

Reform of hospital funding and the introduction of patient choice in England

Andrew Street

Centre for Health Economics, University of York, UK YO10 5DD

ads6@york.ac.uk

English translation of:

Street A. Reforma de la financiación hospitalaria e introducción de mecanismos de elección en el sistema sanitario británico NHS. *Revista Española de Economía de la Salud*, 2006: 5(2) 97-102.

Abstract

The British NHS is currently reforming the system of paying hospitals and allowing patients more choice about where they can receive hospital care. The previous funding system, based on block contracts, is being replaced by case-based payments, a mechanism called Payment by Results (PbR). Introduced in 2003, full implementation is expected by 2007. The calculation of reference costs and the estimation of national tariffs for each Healthcare Resource Group (HRG) are steps towards the goal of making health care delivery more accessible, responsive and productive.

1 Introduction

The National Health Service (NHS) in England is embarking on an ambitious process of reform designed to make health care delivery more accessible, responsive and productive (Department of Health, 2005a). The reform has a number of elements, but two are particularly important: reform of hospital funding and the introduction of patient choice, which allows patients to decide where and when they receive treatment.

The reform of hospital funding follows the lead taken by the USA, Australia and many countries in Europe. England is introducing casemix funding, a system of paying hospitals and other providers on the basis of the work they do (Department of Health, 2002). The key differences to previous contracting arrangements are that prices are fixed nationally, hospital revenue is related to activity, and activity ceilings have been relaxed. Hospitals receive a fixed payment – the national tariff – for each type of patient treated. Termed “Payment by Results” (PbR), the policy rewards hospitals for volumes of work adjusted for differences in casemix.

An overhaul of financial incentives is viewed as essential to improve NHS performance. PbR links hospital income and activity much more closely than previously has been the case. If they receive a standard payment, hospitals should be encouraged to find ways to cut costs and reduce length of stay in order to find capacity to accommodate more patients. Access should improve because hospitals have a direct financial incentive to do more work – they receive extra funds for each additional patient they treat. In the past hospitals may have been reluctant to accept patients not included in their formal contracting arrangements because of the difficulties of dealing with one-off financial matters. The new system is intended to remove these financial obstacles

Funding reform is also central to the policy to allow patients greater choice of hospital. In the past, NHS patients requiring elective (non-urgent) care simply had to wait until their local hospital admitted them. The patient’s general practitioner referred the patient for an outpatient appointment and the patient had to wait for the hospital to contact them with an appointment date. Now a new computer system has been introduced that allows GPs and patients to book an appointment “on line” and choose a specific date from a selection of hospitals. The options include both NHS (public) and independent sector (private) hospitals (Department of Health, 2004).

In this paper, we outline the key features of these elements of these reforms and discuss the challenges associated with implementation of the reform. In the next section, we describe the organisational structure of the NHS and the essential details of contractual relations between purchasers and providers before and after the funding reform. We then review three key components on which contractual relations are based: the description of activity; the method for calculating costs; and the method for calculating prices. We conclude by considering the experience to date and challenges ahead.

2 Changing contractual arrangements

The NHS in England is organised in a hierarchical structure. Every citizen registers with a general practice (GP) and the GP provides the first point of contact with the health service and acts as a “gatekeeper” to the rest of the system. GPs exercise a great deal of influence over health care utilisation, as they provide primary medical services, and patients often require a prescription from their GP if they need medication and a referral from their GP if they require non-emergency hospital care.

All general practices within a geographical area are grouped together to form larger organisations, called Primary Care Trusts (PCTs). There are 311 PCTs and these receive an annual budget from which to fund health care for their resident population. In 2004/05, the average PCT served a population of 164,000 and received a budget of £192 million (Jacobs, 2005). In order to keep within budget, PCTs have to manage the prescribing and referral decisions of their GPs and to have arrangements in place to pay for care from the hospital sector. Hence, PCTs have to ensure that the sum of expenditure on pharmaceutical (E^{Ph}), primary and community care (E^{PC}) and hospital care (E^H) does not exceed their budget. The recent reforms to hospital funding have made the task of keeping within their budget more difficult.

To see why, consider the composition of PCT expenditure on hospital care prior to the funding reform and afterwards. Prior to the introduction of Payment by Results, PCTs and hospitals contracted on a “block contract” basis. For the sake of exposition, assume that each PCT contracts with only a single provider – which, prior to recent reforms, was a reasonable assumption, at least for general hospitals outside major cities. The expenditure E^H of a PCT i can be summarised as follows:

$$E_i^H = \sum_S C_{si} \equiv p_{si} \times Q_{si}$$

Prior to PbR, contracts between a PCT (i) and the provider stipulated a total contract value (C_{si}), usually specified at specialty level (s). The expenditure on hospital care of each PCT amounted to the sum of its contracts across specialties. The PCT would decide how much of its budget to devote to the contract to each specialty and negotiate how much activity (Q_{si}) would be made available. The specialty-level price (p_{si}) was usually arrived at as the by-product of negotiations about total contract value and the volume of activity.

These arrangements allowed for tight control of expenditure, so that the each PCT’s expenditure would be close to the budget it had available. However, “block contracts”

provided little incentive for hospitals to exceed their contracted levels of activity and they did not allow patients much choice about the hospital where they could be treated.

PbR and patient choice changes contractual relationships in three fundamental ways. First, the “unit of activity” is described more accurately. Instead of “block contracts” specifying specialty volumes of activity (Q_s), PbR specifies activity by type of treatment (Q_j). So, instead of describing activity by specialty (eg a patient treated in trauma & orthopaedics), activity is described using Healthcare Resource Groups (eg a patient having a hip replacement). The HRG classification system is described in more detail in section 3.

Second, prices are set nationally, with local specialty-level prices p_{si} replaced by a national tariff \bar{p}_j for each HRG. The method used to set prices is described in section 5.

Third, under “Choose and Book”, patients can seek treatment from different hospitals. This means that PCTs have to pay for care from k hospitals, not just their local hospital.

Thus under Payment by Results, PCT expenditure on hospital care takes the form:
$$E_i^H = \sum_j \sum_k [\bar{p}_j \times Q_{ijk}]$$
 where j represents a HRG and k represents a hospital.

This payment system gives hospitals with low costs strong incentives to undertake more activity, because they are able to increase their revenue in proportion to the growth in activity. However, PCTs have to manage any growth in activity to ensure that their expenditure does not exceed their budget. But under PbR the PCT problem has become more complex than it used to be for two reasons. First, the introduction of “Choose and Book” allows patients greater choice about where and when they are treated. This makes it difficult to specify volumes (Q_{ijk}) in advance with their contractual partners. From the PCT’s perspective, the management of activity takes two general forms:

- PCTs need to ensure that the total *amount* of activity is appropriate and affordable.
- PCTs need to validate the *type* of activity for which they are paying, namely the category (HRG) to which a patient is allocated. Section 3 describes the HRG system.

Second, PCTs are unable to negotiate lower prices, having to pay the set national tariff for additional activity. We describe the approach to costing and pricing by HRG in sections 4 and 5.

3 Describing activity: Healthcare Resource Groups

HRGs are a classification system designed to group together patients with similar expected costs. HRGs are a locally modified version of Diagnosis Related Groups (DRGs), first developed in the United States (Fetter et al., 1980). Originally PbR was based Version 3.1 HRGs, which comprised 565 HRGs (Benton et al., 1998). An updated version has now been introduced, version 3.5, which contains 603 HRGs (NHS Information Authority, 2003). However, the essential features of the classification system are the same for the two versions.

HRGs are formed from electronic patient data. Hospitals are obliged to collect minimum data for all their NHS patients and to supply the information on a quarterly basis to a central data collection point. Each electronic patient record is grouped to a single HRG based on the data contained in the record, including procedures, diagnoses and age. Figure 1 depicts the process by which a patient is assigned to a HRG.

For surgical HRGs, grouping is driven by procedure, not by primary diagnosis. If more than one procedure appears, assignment is determined according to a procedure hierarchy, which assigns a rank to all procedures according to ‘clinical knowledge’ and the relationship with post-operative length of stay.

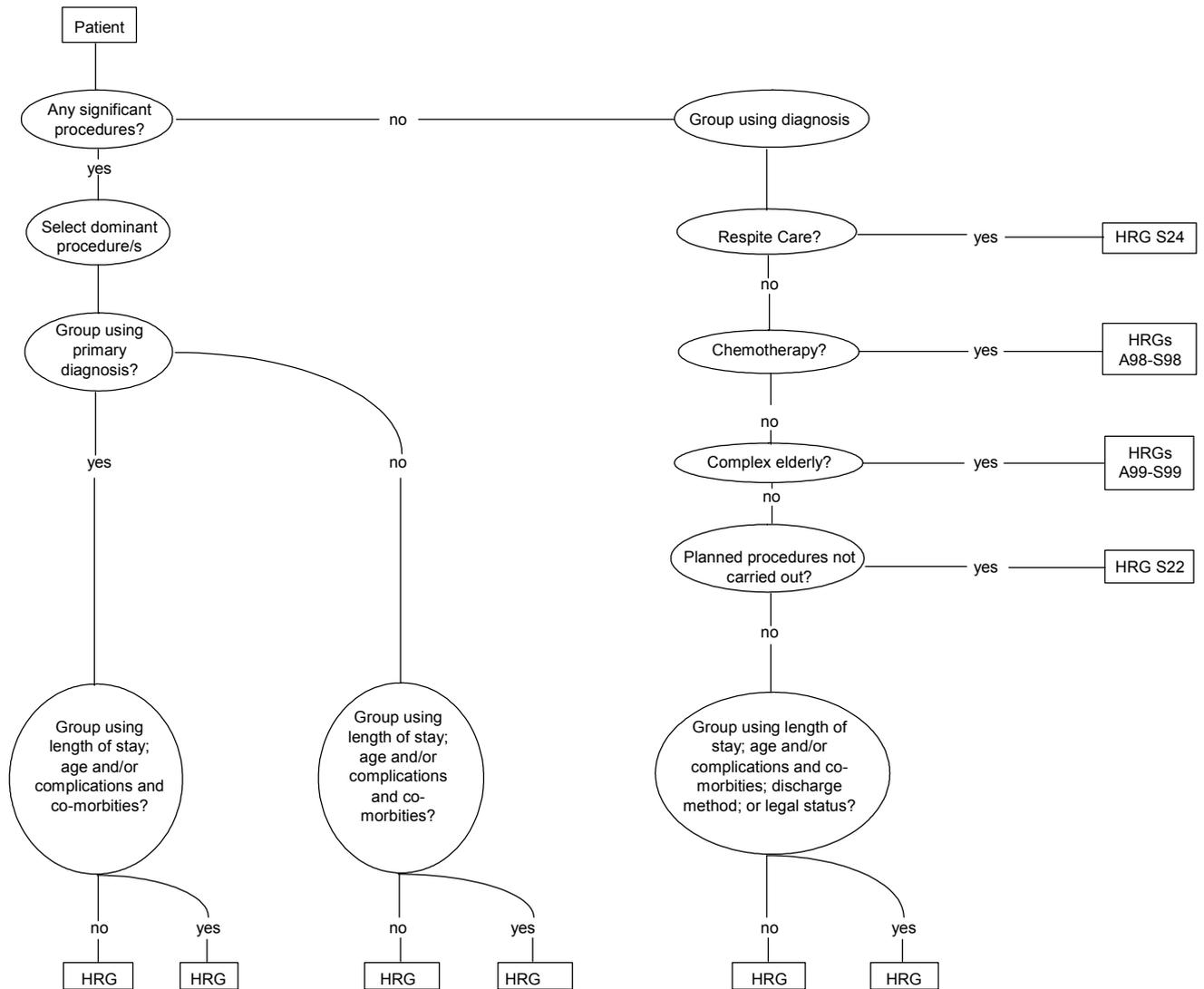
Where only minor procedures or no procedures are recorded, the primary diagnosis is used to determine assignment to HRG. The primary diagnosis may be overridden in the event of holiday relief care, chemotherapy, complex elderly cases (defined as patients aged over 69 with two or more major diagnoses but no significant procedures), and when planned procedures have not been performed. Where information in any field required for grouping is missing or invalid, the patient is assigned to an ‘undefined’ group – called U-codes.

Given that payments are made on the basis of HRG allocation under PbR, there are strong incentives to ensure that all the requisite information to determine appropriate allocation is extracted. But there are two issues with using this classification system that give rise to complaints from hospitals and PCTs.

First, some hospitals complain that HRGs fail to account for the complexity of their patients. Despite being an improvement on describing patients according to the specialty in which they are treated, HRGs are not completely resource homogenous groupings of patients. Within each HRG, there is a mixture of different types of patients. For instance, cardiac valve procedures (E04) is a large amalgam of procedures, each covering more than 50 different operations. Some of the patients allocated to this HRG will be more costly to treat than others. If some hospitals systematically attract these more expensive patients – perhaps because they are specialist hospitals – they will be penalised financially by the HRG-based payment system. There is evidence that the HRG system fails to capture the costs of specialist care accurately enough (Street, 2003). This problem is being dealt with in various ways, including revisions to the HRG classification system and “top-up” payments to specialist hospitals to reflect their more complex casemix.

Second, PCTs find it hard to verify that patients are allocated to the correct HRG. The HRG coding process opens up the opportunity for gaming, because PCTs do not have access to the primary information source (the medical record) from which the electronic data have been extracted. Hospitals have incentives to “up-code” their activity in order to gain higher payment rates. There is evidence from other countries of providers engaging in “up-coding”, extreme forms of which may involve falsifying procedural information or by recording complications that may not have been present (Carter et al., 1990, Hsia et al., 1992, Pitches et al., 2003). The lack of access to the primary source makes it difficult for PCTs to verify that HRG allocations are appropriate. Much effort by PCTs is directed at validating the claims made by providers and arrangements are in place for national auditing of coding practice (Audit Commission, 2006, Marini and Street, 2006).

Figure 1 HRG Classification system



4 Calculating HRG costs – reference costs

Just after the election of Labour in 1997, and a few years before HRGs provided the basis for Payment by Results, the government announced that HRGs could be a useful basis for making comparisons of the costs of treatments in different hospitals. In 1998, the first set of “Reference Costs” was published:

“Reference Costs ... itemise what individual treatments across the NHS cost. By requiring NHS hospital to publish and benchmark their own costs on the same basis,

the new arrangements with give ... a strong lever with which to tackle inefficiency” (p19) (NHS Executive, 1997).

Since 1998, the National Schedule of Reference Costs has been published annually, providing summary details (mean and interquartile range) from every NHS hospital of the cost of each HRG (NHS Executive, 1998). The reference cost data for every hospital is also placed in the public domain.

However, it is not straightforward to draw conclusions about relative efficiency on the basis of these reference costs. For one thing, as discussed above, the HRG classification system may not capture differences in casemix accurately enough. For another, hospitals may have higher costs because of factors that are outside their control (Street and Jacobs, 2002). And for another, reported costs may differ because hospitals may not calculate them in the same way.

Since the mid-1990s, the Department of Health has produced guidance to help hospital finance departments to apportion costs to HRGs (NHS Executive, 1994). However, this guidance has not produced good quality cost information. The NHS lags behind many other countries in terms of the routine cost data collected by health care providers. This reflects the loose historical relationship between hospital budgets and activity and the absence of a substantial private insurance sector requiring a detailed itemisation of resource use for billing purposes.

In countries such as Australia, public hospitals have computerised recording systems which make it relatively easy to determine the type and quantity of resources used by each patient during their hospitalisation (Jackson, 2001). In the US a price is 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, the best known being the Yale Cost Model (Freeman et al., 1986). Reinhardt reports that “contracting and billing departments of US hospitals are huge enterprises, often requiring large cadres of highly skilled workers backed up by sophisticated computer systems that can simulate the revenue implications of the individual contract negotiations” (Reinhardt, 2006).

In contrast, costing procedures in the NHS remain fairly unsophisticated. Even basic data about the number of diagnostic tests, theatre time or nursing dependency are very difficult to obtain. Hospitals 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 England is undertaken on a 'top down' basis (Department of Health, 2005b).

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. This yields the total cost TC of each specialty s .

Hospitals then have to apportion specialty-level costs to patients grouped into HRGs within the specialty. In some specialties – such as general medicine – most costs are driven by length of stay. In other specialties, the calculation of HRG costs may take account of other costs, such as the use of specific types of equipment or staff in the care process. To illustrate the calculation of HRG costs, assume that all variation in in costs is due to length of stay. The total specialty cost (TC_s) is divided by the number of bed days occupied by patients coded to that specialty BD_s , to derive an average cost per bed day c_s^{BD} :

$$c_s^{BD} = TC_s / BD_s$$

This cost per bed day is used to calculate the cost for per patient allocated to each HRG (c_j) as follows:

$$c_j = [c_s^{BD} \times BD_j] / Q_j$$

where BD_j represents the total bed days used by patients allocated to HRG j , and Q_j represented the number of patients allocated to HRG j .

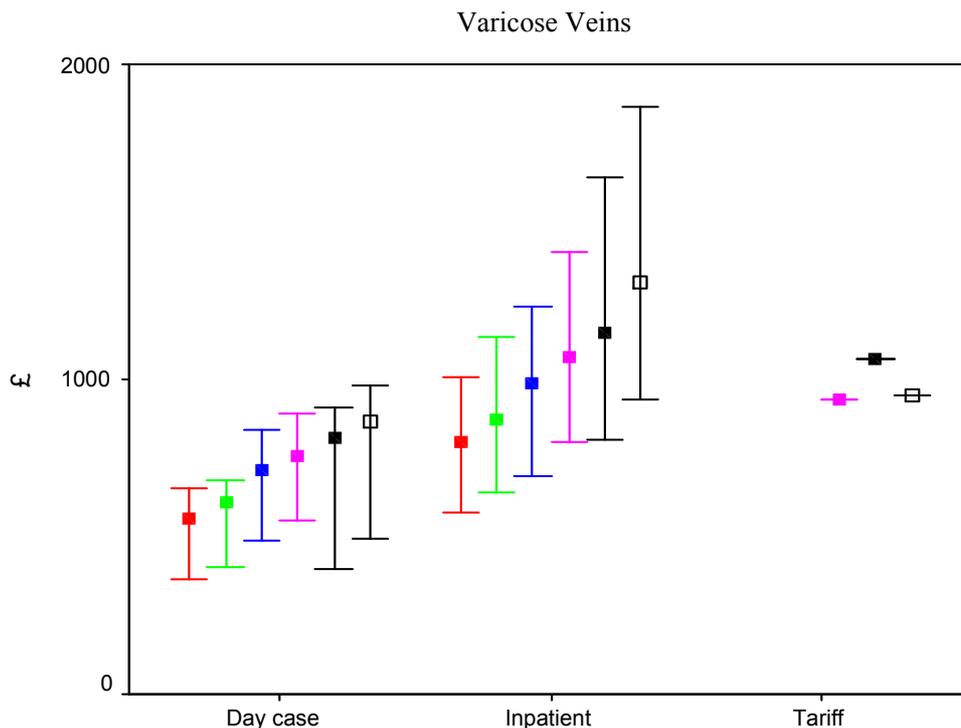
Even though hospitals are supposed to be following a standardised costing procedure, there is wide variation among hospitals in the costs reported for each HRG. As an example, data showing the reference costs for varicose veins are presented in table 1 and summarised in figure 2. The first set of six vertical lines in figure 2 shows the mean and interquartile range in reference costs for varicose veins undertaken on a day case basis, as reported by all English hospitals for each year from 2000 to 2005. The

second set of six vertical lines presents equivalent information for patients who stayed in hospital for at least one night.

Table 1 Reference costs and tariff for varicose veins, £

	Day case			Inpatients			Tariff
	mean	lower quartile	upper quartile	mean	lower quartile	upper quartile	
2000	557	368	652	799	580	1007	na
2001	606	401	681	870	642	1136	na
2002	711	484	842	988	692	1230	na
2003	754	550	892	1069	799	1402	934
2004	816	396	913	1146	810	1638	1063
2005	868	496	978	1309	935	1866	950

Figure 2 Reference costs and tariffs for Varicose Veins



There are two notable features in these two series. First, reported reference costs have been rising over time - this is a typical trend across HRGs and is above the rate of inflation for the economy as a whole. This is partly due to the large year-on-year funding increases made available to the NHS since 2000. The funding increase is associated with higher costs of labour. Wages have increased substantially, particularly for hospital consultants; there has been a wide-ranging exercise in staff

re-grading; and implementation of the European Working Time Directive has placed limits on hours worked, particularly by junior doctors (Maynard and Street, 2006, Singh, 2004, Williams and Buchan, 2006). This has meant that labour costs have increased at a faster rate than increases in activity – so the average cost of a unit of activity has risen over time.

Second, the interquartile range (showing the range in costs for 50% of hospitals) is quite large and does not appear to have narrowed over time. Again, the wide range in reported reference costs is typical across HRGs, and was a prime reason why hospitals were originally made to collect reference costs. When first announcing the collection of reference costs, the government argued that high costs were indicative of inefficiency (NHS Executive, 1997). But cost differences could also be explained by differences in accounting practices or, in some HRGs particularly, differences in the type of patients treated. This is unlikely to be the case for varicose veins, but is more important for some HRGs such as cardiac value procedures, coronary bypass or hip replacements where patients having a great variety of procedures are grouped together into the same HRG. Recognising this possibility, revisions have been made to the HRG classification system, and further revisions are expected in the future.

5 Calculating HRG prices – the national tariff

The final column of table 1 and the cluster of three points at the right hand side of figure 2 show the national tariff for varicose veins in 2003, 2004 and 2005. In England, the tariff is based on the national average reference cost for inpatient care for each HRG, \bar{c}_j , and the average reference cost for day cases, \bar{d}_j , weighting these two averages according to the proportion of activity nationally that is undertaken in inpatient and day case settings respectively. So there are strong incentives under the ‘payment by results’ regime to undertake work in the cheaper, day case, setting. The formula for calculating the tariff looks like this:

$$\bar{p}_j = \delta_j [\rho \bar{c}_j + (1 - \rho) \bar{d}_j]$$

where \bar{p}_j is the tariff price for HRG j , \bar{c}_j is the average inpatient reference cost for HRG j ; \bar{d}_j is the average day case reference costs for HRG j ; ρ is the proportion of elective activity undertaken on an inpatient basis; and δ_j is an “inflationary factor” for HRG j . The inflationary factor is designed to take account of the two-year time delay between the date for the reference cost submissions being returned and the prices being published. This adjustment is specific to each HRG. There is a further adjustment, not shown in the equation above, which may or may not feed through into

individual tariffs, to take account of local market conditions, which may impact on the price of labour. This is supposed to allow for costs in providing treatment that are beyond the control of the hospital.

Interestingly, the published tariffs in 2005 were lower than those in 2004, despite reference costs having risen. It is not clear whether or not it was a deliberate policy decision to publish lower tariffs.

England, then, has adopted a cost-based approach to pricing, with a clear link between reference costs and the national tariff, both of which are published in monetary units. In other countries, the relationship between costs and prices is not so apparent. In Australia and Norway, for example, the relative “values” across HRGs are published as weights rather than monetary prices (Duckett, 1994, Kjerstad, 2003). Moreover, these weights need not be based on the average cost of provision – as in England – but a benchmark price reflecting good practice.

6 Experience to date

The effect of Payment by Results to date is difficult to establish. The funding reform has been subject to phased implementation, with the national tariff applying to only a fifteen HRGs in 2003/04 and around fifty in 2004/05, with local prices applying to remaining activity. It was intended that the national tariff would apply to all activity in 2007/08, but this timescale is unlikely to be realized because of problems in setting the tariff for 2006/07, and particular concerns about the tariff for emergency care. Compared to casemix funding arrangements in other countries, PbR in England is based on a fairly straightforward formulation, which opens up the possibility of unfair reimbursement and perverse behaviour. As PbR is extended, technical aspects of the payment system are likely to improve, including better casemix classification of patients, movement away from simple cost-based pricing, and a more sophisticated payment regime, that allows for marginal pricing above some levels of activity.

Having said this, the general principles underpinning Payment by Results have been accepted across the NHS (Mannion and Street, 2006). PCTs and hospitals have embraced the idea that equal reward should be made for equivalent work and find the transparency of payment arrangements to be an attractive feature of the system. Many hospitals have welcomed the opportunity to expand their activity, knowing that they will be reimbursed at full (national) average cost. For their part, PCTs have recognized the advantage of trying to keep people of hospital, by encouraging GPs to limit their rates of referral or investing in alternatives to hospital care. This investment

has been made possible because they are able to save the full (national) average cost if a hospital admission is avoided.

However, although in theory PCTs have the ability to invest in alternatives to hospital care, in practice their ability to do so may be limited. PCTs have always been in a weak bargaining position with respect to hospitals, because they lack timely information to monitor and hold hospitals to account for their behaviour. PCTs also have few levers to influence referrals by GPs. The main instrument is the introduction of Practice Based Commissioning (PBC) budgets, which is updated version of the general practitioner funding scheme (Department of Health, 2005c). The problem is that, although GPs have incentives to underspend their PBC budgets, there are no penalties for overspending.

The impact of “Choose and Book” remains to be seen, having been implemented only in January 2006. As might be expected, there have been technical problems implementing the computer system but GPs have not embraced the policy as readily as the government would have liked. GPs are now being given an “incentive payment” for each patient they refer using the computer system. This will have to be carefully designed to ensure that these incentives do not undermine PCT attempts to control referral rates. By the same token, close attention to the design of appropriate incentives under Payment by Results is fundamental if the NHS reforms are indeed to make health care delivery more accessible, responsive and productive.

References

- Audit Commission, 2006, Payment by results assurance framework, Financial Health News Spring, 3.
- Benton, P.L., Anthony, P., Evans, H., Light, S.M., Mountney, L.M., Sanderson, H.F., 1998, The development of Healthcare Resource Groups version 3, *Journal of Public Health Medicine* 20, 351-358.
- Carter, G.M., Newhouse, J.P., Relles, D.A., 1990, How much change in the case mix index is DRG creep?, *Journal of Health Economics* 9, 499-504.
- Department of Health, 2002, Reforming NHS Financial Flows: introducing payment by results (Department of Health, London).
- Department of Health, 2004, Choose and book: patient's choice of hospital and booked appointment: policy framework for choice and booking at the point of referral (Department of Health, London).
- Department of Health, 2005a, Health reform in England: update and next steps (Department of Health, London).

- Department of Health, 2005b, NHS Costing Manual: 2005 edition (Department of Health, London).
- Department of Health, 2005c, Practice Based Commissioning: promoting clinical engagement (Department of Health, London).
- Duckett, S., 1994, Reform of public hospital funding in Victoria (School of Health Services Management, University of New South Wales, Sydney).
- Fetter, R.B., Shin, Y.B., Freeman, J.L., Averill, R.F., Thompson, J.D., 1980, Casemix definition by diagnosis related groups, *Medical Care* 18 (suppl), 1-53.
- Freeman, J.L., Fetter, R.B., Newbold, R.C., Rodrigues, J.M., Gautier, D., 1986, Development and adaption of a hospital cost and budgeting model for cross-national use, *Journal of Management in Medicine* 1, 38-57.
- Hsia, D.C., Ahern, C.A., Ritchie, B.P., Moscoe, L.M., Krushat, W.M., 1992, Medicare reimbursement accuracy under the prospective payment system, 1985 to 1988, *Journal of the American Medical Association* 268, 896-899.
- Jackson, T., 2001, Using computerised patient-level costing data for setting DRG weights: the Victorian (Australia) cost weights studies, *Health Policy* 56, 149-163.
- Jacobs, R., 2005, Health Information Service 2003/04 Commentary for The Health Service Financial Database & Comparative Tool 2004 (The Chartered Institute of Public Finance and Accountancy (CIPFA), Institute of Public Finance Ltd, London).
- Kjerstad, E.M., 2003, Prospective funding of general hospitals in Norway - incentives for higher production?, *International Journal of Health Care Finance and Economics* 3, 231-251.
- Mannion, R., Street, A., 2006, Payment by Results and demand management: learning from the South Yorkshire laboratory (Centre for Health Economics, University of York, York).
- Marini, G., Street, A., 2006, The administrative costs of Payment by Results (Centre for Health Economics, University of York, York).
- Maynard, A., Street, A., 2006, Seven years of feast, seven years of famine: boom to bust in the NHS?, *British Medical Journal* 332, 906-908.
- NHS Information Authority, 2003, Version 3.5 Healthcare Resource Group Documentation Set (NHS Information Authority, Winchester).
- Pitches, D., Burls, A., Fry-Smith, A., 2003, How to make a silk purse from a sow's ear - a comprehensive review of strategies to optimise data for corrupt managers and incompetent clinicians, *British Medical Journal* 327, 1436-1439.
- Reinhardt, U., 2006, The pricing of US hospital services: chaos behind a veil of secrecy, *Health Affairs* 25, 57-69.
- Singh, D., 2004, Quarter of hospitals not ready to comply with working time directive, *British Medical Journal* 328, 1034.

- Street, A., 2003, How much confidence should we place in efficiency estimates?,
Health Economics 12, 895-907.
- Street, A., Jacobs, R., 2002, Relative performance evaluation of the English acute
hospital sector, Applied Economics 34, 1109-1119.
- Williams, S., Buchan, J., 2006, Assessing the New NHS Consultant Contract - A
something for something deal? (The King's Fund, London).