# What's stopping efficient change in the NHS – budget silos or power silos?

Andrew Walker
Robertson Centre for Biostatistics
University of Glasgow

Andrew Perry
Global Health Outcomes
GlaxoSmithKline Research and Development,
Brentford

Paul Trueman
Health Outcomes
Johnson and Johnson

**Paper presented at the Health Economics Study Group, Brunel, 3-5 July 2002**

## Introduction: What is the Problem?

At a conference on heart disease, a GP was recounting his experience of trying to implement guidelines on the secondary prevention of heart disease through using statins in people with hypercholesterolaemia. He explained the process of identifying patients from practice records, contacting them, calling them in for review, explaining and starting treatment, checking cholesterol levels and titrating doses, and so on. At the end of this, he asked, who had thanked him? Did the patients thank him? They had no symptoms of high cholesterol levels in the first place, he had changed their medication (with some side-effects) and they felt no better by the end. Did his partners thank him? He had created lots of extra work, tying up the admin staff in the practice for some time. Did the local health board think him? They sent the medical prescribing advisor to visit him and he was asked why his statin prescribing level was three times the local average! His final statement was, "So next time I get sent an evidence-based guideline to implement, why should I bother?"

Seen from the GP's perspective, the incentives to change rest upon how they feel about the effort involved relative to knowing that for every X patients treated, the statistical expectation is that there will be one fewer event over the next five years. The beneficiary cannot be identified. The GP may well conclude this is a relatively weak incentive compared to the more tangible benefit of dealing with the management of symptomatic disease.

This paper describes a project that considered one aspect of the incentive problem, the influence of budget arrangements in the NHS. In other words, can the NHS change the way it arranges its budgets to give a stronger incentive than that described above?

## Budgets and incentives in the NHS

Arrangements for allocating and managing budgets within the NHS impact upon decisions about the allocation of resources. These can be divided into three main issues: inability to switch (or vire) between budgets over time, inability to vire between budgets within the NHS (and public sector) at any given point in time, and the lack of short-term relationship between financial costs and workload. Each of these factors creates different incentives.

### *The 'break-even' requirement*

It is a statutory requirement for each organisation in the NHS to stay within its budget in each financial year. The rationale, of course, is in terms of public accountability for use of government revenue. The incentives created can be referred to as "short-termism" in thinking. For example there is less incentive to spend money in the current financial year that will produce benefits in the future. This will have most impact on programmes in primary and secondary prevention of disease, such as health promotion, screening, and pre-emptive treatment. Unfortunately, viewed from a long-term, societal perspective these are some of the most cost-effective health interventions.

Spending more than the budget is penalised, even if the money is spent on the most cost-effective service available. (Interestingly, spending under the budget level is indirectly penalised if the budget in future years is reduced by this amount.)

***Devolved budgets under rigid headings***

Budgets are devolved throughout the NHS, from government to local health authorities, then to trusts, then to divisions within them, whether these are specialty groups (surgery, medicine, etc.) or functions (pharmacy, nursing, etc.). This allows for local control and flexibility in management.

The problem with fixed budget headings (or "budget silos") is that without a method of transferring savings then even when one sector invests to reduce the workload of another sector (for example primary care prescribing that reduces hospital admissions) the resources freed are retained within that sector and redeployed by those responsible for that budget. If drug prescribing in primary care results in fewer cardiology admissions, the staff time freed up will be redeployed within the ward, or used to take cases off the waiting list at a faster rate. Even if the saving is a financial one, it is still likely to be retained within the budget – for example, early discharge results in prescribing cost savings in hospital, but these are retained and reinvested. This is a problem of virement i.e. a failure to shift savings from one administrative budget to another.

***Financial costs are unresponsive to changes in workload in the short-term***

The NHS is labour-intensive: around two-thirds of expenditure is on staff - in Lothian, the area around Edinburgh, there are 18,000 NHS staff, which is between 2 and 3% of the resident population of all ages). Staff numbers (and hours of work) often do not respond to short-term changes in workload. For example, a cardiology ward generally has the same level of staffing whether bed occupancy is 80% or 85%. As a result, financial costs are fixed to a large degree, at least in the short- to medium-term. Since about half of the remaining non-staff costs are equipment and bricks-and-mortar (also unresponsive to workload changes), considerable attention is therefore focused on prescribing budgets when short-term savings are required.

More generally, the fact that costs do not vary with fluctuations in workload has two consequences:

(i) It is very difficult to save money, although it is often possible to estimate what resources will be freed up for other uses – in other words, there is more likely to be an economic saving than a financial saving. This then begs the question of how decisions are made about how freed resources are reallocated.

(ii) Because it is so difficult to make financial savings through "disinvestment" there is intense pressure on any increases to the budget (which we refer to below as growth money). Some people perceive that the historic allocation is very hard to reallocate and that most priority setting decisions focus on the growth money. Others are more prepared to challenge the status quo, but face considerable difficulties if costs are as fixed.

The lack of responsiveness of financial cost to workload was arguably one of the factors behind the (perceived) failure of the NHS internal market reforms of the early 1990s –

the original principle of "making the money follow the patient" so that hard work was rewarded seems a sound one from an economic efficiency point-of-view (subject to the quality of the product!) However, the inability or unwillingness of trusts to free up resources when workload was moved caused problems for those involved in contracting.

***So what?***

In summary, there is a danger that budget arrangements will inhibit change through creating incentives to ignore (i) impacts elsewhere in the system and (ii) long-term consequences. This is reinforced by the fact that resources freed rarely equate to financial savings, which creates scepticism about the whole concept of "savings" among decision-makers (see below). Finally, this leads to the perception that new services are "not affordable".

The incentives created apply across the NHS. They do not prevent change but they have the potential to significantly slow it. This is highly relevant to assessing the impact of results from economic evaluations to day-to-day NHS resource allocation decisions. Economic studies attempt to take a long-term, societal perspective on decision-making, which is precisely what it has been argued that budget constraints do not allow decision-makers to do. Qualitative research of NHS decision-makers' views confirms that these budget factors are an important (if not the most important) barrier (Drummond et al, Hoffman et al). When asked about barriers to using these results, decision-makers such as prescribing advisors cite as important factors "the savings are not real" and "resources cannot be moved around in the way suggested".

Economists must carry some of the blame for being careless in how the word "savings" has been used, sometimes claiming that new services will "pay for themselves". The distinction between economic and financial savings is one that has eluded many potential users of economic evaluation, as well as economists themselves on some occasions. Experiments with programme budgeting and marginal analysis in the 1990s ran into problems when it came to recommendations about redeployment of resources within a programme (the so-called micro level). To take the example of gynaecology services in Glasgow (Twaddle and Walker), the intention was that change could be achieved "within the existing envelope of resources", in other words that savings could be made and re-cycled within that programme. However, reductions in activity in one area did not realise the cash savings to pay for new members of staff elsewhere. The reason was that in the PBMA exercise the reduction in activity was valued using market prices – while these (supposedly) reflected the opportunity cost of staff time, they did not reflect the cash freed up when one less unit of activity was undertaken. They were the marginal economic cost, but not the marginal financial cost. These "economic savings" did not translate into financial savings because the resources freed were used for other purposes.

## The approach used

To explore the issues further, ways in which the incentives could be changed were drawn up, including consideration of the principles involved, how the option would work in practice and the issues that might arise.

To test the ideas, they were discussed with a sample chosen to represent a cross–section of health service decision-makers. These included medical prescribing advisors, managers, Scottish Executive Health Department officials, and staff involved in primary care. (While the work was carried out in Scotland, many of the issues are common to the whole of the NHS.)

The interviews used a loosely structured approach covering the impact of budgets on decisions, the general area of the impact of economic evaluation, and the list of ways to change incentives. As an example to help focus the discussion, the interviews considered the problem of a drug prescribed in primary care that prevented admissions to secondary care in the future. The use of statins for hypercholesterolaemia in secondary prevention of coronary heart disease might be one example. The incentive issues listed above would suggest that the take-up of such a drug would be slow as GPs were asked to incur extra prescribing costs to achieve "savings" (in the form of freeing-up valuable resources for other uses) in the hospital.

**What can be done?**

Some of the problems are less easy to change than others: for example, the statutory duty to manage within a given budget each year is taken very seriously. On the other hand, the NHS in Scotland now announces what the minimum budget for NHS boards will be for the next three years to allow some stability in planning. In England foundation hospitals appear to have greater powers to borrow money (in a similar way to NHS trusts in 1991); however, a commercial loan might require a heavier emphasis on making financial savings than economic savings (freeing up resources for other uses).

The main focus of this work was smaller scale change. This paper considers seven possibilities, as follows:

1. There should be a unified budget at local level – extra costs and "savings" would thus come under a single heading, addressing the virement problem.
2. The GP would be given an incentive to prescribe by being given some influence on how the resources freed up would be re-deployed.
3. There would be a medium-term plan to "cash-in" the resources freed up as a result of the prescribing with savings being returned to the GP.
4. Costs could be made more responsive to changes in activity in the short-term so that increased prescribing led to a short- to medium-term pay-off.
5. The GP would be protected against criticism from budget controllers if they overspent by being able to demonstrate that this was as a result of expenditure on services that were acknowledged to be cost-effective.
6. The local health authority might keep a fund of money for cost-effective schemes where the incentives for the GP to act is weak.
7. The primary care prescriber should not be expected to make judgments on long-term societal cost-effectiveness: this decision should be made by others, who do not face the virement / fixity problems described above, with the GP implementing their guidance.

Of course, this list is not exhaustive and the options are not mutually exclusive.

### *1. A truly unified budget at local level to resolve virement problems*

*In principle*
Under this option, the local health service would have a single, unified budget – extra spending would thus come back to the same fund. This should help to address virement issues in particular. Local decision-makers are given a financial incentive to act in an efficient way because the resulting savings can either be reinvested or converted into other benefits such as improvements to a GP surgery.

This could take several different forms:

- Primary care led commissioning – statin prescribing from this budget would lead to fewer cardiology admissions being purchased so savings would result. This is the model used in the NHS in England and Wales.
- A single budget for the local health service – all local health services are united under a single body with a single budget. Prescribing statins leads to savings elsewhere but these will come back to the unified budget. This is the model used in the NHS in Scotland.
- A budget based on some other grouping, such as a network of clinicians managing a particular disease, akin to the programmes in a programme budget. Statin prescribing and reduction in future resource use would come under the coronary heart disease network or programme so again savings would accrue to the same budget as the costs.

*To make it work*
In the commissioning example, there will only be an incentive from savings if providers are remunerated in a way that means expenditure varies directly with workload. Examples might include contracting on a case-by-case basis or fee-for-service, but both are administratively expensive. The alternative is that negotiation must take place on how costs will be reduced (see the fourth option in a later section). The unified budget at local level introduced in Scotland avoids some of these problems (but the incentives will be weaker as a result).

In terms of the disease budget, NHS Scotland is developing managed clinical networks. These link together clinicians managing a disease across specialties, trusts and NHS boards if necessary. Early experience is with management of different cancers and to date no recurring budget has been devolved to these groups.

*Issues*
This set of solutions address the virement issue, but not resource fixity. There will only be cash savings from a reduction in staff – for instance when a consultant cardiologist post became vacant, it would be scrapped. Given the priority attached to cardiovascular disease, this is unlikely. If cardiology admissions are reduced, the hospital still has the

same total costs because these are fixed in the short-term; the average cost per case thus rises by a small amount for all remaining cases.

There is also an issue about the strength of the incentive created. In the commissioning case, the GP might be incurring extra costs so that the primary care organisation as a whole can benefit; this might seem a little abstract to the individual in the context of a multi million pound budget for each primary care trust. That GP would be benefiting the "greater good" but might see very little change in the resource constraints they face. A similar argument applies in the Scottish case where the unified budget is held at local health board level.

The disease budget idea would help to encourage investments in the early stage of the disease if the resources freed were retained. However, some people in the network would relinquish control of resources to others. In the case of diabetes, for example, vascular surgery would be reduced to pay for more chiropody (where preventive work on caring for toes and feet can prevent the need for future amputations by vascular surgeons). This would probably not be acceptable to those involved. Another problem is that many health care workers manage several different diseases (e.g. general surgeons) while others (such as GPs) have to manage every type of disease. They would thus be required to be in several networks simultaneously, making planning impossible from their point-of-view.

*Comments from sample*

There was optimism about the potential for this option, especially amongst interviewees who took a strategic perspective of the NHS, although with the proviso that changes in thinking produced by this budgetary arrangement might not occur "overnight".

Most respondents focussed on the primary-secondary care interface as the issue they would most like to see this type of option address. Some saw the potential for savings while others were concerned that secondary care might actually be destabilised by the changes. In one health board area, the example was given of warfarin monitoring having passed from secondary to primary care in the recent past without any resource moving as well; by contrast, the new unified budget arrangement meant that as the bulk of diabetes care was transferred to primary care, resource of several hundred thousand pounds (in the context of a population of around 700,000) would follow. Despite this ability to transfer resource there was still a concern that the vision for the future of health services (in terms of a strategic plan) was lacking. There was some suspicion, however, stemming from an earlier scheme in Scotland, the Joint Investment Fund. (This was a mechanism to encourage discussion about redesign of services across traditional boundaries. While it was not essential to the concept that money was attached this is the way that many people perceived it. Therefore, in the minds of many, the lack of perceived progress with the JIF has created scepticism about unified budgets at "grassroots level").

It was suggested that Scotland had a unique experiment in progress in that the Executive has created unified budgets *and* unified health boards (i.e. a single local health authority board that combines health board commissioners, trust providers and local authority

councillors). This was thought by some people to create opportunities that did not exist in England and Wales.

### *2. Incentives to change via influence on how economic savings are redeployed*

*In principle*
This option accepts that financial savings are unlikely because costs do not respond to changes in workload. Instead, it would create an incentive for prescribers by giving them some influence over how the freed resources are redeployed. In the statin example, free cardiology resources might be used to reduce waiting times for other procedures or to start an open access diagnostic service. Prescribers, in this case GPs, would have some say in this decision. This would acknowledge the virement and fixity issues by letting the freed resources stay in their present environment but directing their use.

*To make it work*
Agreement would be needed about the timing and nature of resources freed – this might be possible in the case of statins where a reasonable evidence base is available, but might be more difficult for other interventions. Those involved (in this case GPs and cardiologists) would also have to agree to the changes.

*Issues*
It is unclear how strong an incentive this would be for GPs. While some are undoubtedly motivated by the provision of a high quality service, it is not clear whether this promise of improvements at some point would be sufficient to induce widespread behaviour changes.

There is also a danger that a single provider dealing with different primary care organisations might have to enter separate agreements with each so that variants on the service existed, with the patient's postcode determining access.

*Comments from sample*
Those who were more positive about this option saw some parallels with GP fund holding. Where this had matured, the 'power game' aspect had been overcome and sensible discussions focussing on meeting needs had ensued: money and contracts were then a means to the end of better quality care.

In general, however, this option attracted scepticism. It was anticipated that there would be a lot of cynicism about schemes such as this among doctors in general and GPs in particular. This was confirmed by primary care staff interviewed: a non-specific promise of a better quality service at some point in the future would be a weak incentive to change – only savings that could be reinvested would be sufficient to induce a serious change in behaviour. Even if this could be overcome, it was anticipated that there would be a power struggle with secondary care to achieve change.

### *3. A medium-term plan for change allowing planned overspends in short-term*

*In principle*
This option insists that financial savings from a change are realised, although this might take some time. For example, there might be an agreed five-year plan to recoup some of the 'savings' of statin treatment, based on the predictions from an economic model on falls in event rates and consequently on resource use. Providers would have to find a way to manage these cash releases, either through reducing staff numbers or by other efficiency savings.

*To make it work*
A reliable model of the impact of additional prescribing is required, adapted to reflect local circumstances as far as possible. The provider must also agree to release cash savings as a result and have a plan on how to achieve this.

*Issues*
There is also considerable uncertainty about modelling treatment effects. Some are widely recognised, such as predicting effects beyond the follow-up of clinical trials, but others are quite pragmatic in nature e.g. determining what proportion of patients eligible for statins are receiving them at the outset.

Some form of monitoring would be required to ensure that the predictions of the model are roughly accurate. The agreement would have to take account of what would happen if the predicted falls were not seen – would the provider still be expected to make some proportion of the savings? There are analogies here with the recent risk-sharing scheme for multiple sclerosis drugs.

The planning required to achieve this might be expensive in terms of local management time – for example, few local commissioners have economists capable of devising the model required or adapting existing ones. The plan might also have to be renegotiated with significant new entrants, changes in the evidence base, drug prices changing, and so on.

More broadly, there is the issue of how the overspend on prescribing in the short-term would be financed. Local health authorities would have to choose which other programmes were to be sacrificed or delayed to achieve the change in the first place.

*Comments from sample*
The pre-requisite here was thought to be that all of the parties, (most notably the trust that would release the savings) would have to be "signed-up" to the idea in the first place. It might work where there was a genuine appetite for partnership, but would run into problems otherwise.

### *4. Make costs more flexible in the short-term*

*In principle*
This would focus on the issue of fixity, aiming to make costs more responsive to workload. A measure that reduced activity in some respect would realise financial savings as well as economic savings, at least in the medium-term.

*To make it work*
In the most literal interpretation, this would work by reducing the length of job contracts and reducing employment protection. It is slightly more realistic to think in terms of greater reliance on agency staff hired on a sessional basis, leasing of equipment, emphasising transferability of staff and equipment within an organisation, etc. The main problem with "making it work" would be the reaction of staff and unions to such moves.

*Issues*
Deciding on a day-by-day (almost hour-by-hour) basis how many resources (staff, equipment, etc.) would face a host of problems. First and foremost, employers have a duty (moral and statutory) to staff in terms of good employment practices; these could well be breached by such developments. There would be strong objections from unions, not least because of the impact this would have on quality of care if all continuity of staffing were lost. For example, there would be great pressure on staff to be generalists rather than specialists, as the former would maximise their employment prospects. The longer-term impact through damage to staff morale should also be considered. Secondly, some types of work in the NHS are not sufficiently predictable to say that a sudden increase or decrease in work is going to be permanent.

It also ignores the existence of excess demand for health care; any reduction in one type of work is likely to be replaced by other work. If statin use reduces the number of bed-days managing myocardial infarction, cardiology resources might be reallocated to carrying out more angioplasties (PTCAs) – if the latter use coronary artery stents (as NICE recommends!) at £500-£1,000 each then total costs will actually increase. All the staff are still employed, but the stent budget has now gone up as well.

*Comments from the sample*
On one level, this was thought to be happening already: for example, in primary care traditional professional boundaries were being blurred and 'multi-skilling' was becoming more common. (Authors' note: This is undoubtedly contributing to improved efficiency and quality of care, but it is not clear that this is realising savings). In hospitals, flexibility was achieved through using agency staff to cover workload peaks, although this carried its own problems. On a more fundamental level there was scepticism because most costs are tied up in health service institutions such as hospitals. Interviewees from strategic levels reported interest in the area, but awareness of the practical difficulties involved.

It was also pointed out that even staff appointed to address one problem could create another. As an example, specialist nurses employed in the community to manage disease X (such as asthma) might be better redeployed to another disease in the medium-term,

but if they are regarded as “disease X specialists” they cannot be shifted. The ‘silo’ is completely artificial but an institution is nonetheless created.

### *5. Incentives to change irrespective of overspending so long as this results from efficient practice*

*In principle*

Accepting that savings are unlikely to occur, this option sets out at least to ensure that the GP is not penalised for acting in a cost-effective way. For example, if a local health authority had acknowledged that statin prescribing as secondary prevention were cost-effective, then a GP who had overspent their prescribing budget would be excused if they could demonstrate that this was as a result of pursuing prescribing in this area.

*To make it work*

A recognised source of agreed and reliable evidence would be required. However, health authorities might pay particular attention to the budget impact of a new technology before it was accepted as cost-effective.

*Issues*

There are two main problems with this option. First, it does not address either of the problems of virement or fixity. Secondly, there is a lack of cost-containment. Once a technology has been accepted as cost-effective there would no longer be a budget constraint on its use, so long as this is in line with indications. In addition, it is not clear where the money to fund any over-spend should come from – local commissioners may become very defensive and try to keep a substantial local contingency reserve to cover end-of-year overspends.

More pragmatically, it might be difficult to pin an overspend down to a single drug in the way this option assumes. It is not clear whether the existing data sources could become capable of such a fine level of analysis. For example, could they distinguish prescribing as secondary prevention and as primary prevention? Could they spot prescribing for secondary prevention that meets accepted guidelines from that which does not?

*Comments from sample*

The view from primary care was that a fundamental requirement was to stop penalising people who were “doing the right thing”. GPs often felt that they were penalised when they were trying to do their best for their patients, and this did not create an atmosphere conducive to future change. Others interviewed pointed out that national standard-setting bodies aimed to give individuals and local health services a basis for evidence-based action that could be used to get managers to release resource.

Again, there were concerns about how this worked in practice. One interviewee pointed out that in practice it is hard to pinpoint the cause of an overspend because existing information systems just aren't good enough. From a prescribing advisor's point-of-view there is a danger that any overspend can be 'justified' by appeal to a particular guideline!

Another approach was to improve the dataset available to managers by using performance indicators that show whether the overspend is based on high quality care. Experience suggested that GPs do respond positively to these and that they have to be carefully chosen.

### *6. A dedicated fund for investment in interventions of proven cost-effectiveness*

*In principle*

In this option the local health service would establish an 'economic efficiency fund' (possibly a sexier title could be devised!) by protecting some amount of growth monies or money that would otherwise have been allocated to trusts ("top-sliced"). This would be used for the introduction of interventions that are cost-effective but do not yield financial savings, short-term savings or savings within the same care sector.

*To make it work*

Rules would have to be agreed on the size of the fund, the services that qualified (through having good quality supporting economic evidence) and when funding for these new services should shift to "mainstream" funding.

*Issues*

The fund would be financed from the top slicing of revenue funds so local services would bear a hidden opportunity cost since funds that would have been allocated to them are now diverted.

It does not address the problems of virement or fixity, but uses new money to steamroller them.

The size of the fund may be uncertain from year to year, especially if it represents the amount that is left after all other essential calls on the available budget have been made. Allocating a given percentage of the budget allocation to the fund would help, but might destabilise providers in the process.

*Comments from sample*

To an extent this was felt to be what was happening already when local health authorities held service development funds that could be invested in different services. In Scotland, this was also the thinking behind the Health Improvement Fund (which was set up as a means of investing revenue from the so-called "tobacco tax" in public health initiatives). There was scepticism among the sample about the political will to forego short-term benefits in order to invest in prevention for long-term gains.

### *7. Cost-effectiveness judgements are made at a national level, with local services merely implementing this guidance*

*Rationale*

In practice, local decision-makers are constrained from taking a long-term, societal view on the development of new services. National decision-making bodies (such as NICE) do not face the same constraints so they are better placed to determine efficient service changes. It is then the job of the local health service to implement this guidance rather than making decisions themselves.

*To make it work*
This option needs a national decision-making body that makes recommendations on the basis of economics evidence. It then needs to disseminate this guidance, convincing local services to follow it either because of their respect for the national body or because they are forced to do so through a management hierarchy. This option would also require a system of monitoring to ensure that the guidance was being followed.

*Issues*
Such a system might work reasonably well for well-defined 'one-offs' (such as NICE HTA guidance) but it is less clear how it will influence day-to-day decisions (e.g. NICE clinical guidelines). Even for 'one-offs' the system is imperfect (see Cookson et al).

In itself, this option does not resolve the problems of fixity and virement. Rather than addressing the disincentives it places a duty on people at a local level to act despite them. The 'price' to be paid is that money to fund these services must then come from elsewhere in the local health service so there will be opportunity costs in terms of other developments (and their consequent health benefits) foregone.

*Comments from sample*
This option was welcomed in terms of controlling entry of expensive new technologies, although it also created uncertainty in terms of when guidance would be launched and what its budget impact would be. However, by making guidance mandatory, the NHS has reduced the ability of local health authorities to manage the introduction of new technologies to fit the existing financial framework. The point was made several times that national decisions such as this distort local priorities.

Many interviewees commented on implementation problems. One familiar analogy likened the NHS to a super-tanker, slow and unresponsive in its handling. The extent of responses to national guidance, if any, were thought to depend upon the objectives and attitudes of clinicians, which are slow to change. The point was made that while fixed budgets were a reality to management, at practice level the focus was still on the needs (or, perhaps more accurately, the demands) of the individual patient.

Views varied on how to respond to clinicians who did not agree with the guidance. One point-of-view held that some form of credible threat was required, while others felt that an appeal to "best practice" had a more positive image (although this could be backed up with an unspoken threat in that trusts controlled budgets).

**Discussion**
All of the interviewees recognised the issue of budget silos as a potential barrier. Problems in shifting work and resources from secondary to primary care are widely recognised. However, interviewees identified two areas in which the argument should be extended:

- the silo problem now goes beyond the NHS context in which it was discussed (mainly prescribing) as the new "Joint Futures" initiative that moves towards joint management of health and social care covers 15% of the NHS budget. The "silo" problem is common to both agencies. From the NHS point-of-view, acute beds are used by patients awaiting discharge to nursing homes or supported living in the community, both of which are paid for by social work. From the local government point-of-view, the NHS advocates a variety of ways to improve health (such as improved housing) that will involve costs that result in economic savings for the NHS. In both cases, the incentives to act are poor.
- secondary care develops services that take pressure away from primary care such as open access service or better A&E facilities. There is no expectation that resources will then switch from primary to secondary care!

***Which way to go?***
None of the options considered provides a complete solution. The first three options address virement issue – these were the unified budget, influencing redeployment of freed resources, and a five-year plan to realise savings. However, these tended to produce relatively weak financial incentives, unless the option of GP commissioning at practice level is considered. For these options to work, therefore, GPs would have to be quite strongly motivated by quality of service issues and to have faith that deferred gratification would eventually occur. Some interviewees were especially sceptical on the second count.

The fourth option addressed fixity. The practical difficulties are enormous, and there is a threat to quality of care in the process. While interviewees felt that flexible working practices were one step in this direction, it could be argued that these were actually addressing virement issues if it led to staff working across care settings.

The final three options try to push through cost-effective change in the face of virement and fixity problems by either making change a managerial imperative (in the case of national guidance) or by 'ear-marking' new money (in the case of the economic efficiency fund). The seventh option attempted simply to stop individual clinicians being penalised for acting in line with economic evidence, but the practical difficulties were considerable.

Possibly the most promising option was the five-year plan – this addressed silos and cost fixity to some extent, although the key was negotiation and co-operative working.

***Two key issues***
The solutions above address disincentives created by budgets through budgetary reform. But will this bring about the desired change? The scepticism of the sample might reflect two things. First, the solutions may not have adequately analysed what motivates people and organisations to change their behaviour. The second (and related) issue is the existence of "power silos" alongside the "budget silos" described. A "power silo" is a concentration of resources and influence in an organisation, team or individual. It may be that barriers to change (specifically on the basis of cost-effectiveness data) that are attributed to budget silos are in fact manifestations of a power silo problem.

***Motives for change: the utility function of the individual or organisation***
A general observation was that the interviews revealed a lack of clarity about what motivates individuals and organisations (or utility functions). This came out in a number of ways:

- At health board level the popular view was that the important motivators are (i) a recommendation of an organisation such as direction from the NHS Executive to implement NICE guidance, (ii) more generally, an imperative to show measurable change in achieving government targets, and (iii) on specific issues pressure from politicians (such as a question in Parliament). This must be kept in context, however: one interviewee reported that drug treatment for hepatitis C had been funded because (i) there was a trusted local doctor prepared to work to a tight protocol, (ii) the acute trust was "on board" with the development, and (iii) there was a good evidence base. In this case, the cost was thought to be justified despite the difficult overall financial situation at the time.

- At acute trust level, delivering on short-term political imperatives was very important. One reported difficulty was that there was a lack of appreciation amongst politicians and the public that policies such as having local access to specialist facilities came at a cost – one of the examples given was the development of local renal dialysis services.

- Within primary care organisations, peer pressure was seen as important: one board top-slices money and releases it mid-year in line with performance indicators but pays out by local health care co-operative (LHCC), not by individual practice, thus creating peer pressure within the LHCC.

- At GP level there was thought to be a spectrum of motivators, from 'pure' financial concerns to 'pure' quality concerns irrespective of cost. Interviewees felt that most GPs fell somewhere in between these two extremes in an approximate normal distribution. The seven options above were thought to be deficient because they focussed on financial motives alone. Schemes such as a quality prescribing award and other means of peer recognition & approval were felt to be at least as powerful as money. One interviewee felt that quality improvements to the service were the key to raising GP morale. However, if mild

'threats' were required then clinical governance and risk management were quite persuasive ideas.

The research was mainly targeted at local factors but for completeness the motivators for the Department of Health (and for the government) would also have to be considered.

Given the disincentive problems listed, some of the later interviews also sought reasons why change *did* occur. For example, the question was posed, "Given the range of disincentives to use statins that we have talked about why is prescribing not 0%?" The reasons listed were:

- Commercial pressures from manufacturers, based on good quality clinical evidence of efficacy in promotional material for GPs.
- GPs desire to provide a good quality service
- Conversion of cardiologists to statin prescribing, with a knock-on effect in terms of influence on GPs.
- A "new blood" effect whereby natural staff turnover resulted in recently trained staff who are more familiar with the recent evidence and have less tradition in terms of the way they practice.
- Prescribing advisors 'badgering' people who prescribe at a very low level

This suggests that the general wish to see a quality service has an impact but it can be slow, even when the evidence from trials is as unequivocal as in the case of statins – back to the analogy of the super-tanker. Financial disincentives seem to slow this process rather than completely inhibiting it.

***Power silos – the real problem?***

While decision-makers at different levels may face disincentives, this only seems to become a problem when they also have the power to resist change they do not desire. Clinicians being "guided" how to treat particular conditions is one example – the key issue is then how much power the government and managers have to impose change. It may be that to even think of "imposing" change is the wrong route – many of the interviewees, especially those involved in primary care, emphasised that persuasion and co-operation was a more fruitful way forward.

The later interviews implied the existence of what might be called "power silos" to parallel the "budget silos" usually referred to. Barriers are not just administrative budgets, but also people, cultures and history. Traditionally, these have been notable in the acute sector where beds and theatre time were jealously guarded and not surrendered without a fight. In other words, the power silos reinforce the extent to which costs are fixed – long-term contracts suit those who hold the power very nicely. Other interviewees were surprised that anyone in the system would volunteer a share of their budget in savings and contended that everyone was looking for a bigger slice of the cake.

There was a feeling that some recent reforms might have unintended effects, by creating institutions that focused power. For example, in the past, health boards and trusts have been able to shift work to primary care without any great investment. Primary care

organisations now give GPs a collective voice to resist and bargain for a better deal from their point-of-view.

The extent of negotiation required to make any of the options work reinforces this view. This was illustrated in numerous ways in the discussion of the seven options above, ranging from negotiation on how resources freed were to be redeployed through to the extent of cash savings to be realised after five years an dhow these were to be achieved. All the time, those who control power silos will be aware that fixed costs create barriers to change, giving them a powerful block on change.

***Implications for economists and the NHS***

It is our hypothesis that the two main barriers revealed in this paper are cost fixity compounded by power silos. So far as barriers to using economic evaluations are concerned, economists can work on the elements of the problem within their own control (e.g. quality of economic studies, availability) but can do little about these more fundamental problems in the short-term. In the longer-term the possibilities are to either help devise policies that address the fundamental problems (of which the internal market was one, arguably naïve, example) or to change evaluation methods to fit with the constraints on resource movement. A possible example of the latter approach is the outcome maximisation approach discussed by Sutton.

For the NHS, the problem is a wider one – how to manage 95% of its budget that is committed each year in line with historical allocations. Our hypothesis would suggest that this money is allocated each year to budget headings that are protected by power silos. The nihilistic view would thus be that NHS strategic policy-makers only have influence over the growth monies each year, and that this is under considerable pressure from different priorities and policy commitments. How to control the 95% is a perennial issue, currently being discussed under the label of "disinvestment". The recent Wanless Report proposed NHS spending should nearly triple over the next twenty years. If care is not taken to direct **and control** this spending in new ways there is a danger that it will not be used efficiently.

**Possible discussion questions for HESG**

Is our diagnosis of the problem accurate?

The work was carried out in Scotland – are there other factors in England & Wales that we should be aware of?

Our preparatory literature search focussed on barriers to using economic evaluations – are there other relevant areas (e.g. on property rights, management studies or politics) that we should be aware of?

**Conflict of interest**
This study was sponsored by an educational grant from Glaxo-Smith-Kline.

**Acknowledgements**
The authors would like to thank all of the interviewees who gave their time. The sample were sent an earlier draft of the paper for comment, but the final views expressed should only be attributed to the authors.

The sample included: John Aldridge, Director of Finance, Scottish Executive Health Department; Marion Bennie, Consultant in Pharmaceutical Public Health, Lothian NHS Board; David Bolton, Director of Primary Care, Lothian Primary Care Trust; Fraser McLeod, Lead GP, Newton Mearns LHCC, Greater Glasgow; Pat Murray, Chief Pharmacist, Lothian Primary Care Trust; John Mungall, Assistant Director of Commissioning, Argyll and Clyde NHS Board; Dr. Andrew Power, Medical Prescribing Advisor, Greater Glasgow Primary Care Trust; Dr. Phillip Rutledge, ?, Lothian NHS Board; Angela Timoney, Consultant in Pharmaceutical Public Health, Tayside NHS Board; Dr. Hugh Whyte, Primary Care Advisor, Scottish Executive Department of Health.

**References**
Cookson R., McDaid D. and Maynard A. 'Wrong SIGN, NICE mess: is national guidance distorting allocation of resources?' ***BMJ*** 2001; 323: 743-745.

Drummond M., Cooke J. and Walley T. 'Economic evaluation under managed competition: evidence from the UK' ***Social Science and Medicine*** 1997; 45: 583-595.

Ham C. 'Improving NHS performance: human behaviour and health policy' ***BMJ*** 1999; 319: 1490-1492.

Hoffmann C, Graf von der Schulenburg JM. 'The influence of economic evaluation studies on decision making. A European survey. The EUROMET group.' ***Health Policy*** 2000; 52: 179-192.

Sutton M. 'How to get the best health outcome for a given amount of money' ***BMJ*** 1997; 315: 47-49.

Twaddle S. and Walker A. 'Programme budgeting and marginal analysis: application within programmes to assist purchasing in Greater Glasgow Health Board' ***Health Policy*** 1995; 33: 91-105.