

THE UNIVERSITY *of York*

CENTRE FOR HEALTH ECONOMICS  
YORK HEALTH ECONOMICS CONSORTIUM  
NHS CENTRE FOR REVIEWS & DISSEMINATION

THE YORK SERIES ON THE NHS WHITE PAPER -  
A RESEARCH AGENDA

## **Reference Costs and the Pursuit of Efficiency in the ‘New’ NHS**

*Diane Dawson  
Andrew Street*

***DISCUSSION PAPER 161***



# **REFERENCE COSTS AND THE PURSUIT OF EFFICIENCY IN THE "NEW" NHS**

*Diane Dawson*

*Andrew Street*

**October 1998**

## **CENTRE FOR HEALTH ECONOMICS DISCUSSION PAPER SERIES**

The Discussion Paper (DP) Series was established in 1984 to make research available quickly to a broad readership drawn from the academic community, professional health care staff and policy makers, both within the UK and internationally. Material published in DP form ranges from work in progress, to edited versions of completed project reports, from papers intended ultimately for journal publication, to those which seek to stimulate immediate debate. Hence the variation in length and complexity of DPs.

In the past, DPs have been subject to a minimum of editorial revision, it being principally seen as the responsibility of individual authors to ensure that appropriate standards were met. Since June 1996, a more formal system of internal peer-review has been introduced to consolidate the high levels of quality that have been achieved in the past. That system involves the refereeing of DPs in draft form, with any recommended revisions being verified independently by a senior member of staff.

DPs are not a substitute for publication in peer-reviewed journals. It is to be hoped that journal editors will recognise the nature of the DP series, and accept that this does not compromise the status of those papers published by the Centre which may later be submitted for consideration in more formal publications. Comments and suggestions on DPs encouraged and should be sent directly to the authors.

## **THE AUTHORS**

Diane Dawson is Senior Research Fellow at the Centre for Health Economics. Andrew Street is Senior Research Fellow at the York Health Economics Consortium.

## **ACKNOWLEDGEMENTS**

The authors would like to thank Michael Place for helpful comments on a previous draft. We are grateful for continuing financial support from the Department of Health. The views expressed in this paper reflect the views of the authors and not necessarily those of the Department of Health.

## **FURTHER COPIES**

Further copies of this document are available (at price £5.00 to cover the cost of publication, postage and packing) from:

The Publications Office  
Centre for Health Economics  
University of York  
YORK YO10 5DD

Please make cheques payable to the University of York. Details of other papers can be obtained from the same address or telephone York (01904) 433648 or 433718.

**ABSTRACT**

Both the White Paper, *The New NHS*, and the later consultation document, *A National Framework for Assessing Performance*, stress the need to develop new instruments to tackle inefficiency in the NHS. Among these instruments it has been proposed to benchmark Trusts, and by association Health Authorities, using Healthcare Resource Group (HRG) costs. The NHS Executive has published plans for a system of 'Reference Costs' that will itemise the cost of every treatment in every Trust. These reference costs will be derived from costing HRGs and are to be used for many purposes: benchmarking cost improvement, measuring relative efficiency, identifying best practice, funding transfers and costing health improvement programmes. This paper examines the construction of reference costs, considers incentives to use the information appropriately and asks whether a single accounting construct, the costed HRG, can be expected to contribute successfully to its many intended functions of regulation and management.

## CONTENTS

<b>I Introduction</b>	<b>1</b>
<b>II The Role of Costing Exercises</b>	<b>3</b>
Remuneration	3
Purchasing	4
Internal management and central monitoring	5
<b>III Cost and Output Measurement in the UK</b>	<b>7</b>
Hospital activity: finished consultant episodes	7
Casemix: Healthcare Resource Groups	8
Costs and cost allocation guidelines	12
<b>IV Climbing the League Table</b>	<b>15</b>
Improving the quality of data used for assignment and costing	15
Increasing throughput	16
<b>V Using Reference Costs</b>	<b>19</b>
Making comparisons	20
Appropriate for intended use	24
Incentives to tackle inefficiency	26
<b>VI Conclusions</b>	<b>29</b>

## I INTRODUCTION

Both the White Paper, *The New NHS* (NHSE 1997), and the consultation paper on performance (NHSE 1998a) stress the importance of finding better measures of performance for the Health Service than the long criticised Purchaser Efficiency Index:

The Purchaser Efficiency Index simply failed to reflect the breadth of what is important in the NHS and created perverse incentives which ran counter to the real priorities for the health service. The balance needs to be shifted so that due weight is given to the things that really matter to patients and the public - the cost and quality of care the NHS delivers and the benefit patients get from their treatment (NHSE 1998a p3).

A range of performance measures has been proposed, most of which will be used to assess Health Authorities (NHSE 1998a). Among these there is to be a set of Reference Costs intended to summarise the efficiency of the hospital sector and, when aggregated over all the providers used by a Health Authority, to assess the efficiency of the Health Authority itself. Reference costs were heralded in the White Paper (NHSE 1997):

the Government will develop a national schedule of 'reference costs' which will itemise what individual treatments across the NHS cost. By requiring NHS Trusts to publish and benchmark their own costs on the same basis, the new arrangements will give Health Authorities, Primary Care Groups and the NHS Executive a strong lever with which to tackle inefficiency (NHSE 1997 p19).

These reference costs will be based on costed Healthcare Resource Groups (HRGs) and will be published in three formats (NHSE1998b):

1. A comprehensive list of the costs of every treatment offered by every NHS Trust in England.
2. The **National Schedule of Reference Costs** (NSRC): which will provide summary statistics (mean, range, inter-quartile range) from all Trusts for each HRG, probably by method of admission.
3. The **National Reference Cost Index** (NRCI): which will report the weighted average of all HRG costs in a Trust relative to the national average.

If efficiency measures based on costed HRGs are to be central to the new system, we need to be clear precisely what aspects of health care efficiency are better understood by availability of this costing information. As information itself does not lead directly to change, we need to examine incentives to use information on HRG costs and whether attempts to use this information to regulate behaviour in the NHS are likely to lead to perverse, beneficial or neutral outcomes.

The paper is organised as follows. In Section II we examine the main purposes for which comparative cost information is generally used in health care systems. In Section III we set out the principles behind the proposed technique for determining the cost of individual treatments described using HRGs. Section IV considers the main strategies Trusts might pursue to improve their league table position while Section V assesses the intended uses of the schedule. In Section VI we offer our conclusions and suggestions as to how the system might be developed.



## II THE ROLE OF COSTING EXERCISES

Information on the comparative cost of different health care providers is used in a variety of ways by governments of different countries. Three applications are particularly common:

- remuneration;
- purchasing; and
- internal management or central monitoring.

The use of comparative cost information to support these activities is outlined below.

### Remuneration

The most common use of comparative cost data is as part of rate setting for remuneration of providers of services. This is basically a form of price regulation but, in setting prices, some information on costs is directly taken into account or implied. Yardstick competition was originally considered as a means of regulating the prices of local monopoly utilities. The objective was to provide incentives for firms to operate efficiently and this was to be advanced by having the regulator set the prices a firm could charge by reference to the costs of the other firms in the industry. Since these prices were independent of the firm's own costs, there were incentives to reduce costs as any resulting surplus could be retained as profit.

In 1980s the US government, concerned about escalating health care costs, applied this model to the health sector when it introduced a system of prospective payment for hospital services delivered to Medicare patients. Each patient discharged was assigned a diagnosis-related group (DRG) code according to either initial diagnosis or primary procedure (Fetter *et al* 1980). A price was set for each DRG episode of care based on the average amount charged by all hospitals for the treatment concerned, with hospitals generally using financial accounting models to determine their charges. These models, the best known being the Yale Cost Model (Freeman *et al* 1986), allocate costs on the basis of DRG-specific service weights rather than by reference to actual patterns of resource use of patients within a DRG.

Techniques for setting DRG related prices to resource use rather than to charges have been developed in Australia where the data is used as part of the system of hospital remuneration. In 1993 the Australian state of Victoria introduced funding arrangements according to which public hospitals are reimbursed a casemix weighted payment for each public patient treated (Duckett 1995). Patients are classified according to Australian National DRGs (AN-DRGs), which are a local modification of the US system. Each treated patient is allocated to one of 526

AN-DRGs on the basis of their principal diagnosis, type of procedure, presence of complication and co-morbidities, age, birth weight, and discharge status, these being factors which explain use of resources in treatment or which justify the length of acute stay.

Each AN-DRG has an associated resource weight, based on information derived from patient level costing systems operating in Victorian hospitals (Jackson *et al*, 1993; Jackson *et al* 1994; Jackson *et al* 1998). These systems aggregate information from departmental cost centres which collect comprehensive patient level data about the intermediate products - the diagnostic tests, radiology, nursing care, operating theatre procedures, etc - provided during the admission. Each product has an associated cost weight, assigned according to relative value scales, most of which derive from studies undertaken in each hospital. For example, in many sites nursing costs are calculated using ward specific patient level dependency systems, which use the results of time and motion studies to detail the amount and type of nursing time required by patients during each shift. Costs are then assigned to individual patients according to the number of shifts spent at each dependency level on a particular ward. Overheads such as infection control, sterile supply and domestic services are incorporated into the cost of producing intermediate products, after departmental shares have been determined using step down methods or simultaneous equations. This approach results in a higher degree of accuracy than can be realised if costs were assigned assuming average resource use for patients grouped together in some way.

Hospitals in Victoria are now reimbursed predominately on the basis of the number and type of patients treated, although funds are also provided for activities such as teaching or to sustain rural hospitals. The AN-DRG system allows funders to vary payments to hospitals according to their levels of activity. Also because payment rates are regularly updated and determined on the basis of whichever is the efficient hospital in each particular AN-DRG, hospitals have an incentive to continually seek efficiency improvements. Further details of the payment system, including methods to control global expenditure, are provided by Duckett (1995).

## **Purchasing**

A second use of cost information is found in countries where purchasers of health care (insurance companies or health authorities) have some freedom to decide which providers to use. Here comparative cost information, if obtainable, may be used by purchasers in negotiating contracts.

Information on variations in treatment costs across providers is of primary use to a purchaser who is in a position to move virtually all of their activity from one Trust to another. While in the UK this applies to general practice fundholders, it is rarely the case that a Health Authority

is able to shift main contracts. This suggests that for the bulk of health care purchasing in the UK, comparative cost information is of limited relevance. It is of little value being told that other Trusts produce similar services at lower cost if the purchaser has no opportunity to redirect activity in order to take advantage of the situation.

If there is excess capacity, comparative cost/price information may be of more use. In the United States, purchasers (health care plans) are often in a position to negotiate charges with providers. Some regulators collate information from published price lists to help inform purchasers. Examples include the Illinois Cost Containment Council documents which publish, for each hospital in each region of the state of Illinois, the quoted price for each of the top 10 DRGs in terms of admissions.<sup>1</sup>

### **Internal Management and Central Monitoring**

A third demand for comparative cost information has come from management consultants and cost accountants who have popularised the idea of cost benchmarking as a tool for managers to use when attempting to improve the competitive position of their firm. Benchmarking has become so fashionable a term that it is used in a variety of ways to describe quite different activities. In some cases it is used simply to describe comparison itself: e.g. my defect rate compared to the industry average. In other cases the term refers to establishing what economists call an 'efficiency frontier' (Barrow and Wagstaff 1989): given data on all producers, what are the lowest feasible capital and labour inputs capable of producing particular levels of standardised outputs? One can then identify the difference between a single firm's use of inputs and the current industry best practice.

But benchmarking can be a more proactive process than comparison alone. The idea is that the management of a firm engaged in a benchmarking process will attempt to:

1. identify an indicator of industry wide 'best practice' (lowest price, lowest cost, lowest defect rate, whatever is of interest);
2. identify the organisational structure or production processes used by the firm(s) achieving best practice;
3. establish whether it would be feasible to import into their own firm the structure or production processes associated with best practice;
4. find ways to implement change.

---

<sup>1</sup> Copies of all these reports can be obtained from the Illinois Health Care Cost Containment Council web site: <http://www.state.il.us/agency/hcccc/ih4cprod.htm>

A recent survey of benchmarking activity in UK organisations characterised the range of meanings for the term benchmarking as starting with the simple 'indicator' or league table benchmarking (simple comparison) and ending with 'ideas' benchmarking (identifying and implementing best practice) (Mayle *et al* 1998). The results of the study show most UK activity is of the indicator/league table variety with relatively little concern with adoption of best practice. One explanation is that implementing change is a much more difficult and idiosyncratic exercise than identifying differences.

The government appears to stress identification and implementation of best practice and sees purchasers as having an active role in the process. It is intended that publication of comparative cost information will "give Health Authorities, Primary Care Groups and the NHS Executive a strong lever with which to tackle inefficiency" (NHSE 1997). Since 1994 Trusts have been required to undertake a rolling programme of costing activity below specialty level. However, in only two Regions (Northern and Yorkshire, and North West) have these data been collated from all Trusts and made publicly available. The difference the new system will make is that all commissioners will now know something about treatment costs throughout the country and can compare these data with the amounts they pay for particular services. The policy presumption is that this increase in the availability of cost information creates incentives for Trusts to change the way they use resources, perhaps under pressure from their purchasers. It will be easier to understand the incentives created by publication of information on costs if we first examine the data and processes used to generate the information.

### III COST AND OUTPUT MEASUREMENT IN THE UK

To be a valid comparative index, a reference costing schedule must be standardised on three bases:

- the means of describing units of activity;
- the method for classifying patients into comparable groups; and
- the approach to determining treatment costs for each group.

#### **Hospital Activity: Finished Consultant Episodes**

The NHS is unusual in its approach to measuring the activities of the hospital sector. Most countries express inpatient activity in terms of the number of patients admitted to or discharged (including deaths) from hospital, so that a unit of hospital activity is said to have been provided when the patient's stay in hospital is complete. In contrast, since 1988/89 the NHS has decomposed the hospital stay to reflect the number of consultants assuming responsibility for the patient. When the patient is discharged from the care of the consultant (or specialty) a completed unit of activity is said to have been provided. This may not correspond with the patient being discharged from hospital. The time between a consultant assuming responsibility for a patient to discharging the patient from hospital or transferring responsibility to a colleague is called a Finished Consultant Episode (FCE). Because a patient may receive care from various consultants during any particular spell in hospital, if NHS activity is measured using FCEs higher levels of activity will be apparent than if the number of discharges and deaths were used.

While this counting method makes international comparisons difficult, it also complicates comparison of hospital activity in the NHS both over time and across institutions. Comparisons over time are made difficult because of the possibility of 'episode inflation' whereby the number of recorded episodes increases at a faster rate than the number of discharges and deaths. This problem may have been further compounded by the introduction of the internal market and the use of FCEs as the currency for contracts between health authorities and hospitals. These changes have created incentives for providers to record as many episodes as possible, however short they may be. Further, the Purchaser Efficiency Index, designed to encourage health authorities to secure non-funded increases in activity from the hospitals with which they contract, expresses desirable activity increases in terms of FCEs. Hospitals could claim activity improvements by simply coding internal transfers more completely.

Cross-sectional comparisons using FCEs are compromised by the amount of local discretion over what constitutes a completed episode. The implications of using FCEs as the basis for comparative cost assessment are recognised in the documentation material accompanying the software for version 3 HRGs:

one hospital may count the admission from A&E as one FCE, the transfer to medical ward as another, and the transfer to surgery for a gastroscopy and then back to the medical ward as two further FCEs; another hospital may class all these internal transfers as one FCE. Needless to say the resource use allocated to the one FCE in the second hospital will be much higher than any of the individual FCEs in the first hospital (National Casemix Office 1997, p21).

In the consultation document on reference costs this problem is recognized and it is proposed that by the second year of publication HRG costs will be based on aggregating FCEs into 'provider spells' (NHSE 1998b). This change is likely to be welcomed but it does create aggregation problems because of the current absence of a unique patient identifier with which to link FCEs forming part of the same spell and in retaining casemix information.<sup>2</sup>

### **Casemix: Healthcare Resource Groups**

The second important condition for accurate cost comparison is that casemix complexity is adequately accounted for. It is difficult to force hospitals to accept they are relatively inefficient without first establishing that their patients are either no different from those treated elsewhere or that differences have been accounted for in determining relative performance. This implies that not only should hospital activity be described in terms of cases treated, but allowance should be made for differences in casemix complexity. If two hospitals are similar in all respects with the exception that one treats a higher proportion of patients with acute myocardial infarction and a lower proportion with varicose veins then the costs of the hospitals will differ. In order to take account of cost differences which are patient related rather than a reflection of the hospital's efficiency, some means is required of describing (groups of) patients according to their expected use of hospital resources.

The proposal to "itemise the cost of individual treatments across the NHS" (NHSE 1997) will entail defining individual treatments in terms of HRGs. HRGs are groups of inpatient FCEs which are purported to have similar health care resource requirements and efforts to encourage their use in the contracting process was first signalled in 1994 (EL(94)51). HRGs are a locally modified version of DRGs, first developed in the United States (Fetter *et al* 1980). The National

---

<sup>2</sup>

See Benton *et al* (1998) for a description of the approach being used by the National Casemix Office to aggregate FCEs into spells.

Casemix Office has been developing HRGs in the UK, and the third version was released in May 1997 (Benton *et al*, forthcoming). The main criteria for local modification and subsequent refinement of HRGs are:

- clinical meaningfulness;
- based on items in the Hospital Episode Statistics (HES) dataset;
- similar resource use;
- relatively few (not more than fifty per specialty).

The application of these criteria has been instrumental in determining the extent to which HRGs can be considered valid descriptions of casemix. Great emphasis has been placed on satisfying the first criteria, and much of the development work has been based around clinical working groups comprising clinical representatives from each relevant specialty.

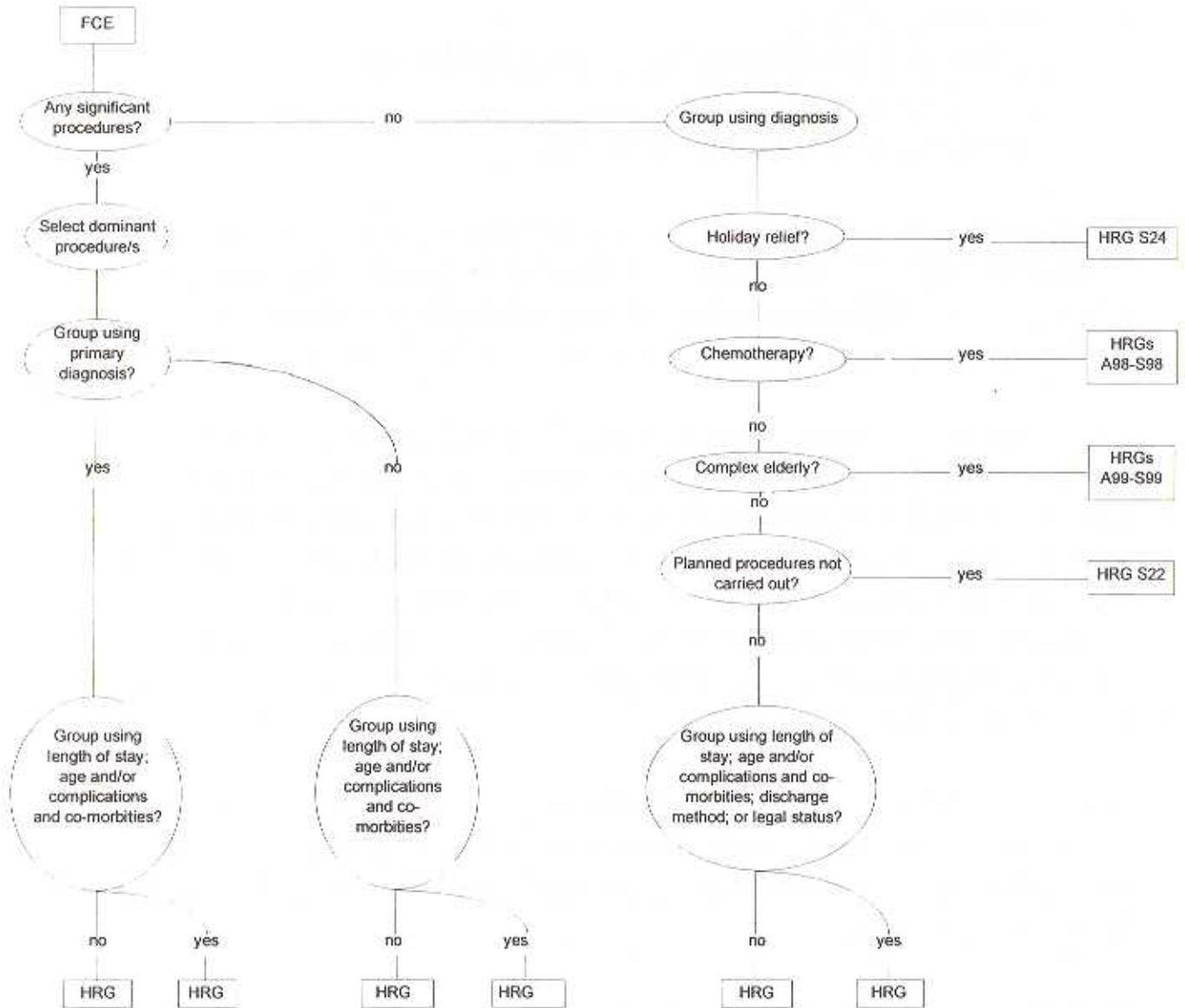
HRGs are formed from routine data about inpatients and day cases recorded in the HES dataset. The clinical working groups are asked to indicate which types of information recorded in the dataset should be used in the refinement process, the aim being to improve the amount of within-group homogeneity in the length of stay of patients classified to the same HRG. This may lead to recommendations that the presence of secondary diagnoses, indicating complications and co-morbidities, should be accounted for, or that an age split should be made. These clinical judgments have been statistically tested using 1993/94 data for every patient admission for England.

Figure 1 depicts the process by which an FCE is assigned to a HRG. Each individual record in the HES dataset is grouped to a single HRG based on the data contained in the record. Where information in any field required for grouping is missing or invalid the FCE will be assigned to one of seven 'undefined' groups.

While previous versions have differed, information used for categorisation in version 3 includes:

- primary and secondary *procedures* (coded using the Office of Population Censuses and Surveys, Version 4). It is hypothesised that the overall cost of a treatment episode is strongly influenced by whether or not surgical procedures are conducted. As such, the grouping process starts by determining whether a major surgical procedure has been

Figure 1: HRG grouping flow chart





•undertaken. This allows records to be divided into surgical and medical groupings. If more than one procedure appears (up to four procedures can be recorded), assignment is determined according to a procedure hierarchy, which assigns a rank (from 0 to 7) to all procedures according to 'clinical knowledge' and the relationship with post-operative length of stay. primary, subsidiary and secondary *diagnoses* (coded using the International Classification of Diseases, volume 10). Where only minor procedures (ranked 0 or 1) or no procedures are recorded, the primary diagnosis is used to determine assignment to HRG. The primary diagnosis may be overridden in the event of holiday relief care (HRG S24), chemotherapy (HRGs A98-S98), complex elderly cases (defined as patients over 69 with two or more major diagnoses but no significant procedures - HRGs A99-S99), and when planned procedures have not been performed (HRG S22). If a record indicates more than one of these events assignment is based on the strict hierarchy depicted in figure 1.

- in some instances, generally where minor procedures are recorded, *length of stay* is also used as a grouping variable to distinguish patients admitted for only a minor procedure from those having the procedure as part of a wider treatment course.
- *age* which separates older from younger patients, although a three way split is occasionally applied. The age groups are divided at 17, 50 or 70 years. Occasionally *gender* is used.
- secondary diagnoses are used to identify the presence of *complications and co-morbidities*, this information being invoked, often in conjunction with age splits, to subdivide 120 HRGs.
- *discharge method* (alive or dead) which is used only when grouping neonates and pulmonary embolus; and
- *legal status*, which indicates whether patients admitted to a psychiatric facility were compulsorily detained and is used to subdivide three psychiatric groupings (schizophreniform psychoses; mania; and depression).

As well as restricting the type of information available for grouping, reliance on the HES dataset has meant that attempts to reduce within group variance has focussed on explaining length of stay as a proxy for cost, this being the only information vaguely related to resource use routinely available, although occasionally it has been supplemented by other types of information and clinical judgment (Sanderson *et al* 1995; Benton *et al* forthcoming). Even the procedure hierarchy, designed to ensure that the most 'resource intensive' procedures are used for grouping, is based not on theatre time or equipment use but on the expected relationship

between the procedure and post-operative length of stay. This approach contrasts with the considerably more accurate individual resource data available to inform the development of AN-DRGs and their associated cost weights in Australia.

Use of length of stay data to determine groupings is not peculiar to the UK (the original version of US DRGs relied on the same (Fetter *et al* 1980)) but strong assumptions must be made about the extent to which length of stay and cost are correlated. Where the correlation is weak, HRGs will poorly reflect casemix complexity. For example, a number of HRGs are formed by subdividing patients on the basis of age, because older patients (generally over 69 years) tend to be hospitalised for longer than their younger counterparts for the same condition. However, their longer hospitalisation might be explained by greater difficulty in making discharge arrangements rather than their having greater care requirements. If so, the marginal resource implications of such delays may be minimal and it may not be appropriate to assume that a hospital treating a higher proportion of elderly patients in HRGs with an age split has a substantially more resource intensive casemix.

### **Costs and cost allocation guidelines**

The third requirement of a reference cost schedule is for 'costs' to reflect resource use. The NHS lags behind many other countries in terms of the routine cost data collected by health care providers. This reflects the historically loose relationship between hospital budgets and activity and the absence of a significant private insurance sector requiring a detailed itemisation of resource use for billing purposes. Even the introduction of the internal market failed to generate patient related costing information.

In contrast to public sector hospitals in countries like Australia, where recording systems make it easy to determine the type and quantity of resources used by any patient during their hospitalisation, even basic data about the number of diagnostic tests, theatre time or nursing dependency are very difficult and extremely time consuming to obtain in the NHS. Trusts rarely have automated information about the resources used by particular patients during their hospital stay. Because hospitals do not collect and record data on the use of resources by individual patients, almost all costing in the UK is undertaken on a 'top down' basis.

This top down process of allocating costs starts with collection of the annual financial returns of a hospital. These contain information on how much was spent on electricity, maintenance of buildings, catering, drugs, salaries, etc. This expenditure is then reallocated to 'patient treatment services': wards, operating theatre, pharmacy, etc. This reallocation may be direct, such as wages for the number of nurses normally staffing a particular ward; or indirect, for example operating theatres may be allocated a proportion of the cost of hospital cleaning corresponding to their share of total hospital area.

These costs, once allocated to patient treatment services, are then reallocated to specialties: paediatrics, general medicine, general surgery, etc. Again, this reallocation may be direct, as where a ward only houses patients in one specialty, or indirect, where the proportion of operating theatre time booked by a specialty is used to determine the share of operating theatre costs apportioned to that specialty.

Once each specialty has been assigned its share of total hospital expenditure, the specialty total is divided by the number of bed days occupied by patients coded to that specialty. This process is the basis for calculating what is essentially the accounting cost of an HRG.<sup>3</sup>

In moving from the allocation of hospital costs by specialty to the sub-division of specialty costs by HRG, a 'care profile' may be constructed by a working group made up of clinicians, nurses and clinical managers. The care profile describes the resources expected for a typical patient and might include estimates of items such as theatre time, consultant time, radiology, significant drugs and consumables.

In general medicine, for example, a relatively simple approach to estimating resource use is recommended. The process of moving from the allocation of hospital costs to general medicine to the sub-division of these specialty costs by HRG can be illustrated using data taken from a National Casemix Office study (Dredge 1997). The total accounting cost of general medicine is divided by occupied bed days in the specialty to derive a cost per bed day:

$$\frac{\text{Total cost of general medicine}}{\text{Occupied bed days}} = \text{£210} = \text{Average cost per bed day of a patient in general medicine}$$

The cost per bed day is then multiplied by the number of bed days attributable to the HRG of interest to derive total expenditure by HRG. This figure is then divided by the number of FCEs assigned to the HRG, to produce the HRG cost per FCE.

$$\text{Average HRG cost per FCE} = \frac{\text{HRG bed days * cost per bed day}}{\text{HRG FCEs}}$$

As table 1 illustrates, the cost per bed day is assumed not to vary across HRGs within a specialty.

<sup>3</sup> Various sources describe the process of cost allocation recommended by the NHS Steering Group on Costing. FDL(97)25 summarises the principles that are set out in some detail in the Costing for Contracting Manual (FDL(93)76). Butler *et al* (1998) provides a case study.

**Table 1 Cost allocation process**

		HRG a16	HRG d25	HRG e38
	bed days	456	850	662
*	average cost per bed day	£210	£210	£210
÷	number of FCEs	68	141	211
=	average cost per FCE	£1,408	£1,265	£659

*Source:* adapted from Dredge (1997)

This example reveals the main mechanisms by which a Trust can improve its position in the reference cost league tables. The three most immediate measures are to reduce length of stay, increase throughput, or reassess cost allocation procedures. These strategies are discussed in more detail in the next section.

## IV CLIMBING THE LEAGUE TABLE

Having consulted the published list of costed HRGs, a purchaser finds that one of its providers reports costs 25% above the lowest cost reported in the National Schedule of Reference Costs for that HRG. The management of the Health Authority or Primary Care Group will be judged partly on the basis of the weighted average of costed HRGs of its providers. Unless the purchaser can get their providers to reduce costs, their own performance looks poor. What happens next?

The purchaser is expected to put pressure on or cooperate with the Trust to reduce costs. There are basically two ways a Trust can improve its position in the league table and these two responses reflect the sources of variation in the published data. First, it can invest in more administrative support to improve its coding and may decide to change its cost allocation procedures. For example, incomplete clinical notes may mean its patients are assigned to relatively less resource intensive HRGs. Second, it can try to alter clinical behaviour and the management of clinical activity. The most direct way of reducing HRG cost is to increase throughput by reducing length of stay and/or treating more patients.

### **Improving the quality of data used for assignment and costing**

Coding to the appropriate HRG relies on the accuracy of information recorded in the HES dataset. This in turn depends on the ability of medical records staff to interpret clinical notes and the extent to which transcription of information is audited and checked. Increasing the accuracy with which support services and the cost of diagnostic procedures are assigned to patients using these services requires more information gathering and monitoring. Devoting additional resources to improving the accuracy with which patient treatment is recorded in hospital information systems and costs are allocated will increase administrative and information and technology costs.

In the past, hospitals and clinicians in the UK have had little incentive to improve the quality of their data collection, let alone report it. In countries like Australia, where reported costs can affect hospital income there is a strong incentive to invest in administrative support for improving the quality of information: medical records personnel are better paid and departments better staffed than in the UK, and Australian hospitals have made greater investments in clinical costing systems.

Of course, there is a danger that hospitals will code patients in such a way that they are assigned to more resource intensive HRGs in an effort to improve their league table position without

addressing efficiency. There is evidence to suggest hospitals in the US adopted such behaviour following the introduction of DRG funding (Ginsberg and Carter 1986). However, if all hospitals act similarly and resource weights are revised periodically, the advantages to each hospital will not persist beyond the short term. In contrast, the health system will benefit from improved information because no hospital will be able to afford to record less completely than the rivals against which it is compared. Regulatory controls can also be introduced to limit coding inaccuracies. For example, the Victorian Department of Health and Community Services undertakes a random audit of 1% of medical records to measure the accuracy of coding, provide advice on coding conventions and to reduce the extent to which casemix payments are subject to gaming (Health and Community Services 1994; Duckett 1994).

UK hospitals may be less likely to respond in the way of their US and Australian counterparts. Unless their income, rather than merely their position in a league table, is directly linked to accurate data, it is unlikely that hospitals in the UK will invest in improved data quality, particularly in view of the continued pressure on management and administrative costs.

### **Increasing throughput**

The White Paper emphasises the government's desire to abandon the Purchaser Efficiency Index which was criticised for, among other things, providing incentives to increase hospital based activity (inpatient and day case) relative to other services (Appleby 1996a). The index was also blamed for diverting attention from securing improvements in health gain and the quality of care. Neither health gain nor the quality of service come into the estimation of HRG costs so the issue is whether use of an index of HRG unit costs is less likely to generate incentives for hospital based activity than was use of the Efficiency Index.

As explained in Section III, 'costing' a treatment in an NHS Trust is the application of a set of accountancy rules for allocating fixed and semi-fixed costs by length of stay and, for surgical procedures, time in theatres. Aside from improving the quality of the data, the 'cost' of an HRG can be reduced by:

1. reducing length of stay and increasing throughput; and/or
2. reducing fixed expenditure and thus the cost per occupied bed day.

The first strategy requires a change in clinical practice and this might be realised because the NHS Executive will have new powers to reduce variation in clinical practice by requiring more general use of treatment guidelines. If the efficiency of purchasers and of Trusts is to be assessed by reference to an index of HRG costs, the short-run response is likely to be to encourage increased hospital throughput and reduced length of stay. Because HRGs currently do not apply to non-acute and non-hospital based activity it is not immediately obvious that

these short term incentives are very different from the actions encouraged by the existing Purchaser Efficiency Index.

Pressure to reduce length of stay has been associated with concern over two problems: cost shifting and reduction in quality of care. A hospital shifts costs on to GPs or the local authority if it discharges patients with a higher level of dependency. There is evidence that this occurred after the introduction of prospective payment in the US (Kosecoff *et al* 1990). If the pressure to increase throughput leads to inappropriate early discharge, the result may be increased readmission rates (Henderson *et al* 1989). The proposed publication of readmission rates may help limit such behaviour, but published clinical outcome data are likely to remain less comprehensive than those produced by other countries. Of those measures proposed (NHSE 1998a), only a subset of about six hospital indicators are expected to be made available, at least in the short term<sup>4</sup>. This compares with 32 indicators currently produced for health districts and hospitals in Scotland (Clinical Resource and Audit Group 1994; 1996). It is important that the disclosure of further information be encouraged. If more information on clinical outcomes were publically available, bad practices might be recognised more quickly. Whistleblowing and articles in *Private Eye* should not be the only means by which disquiet with our public service are voiced.

The second approach to reducing reported cost also involves changing clinical practice but, this time, indirectly. Most costs are labour, primarily nursing, along with diagnostic services, facilities management and estate costs. Given the age of hospital buildings and equipment, case-mix and ward dependency there may be scope for management to reduce the 'cost' of treatments by reducing the use and/or unit cost of these inputs. Fewer diagnostic services could be offered or the assumptions about dependency reviewed to allow reduced nursing costs although, as will be seen below, this approach may prove counter-productive.

Economists are ordinarily interested in costs only to the extent they reflect (or fail to reflect) opportunity costs and thereby influence decisions that will affect overall economic efficiency. Unit costs, calculated in the form of a costed HRG, do not reflect opportunity costs. They do not show what additional resources will be required (or saved) if we increase the number of hip replacements by 5% or if we shift the balance between surgical and medical admissions within a hospital or between hospitals or if consultants in neurology reduce the number of scans they ordinarily require for patients.

---

<sup>4</sup>

Although not confirmed at the time of writing, the indicators are likely to be: 30-day in hospital perioperative mortality by method of admission; 28-day emergency readmissions; discharge home within 56 days of admission with fractured neck of femur, by admission method; 30-day in hospital mortality for fractured neck of femur, by admission method; discharge home within 56 days following emergency admission for stroke; 30-day in hospital mortality from myocardial infarction.

Because HRG costs do not reflect opportunity costs care will need to be taken to ensure that use of the reference cost schedule does not create incentives for Trusts to avoid what would be genuinely efficiency enhancing moves. For example, suppose technical developments within one specialty permit an accelerated move from inpatient to outpatient treatment. If the Trust implemented the new pattern of treatment, all the overheads currently assigned to patient bed days in this specialty would have to be reallocated to the other specialties in the Trust. If the buildings and beds previously used to treat these patients can not be disposed of rapidly, the costing methodology requires these infrastructure costs be reallocated to all other Trust activity. The costs for HRGs in, say, orthopaedics rise because dermatology patients are no longer treated as inpatients but as outpatients. Managers and clinicians may have an incentive to slow down the movement from inpatient to outpatient treatment in order to avoid the opprobrium of being accused of having become less 'efficient' in the treatment areas where HRG costs have risen due to the reallocation of fixed costs.

The type of perverse incentive illustrated here will be present whenever proposals to improve patient care involve movement across institutional boundaries. The movement of patients from acute settings will raise the HRG costs of the patients whose treatment remains in the acute unit. These problems are inherent in any system of costing that is predominately one of allocating fixed and common costs rather than identifying marginal costs. It is like punching a bean bag: the volume is fixed and if you depress it at one point, it pops out at another.<sup>5</sup>

---

5

This point can be illustrated with an example. Radiology is ordinarily a cost centre and it is charged out to specialties at a price of £X per scan or x-ray to cover total cost.  $\text{£X} = \text{cost of buildings} + \text{equipment} + \text{staff} + \text{consumables} \times \text{number of scans performed for all specialties during the year}$ . Assume consumables are 20% of the total cost. If a couple of consultants try to become 'more efficient' by reducing the number of scans they ask for, the only reduction in inputs used will be the consumables. 80% of the inputs they thought they were saving are still employed in radiology and the costs of these inputs will now be spread over fewer scans. Not only will the consultants who thought they were saving resources find the price they are charged for the scans they still use has gone up, consultants in other specialties will find the price they are charged for a scan has risen.



## V USING REFERENCE COSTS

HRGs costs from all Trusts are to be summarised in three information sets:

1. A comprehensive list of the costs of every treatment offered by every NHS Trust in England.
2. The **National Schedule of Reference Costs** (NSRC): which will provide summary statistics (mean, range, inter-quartile range) from all Trusts for each HRG, by admission type.
3. The **National Reference Cost Index** (NRCI): which will provide a cost weighted summary of activity at Trust level, intended to give an indication of the technical efficiency of each hospital.

The consultation document sets out a formidable list of expected uses for reference costs in their various formats (NHSE1998b pp3&6):

Purchasers and Providers are to use them:

- as benchmarks for cost improvement;
- as indicators of technical efficiency;
- as a mechanism for identifying sources of inefficiency;
- as evidence to inform the negotiation of long-term agreements on improving performance;
- as a means of identifying good practice;
- in costing local Health Authority wide Health Improvement Programmes.

Ministers and the NHS Executive are to use them:

- in providing national information for funding transfers;
- in providing a better measure of efficiency;
- in ensuring public accountability for the use of NHS resources.

The public and patients are to use them:

- in assessing the efficiency of the NHS locally.

This array of anticipated use is discussed along lines that reflect relevant economic issues:

1. Can we expect conclusions about relative efficiency to be statistically valid?
2. Are the data appropriate for their intended use?
3. What are the incentives to improve efficiency?

## **Making comparisons**

Can the data used to construct the NSRC and NRCI ensure that conclusions about the relative efficiency of individual Trusts are statistically valid? Even though the NHS has for the last five years provided guidelines for coding patients and allocating costs, large variations in the calculated cost per HRG remain. Table 2 provides an example.

This shows the published prices for four arbitrarily chosen, but fairly common, version 2 HRGs quoted by acute providers in the North West Region for 1996/97. Theoretically, patients in these HRGs are expected to be reasonably resource homogenous, so the variation in costs among patients should not be substantial and should only reflect differences in the costs incurred by different providers to do the same thing. However, as Table 2 shows, the highest price varies from twice to over five times the lowest price quoted by providers for these HRGs. Are we to infer that the efficiency of these Trusts differs by as much as the difference in their published prices (costs) would suggest? Clearly if the differences stem from variations in counting units of activity, deficiencies in casemix measurement, or local interpretation of the costing guidelines, such an inference might be misplaced.

Users of the schedule should be made aware that the costs reported by Trusts might arise from chance rather than being statistically significant indicators of inefficiency. With there being 565 HRGs, there is a danger that individual Trusts will see too few cases within a particular HRG to produce reliable cost figures.

The consultation document on reference costs recognizes this problem and offers several means of reducing its impact. First, it is proposed to exclude from the database "those NHS Trusts which undertake a small volume of activity in a given HRG and those HRGs where, nationally, there is a small volume of activity" (NHSE 1998b). Second, the interquartile range is to be included among the summary statistics as a measure of variation in reported costs.

**Table 2 Variations in HRG prices**

HRG Description (Version 2)	Specialty	In-patient ECR tariff prices, 1996/97				
<b>c26</b> Mouth and throat procedures category 3 without complications or co-morbidities	ENT	24	671	175	379	1,103
<b>e44</b> Varicose veins	General Surgery	30	705	244	257	1,281
<b>h01</b> primary hip and other major joint replacements (excluding knee)	Orthopaedics	30	4,003	695	3,120	6,114
<b>l31</b> transurethral resection of the prostate >74	Urology	30	1,956	1,139	988	6,481
<i>Source:</i> Extra Contractual Referrals: Purchaser Arrangements and Provider Tariffs for 1996/97, North West Regional Office						

However, reported costs are themselves averages and fail to take account of the variation in costs at HRG level within hospitals. It would be preferable to provide confidence intervals based on patient level data rather than on information already aggregated at hospital level (Goldstein and Spiegelhalter 1996; Appleby 1996b). Third, outlier cases, those with exceptionally high lengths of stay for their HRG, are to be trimmed from the returns made by Trusts.<sup>6</sup> This results in lower mean costs, as cases in the upper tail of the distribution are ignored.

Offsetting these adjustments, the problem of small numbers is exacerbated by the NHS Executive's requirement that Trusts subdivide their HRG costs according to admission type and specialty (FDL(97)25). This subdivision seems to have developed in response to

<sup>6</sup> The upper trim point is defined as 1.5 times the interquartile range plus the upper quartile. Because data are positively skewed no low trim point is defined, although day case costs are often separately quoted. Trimming excludes a limited number of FCEs but a high proportion of inpatient days so reconciling HRG unit costs reported in the NRSC back to aggregate expenditure will result in a substantial shortfall. Rather than a single score as currently suggested (NHSE 1998b), the NRCI should report both trimmed and untrimmed costs because, first, failure to account for trimming could, in principle, encourage delayed discharge for patients close to the trimpoint threshold so that they are excluded from the trimmed NRCI calculation and, second, it would avoid erroneous conclusion being made about hospital efficiency which, in fact, stem from a disproportionate volume of outliers. If the proposed single score combines all cases, some means of apportioning costs across the patient stay is required, given that the marginal cost of each additional day is likely to less than average cost.

commissioning requirements. With GP fundholders purchasing elective but not emergency care, Trusts have traditionally subdivided their HRGs into these categories and quoted separate prices for elective and emergency cases. The new Primary Care Groups will be 'total purchasers' so that the earlier need to separate elective and emergency admissions will disappear. If for other reasons the Executive still wants to maintain the separate costing of elective and emergency admissions, a reduction in the robustness of their cost comparison data will be one of the prices to pay. It would be worth ascertaining to which HRGs admission method is an important explanatory factor of cost rather than assuming that a division is appropriate for all HRGs.<sup>7</sup>

A similar problem arises from the common practice for contracts to be specialty specific. Trusts have felt obliged to quote a specialty specific price for each HRG rather than a standard HRG price for the Trust as a whole. For example, for the treatment of varicose veins (version 2 HRG e44), a Trust in North West Region quotes one price in general surgery (£706), another in urology (£794), another in trauma and orthopaedics (£1,040), another in ENT (£422) and yet another in gynaecology (£543) (North West Regional Office, 1997). Clearly, for ostensibly the same treatment, this variation is counterintuitive (quite apart from it being odd to quote prices in some of these specialties!) and is probably nothing to do with differences in actual resource use across specialties. Rather the explanation is either to do with differences in the cost allocation process by which specialty based bed day costs are derived or to differences in length of stay, which may not be systematically related to clinical practice but are random observations from small numbers. It is estimated that 65% of specialty HRGs at individual Trust level contain between 1 and 9 cases per year (Bates 1996). The cost estimates pertaining to such limited samples cannot be expected to be particularly robust.

Assuming we have dealt with problems arising from statistical significance and small numbers, there is the issue of whether remaining observed cost differences are likely to be due to factors other than the relative efficiency of managers and clinicians. Economic theory postulates that if economies or diseconomies of scale exist, efficient small hospitals will have different unit costs from efficient large hospitals. If economies of scope exist, the minimum cost of producing a particular treatment will depend on the set of other treatments provided and hospitals producing different service mixes will have different minimum unit costs for each treatment.

The reference cost consultation paper (NHSE 1998b) recognizes there may be sources of cost

---

7

Emergency admissions are more expensive because they require the provider to carry spare capacity in order to accommodate stochastic variation in arrivals (Posnett *et al* 1998). However, this consideration does not enter into the costing methodology. If an important factor explaining cost (ie length of stay) it could be argued that admission method be used to define HRGs themselves rather than for the subdivision to occur subsequently. The NCO rejected using admission method as a grouping variable because it is not a patient characteristic and coding is open to manipulation.

difference independent of efficiency and proposes that Trusts be clustered, so that comparison is limited to other Trusts within the cluster. Nineteen cluster groups (some of which are further subdivided) are suggested in the consultation document, these having been devised by the NHS Executive and Audit Commission in previous examinations of performance. Trusts are grouped on the basis of:

- 1.size;
- 2.teaching status;
- 3.type of service (acute; children's; orthopaedic; community; learning disability; mental health; or some combination of these);
- 4.whether in London or not.

With the exception of the latter characteristic (arising mainly because of higher factor costs in London), these can be defended on the supposition that, among other things, (dis)economies of scale or scope exist. For example, a recent systematic review of the literature on this topic suggested at least three distinct groups: hospitals with fewer than 100-200 beds, hospitals with more than 600 beds and the rest (Ferguson *et al* 1997).

Clustering further undermines the problem of small numbers for comparison: in the consultation document two clusters contain only five Trusts and one sub-cluster features a single hospital (NHSE 1998b). The extent to which clustering is used is a reflection of the importance of factors (other than casemix) that are expected to effect costs but which are outside the control of management. Unless these factors are accounted for, conclusions about the relative efficiency of hospitals will be spurious. Statistical and econometric techniques are available to identify the extent to which patient treatment costs in teaching hospitals are likely to exceed those in non-teaching hospitals or the extent to which the presence of specialist units is likely to raise costs relative to hospitals offering more generalist treatment. If the government is to succeed in encouraging people to accept the NSRC or NRCI as valid indicators of relative efficiency the techniques used to separate differences in costs attributable to efficiency from other sources of cost difference must be robust and transparent.

### **Appropriate for intended use**

Three different intended uses for data on reference costs appear in the government's list:

- indicators of where problems may exist;
- mechanisms for identifying the sources of inefficiency and good practice elsewhere;
- financial information needed for funding changes in health care.

It is unlikely that a single accounting construct like the costed HRG will be appropriate for such a diverse set of functions.

The basic economic definition of technical efficiency is that inputs are minimised per unit of output of a given quality.<sup>8</sup> If we assume that Trusts are clustered to reflect economies of scale and scope, that the main statistical problems have been dealt with, that coding is fairly uniform across Trusts, that all interpret the costing guidelines similarly and that we are only interested in the existing institutional locus of services, then reference costs should be useful indicators of possible differences in the use of inputs. Whether this difference in use of inputs reflects a difference in technical efficiency requires the further assumption that there are negligible differences in the quality of care among Trusts.

Economists are understandably uncomfortable with policies that equate lower financial costs with higher efficiency. A treatment with higher financial costs per patient may be more cost-effective than an alternative with lower financial costs. The collection of outcome data lags far behind the collection of accounting data and, despite frequent references to the importance of quality and appropriateness, the NSRC will not and cannot relate to the cost-effectiveness with which resources are used and hence to a central part of the idea of technical efficiency. In the absence of these data the financial league tables are likely to be used as if quality were independent of financial cost. Of course, it may be true that quality is independent but there is some concern over creating incentives to respond to financial targets in the absence of evidence that quality is unlikely to be affected. However, careful consideration must be given as to how best to report quality data alongside cost information. HRGs may not be an appropriate level of aggregation at which to report quality data. McKee and Hunter (1995), for example, argue that mortality statistics should be reported by stage of the disease, which allows a more precise description of diagnosis than HRGs provide. It would be a happy coincidence if HRGs, designed to compensate for variations in the expected resource requirements of hospital casemix, were also able to account for differences in expected outcomes.

The published form of reference costs will not be a useful 'mechanism' for identifying sources of inefficiency or best practice. In the absence of information on cost shifting and quality measures at HRG level, the 'lowest' reported cost cannot be treated as an indication that the Trust producing it is the seat of 'best practice'. In order to trace the source of any inefficiency, the published costs will need to be decomposed. For example, a Trust with a published NRCI = 1.3 is reported to be 30% 'less efficient' than the national (or cluster) average while a Trust with

---

<sup>8</sup> Efficiency in *production* of a service requires more in that the mix of inputs must, given factor prices, minimise cost per unit of output.

NRCI= 0.94 is 6% more efficient than the average. The first Trust could have the most efficient paediatrics department in the country while the second might have the most inefficient orthopaedics.

If the NRCI figure is decomposed so that we can identify the orthopaedics problem in the 'efficient' Trust and if that Trust's orthopaedics HRG costs are themselves decomposed, we might find that the 'inefficiency' is due to the use of agency nurses. The published figures will neither enable us to identify the source of inefficiency nor determine whether a hospital is technically efficient. What we might expect is that in the process of assembling the data that must ultimately be reported for publication of the NSRC and NRCI, Trusts will acquire internal information that may help them identify sources of their higher costs. In this sense, the mechanism for identifying possible sources of inefficiency will not be the publicly available data on NSRC and NRCI but the information collected locally.

The closest the economics literature comes to discussing efficiency in the way the NSRC and NRCI address the issue is in the analysis of 'x-inefficiency' (Leibenstein 1966). The assumption is usually that quality is constant but that a degree of monopoly power has permitted managerial slack to go unpenalised so that costs are higher than technically feasible levels. Quality and appropriateness are issues for the consumer and the problem of the management is 'merely' to seek out and implement cost-effective production processes. The problem for the NHS is more complicated because steps to reduce 'x-inefficiency' that have the consequence of increasing the use of inappropriate treatments or reducing quality will not necessarily be obvious to the consumer or regulator.

The one suggested use for reference cost data that is clearly inappropriate is that of financial planning and funding transfers. Health Improvement Programmes (HImPs) are not just to be descriptions of existing services but are to set out changes in service delivery designed to achieve health goals. Costing a programme that involves a change in services requires that we obtain information on marginal costs, not short-run average costs. Short-run average costs are backward looking, representing investments in capacity that were made in the past. Government and the public may want to look backward at how managers and clinicians have used the resources at their disposal much as auditing accounts take a backward look at stewardship and probity. However, financial planning, funding of revenue transfers and Health Improvement Programmes are all forward looking and require that estimates of marginal cost be used in budgeting. The proposed reference costs, based on short-run average cost, will not be particularly suitable for these purposes.

A HImP that includes shifting 30% of maternity services from acute units into the community over the next three years could not be costed using the relevant HRG costs. The funding

requirements of the programme would have to be estimated by asking how many resources would in fact be released by the acute units each year and available for redeployment, and how many resources would be required in the community. The difference is the net cost of the programme and would bear very little correspondence to the present accounting cost of HRGs in obstetrics.

Similar issues arise if HRG costs are used by the NHS Executive for funding transfers. The government has announced its intention of replacing direct payment for what used to be called extra-contractual referrals by 'budget adjustments' for the main purchaser of Trusts providing these services. If a Trust located in one Health Authority permanently holds capacity for treating patients from another Health Authority, then transferring purchasing power from the second HA to the first at a rate that reflects overheads (such as the costed HRG) would be appropriate. However if the activity is temporary, as with a waiting list initiative, then the funding transfer should be at marginal cost.

### **Incentives to tackle inefficiency**

In neither the White Paper nor the consultation document is there discussion of how publishing reference costs will "create a strong lever with which to tackle inefficiency". These documents give the impression that the mere availability of information will make 'high cost' Trusts seek out and change the causes of their inferior performance. Comparative performance information has been available for some time. The Audit Commission regularly publishes material that would enable Trusts to 'benchmark' their performance against others. Trusts that subscribe to the CHKS database use confidential information on comparative performance for internal management purposes (CHKS Limited 1997). Such information has also been available to purchasers. What we do not know is why some Trusts act on this information and others do not. Nevertheless, it would not be difficult to imagine mechanisms for encouraging providers to pursue efficient behaviour.

One possibility would be to provide financial incentives and penalties for Trusts dependent on their position in the league table. Previous government policy was based on an expectation that purchasers would use comparative cost information to direct more of their resources to providers with lower costs but it is unlikely this approach would be favoured by the present government. The mechanism now dubbed 'name and shame' relies on embarrassment or loss of professional esteem as the incentive to ensure someone else appears at the bottom of any performance list that is published.

Another possibility would be a command and control approach where a Trust would be instructed to reduce the distance between some of its costed HRGs and the corresponding



reference costs by a centrally determined figure, say 10% per annum, during the next three financial years. Alternatively the NHS Executive could delegate responsibility to the Health Authority for setting the target cost reduction and this would be incorporated into the long-term agreements between Primary Care Groups and Trusts. Recent announcements suggest Trusts will be set such targets but we should ask what incentive there is to meet the target or what happens to a Trust that fails to meet its target.

To date reported experience with benchmarking in the public sector has concentrated on Next Steps Agencies (Samuels 1998) where Chief Executives are given performance targets by Ministers and the Chief Executives risk dismissal if they fail to meet targets. If this is the incentive intended for Chief Executives (and Chairs) of NHS Trusts and Health Authorities, it will be important that targets are set with reference to activity over which Chief Executives have control or influence. If the variation in length of stay that gives rise to the variation in HRG cost is due to decisions by clinicians and they are not willing to alter their practice, the contracts of consultants protect them from the penalties that can be imposed on managers.

It may be that the main incentive to comply with central directives is the usual one within command and control systems: a subordinate unit that fails to satisfy superiors will lose out in terms of future resources and the opportunity to undertake utility enhancing activities. The NHS Executive now has more control over major investment by Trusts, and Health Authorities have more control over appointment of consultants and developments at specialty level. If a Trust with relatively high reported costs in, say, cardiology HRGs fails to meet cost reduction targets, the opportunity for clinicians in that specialty to develop and expand their activities could be curtailed. In some areas there is excess capacity in the hospital sector and the expectation of major reconfigurations as technology changes. As with any firm trying to protect market share, hospitals or specialties within hospitals will have a strong incentive to avoid being labelled one of the weak performers and, therefore, one of the dispensable units.

A drawback of incentive structures related to 'industry' reconfiguration is that the most vulnerable are not necessarily the least efficient. Particularly in the public sector, hospitals with political influence and/or good news management can survive threats of being closed or having their activities significantly curtailed.

## V CONCLUSIONS

This paper has attempted to examine how the information on reference costs is to be generated, how it is likely to be used and whether the way it is used is likely to contribute to the objective of a more efficient NHS. For years individuals concerned with the efficiency and improvement of health care services in the NHS have complained of lack of data with which to evaluate the performance of this major part of our economy and public sector. While we applaud the new policy of making more information available, it is important that this information is used appropriately and that users are aware of its limitations.

It would appear that the government intends the information on reference costs to be used in part for all of the functions identified in Section II: remuneration, purchasing, internal management and central monitoring. Purchasers are expected to set targets for improvements in cost in their long-term financial agreements with Trusts. Managers and clinicians are expected to 'benchmark' their costs and change the way they organise and provide patient treatment so as to improve cost performance. What is not clear is whether incentives exist for purchasers and providers to change their behaviour in this way. Recent NHS experience suggests purchasers have tended to be weaker than providers and often had difficulty securing desired changes in activity. The creation of Primary Care Groups will further fragment purchasers and the encouragement of mergers will further strengthen Trusts. In view of this it may be that the only incentive structure likely to be effective in securing the government's desired use of reference costs is one of command and control. The increased control Health Authorities and the Department of Health will have over Trust resources could be used to encourage compliance with cost targets.

A problem with the command and control approach is that the data are not sufficiently robust to permit a target cost reduction of 10% for one Trust and 25% for another with any confidence that the consequences will represent efficiency gains. The process of generating the data within a Trust may well help managers and clinicians better understand the use of resources within their organisations but the specific number produced (forty-two?) may not tell us much about the efficiency with which health care is produced in one Trust when compared to another. In the absence of Trust specific HRG level data on quality and appropriateness for each HRG costed, it is not possible to tell whether the Trust reporting the lowest cost is producing a service of roughly the same quality as the Trusts further 'up' the schedule of reference costs. Collection of information on reference costs may well lead to fruitful discussion at the local level (one of the government's aims) but it is unlikely to produce a useful basis for setting targets (also, it appears, a government aim).

Comparison is hampered because of variations in the ways that activity is measured and costs

are apportioned across hospitals. At last calls to abandon FCEs as the main counting method have been heeded (Clarke and McKee 1992, Radical Statistics Health Group 1995), and the introduction of a unique patient identifier will further enhance understanding of activity across health sectors. However, the absence of patient level cost information continues to frustrate attempts to improve casemix measures, and the application of cost allocation guidance is subject to high degree of discretion.

Serious consideration of techniques for identifying the impact on cost of factors other than efficiency is required if the resulting efficiency index is to be credible. If efficiency league tables are to become a permanent part of government monitoring of NHS performance, there must be investment in producing better quality cost information and the information must be generated in a way that permits the average values appearing in the league tables to be accompanied by confidence intervals so that users are better able to judge the significance of a Trust's position in the table. This might increase the co-operation of Trusts as the system would then appear to be fairer than the proposed form of publication. It would also help the centre by making it easier to identify the genuinely poor or very good performers.

Experience with the system can be expected to help identify solutions to some of these problems. It may be worth examining whether the objective of changing clinical practice through cost benchmarking might not be more effectively pursued by using an indicator of resource use other than the costed HRG. As explained in Section III the absence of patient related cost data and the desire to minimise resources devoted to collecting cost data in the NHS means variations in the cost of a particular HRG are virtually all due to variations in length of stay. Why is so much time and effort being invested in translating length of stay information into cost data? This question would not arise if, as in Victoria, we were collecting details of many different resources used by a patient. Translating this information into standardised monetary values is a convenient means of summarising the different dimensions of resource use. But where resource use is proxied by a single dimension (length of stay), nothing is gained from reporting the measurement units on a different scale (money). Rather than providing greater clarity about variations in practice, it may obscure them for the very people we want to influence: clinicians. Peer group/professional pressure to alter practice may be more effectively mobilised if the league tables relate to characteristics of the service recognised to be professionally significant. If it is the managerial model of benchmarking that the government wishes to see adopted in the NHS then perhaps we should be benchmarking length of stay directly rather than indirectly via reference costs.

In its various plans for the NHS the government repeats a concern for quality as well as efficiency<sup>9</sup> but the way cost data are presented and used will influence whether the quality dimension of cost remains obvious and affects interpretation of cost figures. One possibility would be to integrate publication of cost and quality data but the difficulty is that there is very little information on quality at the HRG level. A solution to this problem is to confine efficiency measures to the level of the Trust or at least limit disaggregation to the specialty level. This would increase the feasibility of matching clinical quality indices with costs. Little would be lost by abandoning the plan to measure efficiency at the highly disaggregated level of the HRG. As argued earlier in this paper, given the data and the nature of costing procedures to be used, there is little or nothing that can be inferred about relative efficiency when comparing a costed HRG from one Trust with that in another Trust.

We are promised in the White Paper and various consultation documents that the "New" NHS is to be distinguished from the "Old" NHS by widening the focus on efficiency to embrace quality. A symbol of this is to be abolition of the purchaser efficiency index with its incentives to increase hospital based activity. Unless care is taken in the way the centre uses the NSRC and NRCI, there is a real danger that we will reinstate all the perverse incentives of the purchaser efficiency index when we introduce the new efficiency measures. Given costing methodology, the primary way a Trust can improve its NSRC or NRCI performance is by increasing hospital based activity. Much of the hoped for innovation in patient care pathways and new institutional settings for health care could be inhibited if the proposed efficiency indicator is given sufficient importance in the system of monitoring and setting performance targets that it simply becomes the "New" purchaser efficiency index.

---

<sup>9</sup>

In this context they usually mean 'financial cost' when they use the term 'efficiency' since it is obviously impossible to have economic efficiency unless quality has been specified.

**REFERENCES**

- Appleby J. Promoting efficiency in the NHS: problems with the labour productivity index. *BMJ* 1996a; 313: 1319-21.
- Appleby J. Relegating league tables. *Health Service Journal* 1996b; 14 November: 38-9.
- Barrow, M and Wagstaff A, Efficiency Measurement in the Public Sector: An appraisal. *Fiscal Studies* 1989; 10:72-97.
- Bates T. Casemix grouping for standard GP fundholding. Proceedings of the National Casemix Conference, Birmingham 22-23 October, 1996.
- Benton PL, Azorin D, Sutch S. Linkage of patient episodes into spells. Proceedings of the National Casemix Conference, Brighton 8-9 September, 1998.
- Benton PL, Anthony P, Evans H, Light SM, Mountney LM, Sanderson HF. The development of Healthcare Resource Groups version 3. *Journal of Public Health Medicine*, forthcoming.
- Butler A, Horgan J and Macfarlane L. Case study: Costing, contracting and resource management in Sanderson H, Anthony P and Mountney L (eds.) *Casemix for All*. Oxon: Radcliffe Medical Press. 1998.
- CHKS Limited. *Acute Care 97/98: Healthcare Resource Groups National Statistics 1996/97*, Financial Times Healthcare 1997.
- Clinical Resource and Audit Group. *Clinical outcome indicators December 1994*. Edinburgh: The Scottish Office, 1994.
- Clinical Resource and Audit Group. *Clinical outcome indicators July 1996*. Edinburgh: The Scottish Office, 1996.
- Clarke A, McKee M. The consultant episode: an unhelpful measure. *BMJ* 1992; 305: 1307-8.
- Culyer A. Supporting Research and Development in the NHS. London: HMSO. September 1994.
- Dredge B. The costing of medical specialties in Wolverhampton Hospital Trust. National Casemix Office Occasional Paper 16, July 1997.
- Duckett SJ. *The reform of public hospital funding in Victoria*. Australian Studies in Health Service Administration number 77. School of Health Service Management, University of New South Wales, 1994.
- Duckett SJ. Hospital payment arrangements to encourage efficiency: the case of Victoria, Australia. *Health Policy* 1995; 34: 113-34.
- EL(94)51 Comparative Cost Data: The use of costed HRGs to inform the contracting process.

FDL(93)76 Costing for Contracting Manual.

FDL(97)25 Costing Requirements and the Use of Costed HRGs in 1998/99

Ferguson, B, Posnett, J and Sheldon T. *Concentration and Choice in the Provision of Hospital Services*. CRD Report 8, NHS Centre for Reviews and Dissemination, University of York 1997.

Fetter RB, Shin Y, Freeman JL, Averill RF, Thompson JD. Casemix definition by Diagnosis Related Groups. *Medical Care* 1980; 18(Suppl): 1-53.

Freeman JL, Fetter RB, Newbold RC, Rodrigues JM, Gautier D. Development and adaptation of a hospital cost and budgeting model for cross-national use. *Journal of Management in Medicine* 1986; 1: 38-57.

Ginsberg PB, Carter GM. Medicare casemix index increases. *Health Care Financing Review* 1986; 7: 51-66.

Goldstein H, Spiegelhalter DJ. League tables and their limitations: statistical issues in comparisons of institutional performance. *J R Statist Soc* 1996; 159A (part 3): 385-443.

Health and Community Services. *Casemix funding for public hospitals 1994-95*. Health and Community Services, Melbourne, Victoria, 1994.

Henderson J, Goldacre MJ, Graveney MJ, Simmons HM. Use of medical record linkage to study readmission rates. *BMJ* 1989; 299: 709-13.

Jackson T, Henderson N, Tate R, Scambler D. *Resource weights for AN-DRGs using patient level clinical costs*. Technical Report 3, National Centre for Health Program Evaluation, Melbourne, 1993.

Jackson T, Tate R, Henderson N, Carlin J, Bayliss-McCulloch J. *1994 Victorian cost weights*. Technical Report 4, National Centre for Health Program Evaluation, Melbourne, 1994.

Jackson T, Wilson R, Watts J, Lane L, Bayliss-McCulloch J. *1997 Victorian cost weights study*. Health Services Research Group, Centre for Health Program Evaluation, Melbourne, 1998.

Koseoff J, Kahn KL, Rogers WH, Reinisch EJ et al. Prospective payment system and impairment at discharge. *JAMA* 1990; 264(15): 1980-3.

Leibenstein, H, Allocative Efficiency vs. 'X-Efficiency'. *American Economic Review* 1966;55:392-415.

Mayle, D, Francis, G, Hinton,C and Holloway J, What Really Goes on in the Name of Benchmarking? In Neely and Waggoner (eds) *Performance Measurement Theory and Practice Volume II*. Cambridge: Centre for Business Performance, The Judge Institute of Management Studies. 1998.

McKee M, Hunter D. Mortality league tables: do they inform or mislead? *Quality in Health*

*Care* 1995; 4: 5-12.

National Casemix Office. *Version 3 Documentation Set: Introduction*. NHS Executive G7033, May 1997.

NHS Executive. *The New NHS: Modern Dependable*. Cm3807, 1997.

NHS Executive. *The New NHS Modern and Dependable: A National Framework for Assessing Performance*. A Consultation Document. 1998a.

NHS Executive. *Reference Costs*. A Consultation Document. 1998b.

North West Regional Office (1997) Extra Contractual Referrals: Purchaser Arrangements and Provider Tariffs for 1996/97.

Posnett J, Bagust A, Place M. The rise in emergency admissions: data analysis and modelling projections. York Health Economics Consortium report Z9987, December 1997.

Radical Statistics Health Group. NHS "indicators of success": what do they tell us? *BMJ* 1995;310:1045-50.

Samuels M (1998) *Towards Best Practice: An evaluation of the first two years of the Public Sector Benchmarking project 1996-98*. Cabinet Office, London 1998.

Sanderson HF, Anthony P, Mountney LM. Healthcare Resource Groups - version 2. *J Pub Health Med* 1995;17(3):349-54.

Wilson R. What are we counting with Hospital Episode Statistics? Paper presented to the Third International Conference on Strategic Issues in Health Care Management, University of St Andrews, April 1998.