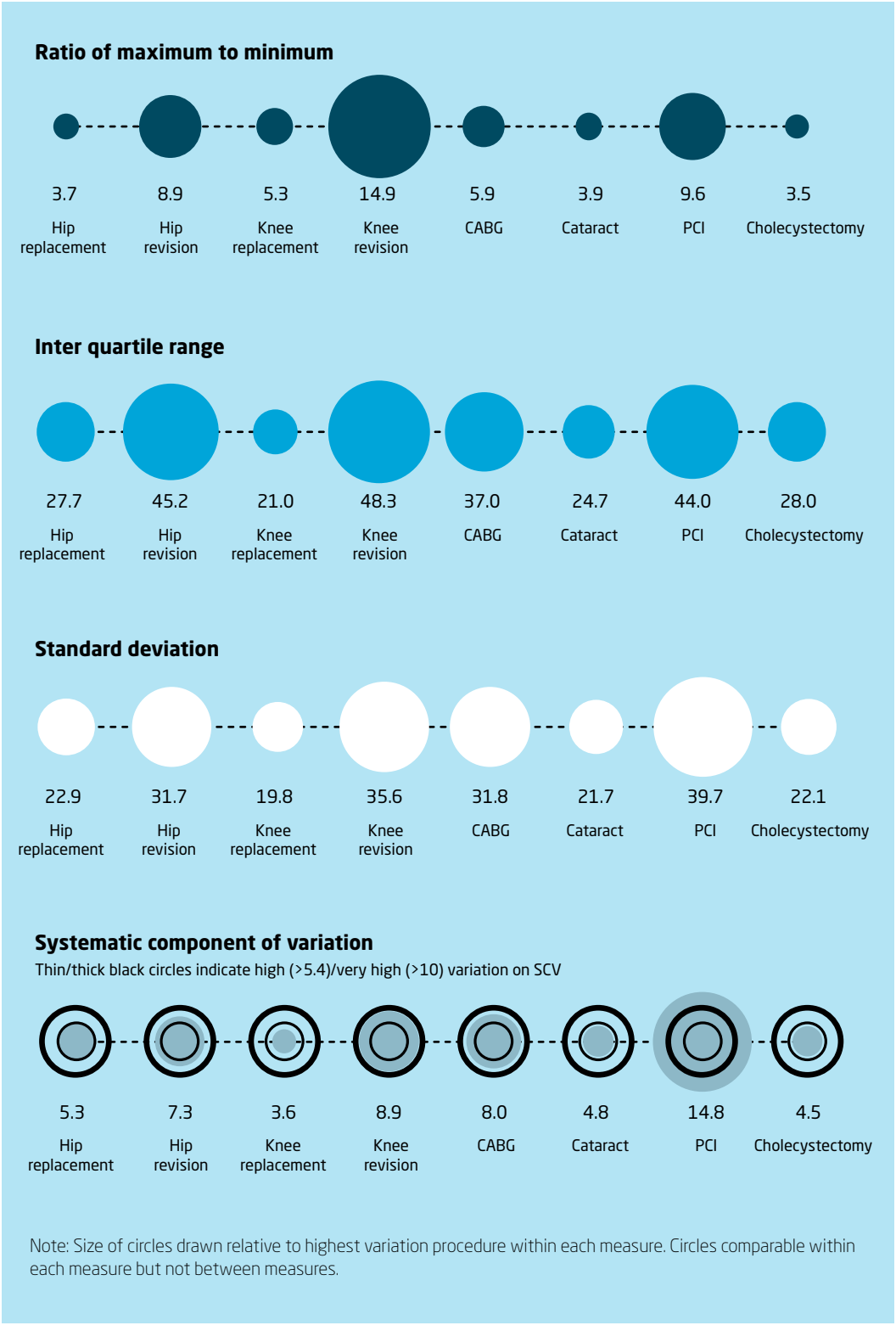
Figure 7 Age-gender standardised ratios for selected elective procedures (2009/10), England=100



© The King's Fund 2011

The variation is apparent across the different types of measures of variation (*see* Figure 8 below), and there is some consistency between the different measures in terms of the relative magnitude of variation observed. In terms of the systematic component of variation (SCV), the greatest variation is apparent for PCI(14.8), knee replacement revision (8.9), CABG (8.0) and hip replacement revision (7.3). All these values place variation in the high or very high category according to McPherson *et al* (1996).

Figure 8 Measures of variation: selected elective procedures (2009/10)



Note: Size of circles drawn relative to highest variation procedure within each measure. Circles comparable within each measure but not between measures.

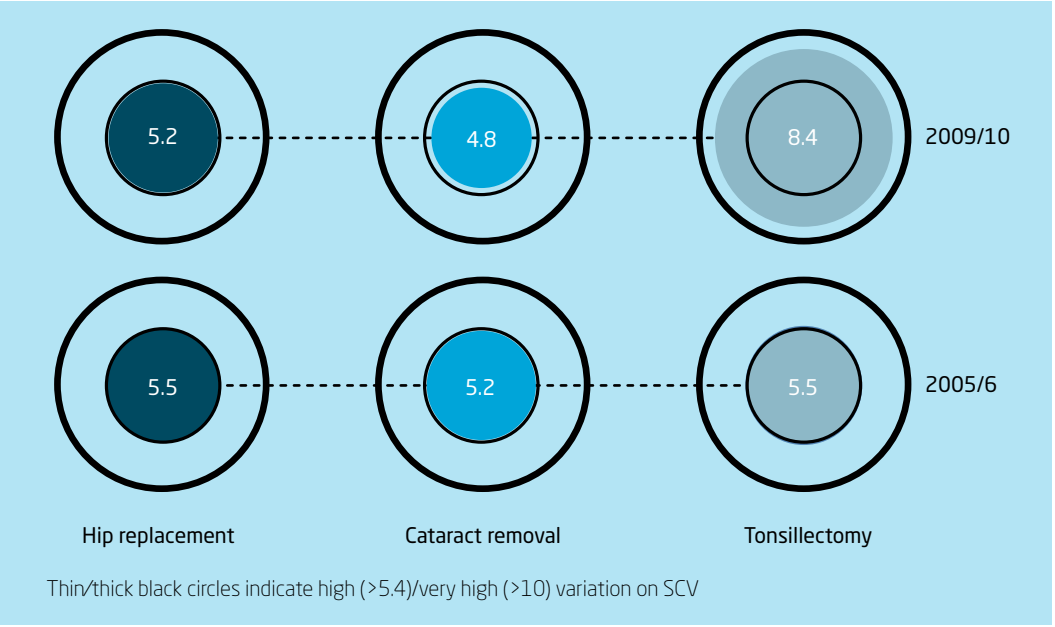
© The King's Fund 2011

Even comparatively commonplace procedures such as hip replacement, cataract removal and cholecystectomy show a four-fold variation across PCTs. Although not directly comparable due to differences in the period covered, the number of PCTs involved and standardisation methodology, Aylin *et al* (2005a) also found large variations in commonly undertaken procedures in their analysis. Over the period 1998/9 to 2003/4, their SCVs ranged between 5.8 and 8.3 for CABG and 4.1 and 5.9 for hip replacements, similar orders of magnitude as presented here.

Persistence: trends in variation

Another striking feature of variations is their persistence over time. For example, as Figure 9 below illustrates, analysis of trends in selected procedures (hip replacement, cataract removal and tonsillectomy) between 2005/06 and 2009/10 shows that geographical variation in utilisation has remained undiminished. The SCV for hip replacement has remained relatively constant over the four years at about 5; for cataract removal it increased in the interval and then dropped back to about 5. Variations in tonsillectomy rates may not now be as high as the 20-fold variation found by Glover (1938) across London boroughs in the 1930s, but they nevertheless remain high, indeed increasing on the SCV measure between 2005/6 and 2009/10.

Figure 9 Variation over time: SCVs for hip replacement, cataract removal and tonsillectomy 2005/06, 2009/10



Aylin *et al* (2005a) noted that several operations showed significantly reduced variation over the period they examined (1998/9 to 2003/4) and concluded that this would be consistent with an increased focus on the use of guidelines, or better and fairer resource allocation. However, there is no evidence of significantly diminished variation between 2005/06 and 2009/10 for the three procedures we examined.

Furthermore, analysis of admission ratios for PCTs for these procedures shows that, while there is year-on-year variation for individual PCTs, overall there is some consistency over time; that is, PCTs with high ratios tend to remain high and similarly for those with low ratios (*see* Figure 10a,b,c opposite). There are of course some changes in the ranking of individual PCTs, but there is little evidence of random variation or regression to the mean overall. Weinstein similarly noted the persistence of geographical variations for joint replacement surgery in the US (Weinstein *et al* 2004).

© The King's Fund 2011

Figure 10a Age-gender standardised ratios for hip replacement (primary) by PCT, 2005/06, 2009/10 (PCTs are ranked according to the ratios in 2005/06)

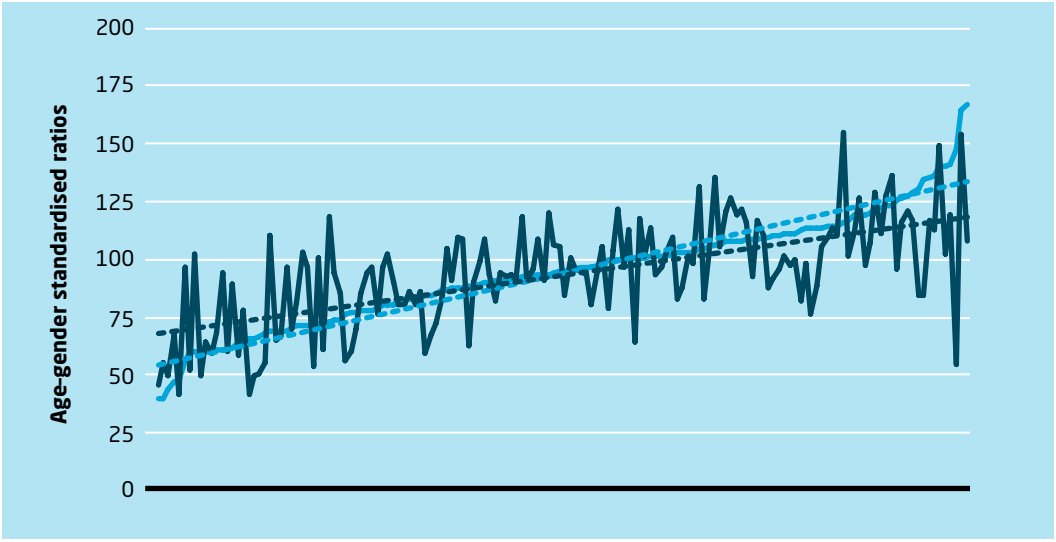


Figure 10b Age-gender standardised ratios for cataract removal by PCT, 2005/06, 2009/10 (PCTs are ranked according to the ratios in 2005/06)

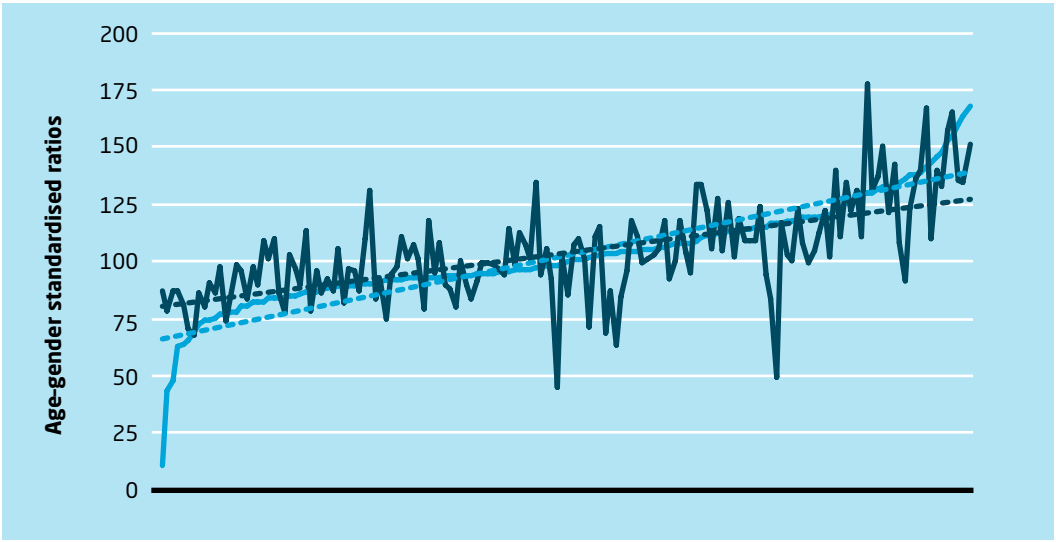
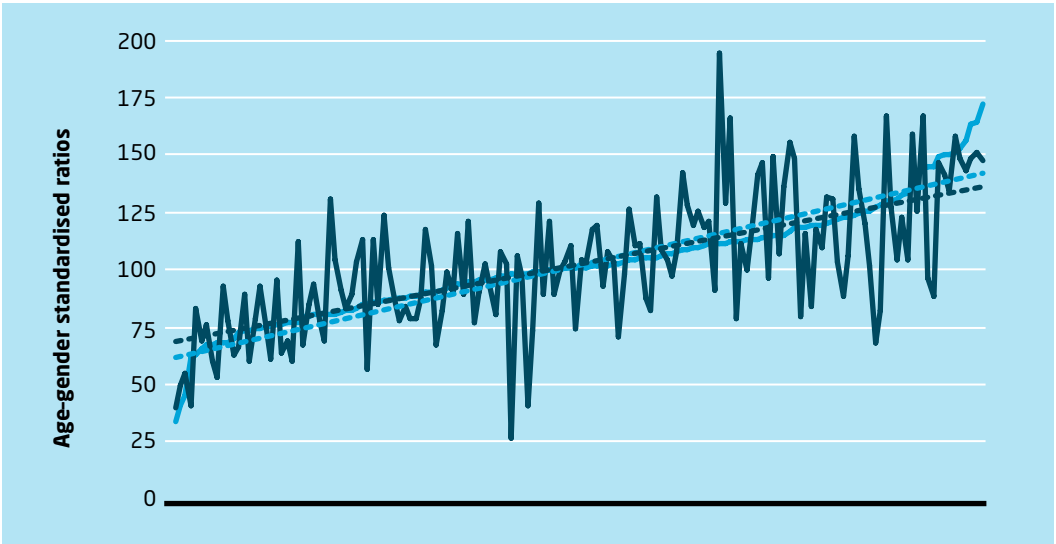


Figure 10c Age-gender standardised ratios for tonsillectomy by PCT, 2005/06, 2009/10 (PCTs are ranked according to the ratios in 2005/06)



© The King's Fund 2011

Efficiency: variation in day case rates

An important aspect of health care variation is the impact on the efficiency of health care services. Changes in surgical techniques and medical technology and knowledge have, for example, allowed more people to be treated in hospital as day cases without the need for overnight stays as an inpatient. According to the British Association of Day Surgery (BADS), patients strongly endorse day surgery as it provides timely treatment, less risk of cancellation, lower incidence of hospital acquired infections, and an earlier return to normal activities. Day surgery is therefore seen as a key element in improving patient experience as well as being more cost-effective. Day case surgery has become more common and is starting to approach 70 per cent of all NHS surgery. Since the 1990s, the Audit Commission has reported wide variation in the rates of day surgery activity (Audit Commission 1992). Following the launch of the day surgery strategy in 2002 as part of the NHS modernisation agenda, there has been considerable growth in the use of day surgery.

In 2000 the Audit Commission, in consultation with BADS, identified a basket of 25 procedures for which day case surgery was indicated as a cost-effective method of delivering care (Audit Commission 2001). The procedures included those that:

- are commonly performed, so account for a large volume of surgery
- are suitable for treatment as day cases
- would not generally be performed as an outpatient case, thus focusing attention on the potential to treat more inpatients as day cases.

This basket of 25 procedures has also become part of the Better Care Better Value indicators, whose aim is to monitor efficiency in the NHS and identify opportunities for improved productivity. Day surgery performance is measured as the number of day case patients treated expressed as a percentage of the number of elective inpatients and day cases combined.

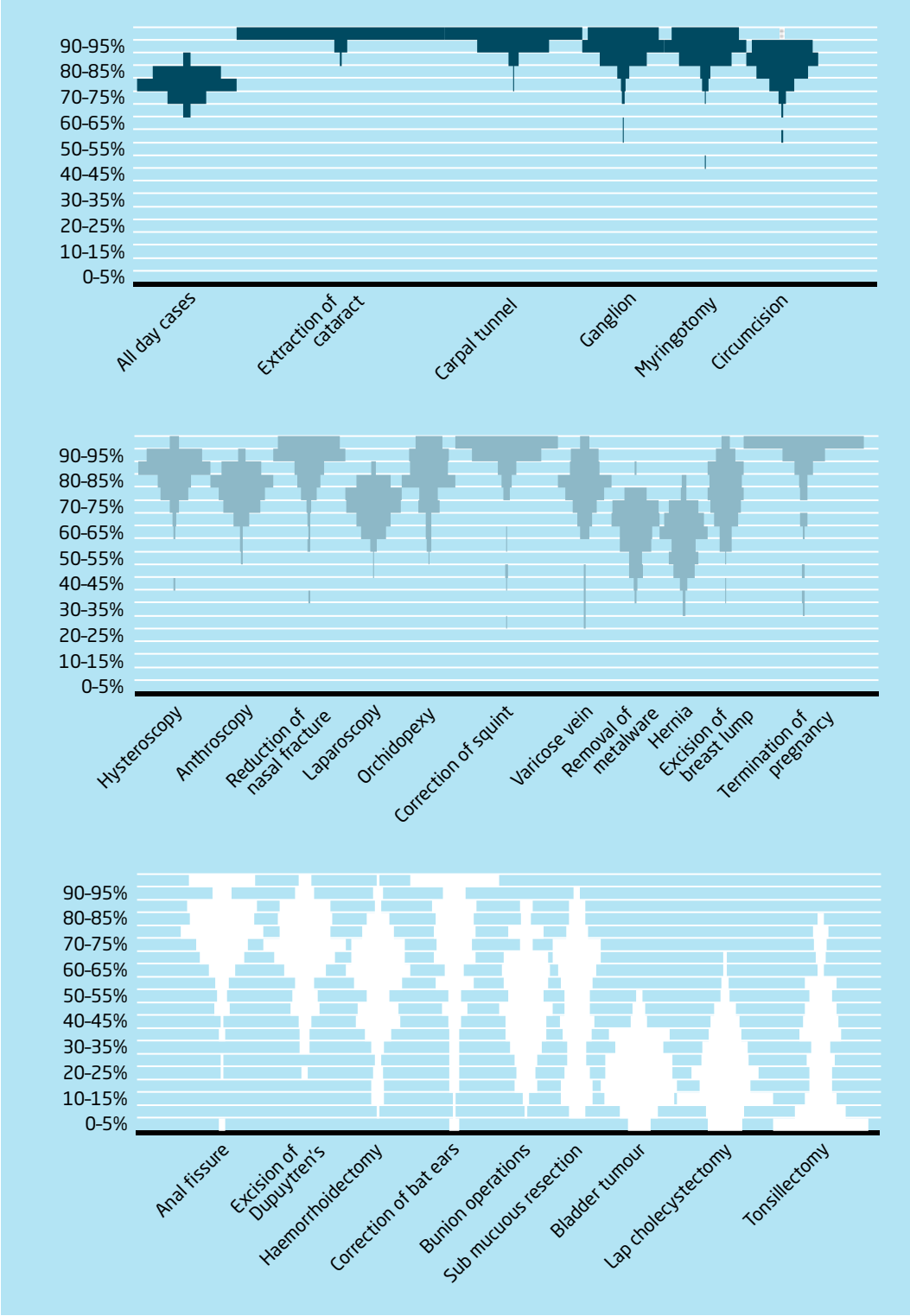
Previous analyses of the uptake of day case surgery for the basket of 25 procedures indicated variable progress across the NHS. For example, the 2005 Healthcare Commission acute hospital portfolio review for day surgery suggested significant trust-level variation for some procedures in 2003/2004 (Healthcare Commission 2005). An analysis from Dr Foster showed significant progress in the proportion of the selected procedures performed as day case surgery between 1996/97 and 2003/04, but also the persistence of inter-trust variations (Aylin *et al* 2005b).

Of the almost 1 million procedures in England in 2009/10 that could potentially have been performed as day cases, just over one-fifth were *not* carried out as day cases. This is equivalent to around 220,000 patients treated as inpatients who could have been treated as day cases, based on the Audit Commission basket alone. Analysis of day case rates by PCT (rather than by hospital where the surgery was performed) reveals significant variations, from 67 per cent to 87 per cent (with a coefficient of variation (CV) of 0.05).

For individual procedures, the picture is mixed. For some procedures (such as extraction of cataracts), day case surgery appears to be the standard mode of delivery, with 88–100 per cent performed this way (*see,* for example, Figure 11 opposite). However, for others there are stark variations in the way care is delivered, with significant under-use of day case surgery in some PCTs. The greatest variation was observed for tonsillectomy (about 40,000 procedures, CV 0.78) and laparoscopic cholecystectomy (about 10,000 procedures, CV 0.61).

© The King's Fund 2011

Figure 11 Audit Commission basket of 25 day case procedures: per cent carried out as day cases by PCT, ordered on size of coefficient of variation (low – black, medium – grey, high – white), 2009/10



As with other aspects of variation, there are a number of explanations for the variations observed between PCTs. For example, for several procedures the proportion carried out as day cases showed a negative association with deprivation. Day case rates for the following procedures were significantly lower in PCTs with higher levels of deprivation than in more affluent PCTs: circumcision, transurethral resection of bladder tumour, excision of Dupuytren's contracture, carpal tunnel decompression, arthroscopy, bunion

© The King's Fund 2011

operations, removal of metalware, cataract surgery and myringotomy. And the overall day case rate for the Audit Commission's basket of 25 procedures was also negatively associated with deprivation (Spearman's rs -0.23, p value 0.004).

It is possible that co-morbidities explain some part of this variation, but the Audit Commission's methodology does not recommend adjustment for deprivation, and the magnitude of variation observed suggests that service factors, such as differential clinical practices and availability of resources and infrastructure within organisations, play a role. The Audit Commission (2001) identified the following barriers to high rates of day surgery:

- inappropriate and insufficient use of day surgery units
- poor management and organisation of day surgery units
- clinicians' preferences for inpatient surgery.

The variation in the Audit Commission's basket of 25 procedures generally represents unwarranted variation. However, the proportion of these procedures performed as day case surgery has continued to show an increase. Among the more numerous of these procedures, the proportion performed as a day case increased between 2003/04 and 2009/10 from 55–80 per cent for varicose vein stripping or ligation, 8–29 per cent for tonsillectomy, 4–24 per cent for laparoscopic cholecystectomy, and 62–80 per cent for arthroscopy (2003/04 figures from Aylin *et al* 2005b). The overall day case rate has increased from 56 per cent in 1996/97 to 67 per cent in 2003/04 to 78 per cent in 2009/10.

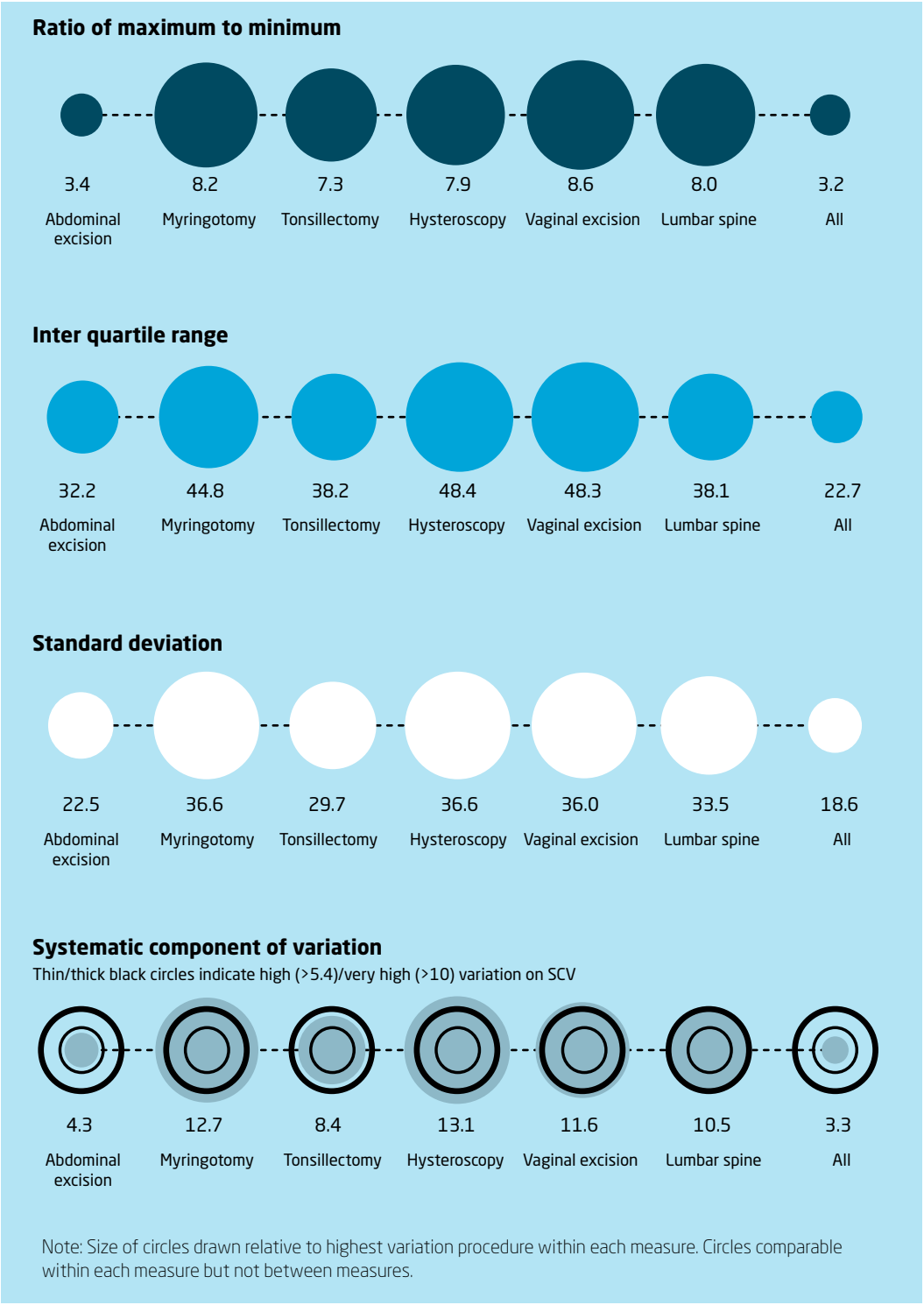
Effectiveness: variation in low effectiveness procedures

Low effectiveness care refers to surgical procedures for which there is substantial uncertainty about clinical effectiveness and that, more often than not, are likely to be carried out inappropriately and with little or no therapeutic value for the patient. Wennberg and Gittelsohn (1982) noted that variations in procedure rates are correlated with the degree of professional consensus about preferred treatment options for the underlying condition. As noted earlier, the case of tonsillectomy illustrates the issue. Glover's (1938) pioneering work demonstrated marked variations in tonsillectomy rates among English children in the 1930s, and the lack of association with any predictive factors. Tonsillectomy is a procedure where medical opinion has varied significantly, although the acceptable indications for the operation have become more strictly defined over time (*see* Department of Health 2006). However, substantial variations in tonsillectomy rates continue to be reported for England (Suleman *et al* 2010; Department of Health 2006).

Our analysis of the basket of surgical threshold indicators comprising six low effectiveness procedures for England in 2009/2010 is in line with these patterns. It shows considerable variation across all the procedures (*see* Figure 12 opposite and Figure 13, p 20), signalling significant overuse in some PCTs. In terms of the SCV, using the thresholds for high and very high variation of 5.4 and 10 respectively taken from McPherson *et al* (1996), the greatest variation is observed for hysteroscopy (13.1), myringotomy (12.7), vaginal excision of uterus (11.6) and lumbar spine (10.5); tonsillectomy shows high variation (8.4), while abdominal excision of uterus shows the smallest variation (4.3). This picture is fairly consistent across the other measures of variation.

© The King's Fund 2011

Figure 12 Measures of variation: selected low effectiveness procedures, 2009/10



Note: Size of circles drawn relative to highest variation procedure within each measure. Circles comparable within each measure but not between measures.

A number of factors could contribute to these variations. The admission ratio data used here includes NHS-funded care provided by the independent sector but excludes privately funded care. If significant proportions of these procedures are privately funded, part of the variation could be explained by the differential distribution of the private sector contribution across England. However, Suleman *et al* (2010) found a seven-fold variation in NHS tonsillectomy rates across local authority areas in England in 2000–2005, and noted that the variation in NHS surgical workload was not attributable to the exclusion of data from the independent sector.

© The King's Fund 2011

Figure 13 Age-gender standardised ratios for low effectiveness procedures, 2009/10, England=100

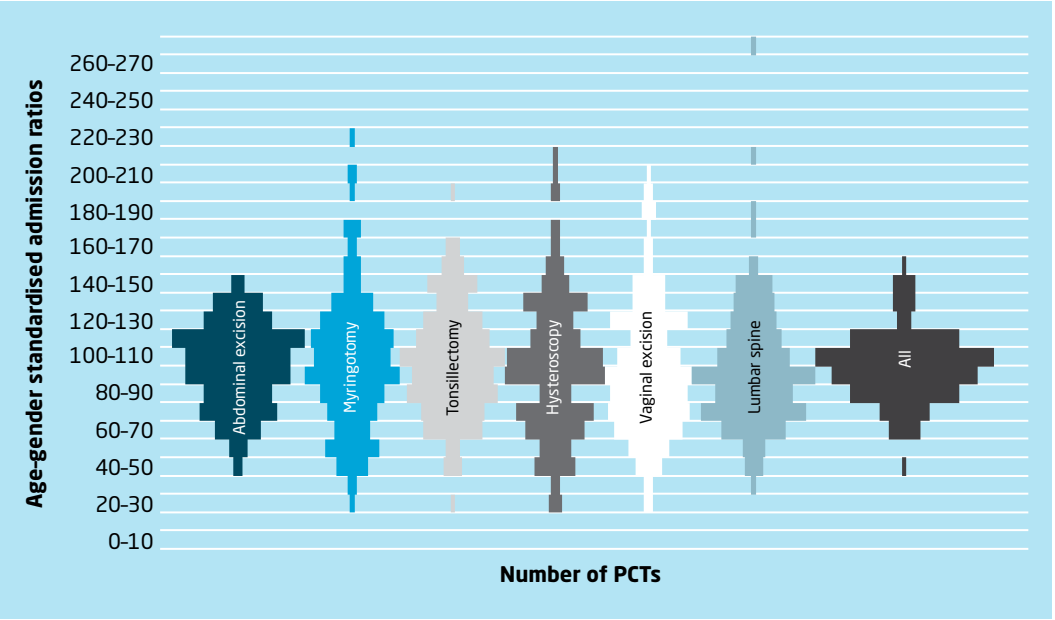


Table 2 Correlation between Index of Multiple Deprivation and ratios of low effectiveness procedures for PCTs, 2009/10

Procedure	Spearman's rs
Abdominal excision of uterus	0.30*
Myringotomy	0.16*
Tonsillectomy	0.26*
Lumbar spine procedures	-0.16
Hysteroscopy	-0.03
Vaginal excision of uterus	0.08
All	-0.02

Note: * Significance at p <0.05

Spearman's rank correlation coefficient ranges from −1 to +1, and is calculated on the rank position of PCTs, not the actual values of the admission ratios and IMD. It is used when the variables being measured are not normally distributed. A positive sign means that as one variable increases the other also increases; a negative sign means that as one variable increases the other decreases. -1 indicates perfect negative correlation, zero indicates no correlation, and +1 indicates a perfect positive correlation.

Evidence suggests high correlation between the degree of variation in the rate for a procedure and lack of consensus on how to treat the underlying condition (Suleman *et al* 2010). The presence of clinical uncertainty offers a basis to interpret these variations as, at least in part, reflecting the impact of physicians' differential preferences for providing these procedures. In addition, even in the presence of guidelines, it is possible that doctors have different attitudes towards them as there is evidence suggesting that doctors interpret (rather than stick rigidly to) guidelines, with a bias in favour of performing surgery.

Deprivation as a marker of need should arguably not impact on the rates of procedures of uncertain clinical benefit. However, there is a statistically significant positive association between deprivation on the one hand, and admission ratios for abdominal excision of uterus, myringotomy and tonsillectomy on the other (*see* Table 2 above). This is consistent with the Department of Health's 2006 report, which noted that the positive social class gradient in tonsillectomy rates observed by Glover had been reversed in the 1950s (Department of Health 2006). The report concluded that, when tonsillectomy was a procedure approved by the medical profession and in the ascendancy, the affluent

© The King's Fund 2011

had the greatest access to it. But with changing clinical opinion and evidence, now that the indications for the procedure are more limited and there is greater awareness of the potential risks, it is the more affluent, empowered, and better informed who have altered their healthcare usage in response. Thus, it is possible that the socio-economic gradient reflects differences in education and availability and use of information across social classes, which result in different attitudes to and degrees of involvement in decisions about treatments (Department of Health 2006).

Equity: variation in pre-operative health

Variations in health care utilisation rates can legitimately reflect differences in population need for health care. However, such variations may not be commensurate with need, and may reflect residual inequalities in access to services or even that the 'inverse care law' applies.

As noted above, there is an apparent negative association between hip replacement admission ratios and one measure of need, the Indicator of Multiple Deprivation (IMD). In other words, areas of *higher* deprivation have *lower* rates of admission for this procedure. An association with deprivation is also reflected in two patient-reported outcome measures (PROMs) for patients about to undergo a hip replacement operation, the Oxford hip score and the generic health-related quality of life measure, the EQ-5D index. Both these measures have been collected from consenting NHS patients in England since April 2009.

Results for April 2009 to April 2010 for pre-operative PROM scores on the Oxford[2] and EQ-5D index[3] measures tended to be lower (reflecting worse self-reported health) in more deprived and higher (better self-reported health) in less deprived PCTs. Moreover, patients reporting worse pre-operative health on these PROMs indicators tended to live in PCTs with *lower* admission ratios for hip replacement operations. The correlation is not large, although it is statistically significant for both PROMs measures.[4,5]

Although knee replacement admission ratios showed no association with deprivation, pre-operative PROMs scores on the Oxford index were again significantly lower in deprived PCTs compared with affluent PCTs,[6] the relationship with deprivation being stronger than for hip replacement ratios. However, there appears to be no significant correlation between pre-operative health status and admission ratios for knee replacement operations.

For hip operations, people living in more deprived areas are less likely to receive surgery than those living in more affluent areas. The fact that patients who *do* receive surgery in more deprived PCTs tend to report worse health just before their hip operation than those in less deprived areas could be due to a number of reasons. It may reflect higher co-morbidities, delays in presentation or more complicated socio-economic phenomenon concerning clinical and patients' attitudes to surgery. However, there is substantial evidence of inequitable access to joint surgery in England for socio-demographic groups, with Judge *et al* noting inequity by age, gender, deprivation, rurality and ethnicity (Judge *et al* 2010; Dixon *et al* 2004).

A study in Canada found that people with lower socio-economic status (SES) had a greater need for, and were equally willing to consider, arthroplasty, compared with those with higher SES (Hawker *et al* 2002). Thus, observed SES disparities in the rates of performed arthroplasties indicated unmet need in those with lower SES. Wide ethnic and geographic variations in joint replacement have been reported also for the US (Fisher *et al* 2010), with the authors noting that these could be influenced by clinical judgment rather

[2] Pre-operative Oxford Hip score vs IMD: Spearman's rs = -0.43, p = 0.000.
[3] Pre-operative EQ-5D index score vs IMD: Spearman's rs = -0.39, p=0.000.
[4] Pre-operative Oxford Hip Score vs admission rate: Spearman' rs = 0.23, p=0.004.
[5] Pre-operative EQ-5D index score vs admission rate: Spearman's rs = 0.25, p=0.002.
[6] Pre-operative oxford knee score vs IMD: Spearman's rs = 0.51, p= 0.00.

© The King's Fund 2011

than patient preference, or by inequitable access to joint replacement. Weinstein *et al* (2004) note that involving patients in choice of treatments (shared decision-making) and outcomes research are promising strategies for reducing unwarranted regional variation.

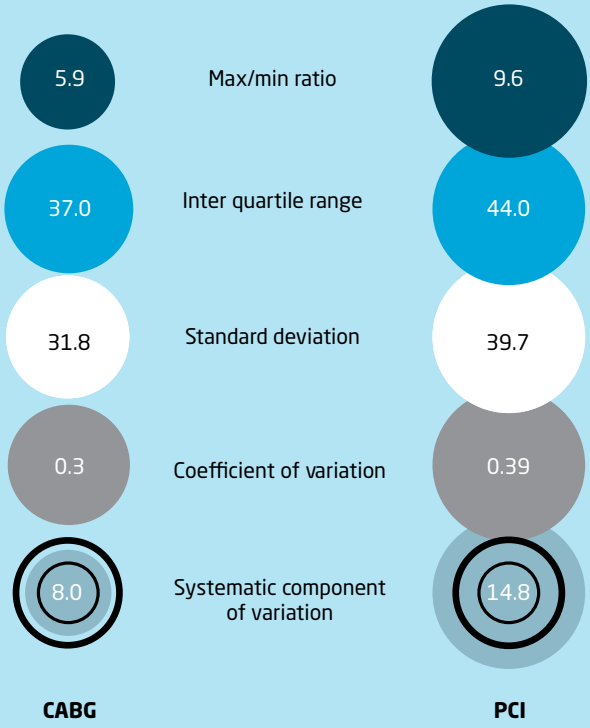
Although we have focused on joint replacement here, there is also evidence of inequity of access for other procedures such as cataract and revascularisation (Majeed *et al* 2002; Hippisley-Cox and Pringle 2000).

Choice: preference-sensitive variations

Preference-sensitive care refers to services that treat conditions for which there are legitimate, alternative treatment options, including watchful waiting, lifestyle changes, drug therapies and surgical and medical options (Dartmouth Atlas 2007). Frequently, these treatment options involve considerable trade-offs in terms of a patient's quality of life. Therefore, the decision about the appropriate course of treatment should reflect patients' preferences and demand for these treatments should reflect informed patient choice.

There are several common conditions with widely varying use of discretionary surgery, for example early stage cancer of the prostate or breast, osteoarthritis of the hip and knee, and gallstones. CABG and PCI surgery also represent typical examples of this type of care as they are two surgical options for the treatment of coronary artery disease for which non-surgical options are also available (Wennberg *et al* 2008). The international literature has repeatedly reported substantial variations in these types of services. The Dartmouth Atlas work, for instance, found striking differences in rates of CABG and PCI (five- and ten- fold differences respectively) across US hospital referral regions (Dartmouth Atlas 2005). Variations have also been reported for England (Department of Health 2006).

Figure 14 Measures of variation for CABG and PCI (2009/10)



Note: Size of circles drawn relative to highest variation procedure within each measure. Circles comparable within each measure but not between measures.

Thin/thick black circles indicate high (>5.4)/very high (>10) variation on SCV

© The King's Fund 2011

Analysis of CABG and PCI shows that admission ratios varied significantly by PCT in 2009/10. This variation is apparent across different measures of variation and the relative magnitude of the variation is fairly consistent across the measures (*see* Figure 14 opposite). In terms of SCV, CABG shows high variation with an SCV of 8.0, and PCI shows very high variation with an SCV of 14.8. For CABG, ratios vary almost six-fold between PCTs, from 33 to 196, while ratios of PCIs vary almost ten-fold, from 27 to 261 (*see* Figure 15 below). Furthermore, there are large statistically significant differences in the CABG ratios (*see* Figure 16 below), and some of these are in neighbouring PCTs.

Figure 15 Age-gender standardised ratios for CABG and PCI by PCT, 2009/10

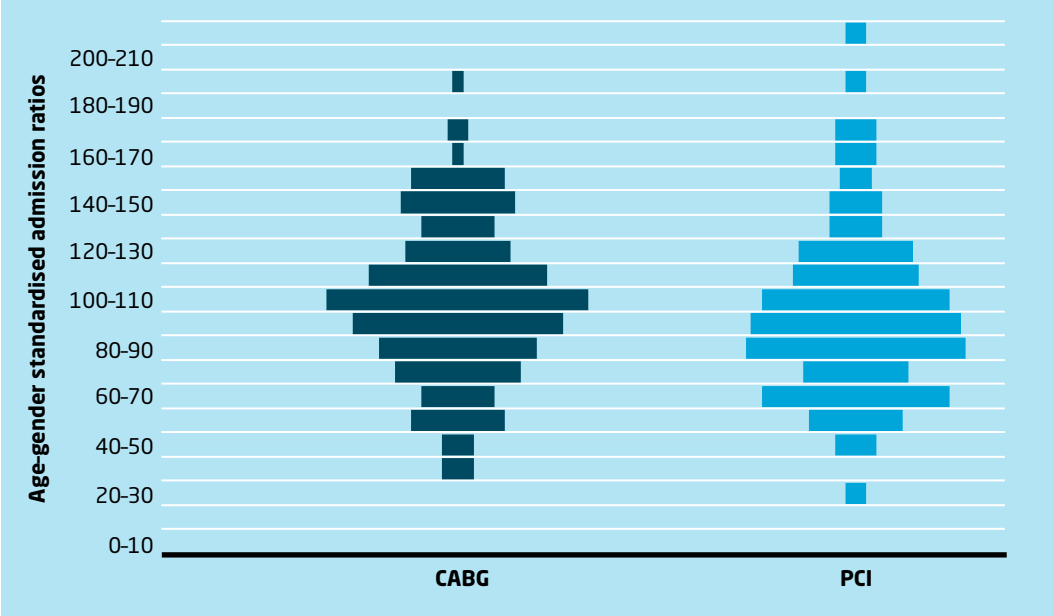
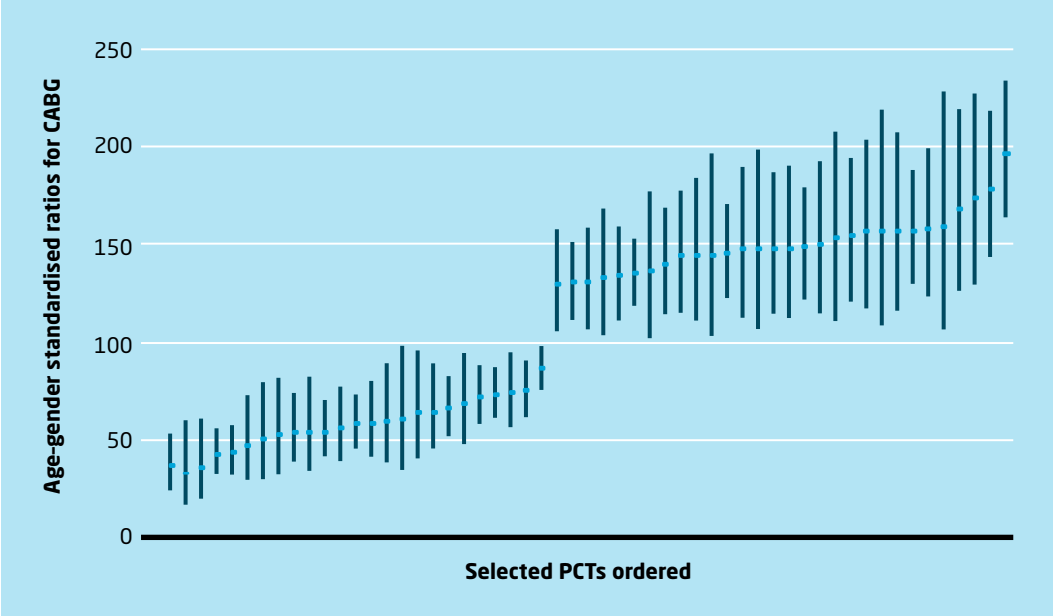


Figure 16 Age-gender standardised ratios for CABG, 2009/10

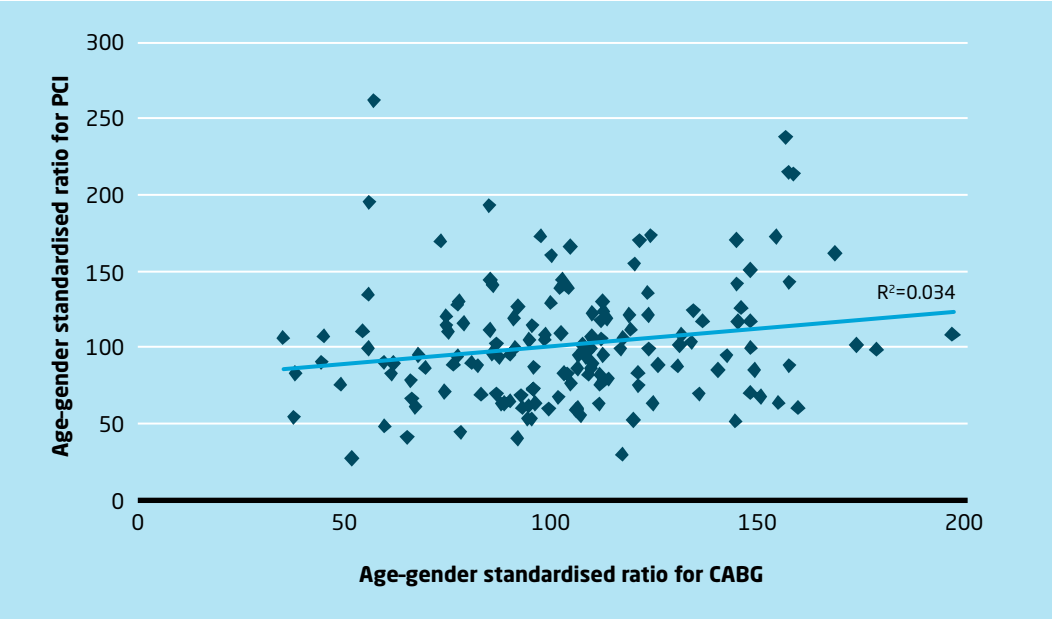


PCI is often considered a substitute for more invasive bypass surgery for some types of coronary heart disease. Even so, Dartmouth found a positive correlation among US hospital referral regions in the rates for these two procedures (Fisher *et al* 2010). We found a weak but nonetheless positive correlation between the ratios for these procedures

© The King's Fund 2011

(see Figure 17 below), and also that significant variations in the ratios of CABGs to PCIs persist, as noted also for previous years (Department of Health 2006).

Figure 17 Age-gender standardised ratios for CABG and PCI, 2009/10



While case-mix issues such as severity and co-morbidities undoubtedly contribute to the variations described, the magnitude of the variation suggests also the impact of other factors. While patient preferences might play a role, it is unlikely that such choices vary as sharply at the level of PCTs as do the surgical ratios. Probably a more important aspect of choice in relation to variations in admission ratios is the choice or clinical discretion exercised by clinicians. This hypothesis has been widely advanced as a cause of variation in ratios for common procedures (Bevan *et al* 2004 Dartmouth Atlas 2007 Wennberg *et al* 2008). Hence, for instance, the major variations even between adjoining US regions in surgery for lumpectomy and mastectomy, which have equal outcomes, in early stage breast cancer.

The Dartmouth work suggests medical opinion and/or physician preferences and attitudes have a substantial influence over which treatment patients will receive and are a major source of such variation (Wennberg *et al* 2008). The literature suggests two main reasons for this form of variation. First, there are variable treatment thresholds, especially in the early stages of the condition, and uncertainty around who benefits most from these treatments and at what stage of the condition. Second, the decision process is flawed either because patients commonly delegate this to clinicians or as a result of the physician's inability to accurately understand patients' values and preferences. Indeed, there is evidence that patients, if fully informed about their options, will often choose differently from their physicians and are less likely to elect for surgery than control groups (Wennberg *et al* 2008).

Previous analysis suggested that revascularisation rates overall were below estimated levels of need, but that it was also important to consider the options for patients as to how it is carried out (Department of Health 2006). The magnitude of variation observed in our analyses suggests variations in both revascularisation provision and relative CABG/PCI rates that in part could reflect the differential effects of clinical decision-making.

© The King's Fund 2011

3 Conclusions and recommendations

As this report has shown, variations in health care continue to exist in the NHS. Some will be good, or warranted, others bad and unwarranted. Such variations are not limited to relatively rare or new interventions. Nor are such variations a surprising new phenomenon, and nor are they simply concerned with the efficient use of scarce health service resources. The existence of persistent unwarranted variations in health care directly impacts on equity of access to services and the health outcomes of populations. But the eradication of all variation is not the task. As Mulley (2010) has noted:

> *If all variation were bad, solutions would be easy. The difficulty is in reducing the bad variation, which reflects the limits of professional knowledge and failures in its application, while preserving the good variation that makes care patient centred. When we fail, we provide services to patients who don't need or wouldn't choose them, while we withhold the same services from people who do or would, generally making far more costly errors of overuse than of underuse.*

The question is, what's to be done? To expand on Mulley's points: identifying variation that can be deemed 'bad' or unwarranted, or conversely 'good', is not necessarily an easy task. As Evans (1990) has noted, trying to set an appropriate normative rate of health care utilisation for, say, a particular population, based on what is known about what proportion of observed variation is explained by the outcomes of, for example, statistical investigation modelling 'need' for health care and other factors that drive legitimate variation, is very hard. Such work, as noted in Section 1 and Figure 1, can to a degree help to expose the 'variations in variations'.

However, as Mulley and others have emphasised, an alternative tack on the problem, and one which, if successful, would work to drive out the bad and encourage good variation, is to focus on the *process* leading to individual clinical decisions – decisions to refer or not refer, to treat or not treat, to treat in one way and not another – rather than attempt to specify the *outcome* of decisions (in aggregate for populations). This is analogous to the concept of procedural justice where the outcome of a decision (for example to resolve a dispute over allocation of resources) will vary from case to case, but in all cases will be based on an agreed fair process. In the case of clinical decisions, while 'fairness' (however conceived) may be part of the decision-making process, the important aspect emphasised by Mulley and others is the sharing of decisions between clinician and patient in a way that, as Elwyn *et al* (2010) state,

> *…clinicians and patients make decisions together using the best available evidence. Patients are encouraged to think about the available screening, treatment, or management options and the likely benefits and harms of each so that they can communicate their preferences and help select the best course of action for them.*

Shared decision-making (SDM) is of course not new, but implementation has been difficult. Elwyn *et al* (2010) note that ready access to evidence about treatment options, guidance on weighing pros and cons of different options and a supportive clinical culture that facilitates patient engagement are critical to successful adoption of SDM.

An important issue in tackling unwarranted variations in future will be the impact of the current reforms of the NHS on the ability of organisations and individuals to make

© The King's Fund 2011

headway with this most persistent of problems. To what extent, for example, will GPs and GP commissioning consortia (GPCC) be better placed perhaps to tackle inequalities in access and implement shared decision-making? How will or could the new roles proposed for the National Institute for Health and Clinical Excellence (NICE) help in generating the right kind of clinical evidence and in particular the guidance Elwyn *et al* note is necessary on how to weigh up the trade-offs such evidence inevitably reveals?

Moves to GP commissioning raise some practical and technical issues with respect to health care variations, not least the unit of analysis. In this report we have used PCTs as the unit of analysis. However, their replacement with GPCCs creates new boundaries that will require reworking of utilisation data to reflect the new groupings of consortia populations – an issue that will also affect the Department of Health's new Atlas of Variations (Department of Health 2010a). The King's Fund will in due course undertake analyses of variations by GPCC.

More broadly, it remains an open question as to whether the new commissioning arrangements (and other aspects of the reforms) will promote greater action on variations than hitherto. In some respects, it could be that the combination of GP responsibility for commissioning through consortia, and tighter funding, may provide incentives to tackle certain aspects of variations, for example primary care prescribing and referrals to secondary care. On the other hand, the scope of health care variations is considerable and leaves responsibility for tackling them somewhat diffuse. Moreover, the financial imperatives for the next few years could mean that variations are addressed in rather crude ways with, for example, above average admission areas aiming simply to cut access rather than seeking to establish appropriate utilisation rates through, for example, shared decision-making and consideration of equity issues.

The first step in addressing unwarranted variations in health care is the systematic and routine collation, analysis and publication of such variations. Although there have been many examples in the NHS of reporting of particular aspects of health care variations, these have tended to be sporadic and piecemeal, without due consideration of the possible underlying causes. The recent launch of the Department of Health Atlas of Variations (Department of Health 2010a) will hopefully act as a real focus for future reporting of variations. An important development for the future will be to report variations for new population groups covered by GPCC and to extend the scope of reporting – particularly into primary care where, as The King's Fund's inquiry into the quality of general practice has shown, significant variations exist (The King's Fund 2011).

While collating and presenting evidence of health care variations is a key first step, and while the broad and in some cases specific causes for variations are known or at least postulated, there is a subsequent need for a programme of work to identify not only the causes of variation at specific local level, but to prioritise those variations and causes that have the most important impact on equity, effectiveness, efficiency and patient health outcomes.

Knowledge does not, unfortunately, always lead to action. Publicising the existence of unwarranted variations and their causes does not guarantee that they will be tackled. There is a further need for local health organisations – both providers of care and commissioners – to be required to publicly justify and explain in a consistent way their relative position on key aspects of health care variation. Further, it may also be necessary to explore the development of harder-edged, locally focused incentives to encourage action to deal with unwarranted variation and to properly engage clinicians with the extent of variations not just at large geographical levels but also between clinical peers.

Most importantly, while publication of variations, the use of incentives, and the development and promulgation of clinical guidelines and other strategies have their place, what is also needed is a much greater encouragement of shared decision-making to establish the *right* level of variation based on patients' own assessments of needs and risk aversion.

© The King's Fund 2011

References

Audit Commission (2001). 'Day Surgery'. Audit Commission website. Available at: www.audit-commission.gov.uk/SiteCollectionDocuments/AuditCommissionReports/NationalStudies/daysurgery.pdf (accessed on 6 March 2011).

Audit Commission (1992). *All in a Day's Work: An audit of day surgery in England and Wales*. London: Audit Commission.

Aylin P, Williams S, Jarman B, Bottle A (2005a). 'Variation in operation rates by primary care trust'. *British Medical Journal*, vol 331, no 7516, p 539.

Aylin P, Williams S, Jarman B, Bottle A (2005b) 'Trends in day surgery rates'. *British Medical Journal*, vol 331, no 7520, p 803.

Bakwin H (1958). 'The tonsil-adenoidectomy enigma'. *Journal of Pediatrics*, vol 52, no 3, pp 339–61.

Bevan G (2009). 'The search for a proportionate care law by formula funding in the English NHS'. *Financial Accountability and Management*, vol 25, no 4, pp 391–410.

Bevan G, Hollinghurst S, Benton P, Spark V, Sanderson H, Franklin D (2004). 'Using information on variation in rates of supply to question professional discretion in public services'. *Financial Accountability and Management*, vol 20, no 1, pp 1–17.

Bloor M (1976). 'Bishop Berkeley and the adenoidectomy enigma: an exploration of variation in the social construction of medical disposals'. *Sociology*, vol 10, no 1, pp 43–61.

Burkinshaw J (1956). 'Is tonsillectomy necessary?' *British Medical Journal*, vol 2, no 4994, pp 713. Available at: www.bmj.com/cgi/pdf_extract/2/4994/713 (accessed on 25 February 2011).

Burton MJ, Glasziou PP (2009). 'Tonsillectomy or adeno-tonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis (Cochrane Review)'. Cochrane Database of Systematic Reviews, issue 1, article CD001802. Available at: 10.1002/14651858.CD001802.pub2 (accessed on 25 February 2011).

Chassin MR, Kosecoff J, Park RE, Winslow CM, Kahn KL, Merrick NJ, Keesey J, Fink A, Solomon DH, Brook RH (1987). 'Does inappropriate use explain geographic variations in the use of health care services?'. *Journal of the American Medical Association*, vol 258, no 18, pp 2533–7.

Dartmouth Atlas (2007). 'Preference-sensitive Care'. Dartmouth Atlas of Health Care website Available at: www.dartmouthatlas.org/downloads/reports/preference_sensitive.pdf (accessed on 25 February 2011).

Dartmouth Atlas (2005). 'Cardiac Surgery'. Dartmouth Atlas of Health Care website. Available at: www.dartmouthatlas.org/downloads/reports/Cardiac_report_2005.pdf (accessed on 25 February 2011).

Department of Health (2010a). 'NHS Atlas of Variation in Healthcare: Reducing unwarranted variation to increase value and improve quality'. Right Care website. Available at: www.rightcare.nhs.uk/atlas/ (accessed on 25 February 2011) (interactive web version of Atlas: http://www.sepho.org.uk/extras/maps/NHSatlas/atlas.html).

Department of Health (2010b). '2009–10 and 2010–11 PCT recurrent revenue allocations exposition book'. National Archives website. Available at: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Managingyourorganisation/Financeandplanning/Allocations/DH_091850 (accessed on 25 February 2011).

Department of Health (2006). 'On the State of the Public Health: Annual report of the Chief Medical Officer 2005'. DoH website. Available at: www.dh.gov.uk/en/Publicationsandstatistics/Publications/AnnualReports/DH_4137366 (accessed on 25 February 2011).

© The King's Fund 2011

Dixon T, Shaw M, Ebrahim E, Dieppe P (2004). 'Trends in hip and knee joint replacement: socioeconomic inequalities and projections of need'. *Annals of the Rheumatic Diseases*, vol 63, no 7, pp 825–30.

Elwyn G, Coulter A, Walker E, Watson P, Thomson R (2010) 'Implementing shared decision making in the NHS'. *British Medical Journal*, vol 341, p c5146. Available at: www.bmj.com/content/341/bmj.c5146.full?ath_user=nhsjappleby003&ath_ttok=%3CTUmCY6Ou1O2RA53bPA%3E (accessed on 25 February 2011).

Evans RG (1990). 'The dog in the night-time: Medical practice variations and health policy' in Andersen TF, Mooney G (eds), *The Challenges of Medical Practice Variations*, pp 117–52. London: Macmillan Press.

Fisher E, Bell JE, Tomek I, Esty A, Goodman D (2010). 'Trends and Regional Variation in Hip, Knee and Shoulder Replacement'. Dartmouth Atlas of Health Care website. Available at: www.dartmouthatlas.org/downloads/reports/Joint_Replacement_0410.pdf (accessed on 25 February 2011).

Glover AJ (1938). 'The incidence of tonsillectomy in schoolchildren'. *Proceedings of the Royal Society of Medicine*, vol 31, no 10, pp 1219–36.

Gottlieb DJ, Zhou W, Song Y, Andrews KG, Skinner JS, Sutherland JM (2010). 'Prices don't drive regional Medicare spending variations'. *Health Affairs*, vol 29, no 3, pp 537–43. Available at: http://content.healthaffairs.org.gate2.library.lse.ac.uk/cgi/reprint/29/3/537?maxtoshow=&hits=10&RESULTFORMAT=&fulltext=Gottlieb&andorexactfulltext=and&searchid=1&FIRSTINDEX=0&resourcetype=HWCIT (accessed on 25 February 2011).

Gray D, Hampton JR, Bernstein SJ, Brook RH (1990). 'Audit of coronary angiography and bypass surgery'. *The Lancet*, vol 335, no 8701, pp 1317–20.

Ham C (ed) (1988). *Health Care Variations: Assessing the evidence*. London: King's Fund Institute.

Hawker GA, Wright JG, Glazier RH, Coyte PC, Harvey B, Williams JI, Badley (2002). 'The effect of education and income on need and willingness to undergo total joint arthroplasty'. *Arthritis and Rheumatism*, vol 46, no 12, pp 3331–9.

Healthcare Commission (2005). *Acute Hospital Portfolio Review: Day surgery*. London: Healthcare Commission.

Hippisley-Cox J, Pringle M (2000). 'Inequalities in access to coronary angiography and revascularisation: the association of deprivation and location of primary care services'. *British Journal of General Practice*, vol 50, no 455, pp 449–54.

Judge A, Welton NJ, Sandhu J, Ben-Shlomo Y (2010). 'Equity in access to total joint replacement of the hip and knee in England: cross sectional study'. *British Medical Journal*, vol 341, no 7770, p c4092.

Leape LL, Park RE, Solomon DH, Chassin MR, Kosecoff J, Brook RH (1990). 'Does inappropriate use explain small area variations in the use of health care services?'. *Journal of the American Medical Association,* vol 263, no 5, pp 669–72.

Majeed A, Eliahoo J, Bardsley M, Morgan D, Bindman AB (2002). 'Variation in coronary artery bypass grafting, angioplasty, cataract surgery, and hip replacement rates among primary care groups in London: association with population and practice characteristics'. *Journal of Public Health Medicine*, vol 24, no 1, pp 21–6.

McPherson K, Downing A, Buirski D (1996). *Systematic Variation in Surgical Procedures and Hospital Admission Rates*. PHP Departmental Publication No. 23. London: London School of Hygiene and Tropical Medicine.

McPherson K, Wennberg JE, Hovind OB, Clifford P (1982). 'Small-area variations in the use of common surgical procedures: an international comparison of New England, England, and Norway'. *New England Journal of Medicine*, vol 307, no 21, pp 1310–14.

Mulley AJ (2010). 'Improving productivity in the NHS'. *British Medical Journal*, vol 341, no 7766, p c3965.

Newton JN, Seagroatt V, Goldacre M (1994). 'Geographical variation in hospital admission rates: an analysis of workload in the Oxford region, England'. *Journal of Epidemiology and Community Health*, vol 4, no 6, pp 590–5.

© The King's Fund 2011

Price CE, Paul EA, Bevan RG, Holland WW (1992). 'Equity and medical practice variation: relationships between standardised discharge ratios in total and for selected conditions in English districts'. *Journal of Epidemiology and Community Health*, vol 46, no 1, pp 58–62.

Suleman M, Clark MPA, Goldacre M, Burton M (2010). 'Exploring the variation in paediatric tonsillectomy rates between English regions: a 5-year NHS and independent sector data analysis'. *Clinical Otolaryngology*, vol 35, no 2, pp 111–17.

The King's Fund (2011). 'Inquiry into the Quality of General Practice in England'. The King's Fund website. Available at: www.kingsfund.org.uk/current_projects/gp_inquiry/index.html (accessed on 25 February 2011).

Weinstein JN, Bronner KK, Morgan TS, Wennberg JE (2004). 'Trends and geographic variations in major surgery for degenerative diseases of the hip, knee and spine'. *Health Affairs* (web exclusive V AR-81) Available at: www.tdi.dartmouth.edu/documents/publications/2004Oct7_HealthAffairs_Weinstein.pdf (accessed on 25 February 2011).

Wennberg JE (2010). *Tracking Medicine*. New York: Oxford University Press.

Wennberg JE, Gittelsohn A (1982). 'Variations in medical care among small areas'. *Scientific American*, vol 246, no 4, pp 120–34.

Wennberg JE, Brownlee S, Fisher E, Skinner J, Weinstein J (2008) ' An Agenda for Change: Improving quality and curbing health care spending: Opportunities for the Congress and the Obama Administration.' Dartmouth Atlas of Health Care website. Available at: http://www.dartmouthatlas.org/downloads/reports/agenda_for_change.pdf (accessed on 25 February 2011).

Wennberg JE, Fisher ES, Skinner JS (2002). 'Geography and the debate over Medicare reform'. Health Affairs website. Available at http://content.healthaffairs.org/cgi/citmgr?gca=healthaff;hlthaff.w2.96v1 (accessed on 25 February 2011).

Wennberg JE, Freeman JL, Culp WJ (1987). 'Are hospital services rationed in New Haven or over-utilised in Boston?' *The Lancet,* vol 329, no 8543, pp 1185–9.

Wennberg JE, McPherson K, Caper P (1984). 'Will payment based on diagnosis-related groups control hospital costs?'. *New England Journal of Medicine*, vol 311, no 5, pp 295–300.

© The King's Fund 2011

Appendix

Methods

This section describes the data used in the study, the analyses undertaken and the analytical methodology.

Analysis of variation in selected procedures

We used the 2009–2010 provisional inpatient Hospital Episode Statistics (HES) dataset, and HES data for 2005–2006, to analyse variation in health care utilisation across the 152 primary care trusts (PCTs) in England. To illustrate the geographical variation that exists across PCTs, rates of selected procedures were analysed by PCT as follows:

1 Rates of selected high volume procedures and/or procedures incurring high levels of expenditure: hip and knee replacements, cataract surgery, CABG and PCI and cholecystectomy (*see* Box 1 overleaf). Trends in rates of hip replacement, cataract removal and tonsillectomy between 2005–6 and 2009–10 were also analysed. For this analysis, the numerator was the number of elective finished consultant episodes (FCEs) (day-cases and ordinary) with an HRG 3.5 code from the selected list of HRGs.

2 The Better Care Better Value surgical thresholds indicator. This indicator includes a basket of six procedures (listed in Box 2 overleaf) for which there is significant clinical uncertainty and evidence of overuse (www.productivity.nhs.uk/). For this analysis, the numerator was the number of elective FCEs (day-cases and ordinary) with an OPCS4 code from the procedures listed in the surgical threshold basket in any procedure field in the episode.

3 Rates of day-case surgery. This indicator includes a basket of 25 procedures (listed in Box 3 overleaf) identified by the Audit Commission as being amenable to safe and cost-effective day case surgery(Audit Commission 2001). This is also a Better Care Better Value indicator. The day-case surgery indicator was derived using the Audit Commission methodology (Audit Commission 2001). It is expressed as the percentage of elective operations from the Audit Commission basket of 25 procedures (identified by OPCS4 codes) performed as day cases.

For the first two analyses listed above, variation in 2009–10 across PCTs was analysed using indirectly age and gender standardised ratios. Age and gender specific rates for England were used as the standard and applied to the mid-2008 PCT resident populations to derive the ratio of the observed to expected number of episodes, given the age and gender distribution of the PCT population. (The 2009 PCT populations were not available at the time of the analysis.) Ninety-five per cent confidence intervals were derived using the NCHOD methodology (http://www.nchod.nhs.uk/).

© The King's Fund 2011

Association with other variables

We examined whether or not PCT variations in the selected procedure rates were associated with:

- The Index of Multiple Deprivation (IMD) 2007
- PROMS pre-operative scores for hip and knee replacements. Pre-operative PROMs scores became available in April 2010 (HES online) and can be used as proxy for a patient's pre-operative health status and need for surgery. We compared pre-operative PROMs scores for hip and knee replacements with hip and knee replacement rates by PCT, to examine whether access was commensurate with need. (PROMs scores are collected for patients receiving hip and knee replacements based on procedures codes rather than HRGs.)

Bivariate associations between procedure rates on the one hand, and the above variables on the other, were examined using Spearman's correlation coefficients and corresponding p values.

Box 1 Selected elective procedures: HRG codes used

Hip primary replacement	H71, H80,H81
Hip revision	H01
Knee primary replacement	H03,H04
Knee revision	H72
Cataract	B13,B14
CABG	E04
PCI	E15
Cholecystectomy	G13,G14

Box 2 Surgical threshold procedures: OPCS 4 codes used

Abdominal excision of uterus	Q07
Myringotomy with/without grommets	D15 (excluding E081,E201,F291,F34,D191 in any position)
Tonsillectomy	F341-F344 (NB F347 part of NICE recommended procedures)
Dilation and curettage/hysteroscopy	Q103,Q18 (004 is not primary diagnosis)
Vaginal excision of uterus	Q08
Lumbar spine procedures	V25,V26,V33,V34,V382-4,V393-5,V433,V473,V485-6,V493

Box 3 Audit Commission basket of day case procedures

Orchidopexy	Bunion operations
Circumcision	Removal of metalware
Inguinal hernia repair	Extraction of cataract with/without implant
Excision of breast lump	Correction of squint
Anal fissure dilatation or excision	Myringotomy
Haemorrhoidectomy	Tonsillectomy
Laparoscopic cholecystectomy	Sub mucous resection
Varicose vein stripping or ligation	Reduction of nasal fracture
Transurethral resection of bladder tumour	Correction of bat ears
Excision of Dupuytren's contracture	Dilatation and curettage/hysteroscopy
Carpal tunnel decompression	Laparoscopy
Excision of ganglion	Termination of pregnancy
Arthroscopy	

© The King's Fund 2011

References

Audit Commission (2001). 'Day Surgery'. Audit Commission website. Available at: www.audit-commission.gov.uk/SiteCollectionDocuments/AuditCommissionReports/NationalStudies/daysurgery.pdf (accessed on 6 March 2011).

© The King's Fund 2011