

TO PUBLISH OR NOT?

EXPERIENCE AND EVIDENCE ABOUT
PUBLISHING HOSPITAL OUTCOMES
DATA

Anne Mason and Andrew Street



Office of Health Economics
12 Whitehall London SW1A 2DY
www.ohe.org

© February 2004. Office of Health Economics.

Any part of this publication may be reproduced without permission for educational and non-profit purposes only if the source is acknowledged. All other rights reserved, including the right of reproduction in whole or in part in any form.

Price £12.50

ISBN

About the authors

Abstract

There is an increasing and controversial policy trend in many countries towards the publication of outcomes data on hospitals and doctors. The policy rationale is that, above and beyond merely measuring and collecting outcomes data, publishing the information spurs quality control and performance improvement. In the UK and the US, a wide range of both clinical and process outcomes is published. Our report reviews the history of hospital outcomes publication on both sides of the Atlantic, noting the differences and similarities in the two systems. We focus primarily on clinical outcomes such as morbidity and mortality, but process outcomes, that are often used either to proxy the former or that are published in conjunction with clinical outcomes data, are also considered. We consider the potential beneficial and harmful effects of publishing health outcome data, presenting the limited evidence that exists. In particular, we consider the risks of making incorrect inferences based on these data and the potential for dysfunctional consequences. Recognising that the public largely mistrusts currently published health outcomes data, we offer some recommendations for the future direction of policy.

THE OFFICE OF HEALTH ECONOMICS

The Office of Health Economics (OHE) was founded in 1962. Its terms of reference are to:

- commission and undertake research on the economics of health and health care;
- collect and analyse health and health care data from the UK and other countries;
- disseminate the results of this work and stimulate discussion of them and their policy implications.

The OHE is supported by an annual grant from the Association of the British Pharmaceutical Industry and by revenue from sales of its publications, consultancy and commissioned research.

Independence

The research and editorial independence of the OHE is ensured by its Policy Board:

Professor Tony Culyer (*Chair*) – *University of York*

Professor Patricia Danzon – *The Wharton School of the University of Pennsylvania*

Professor Naoki Ikegami – *Keio University*

Dr Trevor Jones – *Director General of the Association of the British Pharmaceutical Industry*

Professor Vincent Lawton – *Merck Sharp & Dohme Limited (UK) and President of the Association of the British Pharmaceutical Industry*

Professor David Mant – *University of Oxford*

Dr Nancy Mattison – *The Mattison Group Inc.*

Professor Sir Michael Peckham – *University College, University of London*

Peer Review

All OHE publications have been reviewed by members of its Editorial Board and, where appropriate, other clinical or technical experts independent of the authors. The current membership of the Editorial Board is as follows:

Professor Christopher Bulpitt – *Royal Postgraduate Medical School, Hammersmith Hospital*

Professor Martin Buxton – *Health Economics Research Group, Brunel University*

Professor Tony Culyer (*Chair*) – *Department of Economics and Related Studies, University of York*

Dr Jennifer Dixon – *The King's Fund*

Professor Hugh Gravelle – *Centre for Health Economics, University of York*

Mr Geoffrey Hulme – *Director, Public Finance Foundation*

Professor Carol Propper – *Department of Economics, University of Bristol*

Professor Bonnie Sibbald – *National Primary Care Research and Development Centre, University of Manchester*

Mr Nicholas Wells – *Head of European Outcomes Research, Pfizer Ltd*

Professor Peter Zweifel – *Socioeconomic Institute, University of Zurich*

Further information about the OHE is on its website at: www.ohe.org

GLOSSARY

4

AAMC	Association of American Medical Colleges
AHA	American Hospital Association
AHRQ	Agency for Healthcare Research and Quality (formerly the Agency for Health Care Policy and Research)
AMI	Acute myocardial infarction (heart attack)
BMA	British Medical Association
CABG	Coronary artery bypass graft
CCHRI	California Cooperative Healthcare Reporting Initiative
CEPOD	Confidential Enquiry into PeriOperative Deaths (renamed in 2002 as 'The National Confidential Enquiry into Patient Outcome and Death')
CHOP	California Hospital Outcomes Project
CGR	Clinical governance review
CHAI	Commission for Healthcare Audit and Inspection ('Healthcare Commission')
CHI	Commission for Health Improvement
CHQC	Cleveland Health Quality Choice
CMS	Centers for Medicare & Medicaid Services (formerly known as the HCFA)
CRAG	Clinical resource and audit group
CRSR	(New York State) Cardiac Surgery Reporting System
DH	Department of Health
DHSS	Department of Health and Social Security
FAH	Federation of American Hospitals
GMC	General Medical Council
HAA	Hospital Activity Analysis (replaced by HES in 1987)
HCFA	Health Care Financing Administration (replaced by CMS in 2001)
HEDIS	Health Plan Employer Data information Set
HES	Hospital Episode Statistics
HIPE	Hospital Inpatient Enquiry
HSI	Health service indicator
JCAHO	Joint Commission on Accreditation of Healthcare Organizations
NCAA	National Clinical Assessment Authority
NCQA	National Committee for Quality Assurance (accreditor of health plans)
NHS	National Health Service
NICE	National Institute for Clinical Excellence
NQF	National Quality Forum
PAF	Performance Assessment Framework
PCT	Primary Care Trust
SCTS	Society of Cardiothoracic Surgeons

CONTENTS

5

Chapter 1 – Introduction	8
Chapter 2 – Publication of hospital outcomes data in the UK	10
2.1 Hospital mortality statistics: a story of birth, death, burial and resurrection	12
2.1.1 A birth and brief life	15
2.1.2 Buried under a wealth of information	17
2.1.3 Resurrection	17
2.2 Self-regulation is not enough	18
2.2.1 Medical Audit	19
2.2.2 Disciplinary procedures	20
2.2.3 Bristol: a watershed	23
2.2.4 A political shift	25
2.3 Publication of outcomes data in Scotland	27
2.4 Publication of outcomes data in England	29
Chapter 3 – Publication of hospital outcomes data in the US	39
3.1.1 Codman: a pioneer	52
3.1.2 The HCFA picks up the baton	53
3.1.3 Report cards on hospitals and surgeons	54
3.1.4 Report cards on Health Plans	56
3.1.5 Recent developments	57
Chapter 4 – UK and US: a story of convergence?	59
Chapter 5 – The pros and cons of publication	62
5.1 How might publication be beneficial?	62
5.1.1 Promoting public trust	62
5.1.2 Supporting patient choice	64
5.1.3 Stimulating action	68
5.2 Evidence for the beneficial effects of publication	70
5.3 Data problems	75
5.3.1 The limitations of routine data	75
5.3.2 Problems with the valid interpretation of hospital outcomes	76
5.4 Dysfunctional consequences of publication	79
6 Recommendations	84
7 Acknowledgements	88
8 References	89

LIST OF BOXES

- 6
- Box 1: **Definition of key terms**
 - Box 2: **Scottish hospital clinical indicators (1994)**
Seven acute hospital indicators:
 - Box 3: **Hospital outcomes reporting in the US: the impetus for publication**

LIST OF TABLES

- 7
- Table 1: **An overview of the publication of performance data in the UK, 1860 to 2004**
 - Table 2: **NHS hospital outcomes data: forms of publication and target audience, England 2004**
 - Table 3:
 - Table 4: **Possible dysfunctional consequences arising from the publication of outcomes data**

Since the mid-1980s, the publication of hospital outcomes data has become increasingly popular. Canada and the US operate performance reporting systems, with similar initiatives planned in Australia and New Zealand (Mannion and Davies, 2002). In Europe, outcomes data are collected and published in the UK, Italy, Scandinavia and the Netherlands (Marshall and Brook, 2002).

Whilst political and regulatory attitudes differ across countries, the economic motivations are often similar. Concern with escalating health care costs and regional variations in the quality of care has led countries to examine ways to improve value for money from their health systems, one of which is to place key data in the public domain (Marshall et al., 2003, Davies and Lampel, 1998, Fottler et al., 1987). Publication also forms part of a framework of accountability, along with regulation such as audit, accreditation, licensing and inspections; market (or quasi market) forces; and legislation (Davies, 1999). More broadly, publication is a means of realising the key political and cultural objectives of transparency and openness: publication aims to promote – or restore – public trust (Marshall et al., 2000a, Davies and Shields, 1999). A related motivation is that the public places ever more emphasis on a ‘right to know’ about goods and services generally, a right that is recognised in the Freedom of Information Act 2000. Thus, the availability of information is also valued for its own sake, whether or not this changes behaviour.

The aims of this report are to describe and discuss particular systems of publication of hospital outcomes data in some detail, to present a critique of what these systems, why they were set up and how they operate, and to explore the potential implications for policy and practice in the UK health care system. Box 1 provides some definitions of the key terms used in the report.

Box 1: Definition of key terms

‘Publication’ includes paper and electronic media.

‘Hospital’ includes institutions providing acute and/or mental health services, including specialist services.

‘Outcome’ is defined primarily to mean clinical health outcomes such as morbidity and mortality. However, we also consider those ‘process outcomes’ that are often used either to proxy the former or that are published in conjunction with clinical outcomes data. To artificially abstract the clinical outcomes from their context could lead to a misunderstanding and misinterpretation of findings.

Failures of professional self-regulation in the UK led to the recent surge in publicity for hospital outcomes data. In section 2, we outline the background in the UK, before presenting a more detailed account of recent developments in England and Scotland. In section 3, we provide a similar overview of hospital outcomes publication for the US. There are a number of reasons why the US is interesting for comparative purposes, including its high levels of sensitivity to consumer concerns across many walks of life, the long track record it has in publishing health services data, and the wide range of approaches currently in operation. In section 4, we discuss the similarities and differences between the two countries. The beneficial effects and drawbacks of publication, drawing on the experience of the UK and the US, are covered in section 5. Section 6 draws together the findings from the previous sections and offers some recommendations for the future direction of hospital outcomes publication.

CHAPTER 2 – PUBLICATION OF HOSPITAL OUTCOMES DATA IN THE UK

10

Table 1: An overview of the publication of performance data in the UK, 1860 to 2004

Date	Event
1860	Florence Nightingale proposes the publication of hospital mortality statistics. Data published for the years 1861 to 1865, until Royal College of Surgeons' report (Nightingale, 1860).
1949	Limited hospital activity data collection for all NHS hospitals; for a 10% sample of hospitals, the Hospital Inpatient Enquiry (HIPE) collects more detailed data, published annually in national tables (Aylin et al., 1999).
1969	Data on discharges from and deaths in hospital are collected for all NHS hospitals under the Hospital Activity Analysis (HAA) scheme (Aylin et al., 1999).
1983	DHSS publishes Health Service Indicators (HSIs), which include indicators of clinical activity (Ferguson and McGuire, 1984).
1987	Hospital Episode Statistics (HES) replace HAA. HES attempts to measure all routine hospital activity for inpatients in English hospitals, including whether patients died in hospital (Aylin et al., 1999, Jarman et al., 1999). Standardised mortality rates for health authorities (but not for hospitals) added to the set of HSIs.
1989	Three mortality indicators for each district health authority added to the HSIs: post neonatal mortality rate (ML41); mortality rate for the resident population of children (ML42) and standardised mortality rates (ML61) (Department of Health, 1989).
1991	Patient's Charter published in Scotland, England, Wales and Northern Ireland. NHS to publicise maximum waiting times and quality standards, covering cancelled operations, discharge arrangements (Department of Health, 1991, Scottish Office, 1991).
1992	Standardised death rates indicator for hospitals added to English HSI set and reproduced in The Times (Street, 2002).
1994	Scottish Executive publishes performance (CRAG) data, including emergency readmission with 28 days of discharge, and 30-day survival after emergency admission for stroke (Scottish Office, 1995, Mannion and Goddard, 2003).
1995	Revised Patient's Charter published, with new targets including a guaranteed maximum waiting time of 18 months for all operations (Department of Health, 1995).
1996	HSIs discontinued (Street, 2002).

CHAPTER 2 - PUBLICATION OF HOSPITAL OUTCOMES DATA IN THE UK

11

1998	National Performance Assessment Framework (PAF) consultation document proposes publication of indicators for in-hospital premature deaths (NHS Executive, 1998b). Final publication of Performance Tables using Patient's Charter indicators (Department of Health, 1998).
1999	DH publishes PAF indicators for 1997/98. These feature a set of clinical indicators, including death rates, for acute Trusts (Department of Health, 1999).
2000	NHS Plan outlines reform of PAF: Commission for Health Improvement (CHI) to publish annual 'report cards' for NHS Trusts. Regional offices to rate NHS organisations using 'traffic light' scores. Patient choice agenda and Performance Fund introduced (Department of Health, 2000a).
2001	The Sunday Times publishes Dr Foster's Good Hospital Guide, featuring detailed information on local hospitals (Street, 2002). 'Your Guide to the NHS' replaces the Patient's Charter in England: expected standards of care include waiting times and access targets (Department of Health, 2001b). Department of Health publishes NHS performance ratings for English acute Trusts (Department of Health, 2001a).
2002	National Health Service Reform and Health Care Professions Act 2002 passes responsibility for publishing performance ratings from DH to CHI. Department of Health publishes NHS performance ratings for acute Trusts, mental health, ambulance and specialist Trusts ¹ . Department of Health publishes NHS performance ratings for all NHS Trusts and primary care trusts (PCTs) in England ² .
2003	CHI publishes Performance ratings for NHS Trusts and PCTs in England, including patient survey data ³ .
2003/04	In December, NHS launches 'You're Pregnant!' a new maternity magazine in collaboration with Dr Foster, giving tailored information about local NHS services (Department of Health, 2003a). One month later, 'Your Life!' a new women's health magazine containing celebrity interviews, becomes available free of charge from newsagents, high street shops and beauty salons (Department of Health, 2004a).
2004	British Medical Journal begins a new monthly article entitled 'Dr Foster's case notes', highlighting particular quality issues (Jarman et al., 2004). A similar series from the same source appears in the Health Service Journal. Plans to publish data on risk-adjusted individual surgeon performance from April 2004 (Department of Health, 2002a), have been delayed (Day, 2004). Data will probably include only first time bypass operations, which constitute about two-thirds of a cardiac surgeon's workload (Dyer, 2003).

¹ <http://www.dh.gov.uk/assetRoot/04/05/97/73/04059773.pdf> (accessed 08/11/04)

² http://www.performance.doh.gov.uk/performance/ratings/2002/national_pco.html (accessed 08/11/04)

³ <http://www.chi.nhs.uk/ratings/> (accessed 08/11/04)

2.1 Hospital mortality statistics: a story of birth, death, burial and resurrection

2.1.1 A birth and brief life

The *collection* of mortality statistics in the UK dates back to the early 16th century when weekly 'Bills of mortality' were collected to track the deadly impact of the plague. From the 1600s, collection widened to include hospital statistics (Iezzoni, 1996). In Victorian Britain, when hospitals were essentially charitable institutions, statistics were used to demonstrate value for money to wealthy benefactors and to encourage contributions from other donors.

The *publication* of hospital mortality data can be traced back to the efforts of Florence Nightingale (1820 – 1910) and William Farr (1807 – 1883) in the mid-1800s. Driven by concern about the 'great and unnecessary waste of life' that they attributed to poor hygiene, Nightingale and Farr used hospital statistics to 'wage their civilian reform campaign' (Iezzoni, 1996). In 1859 Nightingale had a set of her 'Model Forms' printed for distribution among hospitals, allowing them to record hospital mortality by disease, injury and type of operation (Nightingale, 1860). Her intention was to use the data:

To ascertain the relative mortality in different hospitals, as well as different diseases and injuries at the same and at different ages, the relative frequency of different diseases and injuries among the classes which enter hospitals in different countries, and in different districts of the same countries.

Cited in (Cook, 1913); (page 430).

Nightingale expressed concern that, rather than benefiting patients, being admitted to hospital was potentially harmful:

It may seem a strange principle to enunciate as the very first requirement in a hospital that it should do the sick no harm. It is quite necessary, nevertheless, to lay down such a principle, because the actual mortality in hospitals, especially those of large crowded cities, is very much higher than any calculation founded on the mortality of the same class of diseases among patients treated out of hospitals would lead us to expect (Nightingale, 1863).

With the support of 'friendly doctors on the medical side' (Cook, 1913), Nightingale encouraged hospitals in London to collect and report mortality data:

Guy's printed a statistical analysis of its cases from 1854 to 1861; St Thomas's of its from 1857 to 1860; St Bartholomew's, a table of its cases for 1860. With regard to the future, a meeting was held at Guy's hospital on June 21, 1861, and it was unanimously agreed – by delegates from Guy's, St Bartholomew's, St Thomas's, the London, St George's, King's College, the Middlesex, and St Mary's – that the Metropolitan Hospitals should ... publish statistics annually (Cook, 1913).

From 1862 to 1866 *The Journal of the Statistical Society of London* published data from metropolitan and some provincial hospitals, with the full series covering 1861 to 1865 (Turner, 1862, Statistical Society of London, 1864, Statistical Society of London, 1865, Statistical Society of London, 1866a, Statistical Society of London, 1866b). In the first year of publication, data for fourteen London hospitals were reported (Turner, 1862). The second publication saw the addition of data for twelve provincial hospitals, as well as a table reporting statistics from St George's Hospital covering thirty-six years, from 1827-62 (Statistical Society of London, 1864). This publication proved to be the most comprehensive though, with fewer hospitals making returns with each successive year.

Each year the data were accompanied by a brief commentary, but rather than concentrating on interpreting the information, comments focussed on how co-operative hospitals had been in supporting the exercise. The lack of co-operation from the Middlesex Hospital and University College Hospital particularly were highlighted – the only set of data to appear for the former had not been provided by the hospital at all but by the Board of Trade (Statistical Society of London, 1866b). In a late effort to encourage co-operation it was proposed that the forms be simplified (Statistical Society of London, 1866b). This proposal failed to engender support and publication ceased entirely in 1866. One of Nightingale's biographers laments:

I do not find that the experiment has been continued. So far from there being any uniform hospital statistics of the kind contemplated by Miss Nightingale, even in London some of the hospitals do not keep, or at any rate do not publish, any at all. The laboriousness, and therefore the costliness, of the work of compilation, the difficulty of securing actual, as well as apparent, uniformity, and a consequent doubt as to the value of the conclusions deduced from the figures are presumably among the causes of the defeat of Miss Nightingale's scheme (Cook, 1913); 433).

Nightingale was aware of the technical problems of measurement and comparison, drawing attention to the complexities of risk adjustment, the possibility of data manipulation, the dangers of a narrow focus on a single outcome measure, and the potential perverse behavioural consequences (Nightingale, 1863). These issues remain a cause for concern today (see sections 5.3 and 5.4).

Technical problems are unlikely to have been the sole cause for the discontinuation of published data. Political influence also played a part. Nightingale's scheme provoked opposition from those whose performance was the subject of public scrutiny. Although the scheme was initially supported by the London hospitals, the medical profession, represented by the Royal College of Surgeons, vetoed greater use of the Model Forms following a critical report by a Committee it had established to consider the matter (Bishop and Goldie, 1962).

Opposition to the publication of data was not confined to the medical profession. Nightingale also encountered hostility from Parliament and government ministries to her proposed statistical analyses. When she applied to Parliament to extend the scope of the 1861 Census to include questions relevant to sanitary improvements, the proposal was heavily out-voted in the House of Commons, with the Home Secretary, Sir George Lewis, explaining that "the question of health or sickness was too indeterminate" for inclusion (cited in Cook, 1913; page 436). Soon afterwards, the War Office and India Office – "furious at her interference" – sought to suppress sections of a report

by Nightingale that detailed incriminating differences in mortality rates between troops in England and India (Vicinus and Nergaard, 1989); (page 242).

The medical profession and government began to view the publication of statistical analyses as dangerous to their own health: it was not in their interests to arm the public and media with data with which to ask awkward questions. It would not be until 1992 that mortality rates for English hospitals were again in the public domain. Strangely enough, though, when they were published once more hardly anyone seemed to notice.

2.1.2 Buried under a wealth of information

In 1992, death rates in English hospitals were included in the Health Service Indicators (HSIs), which reported 1990/1 data (NHS Executive, 1992). The death rates were published annually for the next four years, until 1996 when HSIs were discontinued. However, these English hospital death rates received virtually no publicity and many were unaware of their existence – including the Secretary of State for Health at the time.

The hospital mortality data, published as HSI ML51, reported the number of hospital deaths, standardised against the number expected if the pattern of deaths in England as a whole were to apply to the hospital, with actual and expected deaths adjusted for age, gender, and diagnosis (NHS Executive, 1992). The standardised death rates were reported for only two specialties (general surgery and general medicine) and they never achieved full coverage of all hospitals but, on the face of it, the data appear of sufficient interest to justify commentary within the NHS and among the wider community. The data were easily interpretable, contentious and accompanied by a relative ranking of each hospital. So, for example, Pontefract General Infirmary had a standardised death rate in general surgery for 1990/1 of 195.5, implying that the hospital experienced 95.5% more deaths than expected given its casemix, making it the worst ranked hospital in England against this indicator (Laurance, 1993a).

16 The experience demonstrates that placing data in the public domain does not ensure that the data are publicised. There was widespread ignorance about the existence of HSI data reporting English hospital mortality rates, which appear to have been buried among the wealth of information contained in the HSIs (Davidge and Harley, 1992). This is demonstrated by the reaction when the first series was reprinted in a national newspaper:

The Times received calls from hospitals and health authorities asking where the information had originated. Many were unaware that it was sent routinely to health authorities (Laurance, 1993b).

Even the government was unaware that the data were publicly available. Eight months after the English data had been included in the HSIs, the Scottish Office announced its intention to publish mortality rates for Scottish hospitals. The proposal met with opposition from some members of the Conservative cabinet, embarrassed by the 'precocious Scots' (Brown and Hall, 1994, Kendrick et al., 1999). Virginia Bottomley, then Secretary of State for Health, was particularly opposed to the Scots' proposal (Arlidge, 1994; Brown and Hall, 1994). Questioned as to whether England would be following Scotland's example, a source for Bottomley was quoted as saying "we are not going down that road - we have no enthusiasm for counting dead bodies" (Brown and Hall, 1994). Clearly nobody had informed the Secretary of State that the English were already some way along this road and nor was Bottomley's lack of enthusiasm for counting hospital deaths communicated to those responsible for compiling the HSIs, given that they continued to publish them.

The Secretary of State's disapproval may have ensured that the data were not publicised. No mention was made of the data in scientific medical journals, such as the British Medical Journal or The Lancet, or in journals read by health service managers, such as the Health Service Journal. Nor were the data reported in broadsheet newspapers, with the exception of The Times, which published details of 1990/91 hospital specific death rates in the form of a league table (Laurance, 1993a). This proved to be the first and only time that the death rates incorporated in the HSIs received any publicity in either scientific

17 medical journals or national newspapers (although some local newspapers referred to the information (Jones, 1994)). Indeed, the specialist and news media ignored the HSIs in their entirety, as did most people working in the NHS.

The most plausible reason for this is that the Conservative government directed attention elsewhere. Reacting to the publication of the hospital death rates by The Times in 1993, the then Chief Executive of the NHS, Sir Duncan Nichol, stated that the government remained committed to publishing comparative information (Laurance, 1993b). However, it transpired that the government was not committing itself to publicising death rate data (even though these continued to be produced in the HSI package) but, rather, to a new set of performance indicators. These indicators had been announced in the Patient's Charter (NHS Management Executive, 1991) and were first published in 1994 (NHS Executive, 1994). These were the data that subsequently dominated media coverage and "concentrated managers' minds" (Laurance, 1993b). With hospitals striving to meet new government performance targets for waiting times, day case rates and cancelled operations, the data contained in the HSIs, including hospital death rates, were forgotten and, after 1996, the government decided to discontinue producing the HSIs altogether.

2.1.3 Resurrection

In 1997, there was a change of government. The incoming Labour government announced that mortality statistics for English hospitals were to be published as one of the key measures to monitor clinical behaviour more explicitly. The Labour government had set great store on being the party to protect the NHS (The Labour Party, 1997) and the Prime Minister, Tony Blair, made the NHS his personal preoccupation (Timmins, 2000). The publication of mortality data was seen as a solution to problems of clinical performance, with the government believing that publication would indicate to the electorate that it was tackling the medical profession. Weakened by scandals, the medical profession was forced to acquiesce to the publication of such

18 data. In the next section, we describe how, in the past, successive governments had favoured professional self-regulation over external scrutiny of medical practice. We then explore the factors that led to a change in the traditional relationship between the government and the medical profession.

2.2 Self-regulation is not enough

From early in the sixteenth century, the Royal College of Physicians enjoyed a monopoly in licensing doctors within seven miles of London, which included the great teaching hospitals (which were then virtually the only hospitals in the country). Only sons of the wealthiest families could afford to become members of the Royal Colleges. In 1832, the forerunner of the British Medical Association (BMA) was established in reaction to the exclusive nature of the Royal Colleges. Known as the Provincial Medical and Surgical Association, it functioned from Worcester as a protest group, and, for its first 20 years, specifically excluded London doctors from its membership (Forsyth, 1973, pages 6-7).

In the United Kingdom, the medical profession has long argued the case for self-regulation and, traditionally, governments have acquiesced. This stance was defended on the grounds that the professionalism of doctors would ensure quality (Power, 1997), although this position has not been without critics (Smith, 1989; Brazier et al., 1993). The case for self-regulation was easier to make in an era when patients were considered passive recipients of care and when governments were unlikely to gain much political capital from taking on the medical profession. That era has passed, and one of the reasons for its passing is the failure of self-regulation. A series of well-documented medical tragedies have undermined the profession's claims that self-regulation is in the best interests of patients. In the late 1990s, the government reacted to these tragedies in order to demonstrate to the electorate that the NHS was safe in its hands (The Labour Party, 1997). This section describes how self-regulation failed to protect the interest of patients and the key events that led to a shift in the public's perception of the medical profession, which in turn

forced the government to intervene.

19

2.2.1 Medical Audit

“It makes no sense whatsoever to talk about evaluating the performance of the NHS without also insisting on evaluating the performance of those who take the most important decisions about the use and allocation of resources to individual patients: doctors” (Day and Klein, 1983).

Two main instruments have been employed by the medical profession to evaluate performance, these being medical audit and disciplinary procedures. Medical audit was developed by the profession in the 1980s, encouraged by the Conservative government in recognition that disciplinary procedures are directed primarily at those with extreme levels of poor performance. Medical audit developed as a formal method to improve standards of practice for the profession as a whole. As performance indicators were developed in the 1980s, the Department of Health attempted to encourage the medical profession to engage in medical audit activities. However, the Department of Health conceded that the assessment of clinical performance was exclusively the province of the professional bodies (National Audit Office, 1988). Although the Public Accounts Committee of the House of Commons was of the view that clinicians should be accountable to the NHS rather than merely to their professional bodies (Warden, 1988), the government was unwilling to risk a potentially damaging battle with the medical profession over the issue (Smith, 1998b). The compromise was to encourage clinicians to engage in medical audit on a voluntary basis. The model scheme for medical audit in the NHS was the confidential enquiry into perioperative deaths (CEPOD), which first reported in 1987 (Buck et al., 1987). The CEPOD involved independent assessment of deaths within 30 days of a surgical operation, and relied on the participation of consultants in anaesthesia and surgery (Lunn and Devlin, 1987).

It was hoped that clinicians would recognise the value of medical audit and would be persuaded to participate. But many clinicians

20 consistently refused to take part or withdrew during the process. For instance, during the 1987 CEPOD, 79 consultants withdrew from the exercise (Warden, 1988). Although this represented only 5.6% of the consultants taking part, their withdrawal meant that 500 (12.5%) of the deaths could not be investigated (Warden, 1988). This raises the suspicion that the patients of those consultants who dropped out had above average death rates. Moreover, half of the withdrawals were from three of nearly 100 hospitals involved in the study (Editorial, 1987), suggesting that institutional factors may explain differences in participation and performance. Those conducting the enquiry expressed concern about “the negative, and sometimes obstructive, influence of a very few consultants” (Lunn and Devlin, 1987).

The view expressed by the Royal Colleges that peer pressure would win over consultants who resisted medical audit (Warden, 1988) proved unfounded. More than 10 years after the first CEPOD, up to a third of surgeons and anaesthetists in some NHS regions still did not participate in the various confidential enquiries taking place (Warden, 1998b).

2.2.2 Disciplinary procedures

While medical audit is a fairly recent development, historically much of the profession’s regulatory efforts have been directed at identifying and dealing with unacceptable practice. Assessment of the competence and conduct of the medical profession has focussed on disciplinary procedures. If there is evidence of poor performance, it is dealt with as a punishable offence (NHS Executive, 1999b).

With respect to hospital medical staff, disciplinary procedures have been developed in two localities: in the institution in which the individual works and within the profession to which the individual belongs. The employing institution is responsible for handling problems with the personal conduct of doctors, such as harassment and bullying or inappropriate use of NHS facilities.⁴ When problems concern the competence or professional conduct of doctors, these are

⁴<http://www.scottishexecutive.gov.uk/library3/health/ddit-05.asp> (accessed 08/11/04)

21 subject to national disciplinary procedures, set out by the General Medical Council (GMC), whose statutory duty is to protect the public by regulating doctors in the UK (Stacey, 1992). In extreme circumstances, the GMC holds disciplinary proceedings relating to professional misconduct, but in the majority of cases the employing institution will deal with the matter. Assistance is available from the National Clinical Assessment Authority (NCAA), which was set up in 2001 “to support the NHS deal with doctors and dentists whose performance gives cause for concern”⁵.

This framework suffers a number of shortcomings, not least its almost exclusive focus on determining what constitute punishable offences rather than encouraging all doctors to seek continual performance improvements (Department of Health, 2000b). However, execution of even this limited remit has proved inadequate.

On the face of it, the employing institution is best placed to monitor the performance of its employees. It ought to have accurate, up-to-date information about individual competence and the circumstances in which people work. Often, those working in the health service know which individuals are working at sub-standard levels of performance:

“Unlike the patients they treat, doctors, nurses and others have always been aware of colleagues who perform below standard and of departments that persist in using outdated techniques or perform too few procedures for optimum results” (Dickson, 1995).

There are plenty of anecdotal examples to support this assessment. For instance, long before he was sacked in July 2000, the consultant general surgeon at Pinderfields hospital, Christopher Ingoldby, was nicknamed ‘Dr Death’ by colleagues at the hospital because of the number of his patients who required intensive care after he had operated on them (personal communication). But he was not suspended until after an inquest jury investigated a case brought by the family of a patient who died five hours after Ingoldby had operated on him (Dyer, 2000b). The tragedy that befell this family was not an isolated incident. Ingoldby’s clinical performance led to 84 negligence

⁵<http://www.ncaa.nhs.uk/> (accessed 08/11/04)

22 claims being lodged with the hospital, 13 on behalf of patients who died, and covering an eight year period from 1989 to 1997 (Dyer, 2000b). If his colleagues felt that Ingoldby's competence was sub-standard, why did it take so long for the hospital to suspend him and why was it left to those who had suffered to bring the case?

One explanation for the failure to act is that institutions have an obligation to both their employees and to patients, and loyalty to colleagues may well cloud objectivity. In some cases, rather than upholding the interests of patients, action may be directed instead at protecting poorly performing staff. This was the accusation made against managers following the government inquiry into the serious medical malpractice of consultant gynaecologist Rodney Ledwood, who was struck off the medical register in 1998:

“Powerful constraints against ‘telling tales’ and a culture of treating consultants as gods also contributed to the conspiracy of silence surrounding his [Ledwood's] poor surgical performance. ... The [government inquiry] report details how as early as 1986 senior management was aware of Mr Ledwood's high complication rate” (Roach, 2000).

But internal mechanisms may fail to protect staff as well. In an investigation by the Society of Clinical Psychiatrists into 38 suspensions that were the result of ‘whistleblowing’, fault was proved in only two cases (although some remained ongoing). The Society suggested that motives for whistleblowing include:

“Intellectual arrogance, particularly in inbred departments. Some colleagues regard a different style of clinical practice as incompetence. Other motives include minority intolerance, professional jealousy, and power struggles over control of a hospital department. It is easy to dress up allegations so that an administrator suspends the doctor first and investigates afterwards” (Tomlin, 1998).

Rather than relying on institutional action, a more dispassionate perspective might be expected from the GMC. However, the GMC too has failed to act on information about poor performance. This

23 failure is shown by the case of Richard Neale, a gynaecologist who was allowed to work for 14 years in Britain despite the GMC being aware that he had lost his licence to practise in Canada (Dyer, 2000a). He was struck off the medical register in Britain in July 2000, after being found guilty of “botched surgery, operating without consent, carrying out unnecessary procedures, failing to monitor the condition of patients postoperatively, and duping a patient into paying for a private operation” (Dyer, 2000a).

Examples of medical negligence claims brought against individual doctors such as those mentioned above are numerous. Often there has been an inclination to dismiss medical scandals as isolated incidents, deserving condemnation and, perhaps, a little soul-searching by the medical profession (Ransley, 2000). Rarely have they resulted in a whole-scale re-evaluation of the system by which medical competence is monitored and regulated. The tragedy that befell children and their families at the Bristol Royal Infirmary proved an exception in this regard and precipitated a fundamental shift in the relationships among the public, the government and the profession (Buck et al., 1987; Editorial, 1987; Lunn and Devlin, 1987; National Audit Office, 1988; Warden, 1988; Smith, 1998b; Warden, 1998b).

2.2.3 Bristol: a watershed

In June 1998, a GMC inquiry came to the conclusion that there had been serious professional misconduct by three doctors at the Bristol Royal Infirmary (Dyer, 1998). The verdict resulted in two doctors being struck off: James Wisheart, a former senior surgeon and medical director, and John Roylance, the chief executive. A second surgeon, Janardan Dhasmana, was subject to a three-year ban on performing paediatric heart operations.

In addition to the vitriol hurled at the doctors at the centre of the case (Dyer, 1998), events in Bristol placed the GMC under heavy criticism and eroded public trust in the medical profession (Dunn, 1998; Smith, 1998a; Davies and Shields, 1999). Parents were bitter about how the remit of the GMC inquiry was interpreted and whose

24 interests the GMC represented (Klein, 1998). The inquiry concentrated on only 29 patients, thereby denying many other parents and victims the opportunity to have their grievances heard. Parents believed that the inquiry was set up in the interests of the doctors not the victims. One parent, whose son “suffered severe brain damage which left him screaming in agony for over a year” (Stewart, 1998), saw his charge against Wisheart dropped by the GMC. He complained:

“The only people allowed rights and representation at the GMC were the doctors charged and the GMC itself. My son was accorded no rights, nor was he allowed representation. The High Court in London confirmed this when we took the GMC to judicial review before the start of its inquiry” (Stewart, 1998, page 1581).

This was not entirely the fault of the GMC, which is constrained by the legislation under which it operates and had to consider specific charges against specific doctors (Klein, 1998). A government inquiry would have enjoyed a wider remit and could have investigated more general circumstances that led to the Bristol tragedy, but the government originally decided against convening such an investigation (Klein, 1998), although this decision was subsequently reversed.

The GMC also began to lose support among those it was supposed to represent. Unsurprisingly, doctors at Bristol who were part of the medical hierarchy sought to defend their colleagues, and accused the GMC of using the case to “show to politicians and the public that it could be safely trusted to regulate and discipline the medical profession” (Dunn, 1998, page 1144). But, if this was the GMC’s objective, it failed. Even though it had taken over 150 years to produce and had been introduced only in 1997, the GMC’s system for responding to poorly performing doctors was shown to be inadequate by the Bristol case (Smith, 1998b, Irvine, 1997b). Under the Medical (Professional Performance) Act of 1995, the GMC received powers to investigate a doctor’s performance and, where performance was found to be seriously deficient, to impose conditions on, or suspend, a doctor’s registration (Irvine, 1997a). There was no statutory provision

25 for the GMC to investigate the quality of care in a hospital department, or to inquire into the cause of deaths (Klein, 1998). In the aftermath of the inquiry, the GMC responded to criticism by proposing that doctors be revalidated (Beecham, 1998; 1999). Coupled with other complaints, the suggestion that they face compulsory revalidation led to hospital doctors passing an overwhelming vote of no confidence in the GMC in June 2000 (Beecham, 2000).

2.2.4 A political shift

As well as undermining public confidence in self-regulatory processes, the Bristol tragedy also demonstrated that the transference of the public’s trust from medical professionalism to medical audit processes was misplaced, as audit itself is not a value-free exercise (Power, 1997). This was illustrated starkly when a mother of a girl operated on by Wisheart complained to the Bristol inquiry about the audit system that considered a successful outcome of surgery to be survival after one month:

“My daughter Sophie is still classed as a success even though she cannot walk, see, talk, move, has epilepsy, and can’t do anything for herself. But because she lived for more than 30 days after the operation she is still counted as a success, and I regard that as a travesty” (Dobson, 1999).

The incoming Labour government was keen to demonstrate that it could do something about poorly performing doctors (Smith, 1998b). With the medical profession on the defensive and divided, the Labour government was able to use the Bristol case to mobilise popular support behind its own agenda. The government appeared disinclined to dismiss medical scandals as isolated incidents, believing unsatisfactory performance to be widespread. Various pieces of evidence supported this position. For instance, the report of an expert group chaired by the Chief Medical Officer estimated that 850,000 adverse events are experienced annually in NHS hospitals – 10% of all admissions (Department of Health, 2000b). This implies that little improvement had been made since the publication of the first

26 CEPOD report in 1987, which indicated that at least 1,000 patients a year die as a result of errors by surgeons and anaesthetists (Buck et al., 1987). The errors are costly both to patient health and the public purse - in 1997/8 the NHS paid out £79 million in medical negligence claims (Fenn et al., 2000). These accumulated pieces of information suggest that poor performance was not isolated to a few miscreants, but was a fundamental problem throughout the NHS. The importance of the Bristol case was that it galvanised public opinion, which the government was able to exploit to increase its power over the medical profession.

One response was to start publishing and publicising data and in 1998, Frank Dobson, then the Secretary of State for Health, announced to a meeting in the House of Commons:

“The appalling tragedy in Bristol cannot be allowed to happen ever again. These [mortality] tables will enable us to identify potential problems at an early stage and to act if necessary” (Warden, 1998a).

Subsequently, three sets of mortality statistics were published (NHS Executive, 1999a):

- Deaths in hospital within 30 days of surgery, by method of admission (emergency / elective);
- Deaths in hospital within 30 days of emergency admission with a hip fracture for patients aged 65 and over;
- Deaths in hospital within 30 days of emergency admission with a heart attack for patients aged 50 and over.

These statistics were based on routinely available data and there was no suggestion that they were to be interpreted as capturing all aspects of clinical performance. The mortality statistics were presented, not in isolation, but couched in a set of indicators.

We now turn to a description of the broader context within which publication of outcomes data in Scotland and England took place.

2.3 Publication of outcomes data in Scotland

27

The clinical resource and audit group (CRAG) was set up in 1989 to assess economic and quality issues for the NHS in Scotland.⁶ The Patient’s Charter for Scotland, published in September 1991, specified waiting targets for patients (Scottish Office, 1991). In 1992, the Clinical Outcomes Working Group, a subcommittee of CRAG, was established to produce annual reports on performance. The 1994 report included 17 clinical outcome indicators covering patient populations in the Health Boards of Scotland, acute hospitals and psychiatric hospitals (Scottish Office, 1995, Mannion and Goddard, 2003). Based on data spanning three years (mid 1990 – mid 1993), clinical outcome indicators at hospital Trust level were published (Box 2). Subsequently, these indicators were refined and supplemented and some were dropped. Additional indicators of hospital performance include CABG procedures (1998), mortality rates within 30 days of elective surgery (2002) and 7-day emergency readmission statistics (2002).

Box 2: **Scottish hospital clinical indicators (1994)** **Seven acute hospital indicators:**

Mortality after hip fracture
 Discharge home after hip fracture
 Mortality after admission for AMI
 Emergency readmission with 28 days of discharge
 Re-operation after transurethral prostatectomy
 Survival after admission for stroke
 Discharge home after stroke

Three psychiatric hospital indicators:

All deaths within one year of discharge
 Deaths at ages under 65 within one year of discharge
 Suicides within one year of discharge

The intention of the publication was, and remains, to highlight issues that might need further investigation. A ‘health warning’ accompanies

⁶<http://www.show.scot.nhs.uk/crag/topics/facts/jistory.htm#croc> (accessed 08/11/04)

28 the reports, emphasising that the indicators are not a league table of performance and that no direct inferences about quality of care should be drawn from them, but that the indicators should “provide useful clues and limited evidence relating to quality of care or performance” (Clinical Outcomes Working Group, 2002). The decision to publish data in Scotland involved balancing the possibility that the data would be misinterpreted against that of promoting information flows within the health community and to the wider public (Kendrick et al., 1999). Whatever the relative strength of these arguments, a further risk remained:

“The decisive factor was the acknowledgement that any system circulating identifiable but not fully public outcomes information would inevitably lead to partial leaks and scare stories. Publication would allow proper contextualisation of the information and explanation of the limitations of the indicators” (Kendrick et al., 1999, page 157).

The Scottish experience allayed political fears about how publication would be received. Those responsible for implementing the policy reported that, among clinicians and managers, there was “virtually no simplistic and unquestioning use of the indicators to inform decisions without further enquiry” (Kendrick et al., 1999, page 158). Media coverage was “informed and responsible” (Kendrick et al., 1999, page 158) and there was “very little evidence that publication of the indicators produced public unease or distress” (Kendrick et al., 1999, page 159).

However, an independent evaluation of the system painted a less positive picture. The study was designed to find out how indicators were used and to determine their perceived impact (Mannion and Goddard, 2003). Interviews were held with stakeholders drawn from eight NHS hospital Trusts and their corresponding Health Boards, and with a random sample of general practitioners (GPs). A postal survey of all 16 Health Councils, bodies that represent consumer interests in Scotland, was undertaken to assess consumer use of the data.

The study focussed on two indicators: 5-year survival from breast cancer, reported at the level of the Health Board; and 30-day survival

29 after emergency admission for stroke, reported at the level of the hospital Trust. Low levels of awareness of and interest in the data were found. Hospital Trust staff were generally aware of the data, which they used principally to support funding applications and service development (Mannion and Goddard, 2001). Health Boards rarely used the data, except to identify potential problems requiring further scrutiny. Although one quarter of the GPs interviewed had accessed the data, they seldom referenced it during consultations with patients. However, GPs did use waiting time and audit data to inform their own assessments of local hospital Trusts. Only one of the Health Councils had any experience of consumer interest in the data.

The key lessons distilled from this study were that the perceived quality of the indicators was a barrier to their use and that these needed to be derived from robust datasets. In addition, the context – particularly the incentive structure – within which indicators operated was important to encourage the use of the data for continuous quality improvement (Mannion and Goddard, 2001).

2.4 Publication of outcomes data in England

The Scottish CRAG experience has informed the development of a broader performance assessment framework for the NHS, one plank of which involves the publication of information. One of the Labour government’s principles in modernising the NHS is to restore public confidence, making the NHS “accountable to patients, open to the public and shaped by their views” (NHS Executive, 1997). In England this has led to the development of the National Framework for assessing performance, described as one of the world’s first systematic and comprehensive public reporting systems (Marshall et al., 2000a). Covering six areas, the framework aims to give “a clear signal of what matters in the new NHS” and is designed to tackle unacceptable variations in standards (NHS Executive, 1998b). First published in 1999, the outcomes data are accessible to the public and patients, health care professionals and NHS managers (Marshall and Davies, 2001).

30 The NHS Plan proposed that hospital organisations would be assessed against a “traffic lights” system and, in Spring 2001, hospitals’ cleanliness was rated using the system (Department of Health, 2000a; Allison, 2001). However, when the ratings for acute NHS Trusts appeared in September of the same year, ‘traffic lights’ had been replaced by ‘star ratings’ (Department of Health, 2001a) perhaps because the government did not want the public to associate the stop/go traffic signals as an indication of whether or not NHS organisations were moving forward.

Star ratings were intended to be “not primarily a commentary on the quality of clinical care”, but rather to assess the “overall patient experience” (Department of Health, 2001a). Along a four-point scale, the ratings reflected performance against key targets with the additional criterion that the Trust did not receive a critical clinical governance report from the Commission for Health Improvement (CHI). The targets comprise various waiting time indicators, hospital cleanliness, cancelled operations, the financial state of the Trust and a demonstrated commitment to improve working conditions. Three sets of subsidiary indicators were also assessed, comprising those with a clinical focus (such as emergency re-admission rates); those with a *patient focus* (such as the resolution of written complaints); and those with a *staff focus* (such as junior doctors’ hours).

Trusts with a balanced scorecard – good or high performance in all four areas – were then awarded two or three stars respectively (Kaplan and Norton, 1992). Trusts rated as having three stars benefited from “earned autonomy” that conferred a number of freedoms and rewards, including the freedom to spend their share of the NHS Performance Fund on developing services and rewarding staff.

After 2001, NHS star ratings evolved to include additional – and more sophisticated – indicators. As well as extending the assessment process to include other types of Trust, the rating system placed an increasingly important role on CHI’s clinical governance review (CGR) (Department of Health, 2001a). In 2001, the assessment of star ratings required only that the organisation had not received a critical review; at that time, just 16 acute Trusts and two mental health Trusts had

31 undergone a CHI review. By contrast, when the 2003 ratings appeared, CGRs had been published for 160 acute Trusts and 28 mental health Trusts. The 2003 star ratings for acute, specialist and mental health NHS Trusts were adjusted in accordance with the ‘Finsbury rules’.⁸ In essence, these involve zero-rating any organisation evaluated as having achieved only the lowest standard of performance (level ‘I’) in five or more out of the seven areas of clinical governance assessed, apparently irrespective of the organisation’s performance on key targets or the scorecard. Three stars were awarded only to organisations that had achieved key targets, a balanced scorecard, at least three ‘III’s (the highest standard) and no ‘I’s in the CGR. For Trusts with an older CGR, progress against an action plan informed the ratings.

An overview of outcomes data currently published about NHS organisations in England is given in Table 2. Regulation that is more helpful to patients and users and less burdensome to providers is the goal of the Commission for Healthcare Audit and Inspection (CHAI), or ‘Healthcare Commission’, which took over responsibility for producing ratings from CHI in April 2004. As part of this process, star ratings are to be the primary public measure, to be sensitive and sensible and to reflect the context for service provision, such as local demography. The focus of the assessment has shifted to the quality of care provided to patients and to the capacity of the organisations to deliver services of high quality.

Major revisions to the assessment system are planned for 2005/06, including a set of 24 ‘core standards’, supported by 13 ‘developmental standards’ (Department of Health, 2004b). Organisations’ performance will be judged not just by national targets, but also by whether high quality care is delivered across a range of areas, including National Service Frameworks and NICE guidance. In addition, organisations will be able to set their own targets to tackle local priorities. Hospital Trusts may no longer receive a single summary rating, in recognition of its inability to capture the complexity of

⁷ health improvement, fair access, effective delivery of appropriate health care, efficiency, patient / carer experience, health outcomes

⁸ http://www.chi.nhs.uk/eng/ratings/finsbury_rules.shtml (accessed 08/11/04)

32 health services (Kmietowicz, 2003), although the Secretary of State for Health has indicated that the Healthcare Commission will continue to publish a simple public set of ratings (Reid, 2004).⁹

An overview of outcomes data currently published about NHS organisations in England is given in Table 2. Regulation that is more helpful to patients and users and less burdensome to providers¹⁰ is the goal of the Commission for Healthcare Audit and Inspection (CHAI), or 'Healthcare Commission', which took over responsibility for producing ratings from CHI in April 2004. As part of this process, star ratings are to be the primary public measure, to be sensitive and sensible and to reflect the context for service provision, such as local demography. The focus of the assessment has shifted to the quality of care provided to patients and to the capacity of the organisations to deliver services of high quality.

Major revisions to the assessment system are planned for 2005/06, including a set of 24 'core standards', supported by 13 'developmental standards' (Department of Health, 2004b). Organisations' performance will be judged not just by national targets, but also by whether high quality care is delivered across a range of areas, including National Service Frameworks and NICE guidance. In addition, organisations will be able to set their own targets to tackle local priorities. Hospital Trusts may no longer receive a single summary rating, in recognition of its inability to capture the complexity of health services (Kmietowicz, 2003), although the Secretary of State for Health has indicated that the Healthcare Commission will continue to publish a simple public set of ratings (Reid, 2004).

In addition to these activities, data on performance have been placed in the public domain through other means, usually with government support. First appearing in 2001, the 'Good Hospital Guide' summarises the performance of all public and private hospitals in the UK against a range of indicators, including mortality rates

⁹ <http://www.chi.nhs.uk/ratings/> (accessed 08/11/04)

¹⁰ <http://www.healthcarecommission.org.uk/assetRoot/04/00/00/20/04000020.pdf> (accessed 08/11/04)

(Anonymous, 2001a; Anonymous, 2001b). The Guides are published by Dr Foster, an independent organisation which collects and analyses information on the availability and quality of health services. A range of different guides is available, covering a number of disease areas as well as one that specifically reports on hospital consultants. In 2004, 'Dr Foster's case notes' began to appear monthly in the British Medical Journal (Jarman et al., 2004), while a similar series appears in the Health Service Journal targeted at NHS managers. In collaboration with the NHS, Dr Foster also produces two magazines that are distributed within England. In December 2003, a local maternity magazine for all mothers-to-be was launched. Health Minister Stephen Ladyman commented:

"In our recent Choice Consultation Survey, almost 90 per cent of respondents told us that they needed more information to make decisions about their care and treatment. 'You're Pregnant' is a step forward in providing all mums-to-be with local information about maternity services in their area, so that they and their families can make real informed choices about the NHS services they would like to use. This magazine offers mums-to-be vital information, helping mother and baby to be healthy, before, during and after the birth" (Department of Health, 2003a).

One month later, the collaborators launched 'Your Life!', a "glossy" and "fun" women's magazine, dealing with topics such as eczema, breast cancer and parenthood (Department of Health, 2004a). The magazines adopt a similar format, addressing national issues of general interest and including celebrity interviews and interviews with 'real' people. Information on local services is also presented to allow readers to compare NHS services in their area.

Table 2 NHS hospital outcomes data: forms of publication and target audience, England 2004

Form of publisher Publisher	Format	Type of outcomes data	Location available	Target audience
Healthcare Commission ¹¹	Electronic: http://www.healthcarecommission.org.uk/Homepage/fs/en	Performance ratings Clinical governance reports	Internet Media coverage	General public Any individual or organisation with access to the internet ¹²
Healthcare Commission	Published books and reports	Performance ratings Clinical governance reports	Bookshops Public libraries (some) Media coverage	General public
Your Guide to Local Health Services (the patient prospectus)	Printed leaflets Some guides available on website	Star ratings of local NHS organisations (summary statistics only)	Delivered door-to-door Internet (partial) Public libraries (some) GP surgeries, county, town and parish councils.	General public
Dr Foster Good Hospital Guide and the Good Consultant Guide	Electronic: http://www.drfooster.co.uk/ Published books and reports Selective reprinting in the Sunday Times mail on Sunday	NHS and private hospitals: Facilities and services Transport details Clinic waiting times Consultant detail ¹³ Overall hospital mortality rate ¹³	Internet Bookshops Public libraries Newsagents	Any individual or organisation with access to the internet General public
Strategic Health Authority annual reports	Electronic and printed reports	Star ratings of local NHS organisations	Internet SHA office	General public Distributed to NHS and social services organisations, Community Health Councils, local MPs
SCTS's National Adult Cardiac Surgical Database reports	Electronic and printed reports	Mortality rates for cardiac surgery, reported by clinical and demographic characteristics	Internet (2000/01 report only) Dendrite Clinical Systems Ltd	General public Patients

¹¹ Before April 2004, ratings were produced by the Commission for Health Improvement (CHI). CHI took over performance rating from the Department of Health in 2002 and published its first ratings in 2003.

¹² Therefore excludes those without physical access/ the necessary technical skills to enable access.

¹³ Performance data are not available on the website without subscription.

Dr Foster publishes selected surgeon-level data, but mostly these are not freely available. Although the Labour government had planned to publish data on individual surgeon performance from April 2004 (Department of Health, 2002a), implementation was delayed because the requisite technical infrastructure was not in place in time (Carlisle, 2004; Day, 2004). Whilst the government's plans remain unclear, it seems that data would include only first time bypass operations, which constitute about two-thirds of a cardiac surgeon's workload (Dyer, 2003), with risk-adjusted analysis restricted to lower-risk cases (Keogh et al., 2004). The model for the publication comes from the Society of Cardiothoracic Surgeons (SCTS), which first established a surgical register in 1977 (Treasure, 1998; Keogh and Kinsman, 2002). The National Adult Cardiac Surgical Database allows continuous evaluation of performance which is fed back to hospital units (Keogh and Kinsman, 2002). The data are published at a high level of aggregation in SCTS annual reports; access to more detailed data is restricted, although some hospital units publish performance data on their websites. In line with the National Service Framework for Coronary Heart Disease, responsibility for data collection is transferring from the SCTS to the NHS Information Authority, although the efforts remain collaborative with the Healthcare Commission providing support. Data will be held in the Central Cardiac Audit Database, enabling mortality tracking through the Office for National Statistics (Keogh et al., 2004).

The latest SCTS report, published in September 2004, used crude mortality data to demonstrate that all surgeons met acceptable standards (Society of Cardiothoracic Surgeons, 2004), but the Department of Health is pressing the Society to make individual, risk-adjusted, surgeon-level data available (Hawkes, 2004, Carlisle, 2004). Whether or not the Department is successful in realising its plans, the Freedom of Information Act (which came into effect on 1st January 2005) may mandate the publication of named data, since these are now held by a public organisation.

An important advance over simply publishing information has also taken place. It is of limited value to be told that there are variations in

performance across the country, particularly if people cannot act on the information. But it may reassure the public that publication is accompanied by a plan of action. Recent publications by Dr Foster not only show variations in mortality rates, but also indicate the point at which rates trigger an alert (Taylor, 2004). Interestingly, alerts are not confined to those hospitals that have worse than average mortality rates, as "most trusts in England experience periods of high levels of risk in terms of increased death rates from time to time" (Taylor, 2004). This initiative combines the publication of both hospital outcomes with reassurance that there is a process in place to identify and respond to poor results.

Evaluations of the publication of English hospital outcomes data are scant; in particular, there is little evidence to substantiate government claims that performance is improving (Snelling, 2003). One study of adult critical care found no association between the star rating awarded to a hospital Trust and the quality of care provided (Rowan et al., 2004). A statistical analysis of the relationship between acute Trust characteristics and their star ratings for 2001 and 2002 supported the absence of a link between rating and quality of clinical care. However, other relationships were observed:

"Zero star Trusts do worse than other Trusts across various patient satisfaction measures and financial and efficiency measures. Three star Trusts outperform others on two grounds fairly consistently: waiting times and financial balance, suggesting either more efficient management or fewer capacity constraints. The labour market for consultants and nurses also appears to be utilised in different ways across the groups of Trusts" (Jacobs and Smith, 2004).

Another assessment of star ratings for 2001 and 2002 showed that poor performance was associated with a large 'private sector premium' (i.e. the gap between wages in the private and public sectors) in the NHS Trust's local labour market. If nurses can earn higher wages in the private sector, then NHS Trusts may have difficulty recruiting sufficient nursing staff and this may adversely affect Trust performance (Burgess et al., 2003).

The government maintains that targets have improved access to health care – a top political priority – as well as delivering reductions in mortality rates for both cancer and coronary disease (Reid, 2004). One of the ways in which star ratings may have improved performance is by the close review and reassessment of zero star-rated Trusts. NHS Chief Executive Nigel Crisp has explained that if zero-rated Trusts demonstrate “significant improvement” during the review period, then “current management arrangements will remain.” However, if only “early improvement” is demonstrated, then “franchising the Chief Executive posts offers the best opportunity to secure the most effective managers to lead these Trusts” (Department of Health, 2003b). In simple terms, this means that unless the Trust’s chief executive achieves a substantial improvement in the Trust’s performance, he or she may lose their job. Although Mr Crisp claims that this “is not about punishment, but about strengthening these organisations”, the incentive for Chief Executives to ensure that their Trust meets government targets is clear.

While in recent times pressure to publish hospital outcome data in the UK has come mainly from the government, in the US demands for such information have come from a diversity of sources. The development of managed care, replacing indemnity insurance and fee-for-service reimbursement, provided the backdrop for the publication of hospital outcomes data in the US. With primary concerns focussed on escalating health care costs, the need for information on relative costs drove much of the data collection process. However, information on the quality of care was also collected to inform value based purchasing (Berwick and Wald, 1990; Marshall et al., 2003). Growing recognition and documentation of the ‘quality gaps’ within the US health care system echo UK health policy concerns. The US National Committee for Quality Assurance estimated the consequences in terms of annual avoidable deaths:

“It is clear that 57,000 lives are an unacceptable price to pay for the system’s failures. Moreover, it is a price we need not pay. The proof of this is in the performance of certain health plans and medical practices – in certain parts of the country more than others – where people do, by and large, get the care they should get. In these organizations and regions, care is delivered consistently at a very high level, and according to well-specified guidelines based on medical evidence” (National Committee for Quality Assurance, 2003).

To close the quality gaps, the NCQA argued, would necessitate major reform of the health care system:

Dramatic and system-wide changes are needed for all Americans to obtain the health care we expect and deserve. The keys to closing the quality gap are:

- Investment in technology and systems to support evidence-based care
- Increased collaboration between health care professionals
- Payment systems that reward excellence

- More consumer engagement in provider selection and care decisions
- Greater transparency, to allow consumers and others to see and compare quality information”

(National Committee for Quality Assurance, 2003).

Table 3 An overview of key US performance reporting systems, 1987 to 2004

Name of system	Coverage	Participation	Date of publication	Outcome measure	Data source	Unit(s) of analysis	Comments
Health Care Financing Administration data release	National ¹⁴	Mandatory	1987-1993	Risk-adjusted 30-day mortality rates. Overall hospital rate and 16 diagnostic -related rates (Fottler et al., 1988)	Routinely collected Medicare inpatient data	Hospital	<p>First US system to be published (1987) following changes made in 1985 to the regulatory guidelines governing the Freedom of Information Act (1966) (Fottler et al., 1987)</p> <p>Data were originally collected as part of a quality improvement programme (Fottler et al., 1988)</p> <p>A logistic regression model was used to estimate rates. Data included hospital and patient characteristics, including age, sex, source of admission and co-morbidity</p> <p>Use of routinely collected data allowed only inadequate risk adjustment (Fink et al., 1989; Berwick and Wald, 1990). Mistaken designation occurred in about half of all hospitals with ‘high-mortality outlier’ status, due to inadequate case-mix adjustment (Green et al., 1991)</p> <p>Withdrawn in 1993 following criticism of data quality (Marshall et al., 2003)</p>
New York State Cardiac Surgery Reporting System (CRSR)	State-wide (New York)	Mandatory	1990	<p>Risk-adjusted mortality rates for CAGB surgery (from 1990)</p> <p>Risk-adjusted mortality rates for angioplasty (from 1996) (New York State Department of Health, 1997)</p>	Routine and CRSR-specific data from medical records (Green and Winfield, 1995)	Hospital Surgeon	<p>In 1991, a local newspaper sued the NY Department of Health under the Freedom of Information Act to force publication of surgeon-level data (Dziuban et al., 1994)</p> <p>Data include demographics, preoperative risk factors, intra-operative data, post-operative in-hospital morbidity and mortality, synthesised using a multivariate logistic regression model (Green and Winfield, 1995)</p>

Name of system Coverage Participation	Date of publication	Outcome measure		Data source Unit(s) of analysis	Comments
					<p>Risk-adjusted mortality rates declined in NY to below the national average (Burack et al., 1999)</p> <p>Poor performers improved quality through tailored processes and structural changes (Dziuban et al., 1994; Chassin, 2002)</p> <p>Reduction in mortality rates may be due in part to 'out migration' of high-risk patients to Ohio (Burack et al., 1999) and to exodus of poorer performing surgeons (Chassin, 2002)</p> <p>The validity of surgeon-specific measures is unclear (Green and Winfeld, 1995) and the predictive power of the model has been shown to be flawed (Burack et al., 1999).</p> <p>Data include primary diagnosis, up to 15 secondary diagnoses and severity of illness, with rates derived using a logistic regression model (Localio et al., 1997)</p> <p>Some self-reported evidence of providers improving quality in response to data publication (Bentley and Nash, 1998)</p>
<p>Consumer Guide to CABG surgery</p> <p>State -wide (Pennsylvania) Mandatory</p>	<p>1992</p>	<p>Risk-adjusted inpatient mortality rate for CABG</p> <p>Charges, length of stay and re-admission rates for CABG (Pennsylvania Health Care Cost Containment Council, 2004)</p>		<p>Routine administrative data from medical records</p> <p>Hospital surgeon</p>	<p>A survey in 1995 found that cardiologists and cardiac surgeons were aware of the report cards but made little use of them (Schneider and Epstein, 1996). However, 63% of surgeons said they were less willing to operate on severely ill patients. A patient survey in 1996 revealed that 12% of CABG patients were aware of the report card before their surgery and less than 1% knew the rating of their surgeon or hospital (Schneider and Epstein, 1998)</p> <p>Annual volumes of CABG surgery are too small for meaningful mortality comparisons and this problem is more severe for surgeon level data: some 'high-mortality outliers' had been misclassified as a result of the inappropriate use of conventional statistical tests and methodology (Localio et al., 1997)</p> <p>Hospital and patient level analyses suggest that providers selected healthier patients for CABG surgery, to the detriment of sicker patients' health (Dranove et al., 2003)</p>

Name of system					
Participation	Date of publication	Outcome measure		Data source Unit(s) of analysis	Comments
Health Plan Employer Data Information Set (HEDIS) National Voluntary	1992	Unadjusted data, covering large number of process measures, including preventative care (Bost, 2001), prescribing indicators (McCormick et al., 2002) and patient access (Epstein, 1998)		Medical record reviews, administrative databases and patient surveys (Marshall et al., 2003), but source varies by health plan Health plan (Epstein, 1998)	From 1997, summarised in the 'Quality Compass' (Marshall et al., 2000a) The National Committee for Quality Assurance (NCQA), an independent, not-for-profit organisation, collects data on process measures Although the Quality Compass is marketed as the "most comprehensive database of health plan performance data", the voluntary basis of the HEDIS system permits lower-scoring plans to stop participating (McCormick et al., 2002)
Cleveland Health Quality Choice (CHQC) Regional (Cleveland) Voluntary	1992-1999	Risk-adjusted mortality for medical and surgical patients Risk-adjusted length of stay for medical and surgical patients Risk-adjusted caesarean section rates Patient satisfaction		Routine administrative data from medical records Hospital	Initiated in 1989 as a coalition of businesses, hospitals and physicians, the CHQC programme published reports twice a year (Baker et al., 2002). Data were disseminated to participating hospitals (1992) and purchasers (1993) with report cards for consumers (1993) (Rosenthal et al., 1998) Data include demographics, co-morbidity, test results and admission details, adjusted using the CHOICE model (Rosenthal et al., 1998) Using supplementary data, one study found that deaths shifted from in-hospital to post-discharge period, with no net reduction in 30-day mortality for most conditions (Baker et al., 2002)
					No hospital lost contracts because of their poor performance (Baker et al., 2003) and there was a dramatic and significant improvement in the poor performance of one 'outlier' (Baker et al., 2003) Claiming that purchasers were not using the information, several hospitals withdrew from the programme in 1999, leading to its demise (Baker et al., 2003)
California Hospital Outcomes Project (CHOP) State-Wide(california) Voluntary	1993	Risk-adjusted mortality rates for: AMI (Rainwater et al., 1998), cancers, chronic liver disease, chronic renal disease (Green and Winfeld, 1993), CABG and organ		Routine administrative data (Medicare) from medical records Hospital	Data include demographics, principal diagnosis, co-morbidity, type of admission and transfer source. Data are risk-adjusted using two different models (Luce et al., 1996) Serious underreporting of co-morbidity and miscoding of 'urgency of admission' led to erroneous estimates of risk-adjusted death rates (Green and Winfeld, 1993, Luce et al., 1996)

Name of system Coverage Participation	Date of publication	Outcome measure		Data source Unit(s) of analysis	Comments
		transplantation (Broder et al., 2004)			Content with 'adequate' performance, most public hospitals made no effort to improve quality in response to CHOP reports From 2005, reporting of risk-adjusted death rates for CABG will be mandatory and will include surgeon-level data (Broder et al., 2004)
United HealthCare Report card National Voluntary	1993	Report cards cover quality of care, cost, efficiency and customer satisfaction (Chase and Heinen, 1994) Quality of care includes adjusted data on immunisations, screening and survival rates for liver transplantation		Routine administrative data from medical records Health Plan	In 1992, United HealthCare became the first company to produce a report card on health care access, quality and cost measures Findings are targeted at purchasers, policy makers and providers Quality improvement is a central aim, but United Health Care has resisted pressures to produce lists of poor performers, arguing that variable reporting methods limit the options for comparative analysis (Chase and Heinen, 1994) In 1998, the United HealthCare Corporation was renamed the UnitedHealth Group.
Michigan Hospital Profiles Project State-wide (Michigan) Voluntary	1996	Risk and severity adjusted measures (chiefly length of stay and mortality rates) of medical, surgical and obstetric procedures		Michigan Inpatient Data Base Hospital	Michigan Health and Hospital Association produces the reports, which were available on the web, but access is now restricted Data are summarised on a 3-point scale Michigan has also been involved in piloting the ORYX initiative
Coronary Artery Bypass Graft Surgery in New Jersey State-wide (New Jersey) Mandatory	1998	Risk-adjusted mortality rates for CABG		Extensive data collection from medical records Hospital Surgeon	Risk-adjustment is based on the New York system and was developed in conjunction with a panel of experts. Key risk factors include the patient's age and sex; whether the patient has diabetes or kidney failure; and whether the patient has had previous heart surgery, among other factors. A random sample of data is audited Initiated by the Department of Health and Senior Services "to change medical practice patterns in the state and save lives" 18 Reports include surgeons performing over 100 CABG operations over a two-year period. Treatment

Name of system Coverage Participation	Date of publication	Outcome measure		Data source Unit(s) of analysis	Comments
Rhode Island Program: Patient Satisfaction Reports Hospital Clinical Performance Reports State-wide (Rhode Island) Mandatory	2001	Unadjusted process measures of patient satisfaction Hospital compliance with clinical guidelines for AMI, heart failure and pneumonia (as part of the ORYX initiative)		Patient satisfaction data taken from postal survey of random sample of post-discharge patients Clinical performance data taken from routine administrative data Hospital	of 'extremely ill' patients is excluded from the calculations (New Jersey Department of Health and Senior Services, 2003) Rhode Island Department of Health has published summaries (Clinical Performance Reports) in December 2002 and July 2003, as part of an initiative with JCAHO and CMS Reports give highly aggregated summaries of hospital performance for each of the three disease areas For patient satisfaction surveys, a 3-point 'diamond' rating is assigned according to performance against state average (Barr et al., 2002)
Missouri Hospital Association BENCHMARK Project State-wide (Missouri) Voluntary	2004	Details of hospital compliance with clinical guidelines for AMI, heart failure and pneumonia (as part of the Hospital Quality Alliance)			Established in 1994, the BENCHMARK Project indicators included measures in the following categories: ambulatory, cardiac, cerebrovascular, endocrine, maternal/child, mental health, respiratory/pulmonary, surgical, trauma MHA has broadened its range of measures to include those appropriate for the ORYX initiative (JCAHO). Publication of selected ORYX performance measures appeared in March 2004 ²⁰
ORYX initiative National Mandatory (for those seeking accreditation from JCAHO)	2004	Risk-adjusted data on 'Hospital Core Measures', a range of indicators for AMI (including mortality rates), heart failure, community-acquired pneumonia, pregnancy and related conditions, and surgical infection prevention		Routine administrative data from medical records Hospital	Clinical data are adjusted for admission source, demographic characteristics and co-morbidity Part of the accreditation system for the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), the largest accreditor of health care providers in the US (Epstein, 1998) Designed in response to criticism that accreditation was based largely on process measures, ORYX incorporates health outcomes and quality of care measures (Epstein, 1998). JCAHO has piloted the measures in nine states including Missouri and Rhode Island. Following consultations with key stakeholders, measures are refined as part of a continuous quality improvement cycle and aligned with those required by the Centers for Medicare and Medicaid Services (CMS) JCAHO has been using performance data for accreditation purposes since the 1980s (Fink et al.,

Name of system					
Coverage					
Participation	Date of publication	Outcome measure		Data source	Comments
				Unit(s) of analysis	
					1989), with summary assessments made publicly available from July 2004
Hospital Quality Alliance National Voluntary	Not yet reported	Acute hospital performance against 10 recommended guidelines for heart attack, heart failure and pneumonia. Covering screening, medications and tests, these measures are a subset of those used in the ORYX programme Patient experience survey		Routine administrative data from medical records Hospital Routine administrative data from medical records, for all patients (including Medicare patients) Purpose-built survey instrument Hospital	<p>First announced in December 2002, dubbed 'Project Public Trust', then known as 'The Quality Initiative' before being renamed Hospital Quality Alliance</p> <p>A public/private partnership. The American Hospital Association (AHA), the Federation of American Hospitals (FAH) and the Association of American Medical Colleges (AAMC) lead the Hospital Quality Alliance. Supportive roles are taken by the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), the National Quality Forum (NQF), the Centers for Medicare and Medicaid Services (CMS) and the Agency for Healthcare Research and Quality (AHRQ)</p> <p>The aim is to provide the public with "valid and reliable information on which to assess the quality of health care being received by hospitalized patients"</p> <p>Although participation is voluntary, hospitals that do not submit performance data for 10 quality measures will receive 0.4 % lower Medicare payments in the following fiscal year (Centers for Medicare and Medicaid Services, 2004)</p> <p>AHRQ and CMS have produced the patient experience survey to allow providers and patients to compare care at different hospitals. The instrument contains 23 items, including a summary rating on an 11-point scale</p> <p>Outcome measures overlap with those collected by the ORYX initiative and participating hospitals may submit their data via ORYX.</p>

Box 3: Hospital outcomes reporting in the US: the impetus for publication

- In the early 1900s, Ernest Codman unsuccessfully attempts to have his 'End Results Idea' adopted more widely.
- In 1985, changes in the regulatory guidelines to the Freedom of Information Act (1966) mandate the publication of Medicare data that had been assembled for quality review purposes (Fottler et al., 1987). The Health Care Financing Administration (HCFA) publishes the data annually until 1993.
- Whilst most systems are voluntary, some are state-mandated, such as those operating in New York (1990), Pennsylvania (1992) and Rhode Island (2001). In 1991, the Freedom of Information Act is invoked in New York State to force publication of surgeon-level data.
- In 1989, a Cleveland coalition of businesses, hospitals and physicians is convened in response to excessively high local health care costs. The project ends in 1999 after several hospitals withdraw, claiming that purchasers are failing to honour their original commitment to 'Buy Right' (i.e. to choose higher quality providers).
- By 2000, hundreds of reporting systems are in operation, although few provide physician-level data. Information on Health Plans is prolific and many websites provide the public with comparative data. However, because most systems are voluntary, coverage is partial and selective.
- Over recent years, there has been a concerted effort to 'join up' the quality programs and their respective data requirements. Acknowledging the need for public accountability, 'The Hospital Quality Alliance' (dubbed 'Project Public Trust') involves a collaboration of numerous national health-related organisations. Aiming to provide the public with valid and reliable information with which to assess the quality of hospital care, data collection begins in the summer of 2004. Although participation is voluntary, non-participants will receive lower Medicare payments in 2005.

3.1.1 Codman: a pioneer

An innovative Boston surgeon, Ernest Codman (1869–1940) began a systematic follow up of his patients in 1900 (Spiegelhalter, 1999). Aiming to improve the quality of hospital care by learning from mistakes, Codman produced a card index system for each patient.

These cards contained patient-level data on outcomes together with codes designating any errors that could have led to these outcomes. Ten years later, he started to campaign for the widespread adoption of his 'End Results Idea', convinced that hospitals needed to embrace the comparative analysis of their 'results' and to welcome publicity of both their strengths and weaknesses if they were to improve. Although colleagues were initially enthusiastic about his 'End Results Idea', Codman's confrontational approach in promoting his system resulted in him being branded an eccentric. He lost his colleagues' support and, eventually, his job. He remained convinced that his idea would soon be accepted and implemented, but his optimism proved unfounded and it was not until the late 1980s that any comparative data was to come into the US public domain. Since 1997, a large US accrediting body has offered the 'Codman Award' for "achievement by organizations and individuals in the use of process and outcomes measures to improve organization performance and quality of care."²⁶

3.1.2 The HCFA picks up the baton

The Health Care Financing Administration (HCFA), the federal government agency that administered Medicare²⁷ until 2001, was the first US organisation to publish comparative performance data. With quality aims mirroring those of Codman, the HCFA drew up lists of poor performers, based on risk-adjusted mortality statistics, for each US acute hospital. 'Peer Review Organisations', responsible for overseeing the quality of care in hospitals, were to use the lists to identify potential problems for scrutiny. However, in 1986 the press gained access to the lists of poor performers under the Freedom of Information Act. The HCFA released the data, but accompanied by disclaimers: not only did the data have no intrinsic meaning, but the presence of a hospital on the list should not be interpreted as implying that the hospital was a poor quality provider (Fottler et al., 1988). The HCFA failed to brief hospitals about the release, the press apparently ignored the disclaimers, and confusion and criticism resulted (Fottler et al., 1987). The HCFA responded to the criticism by taking steps to

²⁶ <http://www.jcaho.org/accredited+organizations/codman+award/index.htm> (accessed 08/11/04).

²⁷ Medicare is the programme that helps to finance health care for older Americans.

54 revise both how the data were calculated and also its publication strategy:

1. A different type of statistical model was used to estimate the difference between actual and expected death rates;
2. This model adjusted findings for co-morbidity and for the source of hospital admission;
3. Death rates were no longer measured from admission to death or to discharge, but were assessed from 30 days of admission. Overall mortality rates were still published, but the number of diagnostic-related rates published was increased from nine to 16;
4. In contrast to the 1986 release, the HCFA gave hospitals pre-publication copies of their own data, with the opportunity to comment.

In 1987, the HCFA published these comparative data for all US acute hospitals – not just the ‘poor performers’ – treating Medicare patients (Fottler et al., 1988).

It was hoped that “competition between hospitals based upon real and perceived quality [would] accelerate in the 1990s” (Fottler et al., 1988). However, despite ongoing efforts to improve the statistical methods employed, the analysis remained dependent upon routinely collected data. Continuing criticism of the quality of the data led the HCFA to cease publication in 1993 (Mannion and Davies, 2002).

3.1.3 Report cards on hospitals and surgeons

However, the HCFA action led others to release similar types of information (Galvin and McGlynn, 2003). Publicly released performance reports – so called “report cards” – appeared in New York in 1990 and Pennsylvania in 1992 (Schneider and Epstein, 1996). Focussing on coronary artery bypass graft (CABG) surgery, the report cards contained risk-adjusted mortality statistics both for hospitals and for individual surgeons. Mandated by their respective states, both systems still publish data at the time of writing (November 2004). The New York system expanded its remit in 1996 to include mortality

55 rates for angioplasty (New York State Department of Health, 1997), whilst Pennsylvania’s Consumer Guide now reports charges, length of stay and re-admission rates for CABG (Pennsylvania Health Care Cost Containment Council, 2004). In 1997, New Jersey began publishing risk adjusted CABG mortality statistics for hospitals and surgeons (New Jersey Department of Health and Senior Services, 2003).

The Cleveland Health Quality Choice (CHQC) programme was set up in 1989 by a regional coalition of businesses, hospitals and physicians. Although high local costs of health care provided the original impetus, the coalition was driven by a consumerist value-for-money rationale:

“If Cleveland businesses can reliably identify the highest quality, cost-effective hospital services then this information can be used to encourage their employees to choose these institutions for their hospital care. In turn, with the incentive of more patient volume and reliable comparative information, hospitals will strive to maintain or improve their quality of care, while controlling their costs” (Baker et al., 2003).

The ‘Buy Right’ philosophy works as follows. On the demand side, businesses make a public commitment to buy health insurance for their employees from higher quality providers. In response, the supply side of providers work to improve the quality of care they provide. Publication of appropriate health outcomes facilitates the functioning of this deal (Farley et al., 2003). In other words, a ‘win-win’ situation was anticipated. Data were first published in 1992 and report cards for consumers appeared the following year (Rosenthal et al., 1998). However, in 1999 one of the participating hospitals withdrew from the coalition, arguing that employers were making little use of the data other than to negotiate improvements with their existing providers. In view of the costs of collecting and contributing the data, this hospital decided that participation was not cost-effective. Its withdrawal precipitated a collapse in co-operation and, in 1999, the CHQC programme ceased to operate.

In 1993, the Office of Statewide Health Planning and Development published findings from the California Hospital Outcomes Project

56 (CHOP) (Luce et al., 1996). Focusing on risk-adjusted mortality statistics for a range of conditions, the project was criticised on several counts. Serious underreporting of co-morbidity and miscoding of 'urgency of admission' led to erroneous estimates of risk-adjusted death rates (Green and Wintfeld, 1993). A survey of 17 acute public hospitals revealed that most had made no effort to improve quality in response to CHOP reports. This was primarily because adequate performance was considered sufficient, although the cost of undertaking quality improvements was also an important factor (Luce et al., 1996). From 2005, reporting of risk-adjusted mortality rates for CABG will be mandatory for Californian hospitals and will include surgeon-level data (Broder et al., 2004).

3.1.4 Report cards on Health Plans

One of the best known reporting systems is the Health Plan Employer Data Information Set (HEDIS), which summarises unadjusted data from volunteering plans (insurers) in the 'Quality Compass' (Marshall et al., 2000a). Data are collected by the National Committee for Quality Assurance (NCQA), an independent, not-for-profit organisation and the largest accreditor of health plans. NCQA is committed to evaluation and public reporting on the quality of managed care plans (Epstein, 1998). Comprising largely of process measures, HEDIS brings together information from medical record reviews, administrative databases and patient surveys (Marshall et al., 2003). Findings are also reported by state: for example, since 1994 the California Cooperative Healthcare Reporting Initiative (CCHRI), a collaborative of health care purchasers, plans and providers, has produced a state-specific report card. The CCHRI report card includes information on the HEDIS 'effectiveness of care' measures: plans' use of preventative services; use of new technologies for acute conditions; and medical care for chronic conditions. Although the Quality Compass is marketed as the "most comprehensive database of health plan performance data",²⁹ the voluntary basis of the HEDIS dataset

²⁸ http://www.cchri.org/reports/clinical_measures_2004.pdf (accessed 08/11/04).

²⁹ For example, the July 2003 version of the Quality Compass contained data on over 250 plans covering 61 million Americans (<http://www.ncqa.org/communications/news/qc03.htm> (accessed 08/11/04)).

has invited predictable criticism (McCormick et al., 2002).

57

3.1.5 Recent developments

Several hundred reporting systems now operate in the US, using data produced by government bodies, employers, business coalitions and consumer advocacy groups (Marshall et al., 2003). The publication of comparative data is mandatory in some states, with Medicare and Medicaid databases available for public scrutiny (Barr et al., 2002, Mannion and Davies, 2002).

From 1998, the HCFA reported quality standards of health plans for Medicare patients through the mandatory 'Medicare Compare' programme (US Department of Health and Human Services, 1998; Epstein, 2000). In 2001, the HCFA was renamed the Centers for Medicare and Medicaid Services (CMS). CMS administers the Medicare programme, and works in partnership with the states to administer Medicaid and the State Children's Health Insurance Program, which together provide coverage for about one in four Americans.³⁰ CMS engages in national initiatives to improve the quality of care (US Department of Health and Human Services, 2001). Perhaps the most important example of CMS involvement is Project Public Trust, better known as the Hospital Quality Alliance, which is underpinned by a large public/private collaboration (see Table 3). In the words of Rick Pollack of the American Hospital Association (AHA), the aim of the Hospital Quality Alliance is to provide the public with "valid and reliable information on which to assess the quality of health care being received by hospitalized patients".³¹ In their "Call for Action on Collaborative Effort" addressed to hospitals, the three organisations leading the initiative – the AHA, Association of American Medical Colleges and Federation of American Hospitals – explain its context:

"The pressure for more and better publicly available information about the quality of hospital care is coming from every direction.

³⁰ <http://www.cms.hhs.gov/researchers/projects/APR/2003/facts.pdf> (accessed 08/11/04).

³¹ http://www.hospitalconnect.com/aha/key_issues/patient_safety/advocacy/040203hcahpsletter.html (accessed 08/11/04).

The potential to confuse the public with incomplete, poorly analyzed and conflicting or misleading information is enormous. By working together, the organizations [named above] have pledged to coordinate these efforts for all parties involved – hospitals, consumers and purchasers. Hospitals must continue to improve quality internally and be publicly proactive. The issue no longer is whether quality data are to be made public – that is already happening. For hospitals, this is an opportunity to be leaders in forging a shared national strategy for quality measurement and public accountability” (American Hospital Association et al., 2003).

Acute general hospitals are the first to be invited to participate in the project, in which ten measures in three disease areas were to be assessed. These measures are a subset of those already required by the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), the largest US accreditor of health care organisations, under the ORYX initiative (see Table 3). In addition, the CMS works closely to ensure that data requirements for Medicare and Medicaid are aligned with those of JCAHO and the Hospital Quality Alliance. Although participation is voluntary, the CMS offers a financial incentive to encourage hospital participation (Centers for Medicare and Medicaid Services, 2004): hospitals that do not submit performance data on quality measures will receive 0.4% lower Medicare payments in the following fiscal year than they otherwise would (Centers for Medicare and Medicaid Services, 2004).

There are striking differences between the UK and the US histories of publishing hospital outcomes data. The catalysts for change are different: the principal trigger behind modern US publication was statutory right, embodied in the Freedom of Information Act (1966), whereas in the UK, the Labour government’s response to high profile failures of professional self-regulation was the main driver, with political support for greater patient choice being a later development. Reflecting the very different underlying political and health care structures, participation in the reporting systems differs: in the US, participation is generally voluntary and non-punitive (although with notable exceptions), whilst the UK approach is characterised by statute and regulation and is underpinned by government-stipulated incentives.

Nonetheless, the commonalities between the systems are many and a convergence between them is beginning to appear. Some of the key areas of overlap are:

- *Beginnings*: whilst the catalysts for publication differed, in both countries there had been a pioneer whose – apparently unsuccessful – attempts to publish hospital outcomes data arguably paved the way for modern developments. In the US, Codman’s efforts are now recognised through an eponymous award (see section Error! Reference source not found.);
- *Focus*: in both countries, modern publication initially focussed on the assessment of acute hospital care;
- *Obstacles*: both countries have battled with the problems inherent in the use of routine data and the dangers of inducing unintended consequences (see section 5.4);
- *Dynamics*: as a consequence of the difficulties experienced, both systems are characterised by evolution, growth and change. The indications are that this is set to continue as both systems strive to tackle their respective shortfalls;
- *Approaches*: attempts to integrate efforts and to rationalise the measures published are evident in both countries (NHS

Executive, 1998b; US Department of Health and Human Services, 1998);

• *Objectives:*

- Public trust: heightening public trust through publication, demonstrating a willingness to be accountable to the public they serve, transparency, openness and trust are motives frequently cited to justify publication on both sides of the Atlantic (Davies, 2001; Department of Health, 2002a; American Hospital Association et al., 2003). However, both countries have found publication to be a two-edged sword that can undermine public trust by drawing attention to deficiencies and failures;
- Quality control: policy makers in both countries are concerned over regional variations in the quality of care. In the UK, geographical variations in prescribing ('post code prescribing') and other 'unacceptable variations', such as death rates from coronary heart disease, emergency readmission rates and screening rates, were highlighted in the Labour government's white paper of 1997 (NHS Executive, 1997). In the US, the National Committee for Quality Assurance (NCQA) report claimed that 57,000 lives were unnecessarily lost annually due to quality gaps in the health care system (National Committee for Quality Assurance, 2003). Concerns over variation in the quality of care underlie much of the publication movement (Marshall et al., 2000a);
- Cost containment: in the UK, the NHS budgetary allocation places a ceiling on public sector health care expenditure, putting the opportunity cost of care high on the political and managerial agenda, reflected by concerns for 'efficiency'. In the US, cost concerns within managed care and business drove many of the early attempts to publish outcomes data. More recently, the NCQA report highlighted the "wasted billions" of dollars

resulting from substandard care (National Committee for Quality Assurance, 2003). Although rising pharmaceutical expenditure is a concern, both countries recognise the benefits of cost-effective pharmaceutical care. For example, the NCQA report estimated that appropriate pharmaceutical care could save over \$1.6 billion dollars annually for stroke and heart attack patients.

Despite the different catalysts that sparked publication, the principal objectives of publication in both the US and UK are remarkably consistent: to improve quality, to contain costs and to promote public trust. In pursuing these admirable objectives, the two countries appear to be treading similar paths – or at least, moving in the same direction. However, this in itself does not tell us whether the path they are following and the direction they are heading will lead them to their desired destination. What difference does the publication of hospital outcomes make? Do beneficial effects outweigh harmful ones? These are difficult questions to answer because publication takes place in the real, messy world, where confounding influences abound. Controlled trials, where known and unknown biases can be randomised away, are simply not a practicable method to evaluate publication. Consequently, any evidence that we consider will be of limited reliability. With this caveat in mind, we proceed to discuss how publication could theoretically be beneficial before examining the empirical evidence for its effectiveness.

In this section, we consider the purposes served by the publication of hospital outcomes, in contrast to restriction of access to them. Different groups – such as doctors, managers, the government, and the general public – are likely to enjoy different levels of access, and may have different information needs. We consider:

- what might be achieved by publication;
- the needs of the audience;
- the evidence concerning the success of current attempts to communicate hospital outcomes;
- the difficulties in producing valid data; and
- the adverse consequences of publication.

5.1 How might publication be beneficial?

5.1.1 Promoting public trust

Events at Bristol, coupled with many other examples of medical shortcomings, have led to a “crisis of trust in the UK” (Maynard and Bloor, 2003), which has led to an upsurge in “regulation, inspection, target-setting and audit” (O’Neill, 2002). The Labour government has introduced a sophisticated performance assessment framework for the NHS, which is intended to include both management and patient-oriented information about quality of care (NHS Executive, 1998a).

The framework includes the National Institute for Clinical Excellence (NICE), with a remit for developing national plans for the organisation of service provision and for producing guidance for clinicians and patients about the effectiveness of treatment. NICE has made considerable investment to include patient involvement in its work, although its preferred decision making tool for making recommendations about health care provision, the cost per QALY (quality adjusted life year), is perhaps more informative for policy makers than for the general public.

The Healthcare Commission has been established as an independent

statutory body, designed to monitor and improve clinical practice and to act as a ‘quality watchdog’ with the power to intervene when it finds evidence of poor performance. Perhaps the most important aspect of the framework is the requirement that Primary Care Trusts and hospitals introduce a system of clinical governance. Clinical governance recognises that performance must be owned and driven locally, and relies on involving individual health professionals in their organisation’s programmes for quality improvement. As part of the clinical governance programme, rather than being accountable for only the financial and legal affairs of the hospital, chief executives now carry ultimate responsibility for assuring the quality of services provided by the hospital. Furthermore, participation in audit activities, such as CEPOD, has been made mandatory rather than voluntary (Warden, 1998b).

A problem for the government is that, even if change occurs as a result of introducing the NHS Performance Assessment Framework, any performance improvements may not be obvious to the general public, particularly if changes cannot be reported in a simple, summarised format. By reporting hospital mortality tables, the government can demonstrate to the electorate that action is being taken to improve hospital performance, even if this particular performance indicator is not a central component of the performance assessment framework. According to this interpretation, the value of the indicator lies in its ability to act as a symbol for the broader policy programme (Jacobs and Manzi, 2000). Government ministers and departmental officials might publish performance indicators to send reassuring signals to the public, to suggest that ‘something is being done’. By placing performance data on hospital outcomes in the public domain, governments and officials make themselves accountable, showing their dealings to be transparent and open.

A limitation in all these initiatives is their top-down approach to quality. None has started by understanding what aspects of care and treatment patients value to build quality-of-care indicators and it is perhaps unsurprising that the public feels disconnected from current top-down data better suited to managerial goals.

5.1.2 Supporting patient choice

In the US, there is a prevailing belief that informing patients is a key means to improving the quality of care (Baker et al., 2002; National Committee for Quality Assurance, 2003). There has been less of a tradition of this consumerist philosophy in the UK, but there are signs that this is changing. For example, whereas Frank Dobson, Secretary of State for Health from 1997 to 1999, was opposed to the idea that patients shop around for treatment (Department of Health, 1999), his successors, Alan Milburn and John Reid have pushed the 'Patient Choice' agenda (Department of Health, 2000a), which aims to move the emphasis from quality control by government-imposed targets to quality driven by patient demand:

“As targets fall away they will be replaced by the power and preferences of patients and [Trust Chief Executives'] abilities as leaders and clinicians to respond to their needs” (Reid, 2004).

Quality improvements may follow if providers are under pressure to compete for patients, but if no such pressure exists then it is less clear how quality improvements are to be realised. Unlike shopping for consumer goods, there is no price signal to match supply and demand and some provider organisations may find that they have insufficient patient numbers, whilst others are struggling to cope. Although the new financial framework allows money to follow the patient to support their choice (Department of Health, 2002b), capacity constraints limit the number of patients that a hospital can treat in the short term. Ironically, these limitations may increase waiting times if patients are prepared to wait longer for better quality care. Whilst this may be true of elective care, it is less likely to be true where disease progression is an important consideration. In this case, information on the number of cases treated by an individual surgeon may be more important to patients. In some cases of acute care, the patient (or their carer) may feel too vulnerable to make an informed choice (Marshall and Davies, 2001), and there may be a place for establishing “channels of influence”, through which patient representatives can negotiate improvements in quality (Klein, 1980).

Increasing patient choice translates as giving patients more say in how, when and where they are treated (Epstein, 2000). From summer 2004, NHS patients with coronary heart disease, as well as those who have waited over six months for some elective surgical procedures, have been offered a choice of continuing to wait for treatment at their local hospital or having treatment on a specified date in another hospital. The government plans to extend this choice so that patients can:

- book appointments at a time that suits from a choice of hospitals;
- access a wider range of primary care services;
- have more choice of where, when and how to get medicines;
- have more choice of maternity and end of life services.

Whether allowing greater choice is an effective means of improving health services remains an open question, especially as only a minority of patients are likely to exercise choice. Advocates of the choice agenda would argue that even if only small numbers of patients opt to move, this is enough to stimulate general improvements. In other words, any quality improvement is a 'public' rather than a 'private' good because those who do not exercise choice may nevertheless benefit from the fact that others do.

The experience with general practice fundholding can be cited in support of this contention. General practice fundholders were able to exercise much greater influence upon the health service than would be expected on the basis of the proportion of NHS money they controlled. They had much greater freedom than other purchasers to move their money around, which made providers more responsive to them. But detractors would argue that the responsiveness was largely restricted to patients registered with fundholders, and that a two-tier system developed, disadvantaging the patients of non-fundholding GP practices. Similarly, patient choice may undermine equity, with better-educated people more likely to exercise choice. In a capacity-constrained system, this may mean that benefits accrue to better-informed people at the expense of those less-informed, the more seriously ill, or the more vulnerable.

In economic terms this phenomenon is known as 'elasticity', which describes the responsiveness of demand or supply to a variable such as

66 quality of care. If some patients have a higher elasticity of demand with respect to quality (i.e. are more likely to exercise choice and select better quality care) than others, then quality improvements may in reality be a private good, enjoyed only by those who exercise choice.

The patient choice agenda requires an information strategy. For example, to offer patients a choice of maternity services, providers may need to find out what patients want, and ensure that arrangements are in place to facilitate patient choice. To do this, doctors need to act as an agent on their patients' behalf, interpreting, providing or facilitating access to the information that patients need to make an informed decision. However, for various reasons the agency relationship may be imperfect. Publication of outcomes may empower patients to place less reliance on their agents or to make agents better representatives. For patients to make better decisions or to take a greater role in the decision-making process, they will need various types of information, including information on quality (Klein, 1980, Marshall and Davies, 2001, Department of Health, 2002b).

The likely outcome of treatment has obvious potential to be useful information. There is evidence (admittedly after the fact) that patients would have made different choices if better informed. For instance, parents of babies and children who had been operated on by Wisheart in Bristol said that "they would never have given permission for the surgery had they known about the mortality rate" (Jury, 1996). Calls for information about the performance of individual doctors to be provided to patients have also come from within the medical profession. Commenting on the ramifications of the Bristol inquiry in the British Medical Journal, the professor of cardiothoracic surgery at St George's Hospital, London, wrote:

"It appears to be self evident that parents have a right to know the truth from both referring cardiologist and the surgeon. Why are doctors ever economical with it? Is truth thought to contaminate the trust in a relationship?" (Treasure, 1998)

However, 'consultant-level' data may refer to a consultant-led team – and this may be more appropriate for interventions that are delivered on a team basis. Information about surgeon-specific mortality rates is

67 of less interest to individuals whose treatment is largely medical. This may be especially true of patients with chronic conditions, who may have an on-going interest in other types of health information (Marshall et al., 2003). For these patients, choices involve longer-term investments in their health, and it is less likely that data on outcomes for chronic conditions will be satisfactory, particularly as outcomes can be so difficult to define. Rather "it is likely that reporting schemes will have to rely on measures of process, preferably those surrogate outcomes that are known from research evidence or professional opinion to be strongly linked to good health outcomes" (Marshall et al., 2003).

Suppose a patient suffering a chronic condition is offered a choice of provider, either at a specialist clinic at a hospital 30 miles away or at the local hospital. The patient might want to consider factors such as the following in reaching a decision:

- Will the specialist clinic offer better facilities, expertise and access to a greater range of therapies and medications (perhaps as part of a clinical trial)? Which measures of quality are available to make a judgement?
- How much more difficult will it be to reach the clinic and what implications might this have for work and leisure time?
- Will the patient be able to forge good relationships with staff at the specialist clinic and will these be better than those at the local hospital?
- What 'escape routes' are available if the transfer to the specialist clinic is unsuccessful? Will there be discrimination, or even recrimination, by local staff as a result of the attempted transfer?

Perceptions of loyalty and the building of on-going relationships may well outweigh quantitative evidence, based on routine data:

"How often would a patient with stable coronary artery disease be willing to break ties with his physician and seek a new one because the hospital where he might someday be hospitalized with an AMI has a higher than expected mortality [rate]?" (Baker et al., 2003)

68 Reading about performance indicators may provide reassurance about competency but may tell the patient nothing about what sort of people will provide the care. This may explain why patients appear to rank qualitative evidence from informal sources, such as friends and family, more highly than quantitative data from official sources (Mannion and Goddard, 2003).

5.1.3 Stimulating action

One of the main reasons for measuring and collecting outcomes data is to improve the quality of care. Having outcomes data is clearly an important if insufficient step to quality assurance. Drawing lessons from the Bristol tragedy, it has been argued that the principal problem was the failure not to monitor, but to act on available evidence (Willis, 1998). Although independent watchdogs may scrutinise clinical performance, when outcomes data are placed in the public domain the scrutinisers themselves become – potentially at least – subject to public scrutiny. Failure to detect or appropriately manage poor performance can be judged by visible yardsticks. This knowledge may encourage managers and health professionals to act, rather than procrastinate.

Quality of care can be seen from managerial, health professional and patient perspectives, for which the respective informational requirements are quite different. In part, this is to do with aggregation: managers are interested in overall performance at various levels, while patients want to know the range and likelihood of possible outcomes they may experience from health care interventions. The hospital outcomes data described in this report have largely favoured management goals, for use by institutional managers, regulators and referring doctors, such as GPs, who act as agents on patients' behalf. Using these data, performance may be assessed, good practice highlighted, shared and rewarded, poor practice investigated and improved, and instances of unsafe practice exposed. The Bristol enquiry might have been averted had such performance assessment procedures and the mechanisms for triggering timely correction been in place. Cases of deliberate malpractice, such as the notorious actions of mass murderer Harold Shipman, may also be detected and stopped.

69 Performance assessment is about making organisations and individuals accountable for the quality of care they deliver. It is about ensuring that common standards are met and that value for money is achieved. However, these purposes could be achieved without the publication of the data. It is not clear that quality markers for managers need to be published or that they can be readily interpreted by patients, for whom an alternative and more appropriate format or type of information may be more useful. This could involve measuring and collecting very different types of information, which would have its own costs and benefits (rather than simply being incremental to the existing arrangements). These points cast doubt on the possibility that the publication of mortality statistics will prevent tragedies such as Bristol or encourage general performance improvements.

Another explanation for the UK government's publication policy is more generous about the impact that information may have on practice. There is evidence to suggest that managers and doctors do use comparative information to question and challenge behaviour, although changes may take time to occur and may not always be as intended (Marshall et al., 2000a; Schneider and Lieberman, 2001).

Performance indicators have been described as a "safety net" to ensure that performance does not fall below minimum acceptable standards, with managers more concerned about eliminating bad performance than enhancing good or satisfactory performance (Goddard et al., 1999). But it is apparent that, by itself, a formal system of performance measurement will be insufficient to deliver performance improvement. The challenge is to develop a performance framework that both safeguards minimum standards and identifies and encourages good performers so that they continually seek performance improvements. In this regard, the move to a broader analysis of NHS performance is welcome, but careful consideration must be given to the system of rewards and penalties that support the performance framework. If the performance assessment framework can be incorporated within a coherent incentive structure, the evaluation of NHS services may encourage genuine improvements in performance.

5.2 Evidence for the beneficial effects of publication

“Despite the interest in and resources expended on the production of comparative performance reports, there has been remarkably little formal evaluation of their impact on the various stakeholders or the effect of the reports on the processes and outcomes of care” (Marshall and Davies, 2001)

Given the importance placed on publication of quality indicators it is surprising how little evaluation of the positive and negative impacts of publication has been conducted. One review found published evaluations of only seven reporting systems, all of which operated in the US (Marshall et al., 2000b). Relying chiefly on observational study designs and routine administrative data, the reviewers concluded that the impact on health outcomes was uncertain, but that there may be a small positive effect. Consumers and purchasers rarely used the data, but there was evidence to suggest that a minority of doctors used it. The most receptive audience appeared to be the hospitals themselves and the reviewers highlighted a potential for further exploitation of outcomes data for quality improvement. For example, one hospital undertook case studies and an extensive statistical analysis to explore the reasons for its poor performance (Dziuban et al., 1994). A subgroup of patients with particular risk factors who underwent emergency CABG was identified as the cause of the hospital's performance problems. When steps were taken to improve the quality of care for these patients, the hospital's performance improved. The costs of the exercise are not reported, and there is little evidence in general about the economic implications of publication (Marshall et al., 2000b).

A consultation with 18 US experts in the field of performance data publication highlighted three principal achievements attributable to the publication of report cards (Mannion and Davies, 2002). The first achievement is a shift in attitudes towards quality assurance (see also Davies, 2001). Before publication of the report cards, quality assessment was characterised by self-regulation and confidential peer-review; but in the 1990s, public and external scrutiny was not just accepted as a legitimate approach but came to be viewed as an essential

component of quality assessment. A second achievement of the report card movement was public awareness of variations in quality and performance of health care. One expert believed this knowledge shift to be the major achievement of the publication experience. Lastly, a technical shift occurred, with improved tools and techniques for measuring performance.

In the UK hospital context, there are three potential barriers that may limit the benefit of publishing quality indicators.

First, mere publication of data does not ensure that the information is *publicised*. The experience of including mortality statistics among the 1992 Health Service Indicators (HSIs) illustrates the difference. HSIs were not intended for use by the general public but even most of those working in the NHS were unaware of the information they contained, with there being no standard procedures for analysing and acting upon the data and only weak incentives for purchasers and providers to use the information. Generally, the data contained in the HSIs were little utilised (Street, 2002).

Second, where action does occur, it tends to be concentrated among hospitals that have been ‘named and shamed’ and these hospitals tend to adopt defensive strategies. For example, the usual response to the publication of the hospital mortality tables has been to question the quality of the data, often by producing statistics from an internal (and rarely independent) audit (Laurance, 1993a; Milledge, 1993; Platt, 1993; Jones, 1994; Harry, 1995). Hospitals in the ‘comfort zone’ can safely disregard the data, confident that they will suffer no adverse publicity. The media play a crucial role in publicising data (Davies and Shields, 1999), but they are interested only in highlighting the worst (or, occasionally, the best) cases. Hospitals that avoid the extremes of the ‘league table’ tend to receive no coverage. An exception to this rule is the ‘Good Hospital Guide’ reproduced by the *The Sunday Times* from 2001 onwards, which summarises the performance of all public and private hospitals in Great Britain and Ireland against a range of indicators, including mortality rates (Anonymous, 2001a; Anonymous, 2001b).

Third, the presumption that the public would use comparative performance data to put pressure on their local hospital services appears optimistic. In theory, this pressure could be expressed in two possible ways: ‘exit’, where patients change provider to signal their discontent with services; or ‘voice’, where patients, individually or collectively, negotiate quality improvements (Klein, 1980). A combination of the two routes is a third option. In the US, (Marshall et al., 2000a; Schneider and Lieberman, 2001) with a few exceptions, the publication of data appears to have had little impact on consumers or their employers, who are the main purchasers of health care (Marshall et al., 2000a). In the case of the Cleveland Health Quality Choice program, employers, though encouraged to use the ‘exit’ option, in practice favoured ‘voice’, using information on quality to negotiate better deals with existing providers rather than switching allegiance (see section Error! Reference source not found.). Overall, the US experience suggests that “public disclosure of information about the quality of health care is a weak strategy for ensuring quality” (Schneider and Lieberman, 2001). Despite the increase in public awareness, there is little evidence that patients and carers use performance data from report cards to inform their health care decisions (Marshall et al., 2000b). One reason for this is that consumers often do not understand report cards (Hibbard and Jewett, 1997):

- Patients understand and trust patient satisfaction and patient-reported quality markers above the more ‘objective’ measures;
- Patients do not understand what terms mean, whether high or low is ‘good’, or the link between the indicator and quality of care;
- Patients do not appear to grasp the importance of context in interpreting indicators.

Research has also identified a mismatch between what patients say matters in their choice of health plans and which factors actually influence their choice (Hibbard and Jewett, 1996). The challenge of achieving patient advocacy is increased by doctors’ reported scepticism about quality indicators. An assessment of patient and public views of performance indicators in the UK echoed findings from the US

literature (Magee et al., 2003). Although the 50 interviewees were all within the catchment areas for either zero- or three-star rated Trusts, just one was able to correctly report the recently-published rating for their local Trust. Whilst participants generally agreed that performance should be monitored in some way, there was little enthusiasm for hospital league tables. This finding was supported by a MORI omnibus survey, in which 1,749 members of the public were asked a set of 10 questions on the ‘patient prospectus’ published by each Primary Care Trust (PCT) as part of the government’s NHS Plan:

“Using the guide to explain league tables, star ratings, spending plans and future investment plans (and other strategic health service information) is not welcomed by the public. Largely they see this as either too complicated or too political especially where the context to make sense of the data was inadequate and couldn’t easily be linked to local service provision. Presentation of information in this way should be avoided, or at the very least minimised in future guides.”³²

Around two-thirds of the public surveyed (68%) said they did not recall receiving the prospectus. The passive dissemination of information is generally recognised to have little effect on behaviour, although it may be useful as part of a multifaceted strategy (NHS Centre for Reviews and Dissemination, 1999). Of the 429 people who did recall receiving the prospectus, no one identified league tables as the most useful piece of information and 7% identified them as the least useful. The Consumers’ Association reported a similar finding:

“As they are currently published, the Performance Indicators are simply not useful to patients. They are not detailed enough to highlight differences in performance in a meaningful way. But there are moves to change this by the Department of Health, which we welcome. However, even if information about clinical teams and individual doctors were published, patients are likely to have little choice about who they are treated by or where” (Anonymous, 2001c)

³² <http://www.natpact.nhs.uk/uploads/MORI.doc> (accessed 08/11/04).

74 Despite these negative findings, there are at least two reasons why offering choices to UK patients may still be valuable.

Firstly, US experience that patients do not exercise choice may in part be due to the overriding importance of cost issues within the US health care market, coupled with the failure of US health care prices to reflect quality issues (Galvin and McGlynn, 2003). In the UK, cost issues are set to one side: prices are fixed and so cannot convey quality signals to consumers. Moreover, UK patients do not generally incur the cost of treatment and so costs are unlikely to influence patients' decisions. Does this mean that quality issues inform the patient's choice? This does not necessarily follow: if patients believe that quality issues should be settled by an external regulator (so that only care of good quality or better is available), then choice may be informed by other factors. These include: waiting times; the distance to the hospital; transport facilities; hospital facilities such as visiting hours and ease of parking; rates of infection in the hospital; and staff attitudes. Summary data may provide answers to some of these questions, but informal sources may also be used.

A second reason why the UK patient choice agenda may work is that the exercise of choice in health care will develop very slowly. It is clear that a small proportion of the public are aware of published indicators and that an even smaller proportion actually use them for decision making. However, these low levels of uptake may simply reflect a slow 'diffusion curve' (Galvin and McGlynn, 2003). The literature on the diffusion of innovation literature assesses how new ideas or technologies diffuse through the population, examining why change occurs and who influences its course. Once a critical mass of the population has adopted an idea or taken up a new technology, diffusion speeds up and the idea or technology becomes widely adopted. For patients to take a more active role in deciding their treatment pathways, there needs to be a significant shift in the doctor-patient relationship. If it took over a decade for the US public to become aware of variations in the quality of care (Mannion and Davies, 2002), then to achieve such a significant shift in the doctor-patient relationship may take even longer.

5.3 Data problems

75

Although the themes overlap to some extent, there are two classes of problem that arise when assessing hospital outcomes. The first relates to using routinely collected data commonly designed with other purposes in mind and varying in quality. The second relates to abstracting valid, robust interpretations from the data.

5.3.1 The limitations of routine data

Routine data have limited explanatory power (McKee and Hunter, 1995) and are associated with considerable methodological problems. The interpretation that can be placed on routinely gathered observational data may be severely limited (Davies and Lampel, 1998). Despite the quantity of data produced in the US, the validity of routine administrative data for quality assessment remains unclear (Mannion and Davies, 2002). There is some evidence that this also true of routine data collected in the UK (McKee and Hunter, 1995; Bridgewater et al., 2003), although such data may serve as a marker to poor performance requiring further investigation (McKee and James, 1997).

Problems with routinely-collected data include:

- Incomplete or missing data;
- Lack of adequate adjustment for confounding factors;
- Risk of over interpretation of data and failure to understand the play of chance;
- Miscoding / variation in coding practice.

Chiefly for these reasons, the British medical profession has resisted the call for consultant-specific death rates to be made public. In its discussion paper published in response to the clinical indicators, the British Medical Association argued that there was no measure that could provide the required data because of the problems of case-mix or risk-adjustment (British Medical Association, 2000). Without adequate risk-adjustment, doctors may be tempted to avoid treating high-risk patients (Keogh et al., 1998). Nor do individualised,

76 consultant specific, data recognise the collective nature of health care delivery (Dawson, 1998; Egan, 1998).

The experience of publishing individual data in the US gives a salutary warning for doctors in UK. In the US, individualised data were forced into the public domain following the Freedom of Information Act (Chassin et al., 1996). For example, mortality data were published for 88 surgeons performing coronary artery bypass surgery by the New York State Department of Health (New York State Department of Health, 1993). This resulted in the dismissal of some of the doctors identified as having high death rates, although subsequent analysis suggested that “nearly half of the 88 surgeons moved from one half of the table to the other” (Langford, 1997). This illustrates the problem of drawing conclusions about the performance of individual clinicians from such data.

“Even if all surgeons are equally good, about half will have below average results, one will have the worst results, and the worst results will be a long way below the average” (Poloniecki, 1998).

5.3.2 Problems with the valid interpretation of hospital outcomes

The risk of making inaccurate inferences is likely to be lower if the organisations or doctors being assessed perform large numbers of a procedure where casemix is well described. However this is often not the case. For example, annual volumes of CABG surgery may be too small for meaningful mortality comparisons and this problem is more severe for surgeon-level data. Some ‘high-mortality outliers’ in Pennsylvania were found to have been misclassified as a result of the inappropriate use of conventional statistical tests and modelling techniques (Localio et al., 1997). It is vital to apply appropriate statistical techniques that allow for sample size, or else apparently statistically significant differences may actually be spurious (Localio et al., 1997).

Understanding the variation in performance between doctors must disaggregate at least four sources of variance, two of which are ‘benign’, one which is ‘moot’ and one which is the variance we want to measure:

77 ‘true’ differences in the quality of care. ‘Benign’ variance due to differences in patient casemix can be lessened by adequate statistical adjustment for confounding. ‘Benign’ variance due to the play of chance is managed by using appropriate statistical tests, which avoid over-precise estimates. ‘Moot’ variance may arise due to differences in the resources available and/or composition of health care employed, for example in surgical, technical or nursing support, perhaps reflecting local policy variations (Seagroatt and Goldacre, 2004). Once these three are dealt with, what remains may be variance due to performance, assuming the quality marker is an adequate proxy for the quality of care delivered by the clinician team.

Robust, interpretable findings will emerge only from good data that are carefully analysed. ‘Careful’ here supposes considerable knowledge of the data as well as statistical techniques. For example, an analysis of market share following the publication of performance data identified two important factors (Mukamel and Mushlin, 1998). Firstly, the appropriate unit of analysis should be used: in this study, market shares for hospitals did not change significantly, but when the data were analysed at the level of the surgeon a significant effect was found. Secondly, the impact of the publication of performance data tailed off over time, suggesting that the timeframe chosen for the analysis may also affect findings.

Rather than tackle these issues, the quality agenda in health care seems sometimes to confuse the quantity and quality of indicators. Indeed, the complexity, heterogeneity, differential importance and sheer number of different indicators make for confusion, dissipation of effort and the “paralysis of analysis” (Davies, 1998).

“However extensive the data gathering, and however sophisticated the adjustments made, there will always remain considerable uncertainty about the link between true performance and reported health outcomes... thus the possibility of false denigration and false reassurance limit the meaning that can be read into health outcomes” (Davies, 1998)

The potential problems of confounding are illustrated in current

British league tables based on the Department of Health's Hospital Episode Statistics (HES), which provide a record of all hospital admissions in England. Responding to the Dr Foster and Department of Health 2001 league tables, the Society for Cardiothoracic Surgeons (SCTS) acknowledged that HES data were adjusted for a number of factors, such as age, gender, urgency of operation and deprivation. The SCTS nevertheless argued that:

“Surgeons base decisions for surgery on much more detailed and often subtle clinical information. The way the tables have been constructed is the best that can be expected from the data available, but many other conditions such as underlying heart function, lung function, smoking history, diabetes, obesity, high blood pressure, kidney function and other vascular conditions all have an impact on the risk of a heart operation. These factors must all be taken into account when calculating surgical risk, particularly if meaningful comparisons between units or surgeons are to be made. For example, if a patient has had a heart attack which has significantly reduced his heart function to the point where he is very breathless he is ten times more likely to die during an operation than the average patient, yet he may have the most to gain. Such patients may be seen in some units in greater numbers than in others because some units serve older and sicker populations, or because those units have special expertise; unless such factors are taken into account unfair comparisons may be made. This type of important and clinically relevant information is simply not available in the current NHS information systems. In fact, of the four most important risk factors for coronary bypass surgery (advanced age, emergency status, poor heart function and whether this is a repeat operation), only the first two are accounted for in the published tables.”³³

Appropriate information needs to be collected so that performance assessments can take account of known confounding variables. Even if confidence intervals are reported with these data, the danger remains

³³ The Society's Response to the Dr Foster and Department of Health League Tables. Published in part in the Times on 19th November, 2001 <http://www.ctsnet.org/doc/6153> (accessed 08/11/04).

that assessments are given a spurious credibility, with users failing to recognise that important confounding factors may still have been overlooked (Keogh et al., 2004). However, presenting data that are not risk-adjusted may yet be of value, because it emphasises the indicative, rather than designative, nature of the data.

These comments notwithstanding, substantial advances have been made in terms of data collection and statistical analysis, with the SCTS leading the way in the UK. All consultant cardiac surgeons have voluntarily submitted data to the SCTS register and decisions have been made about the definition of the performance measures, the form of risk adjustment and the trigger point for constructive review of performance (Keogh et al., 2004). The Department of Health is pressing the SCTS to publish surgeon-level mortality data (Carlisle, 2004; Hawkes, 2004). At the time of writing, in November 2004, however these data had not yet been published because of concerns among cardiothoracic surgeons about potential unintended negative effects. We now turn to a discussion of these effects.

5.4 Dysfunctional consequences of publication

Data gathering entails an opportunity cost on scarce health care resources (Davies, 1998) and so it is important to assess the benefits and harms that result. The key benefit for publishing outcomes data is to attempt to improve the quality of care. However, there may also be unintended harm arising from measuring, collecting and publishing such data. For example “beating the system, not improving quality, becomes the aim of the game” and “performance measurement ... may pervert behaviour and engender an adversarial and defensive culture detrimental to quality” (Davies and Lampel, 1998). This phenomenon is known as Goodhart's Law':

“Any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes” (Goodhart, 1984; page 96).

A typology, devised by Smith (Smith, 1995), of the dysfunctional consequences that might arise, together with examples, is given in Table 4.

Table 4: **Possible dysfunctional consequences arising from the publication of outcomes data**

Consequence	Meaning	Examples
Convergence	Aiming for average quality, rather than excellence	Hospitals not labelled as a high or low outlier by the New York rating system failed to use the data to "lift themselves from mediocrity to excellence" (Chassin, 2002)
Gaming	Changing behaviour to gain strategic advantage	Use of inflatable tents by ambulance trusts, to provide a "target-free limbo" (McKee, 2004) Miscoding of prevalence data, reducing severity adjusted mortality rates (Green and Wintfeld, 1993).
Misrepresentation	Includes 'data mining', changes in data recording, creative accounting and fraud	Use of 115 as default age where patient data not recorded to improve risk-adjusted score (Luce et al., 1996)
Myopia	Obsession with short-term goals	Diversion of resources in London Accident & Emergency departments during assessment week to meet targets (Mayor, 2003)
Ossification	Reluctance to experiment with innovative technologies to minimise the risk of poor performance	Self-reported reluctance of some cardiac surgeons to operate on high-risk patients in the US (Schneider and Epstein, 1996; Burack et al., 1999) and UK (Keogh and Kinsman, 2002)
Sub-optimisation	Prioritising narrow objectives that are organisation-specific over broader, inter-organisational strategic goals	Patients held back, deferred or not removed from waiting lists to ensure targets were met (Dobson, 2004)
Tunnel vision	Focussing on areas assessed, at the expense of non-assessed areas	The drive to meet new outpatient appointment targets led to delayed follow-up appointments, allegedly resulting in 25 patients losing their sight (Gulland, 2003)

Adapted from Smith, 1995

Interviews with experts in the area of performance assessment identified two particular unintended effects occurring in the US, namely tunnel vision and ossification (Mannion and Davies, 2002).

Another example of unintended consequences arising from the publication of outcomes data comes from the US experience of cardiac surgery report cards. Dranove and colleagues conducted a series of controlled analyses using Medicare data and survey data from the American Hospital Association. Focusing on elderly patients with cardiac disease, the difference in trends between states with mandatory reporting systems (New York State and Pennsylvania) was compared with the difference in trends in control states (all other states) and in neighbouring states (Connecticut, Maryland and New Jersey) (Dranove et al., 2003). Two key assumptions underpinned the analysis: first, that CABG report cards do not affect the composition of heart attack patients who are hospitalised; second, that CABG report cards do affect the treatment these heart attack patients receive. Comparing 1990 (predating report cards) patterns with those in 1994 (when report cards had been available for a few years), the analysis revealed that:

- Illness severity remained stable over time in all the state groups;
- There was a nationwide increase in treatment intensity for this patient population;
- The increase in the quantity of CABG surgery in New York and Pennsylvania was made up entirely of healthier patients, a trend that contrasted with incidence trends in control states;
- For heart attack patients in New York and Pennsylvania, there was a small improvement in health outcomes for the subgroup of healthier patients, but a substantial increase in adverse outcomes for the subgroup of sicker patients;
- There was an increase in expenditure for all groups of patients.

It appears that providers in New York and Pennsylvania were selecting less sick patients for CABG, and that this was associated with poorer outcomes for sicker patients. However, report cards did appear to have

some beneficial effects insofar as patients were being better matched to hospital capabilities: for example, patients undergoing a heart attack were more likely to be treated at one of the large teaching hospitals. Higher costs and poorer outcomes for sicker patients meant that, on balance, report cards had led to a net reduction in welfare. Redesigning the cards to minimise the incentives and opportunities for providers to select patients could redress this imbalance.

Dranove and colleagues' findings contrast with an earlier study that found no systematic bias against operating on high-risk patients amongst states with public reporting systems (Hannan et al., 1997). However, this study considered only patients who underwent CABG surgery and took no account of high-risk patients who were eligible for, but did not receive, surgery.

Moreover, the analysis by Dranove and colleagues is supported by findings from other studies. An anonymous mail survey of New York State cardiac surgeons found that 62% of the 104 responders had refused to operate on at least one high-risk CABG patient during the previous year, primarily because of public reporting (Burack et al., 1999). Similarly, a Pennsylvania survey found that 63% of cardiac surgeons reported that they were less willing to operate on severely ill patients requiring CABG surgery and that 59% of cardiologists reported increased difficulty in finding surgeons willing to perform it for this patient group (Schneider and Epstein, 1996).

Dysfunctional reactions to publication may not be confined to the organisations assessed; there is some evidence that policy makers may also be tempted to 'cheat'. An article in the *Health Service Journal* reported on a series of emails between the Secretary of State for Health's office and the Department of Health's performance development unit (McLellan, 2003). On July 12th 2002, the Department of Health provided the government with a proposed list of 2002 star ratings and in response an aide at the Secretary of State's office requested information on a number of Trusts. The email continued:

“Secretary of State [Alan Milburn] would also identify South Durham as a high profile Trust – given it serves the Prime

Minister's constituency. Why has it fallen from three stars last year to two stars?” (Cited in McLellan, 2003)

The next day, the ratings had been revised and South Durham was awarded three stars, making it eligible for the privileges associated with this higher rating. When the final ratings were published on 24th July, nine Trusts had been upgraded to three-star status and six had been downgraded to two stars. The Department of Health noted that revision of the ratings was normal practice:

“Data being used in the calculating the ratings is subject to regular change throughout the last few weeks leading up to publication. It is a necessary part of the process to ensure the published information is as accurate as possible” (McLellan, 2003).

To explore the Department of Health comment using an analogy with examination marking: whilst revising cut-off points for grading of exams might be acceptable marking practice, would the same be true if sections of an exam were retrospectively given different weightings? There is an issue with the public perception of quality indicators that they are there to serve the politicians' rather than the patients' interests. The appearance of political massaging of the data merely reinforces this perception (Marshall et al., 2000a).

Our findings may seem discouraging. A consumer-led NHS requires the availability of information to enable patients to participate, if they wish, in decisions about their treatment. At the managerial level, credible performance data are needed to inform a dialogue between managers and clinicians to address the quality agenda and target inadequate performance promptly as it emerges. At the policy level, the Department of Health and government need aggregate measures that value NHS health care delivery in a transparent, valid and robust manner so that both health service professionals and patients can have confidence in policy directives. It is disheartening then that health service quality indicators are held in such low esteem:

“Few examples better show the triumph of ideology over evidence than the continuing quest to encourage patient choice by publishing the outcomes of health care providers” (McKee, 2004).

But whatever the force of such complaints, the provision of information to the public can be viewed as good in principle, and that the health service simply must respond positively to demands by consumers for greater information in all areas of life (Freedom of Information Act 2000). If the public has a right to know about health services, the question arises as to how to provide information properly. The main message of the preceding discussion is that information provision can be improved, and probably at relatively little cost. Five pointers are offered to promote progress.

1. Recognise that publishing inadequately constructed, measured and interpreted quality indicators will have at best equivocal benefits.

Drawing on the US experience of publishing outcomes data, there are a number of desirable features for public reporting systems (Mannion and Davies, 2002):

- **Co-ordinated systems with mandatory participation.** In many cases, outcome statistics are constructed from data collected for other reasons. These may not be ‘fit for purpose’ and attention must be given to primary data collection, with a clear articulation of data specification and objectives;

- **Aligned incentives and minimisation of dysfunctional consequences.** There is a need to establish independent systems of reporting that prohibit manipulation by interested parties. As regards hospital star ratings, the consequences of perceived failure are so disproportionate to the quality of the data, the scope for manipulation of the data so great, and the risk of fraud being discovered adequately small, to result in a system that enjoys little public confidence. There is an urgent need to research objective and independent means of assessing health outcomes, supported by a regulatory framework that protects the integrity of the source data;

- **Preserve and nurture trust and an ethos of learning within the system.**

The perception of a ‘blame culture’ in the NHS is damaging and unproductive. Moving to a Codman like system of ‘End Results’, or the ‘near miss’ assessments used in the airline industry will better foster a culture of learning and continuous, supportive quality improvement;

- **Measures that are appropriate for their intended application and audience.**

The assumption that the costs of disseminating information are trivial compared to data collection costs should be challenged. Unless careful thought is paid to how best to disseminate information to meet the specific needs of each target audience, the objectives of the exercise are unlikely to be realised;

- **Ensure that the process is both effective and cost-effective.**

Arguments for the production and publication of ever more information need to be tempered by consideration of the added value of this information. At what point is additional information irrelevant to decision making? Is there a danger of information overload? Rather than indiscriminate publication, policy might be better designed if alternative strategies were first piloted on a small scale.

None of the hospital outcome measures considered in this report systematically addresses each of these points to inform a policy of publication. The Department of Health's new 'developmental standards' should attempt to accommodate these issues (Department of Health, 2004b).

2. Recognise that different users have different informational needs

Increasing levels of aggregation are needed as we move from patient or carer, to health professional, manager, regional office and government. When making choices about consumption we (often effortlessly) make our own aggregation to inform our choice. As health care moves towards a consumerist ethos, we will have to become better at describing the range and probability of potential consequences of treatment if patient choice is to be promoted. Managerial, aggregated measures of outcome are not helpful in this respect. Managers at various levels need appropriately aggregated data on performance. In the past, these data appear to have attracted a halo of irrefutability once published and a far more intelligent approach to aggregate data is required, understanding that the quality of care is only one source of variance.

3. Work with each target group to develop valid quality indicators and determine their use, rewards and sanctions

This concept is fully concordant with the principles of clinical governance. Informed by existing research, consultation, development, feedback and piloting are essential to promote trust and get the users to work together towards shared goals.

If quality indicators are to be used to inform performance-related rewards and sanctions it is important to determine which values these will be based upon and explore positive and negative consequences, particularly in the context of other incentives in the system, such as those contained in the new Consultant and General Medical Services contracts for NHS doctors. Rewards and sanctions could be based on market share or professional minimum standards and either might be expected to interact

differently with publicly and privately funded health care systems with different remuneration systems for clinicians. The timing of assessments needs careful consideration so that the potential for dysfunctional responses is minimised.

4. Understand users' modes of access to information

Assuming we can develop valid information to inform patient choice, measures will need to be taken to ensure that not just the articulate middle classes, but all strata of society are empowered and that the interests of vulnerable groups are protected (Fairfield et al., 1997). The process can be informed by previous research. For example, patients prefer detailed, locally relevant information including involvement in designing output formats, low levels of aggregation of data and access via a trusted intermediary or agent such as the GP, an information officer at PCT level or patient groups (Magee et al., 2003). Innovative approaches, such as the women's magazines produced by the Department of Health and Dr Foster (Department of Health, 2003a; 2004a), should be evaluated and adapted.

5. Resist the temptation to over-simplify

Initiatives such as the UK hospital star ratings are difficult to interpret because there are so many possible reasons for a good or bad rating, not all of which are to do with the quality of care delivered. The very nature of a profession means that there is skill and expertise held by professionals which cannot be encapsulated by simple rules and regulations – that is, their tacit knowledge (Davies and Lampel, 1998). No single approach to performance management is likely to be supreme (Davies and Lampel, 1998). Describing the product of a hospital is complex and a reductionist approach, such as that encapsulated by a star rating, can easily send discouraging, perverse and counterproductive signals to staff and simplistic messages to the public.

ACKNOWLEDGEMENTS

88 We are grateful to Jon Sussex and to three anonymous referees for pertinent and thought-provoking comments. We would also like to thank Suzanne Wait, Ellen Nolte, Philippe Ulmann and conference delegates for their helpful discussions of an early draft of this monograph at the meeting of the European Health Policy Group in Paris, April 2004.

REFERENCES

- Allison, R. (2001) "Dirty hospitals to be named." *The Guardian*, 10 April.
- American Hospital Association, Association of American Medical Colleges and Federation of American Hospitals (2003) "The Quality Initiative: a public resource on hospital performance." Call for Action on Collaborative Effort.
- Anonymous (2001a) "Good Hospital Guide - part 1." *The Sunday Times*, 14 January.
- Anonymous (2001b) "Good Hospital Guide - part 2." *The Sunday Times*, 21 January.
- Anonymous (2001c) "Health league tables 'no use to patients'." *The Daily Mail*, 6 February.
- Arlidge, J. (1994) "Death twice as likely in some hospitals." *The Independent*, 17 December.
- Aylin, P., Alves, B., Cook, A., Bennett, J., Bottle, A., Best, N., Catena, B. and Elliott, P. (1999) "Analysis of Hospital Episode Statistics for the Bristol Royal Infirmary Inquiry." Imperial College School of Medicine, London.
http://www.bristolinquiry.org.uk/final_report/annex_b/images/HospEpisodeStat_Aylin.pdf
- Baker, D. W., Einstadter, D., Thomas, C. L., Husak, S. S., Gordon, N. H. and Cebul, R. D. (2002) "Mortality trends during a program that publicly reported hospital performance." *Medical Care*, 40, 10: 879-90.
- Baker, D. W., Einstadter, D., Thomas, C., Husak, S., Gordon, N. H. and Cebul, R. D. (2003) "The effect of publicly reporting hospital performance on market share and risk-adjusted mortality at high-mortality hospitals." *Medical Care*, 41, 6: 729-40.
- Barr, J. K., Boni, C. E., Kochurka, K. A., Nolan, P., Petrillo, M., Sofaer, S. and Waters, W. (2002) "Public reporting of hospital patient satisfaction: the Rhode Island experience." *Health Care Financing Review*, 23, 4: 51-70.

- 90 Beecham, L. (1998) "GMC sets up steering group on revalidation." *British Medical Journal*, 317: 1406.
- Beecham, L. (1999) "UK doctors' revalidation should be linked to registration." *British Medical Journal*, 1999: 349.
- Beecham, L. (2000) "Consultants attack 'arrogant' GMC." *British Medical Journal*, 320: 1557.
- Bentley, J. M. and Nash, D. B. (1998) "How Pennsylvania hospitals have responded to publicly released reports on coronary artery bypass graft surgery." *Joint Commission Journal on Quality Improvement*, 24, 1: 40-9.
- Berwick, D. M. and Wald, D. L. (1990) "Hospital leaders' opinions of the HCFA mortality data." [Erratum appears in JAMA 1990 Jun 27;263(24):3261]. *Journal of the American Medical Association*, 263, 2: 247-9.
- Bishop, W. J. and Goldie, S. (1962) *A Bio-bibliography of Florence Nightingale*, Dawsons of Pall Mall, London.
- Bost, J. E. (2001) "Managed care organizations publicly reporting three years of HEDIS measures." *Managed Care Interface*, 14, 9: 50-4.
- Brazier, M., Lovecy, J., Moran, M. and Potton, M. (1993) "Falling from a tightrope: doctors and lawyers between the market and the state." *Political Studies*, XLI: 197-213.
- Bridgewater, B., Grayson, A. D., Jackson, M., Brooks, N., Grotte, G. J., Keenan, D. J., Millner, R., Fabri, B. M., Jones, M. and North West Quality Improvement Programme in Cardiac Interventions (2003) "Surgeon specific mortality in adult cardiac surgery: comparison between crude and risk stratified data." *British Medical Journal*, 327, 7405: 13-7.
- British Medical Association (2000) "Clinical Indicators (League Tables) - discussion document", British Medical Association, London.
- Broder, M. S., Simon, L. P. and Brook, R. H. (2004) "Surgical quality: review of Californian measures." *British Medical Journal*, 328, 7432: 152-3.

- Brown, C. and Hall, C. (1994) "Hospital 'death rate tables' split Cabinet." *Independent*, 24 November.
- Buck, N., Devlin, H. B., Lunn, J. N. and Vickers, M. D. (1987) "The Report of a confidential enquiry into perioperative deaths", *King Edward's Hospital Fund for London*, London.
- Burack, J. H., Impellizzeri, P., Homel, P. and Cunningham, J. N., Jr. (1999) "Public reporting of surgical mortality: a survey of New York State cardiothoracic surgeons." *Annals of Thoracic Surgery*, 68, 4: 1195-200; discussion 1201-2.
- Burgess, S., Gossage, D. and Propper, C. (2003) "Explaining differences in hospital performance: Does the answer lie in the labour market?" University of Bristol, Bristol.
- Carlisle, D. (2004) "How the Government broke its Bristol Inquiry pledge." *Health Service Journal*, 114, 5930: 12-13.
- Centers for Medicare and Medicaid Services (2004) "CMS announces guidelines for reporting hospital quality data." *Press Release*, 202-690-6145.
- Chase, D. and Heinen, L. (1994) "Performance reporting of health care delivery systems: will it make the grade?" *Journal of Ambulatory Care Management*, 17, 4: 15-28.
- Chassin, M. R. (2002) "Achieving and sustaining improved quality: lessons from New York State and cardiac surgery." *Health Affairs*, 21, 4: 40-51.
- Chassin, M. R., Hannan, E. H. and De Buono, B. A. (1996) "Benefits and hazards of reporting medical outcomes publicly." *New England Journal of Medicine*, 334: 394-8.
- Clinical Outcomes Working Group (2002) "Clinical outcome indicators: May 2002", *Clinical Resource and Audit Group*, Edinburgh.
<http://www.show.scot.nhs.uk/indicators/Outcomes/OutcomesReport2002.pdf>

- 92 Cook, E. (1913) "The Life of Florence Nightingale", *Macmillan and Co, London*.
- Davidge, M. and Harley, M. (1992) "Purchasing by numbers." *Health Direct, November: 6-7*.
- Davies, H. T. (1998) "Performance management using health outcomes: in search of instrumentality." *Journal of Evaluation in Clinical Practice, 4, 4: 359-62*.
- Davies, H. T. (1999) "Falling public trust in health services: implications for accountability." *Journal of Health Services and Research Policy, 4, 4: 193-4*.
- Davies, H. T. (2001) "Public release of performance data and quality improvement: internal responses to external data by US health care providers." *Quality in Health Care, 10, 2: 104-10*.
- Davies, H. T. and Lampel, J. (1998) "Trust in performance indicators?" *Quality in Health Care, 7, 3: 159-62*.
- Davies, H. T. and Shields, A. V. (1999) "Public trust and accountability for clinical performance: lessons from the national press reportage of the Bristol hearing." *Journal of Evaluation in Clinical Practice, 5, 3: 335-42*.
- Dawson, R. (1998) "Benefits of openness and teamwork must be emphasised." *British Medical Journal, 317: 812*.
- Day, M. (2004) "BUPA to be first medical provider to show success rates of surgeons." *The Daily Telegraph, 18 April*.
- Day, P. and Klein, R. (1983) "Two views on the Griffiths Report. The mobilisation of consent versus the management of conflict: decoding the Griffiths Report." *British Medical Journal Clinical Research Ed, 287, 6407: 1813-6*.
- Department of Health (1989) "Health service indicators guidance: dictionary", HMSO, London.
- Department of Health (1991) "Patient's Charter: raising the standard." HMSO, London.

- Department of Health (1995) "Baroness Cumberlege launches the new patient's charter in eleven languages." *Press Release; 1995/227*.
- Department of Health (1998) "Report on new charter for the NHS launched: performance tables published for last time." *Press Release; 1998/0581*.
- Department of Health (1999) "New figures give picture of clinical care in the NHS." *Press Release; 1999/0355*.
- Department of Health (2000a) "The NHS Plan: a plan for investment, a plan for reform", The Stationery Office, London.
- Department of Health (2000b) "An organisation with a memory", The Stationery Office, London.
- Department of Health (2001a) "NHS performance ratings acute trusts 2000 - 2001", Department of Health, London. <http://www.dh.gov.uk/assetRoot/04/07/76/00/04077600.pdf>
- Department of Health (2001b) "Your guide to the NHS", Department of Health, London.
- Department of Health (2002a) "Government response to Bristol Royal Infirmary Report." *Press Release; 2002/0030*.
- Department of Health (2002b) "Reforming NHS financial flows: introducing payment by results", Department of Health, London. <http://www.dh.gov.uk/assetRoot/04/01/87/04/04018704.pdf>
<http://www.dh.gov.uk/assetRoot/04/01/87/05/04018705.pdf>
- Department of Health (2003a) "New local maternity magazine for all Mums-to-be." *Press Release; 2003/0527*.
- Department of Health (2003b) "Star ratings system for hospital performance has improved services for patients." *Press Release; 2003/0138*.
- Department of Health (2004a) "It's Your Life: NHS launches women's health magazine." *Press Release; 2004/0021*.
- Department of Health (2004b) "National standards, local action: health and social care standards and planning framework

- 94 2005/06–2007/08." Department of Health, London. <http://www.dh.gov.uk/assetRoot/04/08/60/58/04086058.pdf>
- Dickson, N. (1995) "League tables: use for patients." *Quality in Health Care*, 4, 1: 1.
- Dobson, R. (1999) "Survival alone is not adequate measure of success, Bristol inquiry told." *British Medical Journal*, 318: 832.
- Dobson, R. (2004) "Staff may be disciplined over "fiddling" waiting lists." *British Medical Journal*, 328, 7450: 1220.
- Dranove, D., Kessler, D., McClellan, M. and Satterthwaite, M. (2003) "Is more information better? The effects of "report cards" on health care providers." *Journal of Political Economy*, 111, 3: 555-88.
- Dunn, P. (1998) "The Wisheart affair: paediatric cardiological services in Bristol, 1990-5." *British Medical Journal*, 317: 1144-5.
- Dyer, C. (1998) "Bristol doctors found guilty of serious professional misconduct." *British Medical Journal*, 316: 1924.
- Dyer, C. (2000a) "Gynaecologist struck off the medical register." *British Medical Journal*, 321: 258.
- Dyer, C. (2000b) "Surgeon sacked for 'misconduct and misjudgment'." *British Medical Journal*, 321: 321.
- Dyer, O. (2003) "Heart surgeons are to be rated according to bypass surgery success." *British Medical Journal*, 326, 7398: 1053.
- Dziuban, S. W., Jr., McIllduff, J. B., Miller, S. J. and Dal Col, R. H. (1994) "How a New York cardiac surgery program uses outcomes data." *Annals of Thoracic Surgery*, 58, 6: 1871-6.
- Editorial (1987) "Accounting for perioperative deaths." *The Lancet*, 8572: 1369-71.
- Egan, R. (1998) "Concept of collective responsibility is important." *British Medical Journal*, 317: 813.
- Epstein, A. M. (1998) "Rolling down the runway: the challenges ahead for quality report cards." *Journal of the American Medical*

- Association*, 279, 21: 1691-6.
- Epstein, A. M. (2000) "Public release of performance data: a progress report from the front." *Journal of the American Medical Association*, 283, 14: 1884-6.
- Fairfield, G., Hunter, D. J., Mechanic, D. and Rosleff, F. (1997) "Implications of managed care for health systems, clinicians, and patients." *British Medical Journal*, 314, 7098: 1895-8.
- Farley, D. O., Haims, M. C., Keyser, D. J., Olmsted, S. S., Curry, S. V. and Sorbero, M. (2003) "Regional Health Quality Improvement Coalitions: lessons across the life cycle." *RAND Corporation*, <http://www.rand.org/publications/TR/TR102/TR102.pdf>
- Fenn, P., Diacon, S., Gray, A., Hodges, R. and Rickman, N. (2000) "Current cost of medical negligence in NHS hospitals: analysis of claims data." *British Medical Journal*, 320: 1567-71.
- Ferguson, B. and McGuire, A. (1984) "A short history and review of the performance indicators issued by the DHSS." Departments of Community Medicine and Political Economy, University of Aberdeen, Aberdeen.
- Fink, A., Yano, E. M. and Brook, R. H. (1989) "The condition of the literature on differences in hospital mortality." *Medical Care*, 27, 4: 315-36.
- Forsyth, G. (1973) "Doctors and state medicine: a study of the British Health Service", *Pitman Medical, London*.
- Fottler, M. D., Slovensky, D. J. and Rogers, S. J. (1987) "Public release of hospital specific death rates. Guidelines for health care executives." *Hospital & Health Services Administration*, 32, 3: 343-56.
- Fottler, M. D., Slovensky, D. J. and Rogers, S. J. (1988) "HCFA release of hospital specific death rates: where are we and where are we going?" *Journal of the American Medical Record Association*, 59, 9: 25-9.
- Freedom of Information Act 2000. The Stationery Office, London. <http://www.legislation.hmso.gov.uk/acts/acts2000/20000036.htm>

- Galvin, R. S. and McGlynn, E. A. (2003) "Using performance measurement to drive improvement: a road map for change." *Medical Care*, 41, 1 Suppl: 148-60.
- Goddard, M., Mannion, R. and Smith, P. C. (1999) "Assessing the performance of NHS hospital trusts: the role of 'hard' and 'soft' information." *Health Policy*, 48: 119-134.
- Goodhart, C. A. E. (1984) "Monetary Theory and Practice: the UK experience", Macmillan Press Ltd, London.
- Green, J., Passman, L. J. and Wintfeld, N. (1991) "Analyzing hospital mortality. The consequences of diversity in patient mix." *Journal of the American Medical Association*, 265, 14: 1849-1853.
- Green, J. and Wintfeld, N. (1993) "How accurate are hospital discharge data for evaluating effectiveness of care?" *Medical Care*, 31, 8: 719-31.
- Green, J. and Wintfeld, N. (1995) "Report cards on cardiac surgeons. Assessing New York State's approach." *New England Journal of Medicine*, 332, 18: 1229-32.
- Gulland, A. (2003) "NHS staff cheat to hit government targets, MPs say." *British Medical Journal*, 327, 7408: 179.
- Hannan, E. L., Siu, A. L., Kumar, D., Racz, M., Pryor, D. B. and Chassin, M. R. (1997) "Assessment of coronary artery bypass graft surgery performance in New York. Is there a bias against taking high-risk patients?" *Medical Care*, 35, 1: 49-56.
- Harry, N. W. (1995) "Report was intended to promote discussion." *British Medical Journal*, 310: 599.
- Hawkes, N. (2004) "Surgeons are ordered to reveal heart operation failure rate." *The Times*, 4 November.
- Hibbard, J. H. and Jewett, J. J. (1996) "What type of quality information do consumers want in a health care report card?" *Medical Care Research and Review*, 53, 1: 28-47.
- Hibbard, J. H. and Jewett, J. J. (1997) "Will quality report cards help

- consumers?" *Health Affairs*, 16, 3: 218-28.
- Iezzoni, L. I. (1996) "100 apples divided by 15 red herrings: a cautionary tale from the mid-19th century on comparing hospital mortality rates." *Annals of Internal Medicine*, 124, 12: 1079-85.
- Irvine, D. (1997a) "The performance of doctors. II: Maintaining good practice, protecting patients from poor performance." *British Medical Journal*, 314, 7094: 1613-5.
- Irvine, D. (1997b) "The performance of doctors: I: professionalism and self regulation in a changing world." *British Medical Journal*, 314: 1540-2.
- Jacobs, K. and Manzi, T. (2000) "Performance indicators and social constructivism: conflict and control in housing management." *Critical Social Policy*, 20, 1: 85-103.
- Jacobs, R. and Smith, P. C. (2004) "A descriptive analysis of general acute trust star ratings." *Discussion Paper 189. University of York, York*.
- Jarman, B., Aylin, P. and Bottle, A. (2004) "Acute stroke units and early CT scans are linked to lower in-hospital mortality rates." *British Medical Journal*, 328, 7436: 14.
- Jarman, B., Gault, S., Alves, B., Hider, A., Dolan, S., Cook, A., Hurwitz, B. and Iezzoni, L. I. (1999) "Explaining differences in English hospital death rates using routinely collected data." *British Medical Journal*, 318, 7197: 1515-20.
- Jones, M. A. (1994) "Medical anxieties over the league tables for hospitals." *The Times*, 30 November.
- Jury, L. (1996) "Cardiac unit facing inquiry." *The Independent*, 3 December.
- Kaplan, R. S. and Norton, D. P. (1992) "The balanced scorecard - measures that drive performance." *Harvard Business Review*, 70, 1: 71-9.
- Kendrick, S., Cline, D. and Finlayson, A. (1999) "Clinical outcomes indicators in Scotland: lessons and prospects." In Davies, H. T. O.,

- 98 Tavakoli, M., Malek, M. and Neilson, A. R. (Eds.), *Managing Quality: Strategic Issues in Health Care Management*, Ashgate Publishing Ltd, Aldershot.
- Keogh, B. E., Dussek, J., Watson, D., Magee, P. and Wheatley, D. (1998) "Public confidence and cardiac surgical outcome. Cardiac surgery: the fall guy in medical quality assurance." *British Medical Journal*, 316, 7147: 1759-60.
- Keogh, B. E. and Kinsman, R. (2002) "National Adult Cardiac Surgical Database Report 2000 - 2001." Society of Cardiothoracic Surgeons of Great Britain and Ireland, London. <http://www.scts.org/index.cfm?ukcardiacreg=yes>
- Keogh, B., Spiegelhalter, D., Bailey, A., Roxburgh, J., Magee, P. and Hilton, C. (2004) "The legacy of Bristol: public disclosure of individual surgeons' results." *British Medical Journal*, 329, 7463: 450-4.
- Klein, R. (1980) "Models of man and models of policy: reflections of "Exit, Voice, and Loyalty" ten years later." *Milbank Memorial Fund Quarterly Health & Society*, 58, 3: 416-29.
- Klein, R. (1998) "Competence, professional self regulation, and the public interest." *British Medical Journal*, 316: 1740-1742.
- Kmietowicz, Z. (2003) "Star rating system fails to reduce variation." *British Medical Journal*, 327, 7408: 184.
- Langford, I. H. (1997) "Bayesian analysis should be used instead of league tables of performance." *British Medical Journal*, 314: 72.
- Laurance, J. (1993a) "Hospitals explain surgery death rate." *The Times*, 28 April.
- Laurance, J. (1993b) "NHS chief defends use of league tables." *The Times*, 4 May.
- Localio, A. R., Hamory, B. H., Fisher, A. C. and TenHave, T. R. (1997) "The public release of hospital and physician mortality data in Pennsylvania. A case study." *Medical Care*, 35, 3: 272-86.

- 99 Luce, J. M., Thiel, G. D., Holland, M. R., Swig, L., Currin, S. A. and Luft, H. S. (1996) "Use of risk-adjusted outcome data for quality improvement by public hospitals." *Western Journal of Medicine*, 164, 5: 410-4.
- Lunn, J. N. and Devlin, H. B. (1987) "Lessons from the confidential enquiry into perioperative deaths in three NHS regions." *The Lancet*, 8572: 1384-6.
- Magee, H., Davis, L. J. and Coulter, A. (2003) "Public views on healthcare performance indicators and patient choice." *Journal of the Royal Society of Medicine*, 96, 7: 338-42.
- Mannion, R. and Davies, H. T. (2002) "Reporting health care performance: learning from the past, prospects for the future." *Journal of Evaluation in Clinical Practice*, 8, 2: 215-28.
- Mannion, R. and Goddard, M. (2001) "Impact of published clinical outcomes data: case study in NHS hospital trusts." *British Medical Journal*, 323, 7307: 260-3.
- Mannion, R. and Goddard, M. (2003) "Public disclosure of comparative clinical performance data: lessons from the Scottish experience." *Journal of Evaluation in Clinical Practice*, 9, 2: 277-286.
- Marshall, M. N. and Brook, R. H. (2002) "Public reporting of comparative information about quality of healthcare." *Medical Journal of Australia*, 176, 5: 205-206.
- Marshall, M. and Davies, H. (2001) "Public release of information on quality of care: how are health services and the public expected to respond?" *Journal of Health Services Research and Policy*, 6, 3: 158-62.
- Marshall, M. N., Shekelle, P. G., Davies, H. T. and Smith, P. C. (2003) "Public reporting on quality in the United States and the United Kingdom." *Health Affairs*, 22, 3: 134-48.
- Marshall, M. N., Shekelle, P. G., Leatherman, S. and Brook, R. H. (2000a) "Public disclosure of performance data: learning from the US experience." *Quality in Health Care*, 9, 1: 53-7.

Marshall, M. N., Shekelle, P. G., Leatherman, S. and Brook, R. H. (2000b) "The public release of performance data: what do we expect to gain? A review of the evidence." *Journal of the American Medical Association*, 283, 14: 1866-74.

Maynard, A. and Bloor, K. (2003) "Trust and performance management in the medical marketplace." *Journal of the Royal Society of Medicine*, 96, 11: 532-9.

Mayor, S. (2003) "Hospitals take short term measures to meet targets." *British Medical Journal*, 326, 7398: 1054.

McCormick, D., Himmelstein, D. U., Woolhandler, S., Wolfe, S. M. and Bor, D. H. (2002) "Relationship between low quality-of-care scores and HMOs' subsequent public disclosure of quality-of-care scores." *Journal of the American Medical Association*, 288, 12: 1484-90.

McKee, M. (2004) "Not everything that counts can be counted; not everything that can be counted counts." *British Medical Journal*, 328, 7432: 17.

McKee, M. and Hunter, D. (1995) "Mortality league tables: do they inform or mislead?" *Quality in Health Care*, 4, 1: 5-12.

McKee, M. and James, P. (1997) "Using routine data to evaluate quality of care in British hospitals." *Medical Care*, 35, 10 Suppl: OS102-11.

McLellan, A. (2003) "Milburn secured three stars for PM's trust." *Health Service Journal*, London, 18th December.

Milledge, J. (1993) "Deaths after surgery: doubts on value of a 'league table'." *The Times*, 11 September.

Mukamel, D. B. and Mushlin, A. I. (1998) "Quality of care information makes a difference: an analysis of market share and price changes after publication of the New York State Cardiac Surgery Mortality Reports." *Medical Care*, 36, 7: 945-54.

National Audit Office (1988) "*Quality of clinical care in National Health Service hospitals: report by the Comptroller and Auditor General*", HMSO, London.

National Committee for Quality Assurance (2003) "The State of Health Care Quality: 2003." NCQA, Washington. <http://www.ncqa.org/Communications/State%20Qof%20Managed%20Care/SOHCREPORT2003.pdf>

National Health Service Reform and Health Care Professions Act 2002. *The Stationery Office*, London. <http://www.hmso.gov.uk/acts/acts2002/20020017.htm>

New Jersey Department of Health and Senior Services (2003) "Cardiac surgery in New Jersey, 2000. Technical report", New Jersey Department of Health and Senior Services. <http://www.state.nj.us/health/hcsal/cardtechrpt00.pdf>

New York State Department of Health (1993) "*Coronary artery bypass surgery in New York State, 1990-92*." New York State Department, Albany.

New York State Department of Health (1997) "Angioplasty in New York State 1995." New York State Department of Health, Albany. <http://www.health.state.ny.us/nysdoh/reports/angio95.pdf>

NHS Centre for Reviews and Dissemination (1999) "Effective health care: getting evidence into practice." *Royal Society of Medicine Press*, London. <http://www.york.ac.uk/inst/crd/ehc51.pdf>

NHS Executive (1992) "Health Service Indicators handbook 1990/91", *NHS Executive*, Leeds.

NHS Executive (1994) "NHS performance guide 1993-94", *NHS Executive*, Leeds.

NHS Executive (1997) "The new NHS: modern, dependable." *NHS Executive*, Leeds.

NHS Executive (1998a) "A first class service: quality in the new NHS." *NHS Executive*, Leeds.

NHS Executive (1998b) "The new NHS modern and dependable: a national framework for assessing performance - consultation document." *NHS Executive*, Leeds.

- NHS Executive (1999a) "Quality and performance in the NHS: clinical indicators." NHS Executive, Leeds.
- NHS Executive (1999b) "Supporting doctors, protecting patients", *NHS Executive, London*.
- NHS Management Executive (1991) "Patient's Charter." *HMSO, London*.
- Nightingale, F. (1860) "Hospital statistics. Appendix: proposal for a uniform plan of hospital statistics." *Fourth Session of the International Statistical Congress, 63-71*
- Nightingale, F. (1863) "Notes on Hospitals", *Longman, Green, Longman, Roberts and Green, London*.
- O'Neill, O. (2002) "A Question of Trust." *Cambridge University Press, Cambridge*.
- Pennsylvania Health Care Cost Containment Council (2004) "*Pennsylvania's guide to coronary artery bypass graft surgery 2002: information about hospitals and cardiothoracic surgeons.*" *PHC4, Harrisburg*. <http://www.phc4.org/reports/cabg/02/CABG2002.pdf>
- Platt, D. (1993) "Deaths after surgery: doubts on value of a 'league table'." *The Times, 11 September*.
- Poloniecki, J. (1998) "Half of all doctors are below average." *British Medical Journal, 316: 1734-6*.
- Power, M. (1997) "The Audit Society: rituals of verification", *Oxford University Press, Oxford*.
- Rainwater, J. A., Romano, P. S. and Antonius, D. M. (1998) "The California Hospital Outcomes Project: how useful is California's report card for quality improvement?" *Joint Commission Journal on Quality Improvement, 24, 1: 31-9*.
- Ransley, P. (2000) "In need of urgent attention." *The Guardian, 8 August*.
- Reid, J. (2004) "*Speech by Rt Hon John Reid MP, Secretary of State for Health, 10th February 2004: Chief Executives Conference.*"

- Roach, J. (2000) "Management blamed over consultant's malpractice." *British Medical Journal, 320: 1557*.
- Rosenthal, G. E., Hammar, P. J., Way, L. E., Shipley, S. A., Doner, D., Wojtala, B., Miller, J. and Harper, D. L. (1998) "Using hospital performance data in quality improvement: the Cleveland Health Quality Choice experience." *Joint Commission Journal on Quality Improvement, 24, 7: 347-60*.
- Rowan, K., Harrison, D., Brady, A. and Black, N. (2004) "Hospitals' star ratings and clinical outcomes: ecological study." *British Medical Journal, 328, 7445: 924-5*.
- Schneider, E. C. and Epstein, A. M. (1996) "Influence of cardiac-surgery performance reports on referral practices and access to care. A survey of cardiovascular specialists." *New England Journal of Medicine, 335, 4: 251-6*.
- Schneider, E. C. and Epstein, A. M. (1998) "Use of public performance reports: a survey of patients undergoing cardiac surgery." *Journal of the American Medical Association, 279, 20: 1638-42*.
- Schneider, E. C. and Lieberman, T. (2001) "Publicly disclosed information about the quality of health care: response of the US public." *Quality in Health Care, 10: 96-103*.
- Scottish Office (1991) "*The National Health Service in Scotland: the Patient's Charter: a charter for health.*" *Scottish Office, Edinburgh*.
- Scottish Office (1995) "*Scotland getting better: the National Health Service in Scotland: annual report 1994-95.*" *Scottish Office, Edinburgh*.
- Seagroatt, V. and Goldacre, M. J. (2004) "Hospital mortality league tables: influence of place of death." *British Medical Journal, 328, 7450: 1235-6*.
- Smith, P. (1995) "On the unintended consequences of publishing performance data in the public sector." *International Journal of Public Administration, 18: 277-310*.
- Smith, R. (1989) "Profile of the GMC: Discipline I: the hordes at the gates." *British Medical Journal, 298: 1502-5*.

- Smith, R. (1998a) "All changed, changed utterly." *British Medical Journal*, 316: 1917-8.
- Smith, R. (1998b) "Regulation of doctors and the Bristol inquiry." *British Medical Journal*, 317: 1539-40.
- Snelling, I. (2003) "Do star ratings really reflect hospital performance?" *Journal of Health Organization and Management*, 17, 3: 210-223.
- Society of Cardiothoracic Surgeons (2004) "Cardiac surgery standards improve again." *Press Release*, 10th September.
- Spiegelhalter, D. J. (1999) "Surgical audit: statistical lessons from Nightingale and Codman." *Journal of the Royal Statistical Society*, 162, 1: 45-58.
- Stacey, M. (1992) "For public or profession?--The new GMC performance procedures." *British Medical Journal*, 305, 6861: 1085-7.
- Statistical Society of London (1864) "Statistics of metropolitan and provincial general hospitals for 1862." *Journal of the Statistical Society*, 27: 401-9.
- Statistical Society of London (1865) "Statistics of metropolitan and provincial general hospitals for 1863." *Journal of the Statistical Society*, 28: 527-35.
- Statistical Society of London (1866a) "Statistics of metropolitan and provincial general hospitals for 1864." *Journal of the Statistical Society*, 29: 112-21.
- Statistical Society of London (1866b) "Statistics of metropolitan and provincial general hospitals for 1865." *Journal of the Statistical Society*, 29: 596-605.
- Stewart, J. (1998) "A patient's perspective." *British Medical Journal*, 317: 1581.
- Street, A. (2002) "The resurrection of hospital mortality statistics in England." *Journal of Health Services & Research Policy*, 7, 2: 104-10.
- Taylor, R. (2004) "Mortality rates." *Health Service Journal*, 114, 5908:

- 27.
- The Labour Party (1997) "1997 election manifesto", The Labour Party, London.
- Timmins, N. (2000) "A gamble on the talisman of health." *Financial Times*, 23 March.
- Tomlin, P. J. (1998) "Society of Clinical Psychiatrists supports doctors who have been suspended." *British Medical Journal*, 317: 811-812.
- Treasure, T. (1998) "Lessons from the Bristol case." *British Medical Journal*, 316: 1685-6.
- Turner, T. (1862) "Statistics of the general hospitals of London, 1861." *Journal of the Statistical Society*, 25: 384-8.
- US Department of Health and Human Services (1998) "HHS Achievements in 1998 Build a Healthier America." *HHS Fact Sheet*, (202) 690-6343.
- US Department of Health and Human Services (2001) "The New Centers for Medicare and Medicaid Services (CMS)." *Press Release*, (202) 690-6145.
- Vicinus, M. and Nergaard, B. (Eds.) (1989) "Ever Yours, Florence Nightingale." *Virago Press, London*.
- Warden, J. (1988) "All for audit now." *British Medical Journal*, 297: 1569.
- Warden, J. (1998a) "Hospital death rates to be published for England." *British Medical Journal*, 316: 1767.
- Warden, J. (1998b) "NHS hospital doctors face compulsory audit." *British Medical Journal*, 316: 1851.
- Willis, J. (1998) "Case arose through a failure of action, not of detection." *British Medical Journal*, 317: 811.

RECENT OHE PUBLICATIONS

106

The Economics of Dental Care

By Ray Robinson, Darshan Patel and Rowena Pennycate, 2004

(price £12.50)

Developments in Economic Evaluation in Health Care: a Review of HEED. OHE Briefing No.40

By Clive Pritchard, 2004

(price 7.50)

Mental Health Economics and Policy in a Global Context. Seminar Briefing No.5

Edited by Jon Sussex, 2003

(price £7.50 for hard copy, downloadable free from www.ohe.org)

Reducing Harm to Patients in the National Health Service. Will the Government's Compensation Proposals Help? OHE Briefing No.39

By Adrian Towse, Paul Fenn, Alastair Gray, Neil Rickman and

Rodrigo Salinas, 2003

(price £7.50)

To Heal and Harm: an Economic View of Drug Safety

By Jonathan Silcock and Clive Pritchard, 2003

(price £12.50)

Statistical Methods for Cost Effectiveness Research: a Guide to Current Issues and Future Developments

Edited by Andrew Briggs, 2003

(price £12.50)

Quality in Primary Care: Economic Approaches to Analysing Quality-Related Physician Behaviour

By Michael Kuhn, 2003

(price £12.50)

Institutions for Industrial Competitiveness in the International Pharmaceutical Industry

Edited by Jorge Mestre-Ferrandiz and Jon Sussex, 2003

(price £12.50)

RECENT OHE PUBLICATIONS

Public Private Partnerships for Medicines and Vaccines Research and Development

By Hannah Kettler and Adrian Towse, 2002

(price £10.00)

Details of all OHE publications, and how to order them, can be found on the OHE website at www.ohe.org or by contacting:

Office of Health Economics

12 Whitehall

London SW1A 2DY

Tel: +44 (0)20 7930 9203

Fax: +44 (0)20 7747 1412

107

