BRIEFING

The Nuffield Trust
FOR RESEARCH AND POLICY
STUDIES IN HEALTH SERVICES

June 2010

Making progress on efficiency in the NHS in England: options for system reform

Jennifer Dixon

Funding of health care in the UK is under pressure in the short term after 2010/11 because of the need to fund the £167bn public deficit in the economy. This paper has been produced by the Nuffield Trust, and commissioned by Tribal, who have explored the question of preserving a publicly funded health service in an economic downturn. The Nuffield Trust will follow the publication of this paper with a major programme of research and policy analysis, New Frontiers in NHS Efficiency, that will examine how the NHS can become more efficient in a period of financial restraint.

Introduction

Ministers in the new Coalition Government have pledged that the NHS will receive real-terms increases in funding over the next five years. Precisely what the financial settlement for the NHS will be from 2011 onwards may not be clear until the intended comprehensive spending review in autumn 2010. From 1997/98 to 2010/11 the NHS budget grew in real terms at on average 5.7 per cent per year.[1] Over the period of the next spending review – 2011/12 to 2013/14 – the settlement will be far lower, at best between zero and two per cent. The NHS Operating Framework 2010/11 advises that primary care trusts (PCTs) should plan for 'real flat' growth in revenue allocations in 2011/12 and 2012/13 and reduced capital allocations.[2]

Recent analysis by The King's Fund and the Institute for Fiscal Studies suggests that demographic pressures to 2017 would require real-terms increases in funding of about 1.1 per cent per year in order to maintain quality.[1] The gap in funding in England between what is thought to be needed and the resources available is estimated in this study to be in the range £21–30bn – nearly 30 per cent of the current spend on the NHS in England – requiring efficiency improvements of 3.4 to 7.4 per cent per year. Official NHS sources suggest the gap to be £15–20bn between 2011 and the end of 2013/14. While there was an expected surplus of just over £1bn (one per cent of the NHS budget)

TRIBAL

Commissioned by Tribal

at the end of 2009/10, achieving cash releasing efficiency savings on the scale required will be a severe challenge.

In the short term, demands on health services may rise because of increasing unemployment if the economy remains weak and in the medium to long term because of population growth, aging of the population, the availability of new treatments and technologies (many of which may be cost-increasing), and changing expectations by the public of treatment and care, in part fuelled by the better availability of information. Pressure on funding will also grow because of the changing ratio between those in and not in paid employment as the population ages, and if there is continued slow growth in the economy and higher than predicted levels of price inflation.

During the election campaign, the three main political parties made some concrete suggestions as to how efficiencies and savings might be made: cutting the costs of administration in the NHS and Department of Health (DH); cutting the costs of quangos; scaling back the NHS IT programme; and restraining public sector pay, in particular for highest earning managers. These are areas that the government and the DH can influence directly for strategic health authorities (SHAs), PCTs and NHS trusts, although not for foundation trusts. All political parties had other common elements in their plans to reform the NHS to improve the efficiency of providers of care, for example:

- encouraging **preventive care**, in particular preventing avoidable hospital admission for people with longstanding medical conditions, with the assumption that this would reduce the avoidable cost of hospital care
- reducing hospital use by **speeding hospital discharge**
- **devolving budgetary power to front-line clinical staff**
- **encouraging use of non-NHS providers to compete** for contracts from NHS commissioners
- **more choice for patients** (more competition between providers for patients to stimulate efficiency and quality)
- **better financial incentives for professionals** to improve quality and efficiency
- **publishing of detailed performance data** to allow more informed patient and commissioner choice of providers, for example information from the Care Quality Commission
- **boosting the effectiveness of commissioning**.

Many 'macro' policy initiatives in the NHS, and indeed the political focus of reforms, over the past decade have been oriented towards increasing the quality, capacity and efficiency of *elective* care. These include, for example:

- reducing waiting times
- encouraging non-NHS providers such as independent sector treatment centres to provide additional capacity
- encouraging competition among a range of providers and patient choice of provider for NHS-funded care.

Yet the biggest 'efficiency frontier' is where the major costs in health care are: in the management of patients with long-term medical conditions, in the care of older people, in reducing avoidable emergency admissions, and in care for people at the end of their lives. It is here where reform efforts to improve quality and costs have only focused more recently, and should do so far more in the future.

There have been considerable efforts made to try to support the improvement efficiency locally, in providers. For example, through the Quality, Innovation, Productivity and Prevention (QIPP) initiative at the DH, and through the work of the NHS Institute, for example in the 'productive ward' series. Better information systems in particular on costs, for example through service line reporting and patient-level costing in foundation trusts, are allowing much better scrutiny of the costs of care delivered by managers and, crucially, clinicians.[3] For example, risk-prediction models using routine data, such as the Patients At Risk of Rehospitalisation (PARR) tool[4] used by PCTs, are now helping to identify more accurately patients at risk of future admission to hospital and thus enabling more effective case management and support of these individuals in the community to reduce avoidable ill health and hospital costs. There has also been improvement in financial management locally by PCTs and NHS trusts, as reported by the Audit Commission.[5] All this means that the NHS is in a better state than ever to identify where efficiencies can be made, and to identify and reduce large variations in practice.

Yet there is evidence to suggest that there is still significant waste and that it can be reduced. For example:

- productivity has declined over the last decade, in part due to large increases in the numbers of staff without concomitant rises in outputs
- there continue to be large and unaccountable variations in clinical practice

- there have been significant rises in emergency admissions to hospital for patients with conditions amenable to primary care and for admissions with zero length of stay[6]
- there has been no real shift in care from hospital to community settings, whether because of more effective prevention of ill health or substitution of care
- suboptimal care across provider and budgetary boundaries continues to cause avoidable cost through duplication and preventable ill health.

Furthermore there are some areas of care for which there is very little information and scrutiny, for example community health services.

There has been significant progress in the quality of NHS care over the past decade – in part due to the large number of initiatives[7] such as reducing waiting times and improving care for people with cancer, cardiac disease and other long-term conditions.[8, 9] However, measures to assess the quality of care are still underdeveloped, and for most managers and boards of NHS institutions quality as a goal of health care still comes second to balancing budgets. This means that the NHS now enters an era of significant budgetary challenge without routine measures for assessing the impact of making cutbacks on care quality.

Commissioning is potentially a powerful lever for change, but is weak and has shown little impact in stemming avoidable demand and costs of care.[10, 11] This is in part because providers (primary, secondary and tertiary), in particular medical professionals working within them, lack a focus on supporting patients to reduce their dependence on care by supporting them in the community rather than treating their illness. There are promising examples within the NHS, although incompletely evaluated. This reorientation will take time, and will mean aligning financial and non-financial incentives for providers and the staff working in them, so that they match more closely what commissioners are trying to achieve for patients. It will also mean the development of new forms of care stretching across provider boundaries, the like of which the NHS has never seen before. It may also require a much bigger role for patients and physicians in shaping health care, through taking more responsibility over managing health care resources. The immediate challenge is to make the necessary efficiency savings and cuts in such a way that this medium-term 'radical' agenda is accelerated, not set back.

In the short term, more effort to achieve technical efficiency will achieve results, in particular in high-spending areas. It is likely that no one reform, national or local, will achieve this, rather a spectrum of approaches will be needed. But there is also a widespread recognition that incremental changes will not be enough to meet the profound challenges outlined above. In the medium term, more radical change in the health sector is needed, in particular to reduce the ill health, dependency on the NHS and inefficiencies in care for older people and those with chronic conditions. The costs of care for this group as a whole is larger than for any other group, and the opportunities for efficiencies are the highest. Care is suboptimal because it is poorly coordinated for patients across different providers and is still largely *ad hoc* and reactive rather than systematic, proactive and preventive.

Achieving more cost-effective care for this growing group of people will require an intelligent set of coordinated initiatives, which together evolve into a profound change in the landscape of care. It is unlikely that these initiatives can all be thought out centrally. But there is a role for the government and the DH to ensure that wider reforms are coordinated, to ensure they are coherent and aligned in a way that avoids perverse effects locally.

In the wider public realm beyond health care, in the medium-to-long term there may need to be moves towards a different settlement with the public that evolves from the past. The emphasis is likely to be a combination of initiatives towards promoting self-reliance, staying healthy and informed, saving, reducing consumption, more cost-sharing by individual users and engaging of a wider group of people than presently involved to shape the public realm. Nowhere is this more pressing than in deciding how to fund social care for older people.

This briefing outlines how the NHS might respond to the financial challenge in the short-to-medium term. In doing so it draws upon several recently published materials,[12, 13, 14] the QIPP programme,[15] the NHS Operating Framework 2010/11,[2] and a suite of work underway at the Nuffield Trust entitled New Frontiers in Efficiency.[16] There are three main areas to consider with respect to improving efficiency:

- at 'system reform' level (that is externally to NHS organisations such as introducing more competition, and altering the national tariff)

- at the level of NHS commissioners and providers (that is internal to NHS organisations such as improving leadership, information and local incentives)
- at the professional or individual level (for example motivating individuals to work towards increasing efficiency through peer pressure and comparison).

All three levels are critically important and interlinked. The focus in this briefing is mainly on what can be done at *national* level to improve efficiency and reduce costs.

Direct policy levers

The focus on making efficiencies should be on hospital services, which account for approximately 65 per cent of PCT expenditure, and staff costs, which account for 40 per cent of overall expenditure in the NHS and approximately 70–80 per cent in NHS trusts. In contrast, primary care accounts for about 11 per cent of PCT spending, prescribing 12 per cent and community services nine per cent.

Within hospital care the focus should be firmly on how to stem the rise of preventable emergency admissions to hospital. Emergency admissions have risen by 11 per cent over five years, more than can be accounted for by changes in demography and in illness levels.[6] The rise is mostly in short-stay admissions, across all ages and diagnoses, and suggests that they are being inflated more by 'supply side' (provider) factors than 'demand side' (patient) factors.

In the discussion below, the main national levers to influence the NHS are outlined, grouped into where they have effect – the supply side or demand side.

Supply-side levers

Targets and performance management

Targets, with tight performance management by SHAs, and sanctions for those commissioners breaching annual budgets, are by far the strongest tools available centrally to control expenditure in the NHS. 'Top-slicing' the required efficiency savings from budgets by SHAs and PCTs at the beginning of the year may also be an effective, albeit blunt, tool.

It is also likely, given the manifesto pledges, that the new government will reorganise and reduce the parts of the NHS that are most controllable, for example the number of PCTs and SHAs, and arms-length bodies that receive grant-in-aid. Indeed it may be that the already announced requirement for a 30 per cent cut in management costs for the ten SHAs in England by 2013/14 may make some PCTs effectively unviable, and mergers inevitable. It seems highly likely that in future the government would want to see cuts in the budget of large regulators, in particular the Care Quality Commission.

Tariff

The control of the national tariff price is probably the second most powerful lever open to the DH (or future NHS Board) to keep NHS expenditure down. As the Audit Commission recently noted in a paper analysing productivity in the NHS, given the rise in demand for care 'it is a much more secure strategy to set a low tariff than rely on PCTs to cap activity'.[11]

In 2010/11 the uplift in the tariff will be zero and, according to the NHS Operating Framework 2010/11, will be a maximum of zero in the following three years, and will be coupled with efficiency requirements. Reducing regulated prices (national tariffs) is a blunt way of forcing efficiencies, because evidence from econometric studies suggests that if the tariff reduces to below the marginal cost of care in a provider, the quality of care reduces in a competitive environment.[17] If tariffs are effectively reduced without adequate measures of quality in place with which to monitor providers, then quality for some services will reduce, unless providers find lower-cost forms of care that are just as effective. This presumes that providers have accurate information on whether their costs for treatments by healthcare resource groups (HRGs) are above or below the national tariff. With the significant external challenge from the regulator Monitor, plus service line reporting or systems for patient-level costing in place, it is likely that this presumption is sound with respect to foundation trusts, especially for high-volume HRGs. It is not clear with respect to NHS trusts. To help, an early win by the DH would be to require all NHS trusts to introduce service line reporting.

If the price of tariffed activity is held down, the price of non-tariffed activity will also need to be controlled. The DH intends that the zero uplift to the national tariff in 2010/11 will also apply to non-tariffed prices agreed through national and especially local contracts. It is unclear how control can be exerted by the DH to ensure that local prices do not inflate unduly.

Holding the price uplift to zero for tariffed or non-tariffed activity, while effective in the short term, does not address possibly the largest area of potential inefficiency – unaccountably wide variations in clinical practice. Best-practice tariffs in four clinical areas of service will be piloted in 2010/11, which may help to ensure that evidence-based clinical practice occurs in high-volume service areas. But while as yet unknown new technologies may in future allow a significant quality of care to be delivered at lower cost, in the short-to-medium term radical quality-preserving efficiencies are most likely to be achieved through forensic scrutiny of clinical variations in performance and costs by local clinical leaders within provider organisations. (In the future, primary care commissioners, enabled by high-quality information systems and provider incentives, may be able to achieve this level of analysis.)

The importance of service line or patient-level costing in foundation trusts in this regard was noted above. Further, this type of scrutiny will be limited if providers are tied into their current structures – hospitals and general practices for example being entirely separate organisations – since primary, but especially secondary, prevention of ill health through care and other activities in the community will be key in making the greatest efficiencies. To help analyse use, costs and quality across providers, it will be critical for information on these to be shared, linked and analysed across different providers. The DH could accelerate this process centrally, for example by speeding up initiatives by the NHS Information Centre to extract data collected on electronic records in general practice (GP Extraction Service) and community services and facilitate linkage of these data with activity and cost data collected in NHS trusts.

The DH can also use the national tariff to provide disincentives for local providers to provide certain types of care. For example, the tariff for emergency admissions over the planned volume contracted with the local PCT commissioner is 30 per cent of that for an emergency admission within the volume planned. Of particular interest to tariff-setters will be the 12 per cent rise in emergency admissions across England over the past five years, of which approximately four per cent can be accounted for by demographic change, leaving a seven per cent real rise.[6] This seven per cent rise can almost all be accounted for by an increase in admissions where the patient stays less than one day in hospital, and to a much lesser extent admissions lasting one day.

There is likely to be huge scope to reduce these short-stay admissions appropriately. The lower tariff for emergency care should help, but more incentives are needed locally to prompt clinicians to reduce these admissions (see below).

In the past few years there have been other financial incentives to encourage social services to act more effectively to support patients at home to avoid costly long lengths of stay in hospital. The impact of these could be closely reviewed, to see if the incentives could be sharpened. However, the wider effect on social services needs to be monitored, especially if there are even greater pressures on funding in local authorities to cut costs in social care. Early analysis by the Trust is suggesting that people living in nursing homes have lower use of hospital care than people of similar age and with similar medical conditions living at home.[18] The possible substitution of social for health care, and vice versa, needs to be further analysed.

Allowing closures or reconfigurations of hospitals and other services

Over the last 30 years the number of beds in NHS hospitals has reduced by 30 per cent – a trend not unlike that seen in health systems across the developed world. In the last 20 years there have been countless reconfigurations and a marked reduction in the number of facilities. These changes have been driven by new technologies, developments in clinical practice, the need to make efficiencies because of funding constraints, and better information uncovering differences in quality of care and driving a need to concentrate highly specialist care in larger centres. As health care changes, so services have to change. And since, according to Roemer's Law[19] 'a built bed is a filled bed' almost regardless of levels of health need, if there is to be a shift of care appropriately out of hospital into the community then hospital bed capacity must reduce.

Hospital and service closures are clearly highly contentious and politicised decisions. For example, recent analysis suggests that hospitals are less likely to be closed in marginal political constituencies.[20] Despite better guidance, for example on how local health economies should draw up the case for reconfiguration, how effective consultation might best be achieved, regimes for managing failing hospitals, and the existence of a national independent panel to advise on reconfiguration, there is no escaping the fact that these

decisions are always going to be among the most contested and difficult to make in the NHS. No doubt the evidence for service configuration could be vastly improved – in particular the assumptions on which they are based more carefully thought through – and they could be more objective and transparent, and better supported by clinicians. As information improves, in particular on cost and quality of care, it may be that more objective local support for change can be commanded. But ultimately, as many have pointed out, a different order of political leadership – both local and national – will be required to make the decisions that are needed in the short term. While closures are possible and desirable, decisions need to be driven by GPs and local communities, with clear evidence to which commissioners can respond.

Monitor quality

The extent of efficiency savings to be made is much larger than required at any point in the NHS' history, if real-terms funding increases are as small as predicted. Efficiencies or cuts in budgets will need to be made quickly but in a way that preserves quality of care. It is critical that over the next five years cuts are not made without regard to the effect on quality. The new quality accounts for NHS trusts[23] should help, but underlining their importance alongside the financial accounts will require high and persistent pressure from the DH, SHAs, Monitor (in the case of foundation trusts) and local commissioners, to encourage the boards of provider organisations to focus at least as much on quality as costs.

As noted above, it is disappointing that more progress has not been made over the past decade to develop suitable measures of clinical outcomes, particularly relating to hospital care, and to put in place the clinical information systems needed to monitor them. There is now more effort, for example through the new Quality Observatories in each SHA, and in the Better Care Better Value indicators developed by the NHS Institute. Quality accounts are also being piloted for primary and community services in two SHAs. But as the measures of outcomes of care remain relatively underdeveloped it will be very important, in the face of significant budget cuts, to make more use of existing local systems to monitor patient feedback on care received. This should be a central and urgent concern of PCT commissioners, SHAs, the DH, Monitor and the Care Quality Commission. Patient feedback on experience of care is being incorporated into Quality Accounts, will be made available on NHS Choices from December 2011, and will be a feature in the Commissioning for Quality and Innovation scheme. But this feedback should be central to every board's analysis of the impact of efficiency initiatives or cuts in their local NHS organisation, a fact that needs to be repeatedly underlined by the most senior officials at the DH, and the chief executive of the NHS or future NHS Board.

Workforce contracts

The NHS employs 1.6 million people and the pay bill in the NHS is approximately 40 per cent of all costs, rising to approximately 70–80 per cent in acute NHS trusts. Most non-medical staff are on contracts agreed through Agenda for Change,[24] which increases staff costs in real terms by 1.5 per cent per year due to annual increments in salaries that equate to approximately £420 million per year. The three-year pay deals agreed under Agenda for Change contracts expire in 2011, when they can be renegotiated, and the deal is likely to be much tougher and to yield significant savings, due to the large number of staff on these contracts. The government has recently announced a wage freeze in 2010/11 for senior NHS managers, most consultants, GPs and dentists with the minority of GPs and dentists employed directly by the NHS receiving modest rises of one to two per cent.

Significant inflation of the salaries of GPs occurred since the new national GMS contract was introduced in 2004. In 2006 approximately two thirds of practices were funded through the national general medical services (GMS) contract (with Quality and Outcomes Framework payments accounting for about one third of practice earnings), and one third through the alternative personal medical services (PMS) contracts. Between 2004 and 2007, the average pre-tax pay for GP partners (within GMS and PMS practices) increased by 58 per cent,[25, 26] yet over this period GPs reported a reduction in their working week[27] and approximately 90 per cent of practices opted out of the obligation through the contract to provide out-of-hours care.[26] It is highly likely that these national contracts will be reviewed again to require greater productivity and payments to practices held down over the next few years. This way it is conceivable that GMS and PMS contracts may ultimately become less attractive relative to local employment contracts with PCTs, which the so-called 'salaried GPs' have, although currently pay for salaried GPs is significantly lower than for GP partners under the GMS and PMS contracts. The pros and cons of national contracts for general practices must surely come under scrutiny in future for other reasons, for example to

try to remove potential obstacles to the development of more integrated cost-effective care (see below).

The intended benefits of the policy of Agenda for Change were that staff would have more flexibility to work differently to improve productivity for better pay. But as the National Audit Office (NAO) noted, NHS trusts did not develop productivity measures when they introduced changes in the way staff work, and so it has not been possible for the DH to show the contribution that Agenda for Change has made to productivity.[25] The NAO recommendations included that the DH should be more active in encouraging trusts to specify, within business cases for changes to the way services are delivered, how planned changes or productivity would be increased by use of Agenda for Change, and that they should collect this information to show a national picture and specific examples of impact. The latter will be very important to collate if there are a rash of service changes proposed. These recommendations also hold true for the main contracts with medical staff – the national consultant contract and GP contract.

Policy Exchange, a think tank, has suggested that top-performing very senior NHS managers should be able to receive performance-related bonuses above the seven per cent of salary restriction that exists, perhaps up to £30K.[13] The rationale offered is that bonuses of between 50 to 100 per cent of salary are offered to executives in the private sector. But this is an overly simplistic comparison that deserves more scrutiny. For example, the nature of health care is more complicated than an average firm, and the intrinsic motivations of staff working in health are likely to be quite different. Now is an odd time to be considering such bonuses; arguably the next few years will provide ample external challenge on senior managers to make required efficiencies, and the politics of awarding large bonuses to senior managers in the face of what will be a period of cutbacks would be very inflammatory.

The cost of NHS pensions was approximately £12.5bn in 2009/10 with employers' contributions set at 14 per cent of pay.[13] It is unlikely that pension benefits in the short term will alter for existing staff, so the scope for significant reductions in pension costs is small.

Costs of prescription drugs

Prescription drugs account for about 12 per cent of PCT expenditure or £7.5bn. Branded drugs account for 80 per cent of spend and 20 per cent of volume, and generic drugs the remaining 20 per cent of spend but 80 per cent of volume.

Nationally the prices of branded prescription drugs are negotiated every five years between the DH and the pharmaceutical industry under the pharmaceutical price regulation scheme (PPRS), resulting in a voluntary agreement between the DH and the industry. The last PPRS came into effect in January 2009 and included for the first time support for innovation and uptake of clinically and cost-effective medicines, which together are designed to reduce the costs of prescribing in the NHS by five per cent by 2014. This includes an initiative for pharmacists to switch from a branded to generic drug on a prescription unless the prescribing doctor indicates otherwise by ticking a box. In a review of the PPRS in 2007, the Office of Fair Trading recommended the DH take much more active steps to develop value-based pricing, particularly for drugs that create the biggest revenues globally – those used for conditions that are chronic and non-fatal.

But while these national initiatives are helpful in curtailing the price and to an extent the demand for branded drugs, the most effective way of curbing avoidable demand for drugs will be local action to scrutinise and challenge prescribing practices of physicians. In turn this local action will be strengthened by better information, incentives, and leadership (see below).

Procurement

The DH has been active in a number of ways to help increase efficiency within the NHS in back-office functions, IT, use of property and procurement of supplies. Some of the recommendations of the *Operational Efficiency Review* commissioned by the Treasury[28] have fed into directives and guidance. *Smarter Government*[29] emphasised the setting of benchmark comparisons for back-office functions, and signalled significant reductions in spending on IT, external consultancy support, and communications and marketing, which in turn are requirements set out in the NHS Operating Framework 2010/11. The newly set up Commercial Support Units in each SHA may help strengthen procurement by PCTs and NHS trusts, and there have been other structural changes of entities to boost the NHS supply chain efficiency.

However, as the National Audit Office and Audit Commission point out in their recent review of collaborative procurement across the public sector[30] public bodies are still conducting

'expensive procurement exercises rather than using existing framework agreements to buy standard commodities such as stationery, computer equipment and travel services'. As a result, wide variations in the prices paid by the public sector for key commodities were found. Furthermore they found that 80 per cent of bodies surveyed did not measure the costs of letting a contract. A key recommendation was that the Office of Government Commerce should develop a consistent cross-government approach for all spending by the public sector on procurement.

Reduction of central budgets

This option is already taking place, although central budgets are a small proportion of overall NHS spend.

Strengthening commissioning

NHS commissioners, PCTs and practice-based commissioners, clearly have an important role through the contracting mechanism to encourage greater efficiency among providers, in particular hospitals. Yet as noted above, both PCTs and practice-based commissioning groups have been unable significantly to influence rising demand for hospital care, in particular emergency admissions.

The direct central lever to try to boost commissioning has been through the World Class Commissioning initiative; that is, through central command. The initial process to assess PCTs against a range of core competences has provided some information on the strengths and weaknesses to date, showing many PCTs as weak. There has been much analysis of why PCT- and practice-based commissioning has not been stronger, most recently by the Health Select Committee.[10, 31]

While support of PCT commissioners is taking place, for example through the World Class Commissioning assurance process, it is unlikely that national attempts to 'upskill' commissioners can themselves result in significant change in the efficiency and quality of clinical care in the short-to-medium term. Many PCTs are too small to attract the management, analytical and clinical expertise needed, and will need to evolve into larger entities, perhaps initially by sharing back-office functions and other business services. This is happening to some degree across England, and this natural evolution is probably more cost-effective and certainly less disruptive than wholesale mergers (though slower to take effect).

But the weakness of commissioning is fundamentally rooted in the fact that NHS trusts, and the clinical staff working in them, have too few incentives to reduce avoidable admissions. What is urgently needed is to provide stronger incentives to the provider system to orientate health care towards these goals, particularly for patients with chronic medical conditions, thus aligning commissioner and provider objectives.

There have been signs, from the Conservative Party, Policy Exchange and others, of renewed interest in giving GPs incentives to improve ambulatory care and thus reduce the need for costly hospital care. The incentives are to give GPs hard budgets with which to commission care, a reincarnation of something like the GP fundholding scheme pioneered in the 1990s. But the evidence of the impact of fundholding practices, and their variants Total Purchasing Pilots in the 1990s shows that their transaction costs were high, and crucially because of their small size they were unable to make progress in influencing powerful hospitals to shift care into the community[32, 33] because of a lack of purchasing power, and a lack of the management, information and clinical expertise needed. Moving from relatively large to small purchasers in this way is not going to achieve the efficiencies now demanded. A bigger vision is needed.

Demand-side levers

Limiting the benefits available on the NHS or requiring co-payments by patients

The NHS Constitution begins to outline the benefits that are guaranteed on the NHS, although at present these relate more to levels of service (such as waiting times) rather than clinical treatments. The National Institute for Health and Clinical Excellence (NICE) is an institution which helps define the cost-effectiveness of treatments and thus what may or may not be recommended as being funded by the NHS. The role of the proposed NHS Board is yet to be made clear, but one aspect could be to make more clearly the health care benefits that are NHS funded, and thus what must be paid for privately.

Yet defining the clinical benefits more explicitly in this way as a package would be highly controversial. First, there is no purely objective and uncontentious method of defining a package, and most countries that have tried to do so have had to draw back from this policy because of public and thus political discontent. Second, because the founding principle of the NHS is to offer comprehensive care that is free at the point of use, and this principle is strongly supported by the British public. In the early 1990s, there

were attempts by one or two regional health authorities in England to curtail whole treatments such as *in vitro* fertilisation, varicose vein surgery and removal of wisdom teeth on the NHS. Curtailing these treatments would not have saved much money and generated huge controversy to the point that the DH stepped in to overturn this decision.[21] Third, with obvious evidence of waste in the NHS, it may be highly inappropriate to cut back on benefits available in this way until other ways of increasing efficiency have been exhausted. A full analysis of the scope for the NHS to make explicit the benefits available, given international experience, is due to be published by the Nuffield Trust as part of its New Frontiers in NHS Efficiency programme mentioned above.

Similarly there has been much analysis of the use of co-payments in the NHS.[22] The main arguments for these are that co-payments reduce the phenomenon described by economists as 'moral hazard' (whereby a service is overused (and abused) if provided free of charge, because users have little reason to curtail their demand appropriately). The main arguments against are that co-payments are often highly regressive, costly to collect, and deter appropriate as well as inappropriate demand. If appropriate demand is deterred, then greater costs may be incurred later (for example via emergency admission, or an exacerbated condition), not to mention significant ill health. The rationale has been that it would be far more effective to apply co-payments first to *physicians* who are mainly responsible for decisions about treatment and associated costs, in other words to develop disincentives for inappropriate treatment, and incentives for appropriate and preventive care. Possible incentives along these lines are outlined further below.

Indirect policy levers

Clearly the DH has a very wide range of more indirect ways to encourage local action to improve efficiency. These include ensuring that local health economies have information and support on how to make efficiency savings, and ensuring that the incentives in the system are encouraging efficiency and quality to the degree needed.

Supply-side levers

Information, guidance and support

The DH funds the development of a huge amount of information on how to improve efficiency and boost quality, for example in its own work through the QIPP initiative, and in funding the work of NICE and the NHS Institute, the Public Health Observatories, to name a few. Consequently there is no shortage of guidance available, for example the Productive Ward series and information about lean management processes available from the NHS Institute.[34] The savings that could be made through these and other methods are estimated at £5bn. 'NHS Comparisons' developed by the NHS Information Centre allow benchmarking of indicators across the NHS with similar organisations. There is also a wealth of other information from the Treasury which has fed through into requirements as part of earlier Comprehensive Spending Reviews, such as the Gershon Efficiency Programme, and the Value for Money Delivery Agreement.[35]

Similarly the DH has invested, somewhat unsuccessfully to date, in developing an NHS information infrastructure that allows transfer of clinical information across providers, reducing the potential for waste. Locally, especially in NHS trusts there has been less direction by the DH to encourage the use of information systems that yield detailed information at service or patient level on costs and quality. These systems will be critical for providers, and commissioners, to use to make challenging savings in future and preserve quality. Clearly there needs to be a fundamental rethink of the NHS IT strategy along these lines, including the potential of open source software[36] to enable local IT systems to develop with clinical input, and to link with systems in other providers.

There is a serious gap in information on community services, and to a lesser extent in mental health care. At present it is not possible to assess the value for money of these services, and much more needs to be done to improve the data and scrutinise efficiency. But the most pressing issue is how to encourage people (especially clinicians) to use information, in particular the wealth of information that is currently collected and available. The focus of centrally driven reform over the short-to-medium term must be on developing incentives to this end.

Incentives

The incentives in the NHS arising from bearing down on tariff prices, from a stringent financial settlement, and from central performance management have been outlined above. These could be characterised as 'push' incentives. Here, less direct incentives that might help to improve

efficiency are briefly discussed – these could be called 'pull' incentives, as they encourage intrinsic motivation within providers to drive change.

There is no doubt that there are three major and related groups of factors which are powerfully inflating the use of costly hospital care in the NHS. These can be summarised as:

- positive incentives on institutions
- lack of incentives on clinicians, in particular doctors, to scrutinise clinical practice
- difficulties in reducing the existing supply of staff and facilities.

Where the tariff price for a treatment/inpatient admission to hospital is, or is perceived to be, above the marginal costs of care, this effectively results in a positive 'fee for service' incentive for NHS trusts and foundation trusts. There are not enough incentives at institutional level for providers to seek ways to reduce activity – neither from the tariff, nor from performance management via commissioners or SHAs. Clearly allowing no growth in, or reducing, the national tariff to a point at or below marginal costs will encourage activity to be reduced, but this is a crude method since it bears no relation to need for care.

Clinical audit, continuous professional development and staff appraisals for example through the consultant contract are some ways by which clinicians can be encouraged to reflect upon their clinical practice. But these are more orientated towards individual professional development than management of personal or team resources which are crucial if the required efficiencies are to be made in the short term. As noted above, there are signs that, in foundation trusts, the use of service line reporting is engaging clinicians.[3] This, combined with effective clinical leaders to peer review performance, is proving to be a promising way to reduce avoidable cost for example in Darlington.[37] However, service line reporting does not as yet extend outside the walls of foundation trusts. To have real impact all costs along the patient pathway, from primary to hospital care, need to be managed more closely – in particular for high-cost cohorts of patients.

If admissions to hospital, and hospital stays, are reduced then, because of Roemer's Law noted above, any spare capacity will be immediately filled unless wards or other facilities are closed. The inevitable political difficulty of doing this was noted above, but cases are less likely to be contested if they are supported by accurate information on quality and costs and justification by clinicians, with good clear two-way dialogue with the public.

These three factors call for a new approach in reform, which encourages clinicians, doctors in particular, to manage budgets covering care outside and inside hospital, to give positive incentives to hospitals for helping people stay well and out of hospital, and puts clinicians in a more central role (as envisaged in Lord Darzi's Next Stage Review)[23] in deciding and justifying decisions on major service changes. This is discussed further in the next section.

Integrated care

Because of this, there is much talk of vertically and horizontally integrated health care, which may be a very promising way forward.[38, 39, 40, 41, 42, 43] The important point is that integrated care models can effectively align incentives across different institutions, professional groups, budgets and across commissioner/providers to help a population stay well and reduce costs of care. As yet the evidence to suggest their impact on efficiency is underdeveloped but growing – it is the subject of a review to be published by the Nuffield Trust in Autumn 2010. But it is promising enough for the DH to pursue this track of policy more vigorously than it has to date by allowing more radical forms containing the key ingredients to evolve, and to give 'permission' and moral support for the local risks to be taken for evolution to occur. 'Integration' is not necessarily at odds with competition, since integrated networks or organisations may effectively compete for patients.

Integrated care may be achieved through a variety of arrangements, for example multi-professional integrated care teams working to shared goals but employed by different organisations, networks of provider organisations operating under a single integrated budget, or single organisations consisting of merged providers. It is important not to assume that organisational integration (or merger) is the optimal way of achieving integrated care for patients. Integrated care organisations can comprise primary and secondary care clinicians, and payer/provider or just provider delivery systems. The integrated care pilots initiative funded by the DH are a start, and other interesting and more radical plans are being implemented

around England outside the pilot initiative, for example in Trafford, Cumbria and Redbridge. No one model is likely to fit everywhere in England and a plurality of approaches should be championed.

There are similar initiatives to integrated health and social care which show early promise (including Torbay and the Isle of Wight), for example by helping to stem the rise in emergency care for older people by supporting them more effectively at home.[6, 39] Whatever forms of integration develop it will be important that social care is an ingredient. However, forcing health and social care joint budgets together across England through national policy (mooted in government circles earlier this year) is unlikely to be the right approach. In part this is because it tramples on existing arrangements that have developed locally, may be an appropriate lower priority in some areas than others, and diverts management time. It is unclear for example whether bigger gains in efficiency and health might occur from more integration between primary and secondary care, than health and social care. Instead, as in health, an evolutionary and plural approach, with support and nudge from the centre is more appropriate.

Some of the main ingredients of high-performing integrated care organisations internationally are that they have:

- a risk-adjusted capitated budget for an enrolled population for which they assume the financial risk
- primary care at their centre
- information systems that allow detailed scrutiny of cost and quality at patient level
- clinical leaders who conduct effective peer review of performance and
- the ability to remove poorly performing clinical staff
- well-aligned financial incentives within the organisation, to reduce avoidable cost and maximise quality.[43, 44, 45]

These ingredients appear to be needed together, for example it is not enough to give a provider entity hard budgets without ensuring effective clinical leadership, management capacity, or IT systems and capacity for data analysis. More pertinently, as evidenced in the USA,[45] effective and larger forms of integration may take a lot of time and autonomy to develop fully, often a decade or more.

Because of the central importance of a registered population and primary care in developing integrated care, and the weakness of practice-based commissioning, a lot of policy discussion has focused on whether practice-based commissioning could develop into integrated care perhaps initially by granting PBC groups 'hard budgets' to commission non-primary care. This may be the case, but as noted immediately above, a number of important ingredients need to be in place for integrated care to be effective. At present it is clear that most PBC groups lack most of these, and granting them hard budgets in isolation is not likely to be appropriate without significant other support, likely to come from the PCT or third party sources. Some PCTs remain unsupportive of these local initiatives, and again a firm steer from top management in the NHS will be needed to change this.

These significant initiatives will take time to develop and bear fruit. Another approach might be to look at Germany and Holland, where there has been a central policy push towards developing specific disease management programmes, for which large-scale service mergers may not be necessary.[44, 46]

The basic principle that prevention (primary or secondary) can improve health and reduce avoidable hospital costs must be a sound one on which to base policy. Yet as noted above, hard evidence on the impact of integrated care is as yet underdeveloped. This may be because many of the high performing integrated care examples exist in the US and the proprietary nature of these organisations may mean that evidence of impact is not shared publicly. It is also the case that evaluating the impact of complex health service changes is methodologically very difficult, a reason why the evidence available on impact is thin. Since integrated care holds significant promise for the NHS, it will be crucial that it is coolly and rigorously evaluated, especially if a diversity of forms is developed.

New and cheaper ways must be sought to monitor the impact of integrated care on quality and cost, for example which exploit routinely collected electronic data, are methodologically sound, and provide timely feedback to sites and the DH (for example not years after the start of an initiative). This is entirely possible given the data collected in the NHS, recent advances in data linkage at person level (anonymised) and risk stratification to identify control groups, and the state of health services research expertise. The evaluation of the whole system demonstrator programme which is piloting telehealth and telecare to support older people at home is an example where routine

data is being used to assess the impact on NHS use and cost.[47] A 'national tracking' facility could be set up to monitor the impact on service use and cost of all major initiatives across England, providing timely information.

Competition

Encouraging competition between providers for NHS funded clinical care has for some years been DH policy to increase the capacity available for care by drawing in non-NHS providers, allow patients greater choice of facilities for care, and to incentivise greater quality and efficiencies within NHS providers. Nationally, at most one per cent of NHS-funded care is currently provided for the independent sector. NHS-funded independent sector provision is thus still relatively small, is concentrated in few parts of the country, and focused on elective care.

It is likely that policies to encourage competition between providers will develop further in the future. Econometric studies suggest that in a market with regulated prices above the marginal costs of providing treatments, competition increases quality.[17] Two early studies in England are suggestive of the same finding.[48, 49]

But many of these studies examine competition for elective care. As noted above the biggest efficiency gains are likely in the treatment of frail older people, and those with multiple long-term conditions, for whom integrated care holds more promise. Integrated care has clinical collaboration at its heart. It is far less clear that competition can achieve as much gain in this area as in elective services, although it is entirely possible that competition among integrated care networks or organisations can exist and function well.

But integration is likely to be viewed as anti-competitive, if there are sufficient vertical or horizontal mergers of providers, or commissioner providers, and patients are encouraged to have care within this preferred network. This is true in many integrated care organisation in the US context, because quality and efficiency are more easily influenced and understood within the network. Competition occurs less between providers within the network, but between integrated care organisations for enrolees and for contracts from employer-payers. It is not clear from US examples of high-performing integrated care organisations (ICOs) the extent to which competition from other ICOs prompts greater efficiency. In some regions there is far more competition than others, for example more in North and South California where Kaiser Permanente operates, and very little in rural Pennsylvania where Geisinger Health System operates. In Geisinger, factors internal to the organisation itself (for example good information, clinical leadership and peer review, mission), rather than competition, drive efficiency and quality.[52] In an analysis of Kaiser Permanente's superior performance in a number of respects relative to the NHS, experts could not agree why, in particular on the role of competition.[50]

The current policy of free choice of provider for patients may conflict with the aims of integrated care to consolidate vertically to achieve efficiency and quality gains. There may be several options to consider for the future, for example offering people a free choice of facility for elective care only (using the HRG as financial currency outside of the ICO network), and the choice of which ICO to join. These are discussed in more detail elsewhere[40, 41, 51] and some current policies in fact point in this direction.

Conclusion

Because of the pressure on public finances, the financial settlement for the NHS over the next three to five years will be much lower than over the last ten years. The NHS will be forced to make cuts and efficiency savings on a scale not seen before.

This briefing outlines some of the major choices open to the DH in helping to make savings, and supporting NHS commissioners and providers make the right changes at local level. It argues that while incremental improvements in technical efficiency will be important to achieve, the scale of what is now needed requires more fundamental reform.

In the short term, the main levers open to the DH will be:

- cutting administrative costs and the costs of quangos
- performance management against targets
- holding down the tariff
- allowing reconfigurations of services including closures of facilities
- limiting the NHS-funded benefits available and increasing co-payments
- holding down pay settlements and the costs of procuring supplies.

More indirect levers include:

- providing more information and support on how to make efficiencies to NHS bodies

- changing the incentives in the tariff to reduce avoidable costs and poor care and increase preventive care in the community
- testing more radical forms of integrated care
- strengthening competition among providers to prompt better performance.

The focus of reform now and in the medium-to-long term should be how to reduce avoidable ill health and costs to the NHS arising from chronic disease and in particular those of older frail people. This can only be achieved by eliminating the barriers to care that currently exist that lead to uncoordinated care, duplication, reactive rather than proactive care, and dependency on the NHS for care, rather than a greater emphasis on self-management and self-reliance. In short, better integration of care is needed across primary and secondary care, between NHS and social care, and between NHS, social care and self-care. The next stage of NHS reform should encourage and test different forms of integration and their ability to reduce avoidable cost and ill health.

References

1. Appleby J, Crawford R and Emmerson C (2009) *How Cold Will it Be? Prospects for NHS funding 2011–17*. The King's Fund.
2. Department of Health (2009) *The NHS Operating Framework for England for 2010/11*.
3. Audit Commission (2007). *A Prescription for Partnership: Engaging clinicians in financial management*.
4. Billings J, Dixon J, Mijanovich T and Wennberg D (2006) 'Risk stratification in the NHS', *BMJ* 333, 327–331
5. Audit Commission (2009) *Auditors' Local Evaluation and Use of Resources 2008/09*.
6. Blunt I, Bardsley M and Dixon J (forthcoming) *Trends in Emergency Admissions in England 2004–2009*. Nuffield Trust.
7. Leatherman S and Sutherland K (2008) The *Quest for Quality: Refining the NHS Reforms – A policy analysis and chartbook*. Nuffield Trust.
8. Sutherland K and Coyle N (2009) *Quality in Health Care in England, Wales, Scotland and Northern Ireland: An intra-UK chartbook*. The Health Foundation.
9. Department of Health (2008) *Long-term Conditions: A compendium of information*.
10. Smith J, Curry N, Mays N and Dixon J (2010) *Where Next For Commissioning in the English NHS?* Nuffield Trust.
11. Audit Commission (2009) *More for Less. A briefing*.
12. Tribal (2009) *System Transformation in the NHS, Quality, Innovation, Productivity and Prevention (QIPP): The Tribal approach*.
13. Featherstone H and Evans N (2010) *Controlling Public Spending: The NHS in a period of tight funding. Research note*. Policy Exchange.
14. Wood C (2009) *The NHS in an Age of Progressive Austerity. A joint DEMOS/Tribal roundtable*. DEMOS.
15. The Quality, Innovation, Productivity and Prevention programme
    Available at: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Healthcare/Highqualitycareforall/Qualityandproductivity/index.htm
16. New Frontiers in NHS Efficiency
    Available at: www.nuffieldtrust.org.uk/projects/index.aspx?id=49
17. Gaynor M (2006) 'Competition and quality in health care markets', *Foundations and Trends in Microeconomics*, 2, 6.
18. Bardsley M (2010) Personal communication.

19. http://en.wikipedia.org/wiki/Roemer's_law

20. Bloom N, Propper C, Seiler S and Van Reenan, J (2010) *The Impact of Competition on Management Quality: Evidence from Public Hospitals. CEP Working Paper – February 14th, 2010 Draft*. London School of Economics.

21. Dixon J and Welch WG (1991) 'Priority-setting in the NHS: Lessons from Oregon', *Lancet* 337, 891–894.

22. Assato J (ed.) (2006) *Charging Ahead. Spreading the costs of modern public services*. Social Market Foundation. Available at: www.smf.co.uk/assets/files/publications/Charging%20Ahead%20-%20Spreading%20the%20costs%20of%20modern%20public%20services%20final%20draft.pdf

23. Secretary of State for Health (2008) *High Quality Care for All: Next Stage Review Final Report*. Department of Health.

24. Department of Health (2004) *Agenda For Change Final Agreement*. Available at: www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4095943

25. National Audit Office (2009) *NHS Pay Modernisation in England: Agenda for change*. Available at: www.nao.org.uk/publications/0809/nhs_pay_modernisation.aspx

26. *General Practice in England: An overview*, King's Fund Briefing, September 2009.

27. House of Commons Health Committee (2007) *Workforce Planning: Fourth Report of Session 2006-07, Vol 1*. HC 171-1. London: The Stationery Office. Available at: www.publications.parliament.uk/pa/cm200607/cmselect/cmhealth/171/171i.pdf

28. HM Treasury (2009) *Operational Efficiency Programme*. Available at: www.hm-treasury.gov.uk/press_40_09.htm

29. HM Government (2009) *Putting the Frontline First: Smarter Government*. Available at: www.hmg.gov.uk/media/52788/smarter-government-final.pdf

30. National Audit Office and Audit Commission (2010) *A Review of Collaborative Procurement Across the Public Sector*.

31. House of Commons Health Committee: Commissioning (2010). *Fourth Report of Session 2009/10. 30 March 2010*. Available at: www.publications.parliament.uk/pa/cm200910/cmselect/cmhealth/268/268i.pdf

32. Goodwin N, Mays N, McLeod H, Malbon G and Raftery J (1998) 'Evaluation of total purchasing pilots in England and Scotland and implications for primary care groups in England: personal interviews and analysis of routine data', *BMJ* 317, 256–259.

33. Dixon J and Glennerster H (1995) 'What do we know about fundholding in general practice?' *BMJ* 311, 727–730.

34. *The Productive Ward Module Structure*. NHS Institute for Innovation and Improvement Available at: www.institute.nhs.uk/quality_and_value/productivity_series/the_productive_ward_module_structure.html

35. Department of Health (2007) *Value for Money Delivery Agreement 2008 to 2011*. www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_081547

36. Open source software available at: http://en.wikipedia.org/wiki/Open-source_software

37. Ham C (2009) *Health in a Cold Climate: Developing an intelligent response to the financial challenges facing the NHS*. Briefing paper. Nuffield Trust.

38. Rosen R and Ham C (2008) *Integrated Care: Lessons from evidence and experience*. Briefing paper, Nuffield Trust.

39. Ham C (2009) *Only Connect: Policy options for integrating health and social care*. Briefing paper, Nuffield Trust.

40. Lewis RQ, Rosen R, Goodwin N and Dixon J (2010) *Where Next for Integrated Care Organisations in the English NHS?* Nuffield Trust.

41. Smith JS, Wood J and Elias J (2009) *Beyond Practice-based Commissioning: The local clinical partnership*. Nuffield Trust.

42. Gleave R (2009) *Across the Pond: Lessons from the US on integrating health care*. Nuffield Trust.

43. Dixon J, Lewis R, Rosen R, Finlayson B and Gray D (2004) 'Can the NHS learn from US managed care organisations?' *BMJ* 328, 223–225.

44. Rosen R and others (forthcoming) *Integrated Care: Lessons for successful implementation from international case studies*. Nuffield Trust.

45. Minott J, Helms D, Luft H, Guterman S and Weil H (2010) *The Group-employed Model as a Foundation for Health Care Delivery Reform*
Available at: www.commonwealthfund.org/Content/Publications/Issue-Briefs/2010/Apr/The-Group-Employed-Model-as-a-Foundation-for-Health-Care-Delivery-Reform.aspx

46. 'How Germany is reining in health care costs: an interview with Franz Knieps', *McKinsey Quarterly*, February 2010.
Available at: https://www.mckinseyquarterly.com/Health_Care/How_Germany_is_reining_in_health_care_costs_An_interview_with_Franz_Knieps_2534

47. *Evaluation of the Whole System Demonstrator Project*. Nuffield Trust.
Available at www.nuffieldtrust.org.uk/projects/index.aspx?id=294

48. Gaynor M, Moreno-Sienna R and Propper C (forthcoming) *Death by Market Power: Reform, competition and patient outcomes in the National Health Service*. CMPO.

49. Cooper Z, Gibbons S, Jones S and McGuire A (2010) 'Does hospital competition save lives? Evidence from the English NHS Patient Choice Reforms', *LSE Health Working Paper No 16*.

50. Feachem R, Sekhri N and White K (2002) 'Getting more for their dollar: a comparison of the NHS with California's Kaiser Permanente', *BMJ* 324, 135–143, with subsequent commentary by Dixon J, Berwick D, and Enthoven A.

51. Dixon J, Le Grand J and Smith P (2003) *Can Market Forces be Used for Good?* The King's Fund.

52. Steele, G (2009) Personal communication.

The Nuffield Trust

The Nuffield Trust is a charitable trust carrying out research and health policy analysis on health services. Our focus is on the reform of health services to improve the efficiency, effectiveness, equity and responsiveness of care.

Key current work themes include:

- new forms of care provision
- commissioning
- efficiency
- national and international comparisons
- competition policy.

This briefing, which was commissioned by Tribal, forms part of our efficiency work theme. A major programme of Trust work in this area, New Frontiers in NHS Efficiency, is currently underway.

For more information on our work programme and to sign up to receive our regular e-newsletter, visit www.nuffieldtrust.org.uk

Published by the Nuffield Trust.

Available to download at www.nuffieldtrust.org.uk

The Nuffield Trust
FOR RESEARCH AND POLICY
STUDIES IN HEALTH SERVICES

The Nuffield Trust
59 New Cavendish Street, London W1G 7LP
Tel: 020 7631 8450
Email: info@nuffieldtrust.org.uk
Fax: 020 7631 8451

Printed on paper from renewable sources