Improving the allocation of health resources in England
How to decide who gets what

Key messages

- Health resource allocation decisions – how much money clinical commissioning groups and local authorities get from the Department of Health, and on what basis – will soon come under increasing scrutiny as NHS funding slows, reforms are implemented, and local services are held more accountable.

- While the principles behind and methods of resource allocation in England have changed little since the mid-1970s, the NHS has changed considerably. This paper argues that the resource allocation system needs to change too, not only to reflect the challenges facing the NHS today but to ensure that it is fit for purpose for the future.

- We review how health resource allocation in England and the formula on which it is based have changed over time, with a detailed critique of how decisions are currently made and how the process will change from April. We also explore how politics has influenced resource allocation through the pace of change policy, as well as some of its unintended impacts.

- We argue that the coalition government’s reforms risk creating a more complex and fragmented resource allocation process. We suggest some improvements to the current process and also explore how it could be used to support alternative visions for the future of the NHS: clinically led, driven by outcomes, more integrated and provider led, or more integrated with other services through a ‘single budget’.

- We welcome the NHS Commissioning Board’s recent announcement of a fundamental review of the formula and believe it should be well resourced, independent, and sufficiently broad in scope (including public health resource allocation). But we argue that it needs to go further, and provide a more fundamental reassessment of the objectives, processes and methods of health resource allocation in England. In doing so, it must address two key questions. Should resource allocation be a neutral way of allocating funds to the NHS and local government? Or a policy tool to support the NHS in delivering its mandate and changes to the wider system?
Introduction

The NHS faces one of the longest sustained periods of low or no real growth in funding in its history. In all probability, this situation will continue beyond 2014/15 (Appleby 2011) and the Secretary of State has refused to rule out budget cuts for the NHS in future years (The Spectator 2012). The NHS is being asked to find £20 billion of efficiency savings by 2014/15, within the broader context of low or declining economic growth. Public sector spending overall is taking an even larger hit, so despite the funding squeeze facing the NHS, the proportion of public service spending allocated to health will actually rise from 18 per cent in 2010/11 to 28 per cent by 2014/15 (Crawford and Johnson 2011). Given the unprecedented financial challenges ahead, we believe there will be even closer scrutiny to ensure that health resources are spent wisely. In these times of austerity, it is crucial that health spending gets to the places where it is needed most if the NHS is to improve efficiency and equity.

But amid the heated debate on health and social care reforms, important changes to the resource allocation process – and who is responsible for it – have received relatively little attention. As the reforms begin to be implemented, giving greater decision-making power to local bodies, and as health funding growth grinds to a halt, key stakeholders including the public are likely to want answers to pressing questions such as ‘Why does my area get X amount while the area next to us gets Y?’

This paper aims to shed light on this and other important questions about changes to health resource allocation. The current principles, methods and decision-making responsibilities for health resource allocation have remained essentially the same (albeit with many technical improvements) since the mid-1970s. Here, we assess the impact of the coalition government’s NHS and public health reforms on resource allocation and whether the methods of resource allocation are likely to be fit for purpose for the future.

The NHS Commissioning Board has recently announced a fundamental review of resource allocation. We hope the issues explored in this paper will inform the scope and implementation of that review. We believe there are some incremental improvements that can be made in the short to medium term. However, there are a number of emerging issues that call for a more fundamental rethink of the objectives and methods of resource allocation. We argue that some very different approaches may be required as the focus of commissioning shifts to outcomes, as decision-making and accountability are increasingly devolved to local areas, as more integrated delivery systems develop, as the relationships between commissioners and providers change, and as and when attempts are made to define benefits more explicitly. We welcome the opportunity provided by the review and encourage it to be true to its name and go back to fundamentals.

The paper is structured as follows. First, we explain the concept of resource allocation and how it has developed in England over time, and describe how the resource allocation system currently operates. We then assess changes to resource allocation as a result of the coalition government’s reforms, and the risks and opportunities they pose. This is followed by a critique of the current approach, exploring opportunities to use resource allocation to support more radical change. Finally, we discuss how the current formula-based approach can be improved and set out some alternative approaches that the NHS Commissioning Board may wish to consider as part of its fundamental review.

\footnote{This paper was written before 1 April so refers to the responsibilities before this date.}
What is health resource allocation?

At its most basic, health resource allocation, in England, is the process through which money gets from the organisation that collects national tax revenues (the Treasury) to the organisations that purchase and provide health care on behalf of patients (health commissioners). Resource allocation is different from purchasing, which concerns the transfer of money to those responsible for delivering or providing services (see Appleby et al 2012).

The process of health resource allocation differs across countries and is fundamentally dependent on how the delivery of health care and related services is organised. Table 1, below, shows how different European countries made their health resource allocation decisions in the 2000s.

### Table 1  Primary health resource allocation processes in Western Europe

<table>
<thead>
<tr>
<th>Competitive social insurance plans</th>
<th>Employer-based insurance plans</th>
<th>Public sector: devolved</th>
<th>Public sector: centralised</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>Austria</td>
<td>Denmark</td>
<td>Ireland</td>
</tr>
<tr>
<td>Germany</td>
<td>Austria</td>
<td>Denmark</td>
<td>Portugal</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Germany</td>
<td>Finland</td>
<td>Spain</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Denmark</td>
<td>Norway</td>
<td>UK</td>
</tr>
<tr>
<td>Belgium</td>
<td>Austria</td>
<td>Denmark</td>
<td>Ireland</td>
</tr>
<tr>
<td>Germany</td>
<td>Austria</td>
<td>Denmark</td>
<td>Portugal</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Germany</td>
<td>Finland</td>
<td>Spain</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Denmark</td>
<td>Norway</td>
<td>UK</td>
</tr>
</tbody>
</table>

Source: Rice and Smith (2002)

In countries where health care is dominated by private or social insurance systems, there is no externally set budget from which to allocate resources; the resources available to spend are simply the sum of individual premiums or contributions paid to health insurers, who then use this money to reimburse or purchase services.

However, few social insurance funds are self-financing; there is usually some element of pooling at a national level (Rice and Smith 2002). In social insurance systems, resources are (re-)allocated between funds to ensure that the money they receive is a fair reflection of the risks of their members. In Germany, for instance, a risk equalisation scheme is used to reallocate funds between competing social insurers to ensure that those with higher risk, and therefore higher cost, receive a fair allocation (see Van de Ven et al (2007) for a discussion of risk equalisation schemes).

In some countries such as Sweden, Denmark, Australia and Canada, where the health system is decentralised or federalised, national revenues are allocated in the form of a block grant to the provinces, states or local governments that are responsible for health care. This central funding supplements local budgets in order to ensure equity of access to health services. In Portugal and Spain, health care is largely funded by general taxation collected at national level, and allocations are made to devolved entities that purchase services on behalf of geographical populations.

In the United Kingdom, resource allocation takes place at two levels: funds are first allocated between the constituent countries of the United Kingdom, as health care is a devolved matter, through what has become known as the Barnett formula. Second, resources are allocated within England to local purchasers.

The Department of Health determines how its annual budget (currently just over £100 billion) gets divided up between commissioners responsible for purchasing services.

---

2 The Barnett formula has been heavily criticised for being out of date and not reflecting population differences or changes in needs in the devolved administrations since it was instituted over 30 years ago (see House of Lords Select Committee on the Barnett Formula, 2009) but this is not the focus of our paper.
in different areas of the country. The next section describes how it does this and how the process of resource allocation has developed over time. But, as we have shown above, how it chooses to do this is intimately related to the overall organisation and structure of the health care system in which it is embedded. As the reforms change the nature of the system, a good resource allocation system should not be preserved in aspic, it needs to adapt as the nature and objectives of the system change.

Resource allocation in England: how the process has changed over time

The vast majority of resources for health in England are raised through general taxation. The proportion of the national budget that is allocated to health (as opposed to other priority services like education and defence) gets decided during periodic spending rounds (usually every three or four years), led by the Treasury, and subject to considerable political bargaining.

The Department of Health receives a set budget, the Departmental Expenditure Limit, which cannot be breached. The Department then has to make two decisions: how much to allocate to NHS commissioners and other bodies; and then how much to allocate to each commissioner. The responsibility for these decisions currently lies with the Secretary of State for Health.

There is a long and well-documented history of the way health resources have been allocated in England; familiarity with this history is important to understand how resource allocation works today and why the current system may not be fit for the future. This section therefore discusses the core principles of health resource allocation in England, established in the 1970s. Although the core principles remain central to health resource allocation today, the details of resource allocation have been constantly shifting as the Department of Health and its advisers have sought to improve the approach and expand its coverage over larger areas of funding. The history of resource allocation is revealed as a constant interplay between what successive formulas (devised by academics and practitioner advisers, and managed by the Department of Health) say about where the money should go, and how this is moderated by politicians’ judgements, which moderate the impact of the formulas, and what is ultimately decided on where the money actually goes.

Establishing the principles for health resource allocation: Crossman and RAWP

When the NHS came into being 65 years ago, more than 1,700 hospitals and almost 430,000 beds were transferred from local government to the new service, along with another 1,300 or so hospitals and almost 120,000 beds from voluntary hospitals (Department of Health undated). Despite the important principles of equity underpinning the NHS at its inception, these institutions continued to be largely funded on the basis of their historic running costs. There was little consideration as to whether hospital stock and other services were in the right places, and whether they had kept pace with population mobility, demographic changes and health care needs, or changes in the ability to treat those needs.

By the late 1960s and early 1970s, it was becoming clear that this was a substantial problem (Tudor-Hart 1971). The 1970 Labour Government’s Green Paper on NHS reorganisation set out a commitment to a new method of resource allocation, moving...
beyond the accurately parodied system of ‘What you got last year, plus an allowance for growth, plus an allowance for scandals’, as Maynard and Ludbrook (1980) memorably put it.

This was brought to life as the 'Crossman formula', named after the Labour minister who devised it (although it was actually implemented by the incoming Conservative government). Reflecting the equity principle underpinning the NHS, the Crossman formula aimed to remove regional inequalities in access to hospital services over 10 years by reallocating funding, sending a higher proportion of money from the NHS pot – however much was in it – to areas with the greatest needs. It established the principle of defining 'target allocations' for each area – a principle that remains in place today.

These target allocations were based on a formula that took into account each regional health authority’s population size, controlled for age and gender distribution (with more money going to areas with more older people), and the volume of patients and number of beds (reflecting differences in utilisation rates). A second principle – that of a sum per head of population adjusted for factors that should reflect need, or ‘weighted capitation’ – was also established.

In 1975, in response to criticisms that the original formula was not sufficiently based on need and the weights it used were arbitrary, the Department of Health and Social Security announced a review by the Resource Allocation Working Party (RAWP). A third principle was thus established: that the formula be set and updated by independent technical experts.

RAWP expressed the equity principle behind resource allocation much more explicitly than the Crossman formula did. Its objective was to allocate NHS funds to local areas so that ‘…there would eventually be equal opportunity of access to health care for people at equal risk’ (Department of Health and Social Security 1976). This fourth principle, too, has stood the test of time, and remains the fundamental objective of health resource allocation in England today.

Successive governments adhered closely to RAWP’s interpretation of the equity principle for resource allocation until the late 1990s. Then, the incoming Labour government added a new objective: ‘to contribute to the reduction in avoidable health inequalities’. This reflected its view that the NHS should do more than provide equitable care when people are sick; it should also help reduce variations in people’s need for care in the first place, and be given the funds to do so. The inequalities adjustment was designed to address the cliché that the NHS should be a health service and not a sickness service. Areas with more unmet need – and more unexpressed need – would therefore get more funds. The resulting changes were introduced in 2003/04, accompanied by policies to reduce health inequalities pursued through a target regime and other initiatives (Department of Health 2009). This additional objective remains, although as will be discussed later, the coalition government has reduced its weighting.

The RAWP formula proceeded to allocate resources for hospital and community health services (HCHS) on the basis of an amended weighted capitation formula, based on bed utilisation rates and standardised mortality ratios (SMRs) as proxies for ill-health and therefore the need for health care in different areas. RAWP also made adjustments for unavoidable input cost differences, including higher wages in places such as London, with more competitive labour markets and cross-boundary flows of patients between different areas. For good measure, it also took into account the differential costs of medical education and capital investment (although the accuracy of this and subsequent adjustments for the costs of teaching remain controversial (Bevan 1999)).

The results of this process showed a large difference between the amount of resources areas were actually receiving and what they should be receiving based on RAWP’s
assessment of differences in need – each area’s ‘target funding’. In particular, the Thames regions (encompassing London) and Oxford were relatively over-funded compared with the rest of England, which was relatively under-funded (see Figure 1 above).

RAWP stated that: ‘We consider these large disparities to be indefensible and that they should be removed’. But the working party was also acutely aware – as the architects of the previous Crossman formula had been – that areas could not be moved to their target allocation immediately. This would mean pulling funding out of areas defined as ‘above target’, putting existing services for patients in those areas at risk – something that would be unpopular whatever government was in power – and in the working party’s view ‘…there are limits beyond which individual RHAs [regional health authorities] could not, with efficiency and effectiveness, accommodate an unprecedentedly high growth rate’ (DHSS 1976, p 29). For instance, moving too quickly to the target allocation could lead to some of the additional resources being absorbed by grade inflation linked to national pay deals or other inefficiencies in purchasing services.

There was therefore a need for a ‘pace of change’ policy that would, over time, move areas to their target funding, according to the formulas – establishing yet another principle that has endured until today.

RAWP did argue that within the pace of change policy, those areas most under target should be moved closer to their target more quickly to ensure that they did eventually catch up. But only when resources became very stretched did RAWP consider it justifiable to actually reduce funding to those areas that were over target. Finally, RAWP recognised that decisions on how quickly areas were moved to their target allocations were essentially political, and so it left ministers with this responsibility. Again, this principle – of a separation between technical advice from experts on the formulas for resource allocation and political judgements about how fast the Department of Health shifts relative funds between areas that are under or over target – remains in place today.

Thus Crossman and RAWP in the 1970s put in place the essential building blocks of the way the Department of Health in England currently allocates resources within the NHS.
The approach, based on the principle of securing equity (defined as funding for ‘equal access to NHS services for equal need’) remains the basis of our current system. This is operationalised through weighted capitation formulas set by independent technical experts that calculate target allocations for each area, and political decisions about how fast relative funding to areas can be changed to meet their target allocations.

Refining the formulas over time

Within this framework the approach to health resource allocation has been constantly refined since the 1970s; leading to a more complex system. These refinements include increasing coverage beyond hospital and community health services, the use of more detailed and granular data sets, more precise formulas for different services, changes in how the population denominator is calculated, and more adjustments for input costs.

Initially, the Crossman and RAWP formulas applied to hospital and community health services only, reflecting the primary objective of removing regional inequalities in access to hospital services. Over time, the formula approach has been extended to cover prescribing and GP services. Figure 2, below, shows what impact this had on one area (County Durham) as an example.

Over the past 35 years, the formula has been influenced by much greater availability of data at lower levels of granularity so that today, the records of millions of individual patients are used to help make the formula more precise. In the early years of RAWP, by contrast, data were only available from the 14 regional health authorities in aggregate form.

Figure 2  How extending coverage of the resource allocation formula affected County Durham’s allocation from 2009 to 2011

HCHS: hospital and community health services; PMS: personal medical services; M&D costs: medical and dental costs; EACA: emergency ambulance cost adjustment

Source: Thompson (2010)
Alongside the extension of coverage and use of ever more detailed data, there has also been a proliferation of formulas within each of the services covered. This aimed to take more precise account of the different factors driving relative needs for funding for specific services such as maternity services within hospitals, for example.

There has also been a constant drive to improve the accuracy of the baseline population data. At the level of the old regional health authorities, census information was reasonably reliable, although there were problems with undercounting, and the further from the actual year of the census, the less reliable population data were, and continue to be. The preferred option is to use GP-registered populations, but GP lists are widely known to be inaccurate and to over-count – a problem known as ‘list inflation’. While efforts are under way to improve this situation and the longer-term aim is to use registered populations for the formula, the latest system for resource allocation uses both sources of data.

Adjustments have also been made in an effort to improve the accuracy of estimates of input costs. RAWP initially only allowed for a wages weighting for London to reflect the more competitive labour market for NHS staff in the capital. Over time, this has developed into the market forces factor (MFF), which was designed to take into account competition in the general labour market for the skills of NHS staff across England. However, there are concerns that the MFF adjustments remain insufficient, particularly in high wage areas, leading to poorer quality staff, with an impact on patient outcomes (see Hall et al 2008).

The history of resource allocation in England has been driven by a genuine desire to fund the NHS equitably so that it can deliver health care in line with its founding principles. Ongoing technical innovations have sought to help to deliver this aim, including identifying better proxies for need for specific services, using more detailed data, improving the accuracy of population data, and getting more accurate estimates of variations in input costs. But while these changes have sought to improve the resource allocation formula, they have also made it more complex. Later, we discuss the disadvantages of this greater complexity and put the case that the resource allocation process needs to be simplified. But first, we describe the practical and political reasons why progress in the actual reallocation of resources has been slow.

The politics of health resource allocation: the pace of change

As already noted, actual health resource allocation is a constant interplay between the advice of technical experts developing formulas and the judgements of politicians.

For example, in the mid-1980s, as one of the periodic reviews of the RAWP formula was announced, Birch and Maynard (1986) made a strong case for extending the formula to primary care. They argued that it made little sense for the distribution of HCHS funds to be subjected to a rigorous approach when primary care was not.

Their simulation of the consequences of ‘RAWPing’ primary care was unsurprising; it would give a higher proportion of available primary care resources to poorer areas and out of London – exactly what RAWP did for HCHS services. However, Birch and Maynard were acutely aware of the political obstacles involved in implementing their proposals: ‘… the contractual status of general medical and dental practitioners, with lifelong contracts and payment according to centrally determined, non-capped criteria. Including them in RAWP would have to mean a change in those contracts and an overall limit to the budget for primary care by definition.’ It took until the late 1990s for primary care to come under the scope of the formula approach, when it was bundled as part of an undifferentiated allocation.
The time it took to extend the formula to primary care is indicative of the fact that politics and political judgements play a clear role in influencing the scope of the formula. They also influence the pace of change.

As RAWP argued, there are very good practical reasons why – even in the absence of politics – a pace of change policy is needed. Changing relative funding quickly, at least in a knee-jerk, annual un-phased response to the formulas, could threaten the stability of services and medium-term planning. In addition, the details (if not the principles) of the formula are constantly under review; with each new iteration, baselines are reset and estimates of populations and underlying needs change. These constant background changes mean that the target funding for each area is always slightly out of alignment with where the latest formulas suggest they should be.

However, to these practical reasons must be added more political motivations. First, there is the temptation – although impossible to prove – to shift resources at the margin towards areas based on favouritism, special cases or, at the extreme, even party affiliation or loyalty. For example, in the mid-1990s, a decision was taken not to apply an updated weighting for need across all services, in particular not applying it to community health services. This was criticised for limiting the scale of redistribution to areas with more material deprivation, supposedly to benefit Conservative constituencies (House of Commons Health Select Committee 1996; Carr-Hill et al 1997).

The second, and materially more significant, political motivation is that all political parties have abided by the unwritten rule that no area should receive a real-terms cut in NHS funding as a result of resource allocation decisions. The political fallout from this would be immense and no party has had the courage to take such decisions. The amount of funding available for reallocation is therefore much smaller than first appears. It is only ever additional NHS spending that is available for reallocation; areas keep what they had historically, up-rated for inflation. It is therefore no surprise that under-funded areas have never fully caught up. Furthermore, the rate of catch-up is faster during times when the NHS budget is increasing, as there is more growth available to re-distribute. In the recent context of year-on-year growth, this has not been too difficult. Looking ahead though, it is hard to see how this can be sustained.

How health resources are currently allocated

The Department of Health’s funding from the Treasury, its Departmental Expenditure Limit, covers everything the Department is responsible for delivering, which includes NHS care but also the central programmes delivered by the Department and its arm’s length bodies and agencies. Until April 2013, this includes strategic health authorities (SHAs). It also includes the various regulators and advisory bodies such as the Care Quality Commission (CQC), Monitor, and the National Institute for Health and Clinical Excellence (NICE). Funding to these bodies is determined by the Department. Once this is accounted for, by far the largest amount of planned spending goes directly to the NHS according to the current resource allocation formula and politicians’ judgements on pace of change.

Figure 3, overleaf, shows overall funding flows in 2012/13. The large majority of the Department’s budget is allocated to primary care trusts (PCTs) based on the current formula and pace of change. In 2012/13, PCTs received around £94 billion plus some adjustments (including an under-spend from 2011/12 and other income). From their individual allocations, PCTs then make their own commissioning and spending decisions, contracting services from GPs and from acute, mental health and ambulance trusts.
Figure 4, opposite, sets out the main features of how each individual PCT gets its share of the £89 billion total allocation, and the weighting given to each. There are three separate formulas: HCHS spend – where RAWP started all those years ago – is still the most significant, but prescribing and primary care now account for around 20 per cent of the overall allocation. Each of these formulas includes the age profile of the PCT’s population, which dominates, but other ‘additional needs’ are also important such as those driven by deprivation. Although age is the dominant factor in determining the level of allocations, deprivation varies much more between areas than age profiles do, so its importance in explaining differences between areas is magnified. Each of these elements in turn currently attracts an additional health inequalities component, for the reasons set out above.

Beyond this, there are other adjustments to ensure that PCTs are compensated for unavoidable cost factors such as different wage rates between urban and rural areas. In practice, though, the pace of change policy leads to considerable differences between what the formulas predict each PCT should receive in terms of relative needs and what they actually received in 2011/12 (see Figure 5 opposite). As Figure 5 shows, some PCTs are still receiving allocations that are below target, but a small number are well above target; clearly, RAWP’s ambition has not been met in the many years since it was established.

If PCTs received the funding that the formulas say they should, then all areas would lie on the 100 per cent line in Figure 5. The PCT most under target was North Somerset, while Stoke-on-Trent was also 5 per cent below where the target suggests it should have been. Forty-four PCTs are between 5 per cent and 2.5 per cent below target, and a further

---

4 Data for 2012/13 are available but in that year, there was an essentially uniform uplift for all PCTs (albeit there were some very minor changes for non-recurrent funding such as GP practices changing PCT boundaries), making no essential difference in terms of relative allocations from 2011/12.
Figure 4  The primary care trust allocation formulas, 2011/12

Note: The weightings applied are for 2011/12 allocations.
Source: National Audit Office (2011)

Figure 5  Target and actual allocations for PCTs, 2011/12 – the impact of the pace-of-change policy

42 are up to 2.5 per cent below their target allocation. Most of the rest receive more than the formulas suggest they should. At the top end, four PCTs are between 15 per cent and 20 per cent above target, while Kensington and Chelsea and Hammersmith and Fulham are both more than 20 per cent above target.

Figure 6, below, uses recent data to make (as far as possible) a comparison to the mid-1970s position (as shown in Figure 1, p 6); it shows that relative under-funding and over-funding has reduced for most areas of England. The Thames regions (which subsumed London) were by far those most over target in RAWP’s time, and London remains the area most over target today. Although as a whole London is only 6 per cent over target, the PCTs above remain more than 20 per cent over target.

**Figure 6**  Target and actual allocations for strategic health authorities 2011/12 – the impact of the pace-of-change policy

The coalition government’s first steps on resource allocation policy were to reduce the inequalities weighting, from 15 per cent in the last year of the outgoing Labour government to 10 per cent (as reflected in Figure 4, p 11). In terms of target allocations, this makes a substantial difference. For 35 out of the 151 PCTs, it changes allocations by more than 5 per cent, and in relative terms, shifts funding to southern and eastern England from the north and from inner London (National Audit Office 2011). In practice, though, any effects from the change to the inequalities weighting – whether good or bad from an individual PCT’s perspective – have been cancelled out by the fact that the pace of change policy has ground to a halt because of the squeeze on overall NHS funding.

Like previous governments, the coalition has been unwilling to cut allocations to over-target areas and redirect them to under-target areas to get them closer to target more quickly, resulting in little actual change in allocations. The effects of changing the inequalities weightings are therefore primarily symbolic, at least in the short term. But they do reflect a real change in view about the core role of the NHS in reducing health inequalities.
How the coalition government’s reforms affect resource allocation

Tucked away in the Health and Social Care Act and related policy documents are key changes, set to come into effect from April 2013.

The coalition government’s reforms affect three big decisions about health resource allocation. First, the Secretary of State for Health will make a new allocation decision: how much should be spent ‘on the NHS’ overall, and how much ‘on public health’. Two further decisions flow from this one: how then to allocate NHS funding and public health funding.

From April 2013, and for the first time since the NHS was established, someone other than the Secretary of State for Health will decide how NHS resources – totalling more than £95 billion in 2013/14 – are allocated. The reforms hand responsibility for this decision to the new national NHS Commissioning Board. But while the Secretary of State loses the power to make one key decision, he takes on new responsibility for another: how to allocate resources for public health.

NHS allocations

The draft mandate between the Secretary of State for Health and the NHS Commissioning Board establishes the Board’s duties in relation to resource allocation:

In future the Board will be responsible for allocating the budget for commissioning NHS services, this will prevent any perception of political interference in the way that money is distributed between different parts of the country. The government expect the principle of ensuring equal access for equal need to be at the heart of the Board’s approach to allocating budgets. This process will also need to be transparent, and to ensure that changes in allocations do not result in the destabilizing of local health economies.

While decisions about allocating funding to commissioners are for the Board to make, the Department is responsible for managing the overall NHS budget, and therefore will need to understand the Board’s approach to allocations. The Board should share its approach with the Department as it is developed, and should involve the Department at key stages in the process.

(Department of Health 2012c, pp 32–3)

These highly crafted paragraphs emphasise the goal of equal access for equal need. They also make clear the need for the Department to ‘understand’ the Board’s approach as it develops (although the last paragraph was omitted from the final mandate (Department of Health 2012d), presumably for reasons of economy of space). How the relationship between the Department and the Board actually plays out in practice, and how much ‘soft control’ the Secretary of State will seek to exert, is difficult to predict. Arguably, the then Health Secretary Andrew Lansley began this process with a speech early in 2012 arguing that age was a far more important determinant of NHS need than deprivation, comments which led to a heated debate (Asthana 2012; Campbell 2012; Gainsbury 2012).

What is certain is that the mandate signals the coalition’s return to RAWP’s view, some 35 years ago, that the objectives of NHS resource allocation are primarily to meet treatment need equitably, rather than to prevent inequalities in health – back to a sickness service rather than a health service. This approach is consistent with the new duties placed on the NHS, which are restricted to addressing inequalities in access to and outcomes of NHS care, explicitly ruling out a general duty to improve health outcomes. This downplays the broader role played by the NHS in preventing health inequalities, both through its clinical prevention work but also the wider determinants of health, as the country’s largest economic entity, purchaser and employer.
Public health allocations

At present, each PCT makes its own decisions about how much to spend on public health, out of their overall allocation. Under the reforms, a range of public health functions transfer to local authorities, and a new executive agency of the Department of Health – Public Health England (PHE) – has been set up. The Secretary of State will decide how much to spend on public health overall, and with the advice of PHE, allocate it to different functions. This includes a new ring-fenced grant to each local authority to fulfil their responsibilities.

The Secretary of State for Health asked the Advisory Committee on Resource Allocation (ACRA) to devise an approach to public health allocations. It reported back in summer 2012, arguing that the majority of funding should be allocated on the basis of each local authority’s under-75 years standardised mortality ratio (SMR<75), an echo of RAWP’s early work on NHS allocations. Areas with higher rates would see this reflected in higher relative funding to tackle the higher prevalence of early, preventable deaths and other problems.

The Department accepted this recommendation in principle (Department of Health 2012a) and, following subsequent amendments in response to consultation, has issued allocations to local authorities for 2013/14 and 2014/15 (see [www.dh.gov.uk/health/2013/01/ph-grants-las/](http://www.dh.gov.uk/health/2013/01/ph-grants-las/) for all relevant documentation).

A critique of the current approach to resource allocation

The changes to the system of resource allocation arising from the health and social care reforms carry significant risk but they also offer a real opportunity to address some longstanding problems. This section considers the main risks, namely an increasingly fragmented and complex decision-making process, lack of transparency and consultation, inadequate measurement of need, and the slow pace of change. It also suggests some incremental improvements that could be made to the system to address the longstanding problems.

More fragmented and misaligned resource allocation and commissioning decisions

An inevitable corollary of the changes to the structure of the NHS and the public health system is that decisions about resource allocation will become more fragmented, as there are more decision-makers involved and more decisions to be made.

The resource allocation decision-making process differs in key respects under the new system architecture, compared with 2012/13 (see Figure 7, opposite). First, the Department will decide how much to allocate to public health and how much to the NHS. For 2013/14, this decision has already been made: the NHS Mandate allocates a budget of £95.6 billion to the NHS Commissioning Board. But it is not clear how this decision was made, or what factors informed it.

First, the Secretary of State will decide how much to allocate to public health and how much to the NHS. For the NHS, the NHS Commissioning Board will make four key decisions: how to divide up its budget between its own commissioning responsibilities (and those of clinical commissioning groups); how much to spend on primary care services; how much to spend on specialised services; and how much to allocate to each clinical commissioning group for secondary and community services.

For public health, the Secretary of State (guided by Public Health England) will make two key decisions: how to divide up the public health budget between Public Health England
and local authorities; and how to allocate public health budgets (£2.7 billion for 2013/14 compared to the simulated £2.1 billion in Figure 7 for 2010/11) to each local authority.

There are then additional decisions that can be characterised as quasi-allocation and quasi-commissioning decisions between the new bodies that have been created. The most important of these is how much funding will be allocated to the NHS Commissioning Board from Public Health England’s budget for public health services to be provided by the NHS. This has been set at £1.8 billion for 2013/14 (NHS Commissioning Board and Department of Health 2012).

It is striking just how much more fragmented this process is compared to the system it replaces.

The new approach also runs counter to other government policies and evidence, both of which increasingly emphasise that integrated care is the key to improving health and meeting the productivity challenge. The NHS in England, like its counterparts in other developed countries, is facing two major, interlinked challenges: an increasingly frail older population with complex care needs, and public health problems associated with unhealthy lifestyles. Addressing these challenges requires a more integrated approach to commissioning across public health, health care and social care – something that present and previous governments in the United Kingdom have acknowledged.

However, as noted above, the reforms create multiple funding streams and dramatically increase the complexity of subsequent commissioning. We are moving away from a system where PCTs, whatever their faults, had population-based budgets that covered all the needs and associated costs for their population, and were held accountable for keeping expenditures in line with their budget.

The new system fragments this into clinical commissioning group budgets for secondary and community care, and the NHS Commissioning Board for primary care and highly specialist services, while public health budgets are split between Public Health England, local authorities and the NHS. This will make it more difficult to commission integrated forms of provision. It will also make it more difficult to maintain a financial grip as
multiple budget-holders shift costs to each other. Overall, there is also the risk that the new allocation and commissioning system will make it more difficult to ensure allocative efficiency and equity in resource allocation locally, as it will be more difficult to move funding upstream or to other parts of the pathway where care can be delivered more efficiently.

Short of further structural reorganisation, the alignment of clinical commissioning groups and local authority boundaries, together with a stronger role for health and wellbeing boards, could compensate for this increasing fragmentation. However, it remains to be seen whether the boards will consider it part of their role to bring these allocations together, and if so, whether they will have the capacity and capability to do so.

A more complex process

A formula-based approach has great advantages. As long as it is set independently, it is objective and fair. The technical development of the formulas over time has also been a major intellectual achievement, engaging some of the most capable statisticians, econometricians and health service practitioners in ACRA and its predecessors, back to RAWP.

The fact that the resource allocation formula has both extended its scope and become more specific over time – with separate formulas for various elements of need – brings the benefits of precision and fairness. But it also significantly increases complexity, making the outputs more difficult to understand and for politicians and other decision-makers to question the basis of the formula.

More than 10 years ago, Rice and Smith (1999) undertook an in-depth survey of other countries that used versions of weighted capitation to allocate health care resources. All the problems that have been identified with the resource allocation process in England (if not the exact pattern) are evident elsewhere. But Rice and Smith expressed concern about England’s path towards an ever more complex process compared with other countries. They recommended that the Department of Health undertake a thorough review of the totality of the adjustments made to the formula over time, to ensure that it is no more complex than it needs to be.

But no such review was undertaken, and as we have documented, the formula has become even more complex. As Bevan (2008) argues persuasively: ‘There seems to be a natural history of development of formulas: the initial formulation of a simple formula driven by a few principal elements; this is followed by an attritional process of demands from the NHS for special allowances, which results in various elements being added.’ As RAWP stated, ‘the working party rejected many approaches which might have made the criteria more sensitive, but which on examination would have led to much greater complexity with little significant gain in the result’ (DHSS 1976, p 8).

Bevan supports his argument by simply plotting the relative PCT allocations that would be made if the hospital and community health services (HCHS) formula were applied to the whole of the budget (including the additional prescribing, primary medical services, pharmaceuticals and HIV and AIDS allocations derived from their separate formulas) against the relative allocations to PCTs from the HCHS formula alone. The result is a virtual 45 per cent degree line, which prompts the obvious question: ‘What is gained by detailed weightings for the smaller components of expenditure, which made the formula more complex and hence less transparent?’

An update of Bevan’s analysis on PCT allocations is shown in Figure 8, opposite. The result, again, is a virtual straight 45 per cent line with a virtually perfect $R^2$ of 0.9993; adding the other formulas makes next to no difference, compared to relying on the HCHS alone.
The NHS Commissioning Board needs to take a fresh look at just how complex the allocation formulas really need to be. If this can be done without losing reasonable precision, this would clearly be preferable to where we are now. The range of stakeholders interested in, and affected by, resource allocation policy and decisions is widening fast as budgets tighten and the reforms are implemented. There is a strong case for greater simplicity and transparency in the formulas.

Lack of transparency and consultation

Anyone with the requisite specialist expertise and time can peruse ACRA’s research and deliberations, as well as the Department of Health’s final decisions and how they affect individual PCTs, as set out in the Exposition book (Department of Health 2011a). But these research papers and the Exposition book are only made available to the public long after ministers have taken decisions based on their content. Furthermore, the increasing complexity of the formula means that it is difficult for anyone not involved in designing it to understand how it actually works. The reforms (which, as noted, increase the range of stakeholders involved in decision-making) and the tightening of the health budget mean that an ever-growing audience of local citizens, politicians and clinicians will demand more clarity so that they can understand why their areas receive more or less than neighbouring areas. The current approach is anything but clear.

The NHS Commissioning Board could contribute to making the resource allocation decision-making process more transparent; its mandate is both explicit and encouraging in this respect. The Board could send a very positive signal about transparency by holding ACRA meetings in public, or at the very least ensuring that deliberations and papers are released before all the relevant decisions they need to inform have been taken.
Consultation will also be critical to confer public legitimacy on the transfer of control over resource allocation decisions to the Board. The Department makes no pretence about NHS resource allocation being a consultative process. It has argued that ACRA needs space to be independent and protected from constant special pleading by PCTs. However, it has also taken an important step by allowing ACRA to consult on its work on public health allocations. It is now time to ensure that its deliberations for the NHS are also open to consultation, as is already the case for the equivalent decisions in Wales, Scotland and Northern Ireland (Bevan 2008).

Inadequate measurement of need

The successors to RAWP have long grappled with the fact that the health allocation formula remains essentially utilisation-based and does not distinguish sufficiently between the need and demand for care. And as we will see, this is one of the concerns about the latest iteration of the formula process, the person-based resource allocation (PBRA) approach, developed in the context of GPs holding budgets for acute services. PBRA is based on the expected utilisation of health care of anonymised individual patients estimated from millions of records from Hospital Episode Statistics (HES) and other sources (Bardsley and Dixon 2011). This allows funding to be estimated for each general practice based on the needs of people registered on their list.

The fact that, in England as elsewhere, greater supply of health care tends to increase demand has been well documented. Ensuring that allocations based on the formula approach do not reward ‘supplier-induced demand’ is difficult, given that the formula is still based on data on people's actual use of services, rather than direct measures of need. While the advisory body, ACRA, and its predecessors have worked hard to adjust their models for potential supplier-induced demand, it is not clear how successful they have been.

To be fair to the Department of Health, it is well aware of this issue, stating that, ‘Observing additional need directly has not proved possible to date’ (Department of Health 2011b, p 10). Commentators such as Sheldon and Carr-Hill (1993) and Asthana and Gibson (2008) have questioned the legitimacy of deriving estimates of health care needs from utilisation-based information, regardless of how sophisticated statistical adjustments are. Sheldon and Carr-Hill (1993) reminded us that RAWP ruled out utilisation-based measures because of the problems of disentangling them from supply-side effects, relying on standardised mortality ratios instead.

The development of the PBRA approach is a major achievement, but there are some issues beyond specific criticisms of the modelling (see Stone 2013 for a detailed critique of the coefficients, specifically for dementia, and his view of the implications). PBRA may be more influenced by past utilisation of health services as opposed to current or future need than it seems. This is because patients’ past diagnoses – one of the prime drivers of funding under PBRA – are assumed to be independent of supply. However, Bevan (2011) argues that diagnosis rates are not in fact independent from supply and can reflect differences in diagnostic practices, intensity of observation, and ease of access.

So, despite excellent analysis, there remain contested major challenges in correctly assessing the need for care. After carrying out an in-depth assessment of the pros and cons of various options (including the English resource allocation system), Wales adopted a different approach (National Assembly of Wales 2001). Need is now assessed directly based on the reported prevalence of 17 health conditions measured in the Welsh Health Survey. However, this direct approach is also not without its problems. For instance, it is no easy matter to reliably measure the burden of ill-health (although more physiological measurements are increasingly used in the Health Survey for England and others to help triangulate with self-reported data); and it is difficult to then establish what resources...
should be used to meet this need (McConnachie and Sutton 2004). The use of direct measures in Wales has also not been totally divorced from utilisation, since needs are then ‘scaled’ against expenditure to give a sense of the overall resources required to meet health care needs, to stick within available budgets. This scaling, of course, relies on existing patterns of spend and utilisation.

Nonetheless, the Welsh approach is a promising one and other modifications could theoretically strengthen it further, including scaling against ‘best practice’ as opposed to actual practice. Asthana et al (2007) assessed the feasibility of a similar approach for 14 programme budget areas in the east of England using data from the Health Survey for England. They reported that it performed better than the allocation formula in predicting actual programme budget spend and prescribing costs for the PCTs concerned. Scotland and Northern Ireland also take subtly different approaches to England, although they too essentially rely on utilisation-based approaches and attempt to statistically adjust for the effects of supply factors. One significant difference is that Scotland takes into account the costs of rurality in their formulas, whereas England does not – something that Asthana et al (2009) are particularly critical of.

It is our view that there is a strong case for the NHS Commissioning Board to investigate the merits of approaches being used in the other nations of the United Kingdom, and in particular, whether the Welsh approach to measuring need is preferable to indirect measurement and complex statistical adjustment, which the English system has always relied on.

**Slow pace of change**

Politics clearly and legitimately influences the overall objectives of the resource allocation process. It is also legitimate for politics to come into play in terms of pace of change. But this also has disadvantages. Principal among these is that areas remain significantly under-funded over time, and because politicians have been unwilling to cut funding to over-funded areas in times of low overall growth, under-funded areas suffer relatively more when overall resources are tight. So in times of economic hardship, when under-funded areas are most likely to need more resources, how pace of change policy has been implemented ensures they do not receive them.

The political risks implicit in pace of change – whichever party (or parties) are in power – do not incentivise much-needed reconfiguration. Those developing the formula have attempted to ensure that allocations are increasingly based on need and not demand, but the smoothing effect of pace of change means that difficult decisions to change models of care and downsize hospitals – particularly in traditionally over-funded areas such as London – are often delayed, since the funding crunch has always remained just beyond the horizon. The politics of pace of change have thus undermined the urgency of much-needed decisions to re-balance funding.

In retrospect, whatever view is taken on the precise approach to determining need, it is clear that health ministers in the previous Labour government should have taken a far more radical approach to shifting resources between parts of England that were over- or under-funded when NHS funding was more favourable. Now, with the likelihood of no real growth for some years ahead, pace of change has ground to a halt. No Secretary of State, of whichever party, has been prepared to take any baseline funding from over-target areas to give to under-target areas. Pace of change has, in effect, institutionalised the mis-funding of local health services in England, slowing progress towards a truly equitably funded NHS.

There is a view that this is precisely the wrong time to be talking about radical changes to the pace of change policy. But we disagree. There is, we think, a welcome new realisation that as spending growth stalls, local areas will need to look at much more radical
solutions to meeting the needs of their populations, including finally getting serious about reconfiguration (Imison 2011), the rationalisation of hospital stock, and, where appropriate, shifting services into the community.

The reforms do actually give clinicians more power to begin doing this. Integration of services and care is at last being considered seriously at all levels, and the former Secretary of State showed strength in accepting several recent decisions by the Independent Reconfiguration Panel (Laja 2011), which The King’s Fund has also welcomed (The King’s Fund 2011). In this context, allowing pace of change to be frozen sends exactly the wrong signal, since it will continue to reinforce the status quo rather than support long-overdue decisions to reconfigure services.

Because of its new independence, partially insulated from party political and broader political considerations, the NHS Commissioning Board is in an excellent position to ensure that its resource allocation decisions work to support reconfiguration rather than stymie it.

### Improving the current approach

We believe a number of incremental improvements could be made to improve the health resource allocation process by addressing some longstanding problems. These include:

- greater transparency in the work of the Advisory Committee on Resource Allocation (ACRA), with greater opportunities to consult on proposals for future revisions to the formulas
- simplification of the formulas and a test of materiality to be applied to any future refinements
- further work to establish valid measures of need and to minimise the supply-side effects of using utilisation-based data.

However, the combination of reforms and austerity have brought with them risks that need to be addressed. There is also an urgent need to:

- identify how fragmented funding streams can be reintegrated at local level through health and wellbeing boards, joint commissioning budgets and other mechanisms
- review the pace of change policy, given the prospects of little or no real growth in budgets for the foreseeable future.

### Time for a fundamental look at resource allocation

**Deciding whether resource allocation is simply a way of allocating funds or a policy tool to support wider objectives**

Our discussion of health resource allocation reveals an evident, but largely unacknowledged, tension between resource allocation being seen as a purely passive device to get ‘fair’ allocations to where they are needed, and as a more active tool that supports wider policy objectives.

In our view, the NHS Commissioning Board, Public Health England and the Department of Health need to address this tension head-on and be clear about what the real purpose of resource allocation is. Is it first and foremost a neutral mechanism to get resources from the centre to local areas in the most equitable manner? Or is it an important tool, along with others, to support the government’s health objectives – for example, to drive improvements in outcomes, as set out in the NHS Commissioning Board’s mandate?
RAWP’s stated intention more than 35 years ago was clearly the former; it explicitly ruled out considering how its work developing the formula should be linked to operational policy. On the other hand, the introduction of an element to tackle health inequalities by the previous Labour government is an example of resource allocation being used to pursue additional policy objectives.

Commissioners’ overall priorities for health spend have been driven by other national policies and targets, not by the objectives that underpin the formula; this separation can result in bizarre misalignments. One example under the previous government is that just as the Department of Health was holding the NHS to account and performance-managing health inequalities targets expressed in terms of life expectancy improvements, the Department accepted ACRA’s recommendation to fund the NHS for reducing inequalities on the basis of disability-free life expectancy. The NHS was therefore expected to deliver improvement on the basis of one measure while simultaneously being allocated money on the basis of another.

More broadly, two key factors have limited the policy impact of the current resource allocation process: relying on utilisation as a proxy for need, and the slow pace of change. Together, these explain why resource allocation has failed to incentivise the reconfiguration of local services. If allocation cannot be sufficiently disentangled from utilisation, it will institutionalise existing patterns of supply, and do little to incentivise disinvestment in areas of oversupply. And if areas which are over target continue to receive more funding relative to need than other areas, commissioners in those areas will have less incentive to make difficult disinvestment decisions or support changes to deliver a more efficient model for the supply of care.

Pace of change has worked in practice to blunt the impact of the formula, helping to sustain distributions of hospital estate and models of care that are clearly sub-optimal from an equity perspective. Looked at this way, the current resource allocation process is very much an active policy tool, but a covert one, supporting stagnation rather than change.

Clearly, there are pros and cons of using resource allocation as a policy lever. But whatever the choice, we believe it should be explicit rather than covert, as has been the case up until now. Being more explicit about the policy role of resource allocation would send a strong signal to the NHS that money and resources will follow policy priorities. It would also give the NHS Commissioning Board a much stronger lever to pursue its objectives as set out in the mandate, and the NHS locally would know that central priorities would come with more appropriate, consistent and longer-term funding attached.

But there is also a case against a more proactive policy role for resource allocation. It would certainly go against the grain of local decision-making and prioritisation, especially in the reformed NHS, which emphasises local joint strategies agreed by clinical commissioning groups and health and wellbeing boards. Where local rather than national priorities take precedence, it is arguably more appropriate for NHS resource allocation to be more neutral, enabling local decisions rather than being tied to national policy imperatives.

We therefore welcome the NHS Commissioning Board’s recent announcement of a fundamental review of health resource allocation (Kaffash 2012). But we urge the Board to be explicit and transparent, and consult widely about whether resource allocation should be simply a way of moving funds from some areas to others to give a ‘fairer’ allocation, or a more proactive policy tool to support the delivery of its broader mandate and much-needed changes to the system (Ham et al 2012).
Using resource allocation to support more radical change?

The NHS has changed beyond recognition since the mid-1970s when the current resource allocation system was established, and is set to change even further as a result of the current reforms and the demands placed on it in terms of available finance, changing needs, its workforce and technology (Ham et al. 2012). As we showed at the beginning of the paper, resource allocation systems differ according to the characteristics of the systems in which they sit. Given the reforms will change our health system, how should the resource allocation system change in response?

We believe there are a number of emerging issues that call for a more fundamental rethink of the methods and objectives of health resource allocation in England, particularly in relation to using allocation to support improved outcomes, localism, more integrated provider organisations, and any move to mandated services.

First, since the system in future is to be held accountable for improving outcomes as set out in the three outcomes frameworks for public health, social care and the NHS, should local authorities and clinical commissioning groups be allocated resources in a way that reflects outcomes? Second, with a strong emphasis on localism and the importance of health and wellbeing boards in setting local priorities, does this require local authorities to have more flexibility to reallocate resources across health and social care and beyond? Third, with increasing moves to create larger integrated providers of care and to shift risk to providers by changing the way that care is paid for – such as bundled payments or capitation – what are the implications of changes in the purchaser-provider relationship for how resources are allocated? Finally, in public health, there is already a move to define mandated services, and similar ideas to define NHS benefits more explicitly have been mooted (Rumbold et al. 2012); is there a case for allocations to be more clearly linked to the costs of delivering these services?

Below, we discuss the implications for resource allocation under these different possible futures for the NHS and public health, and begin to highlight the advantages but also the likely challenges and disadvantages of implementing these more radical approaches. Our discussion is exploratory and we recognise that they represent a major departure from the current approach which has been developed and refined over nearly 40 years. We do not suggest that any of these changes should be embarked on lightly, if at all. Our intention is to encourage broader thinking.

Supporting outcomes

One of the revolutionary changes that the reforms bring is an increasingly clear, if still evolving, view of what is expected of the NHS in terms of the health outcomes it needs to deliver. This is being formalised in the NHS Outcomes Framework, the mandate, and supporting frameworks such as the outcomes indicator set. While we have reservations about the specifics – particularly having three separate frameworks for the NHS, public health and social care, which will make integration more difficult – they provide a welcome clarity about what the NHS should be held to account for delivering. If politicians are now clearer and more explicit than ever before about what outcomes they expect the NHS to deliver, should this not have implications for decisions on how much should be allocated and where, in order to achieve those outcomes?

Of course, it is easy to propose this in principle. However, there would clearly be practical problems to be overcome if allocations were to be related to the outcomes frameworks. How would outcomes be attributed? What weight would be given to different outcomes? How would services not covered by the outcomes framework be resourced? And what would be the right balance between sending more money to areas with poor outcomes versus rewarding those that make progress?
But these are not insurmountable and by creating the quality premium for clinical commissioning groups and the health premium for local authorities, the government has already taken some funding out of the baseline allocation to reward improvements linked to specific outcomes. The Department has also completed relevant work along these lines, albeit for another purpose. In the draft mandate, the NHS Outcomes Framework was accompanied by an extensive technical annex (Department of Health 2012b) which showed detailed work on what the NHS was expected to contribute to each outcome indicator. This was dropped from the final mandate, but it could inform resource allocation baselines and rewards.

Questions of feasibility aside, if the NHS is being asked to deliver outcomes through the mandate, assessing how resource allocation can be used to support it is a legitimate topic of inquiry for the Board’s fundamental review.

Supporting localism

Two aspects of the reforms – the introduction of health and wellbeing boards and local authorities gaining responsibility for public health – significantly strengthen localism. Increasingly, the boards are being seen as having the potential to be an important commissioner (Dorrell 2012) and also potentially the holder of a single health and social care budget (Wiggins 2012). The possibility of such a radical change has been given much greater impetus by Labour’s Shadow Health Secretary, Andy Burnham, who recently set out such a vision in a speech at The King’s Fund in January 2013 announcing the terms of Labour’s health policy review (Burnham 2013). While the NHS Commissioning Board’s fundamental review of resource allocation has only just started, and many questions remain unanswered, it is clear that there is a growing political, clinical and financial imperative towards more integrated allocations and budgets across public services. But as the scale of joint commissioning grows, anomalies in how total public resources are allocated to different areas, populations and services will become starker.

Progress could continue along the current trajectory of bringing separate sources of funding (from the NHS, the local authority and others) together to jointly fund defined services. But under a future scenario where the NHS becomes led and directed by local priorities and stakeholders, through the health and wellbeing board, it makes sense to look much more closely at how NHS resource allocation decisions dovetail – or not – with other public service allocations that contribute to health and wellbeing. Calls to align or bring together allocations for local areas will only intensify as discretionary local public services face shrinking (or at best stagnant) budgets for the foreseeable future.

The National Audit Office and Public Accounts Committee have recently reviewed how more than £150 billion of government spending (including education, police, fire and other services as well as health) is allocated to local authorities and other local bodies (House of Commons Committee of Public Accounts 2011). Compared with other services, the Department of Health’s process for resource allocation came out very well. Despite the many problems discussed here, its approach was judged more independent, transparent and rational than that of the other major public service spending departments.

However, the Public Accounts Committee found scant evidence that departments were considering how their individual approaches to allocations contributed to equitable public sector funding allocation in local areas as a whole. Without urgent work to align how budgets are defined across public services, the shared pools of resource available will be increasingly out of sync with the shared needs to be met. As a minimum, the NHS Commissioning Board should align its periodic reviews of health allocation approaches as closely as possible with those of other government departments.

A more radical approach would be a single allocation process that devolved all public sector resources down to local areas. Within that, there are many models that could
be considered, from a wider rollout of community budget pilots to Labour’s emerging single budget vision and the call from the largest cities outside London for an extension to whole place budgets (Johnstone 2013). All of these models would imply a significant change to the current resource allocation process; the Board’s fundamental review will have failed in its title if it doesn’t bring an assessment of the benefits and risks of these issues within scope.

Supporting integrated provision

The rapidly ageing population and its increasingly complex health and care needs have led to a clear cross-party consensus on the need for more integrated care, and services and models that will deliver it (Ham et al 2012; Williams 2012). As part of this, there is growing interest among commissioners and providers in changing the relationship between them, from one where commissioners pay for individual items of care or episodes through Payment by Results (PbR) to one where they hand over capitated amounts to groups of providers or integrated care organisations who are then held accountable for delivering outcomes within the budget. Any savings on the amount generated by the providers are either retained by them or shared with the commissioner on a basis agreed in advance. The idea is that integrated providers may be better placed than commissioners not only to integrate care for patients but to do so more efficiently by allocating resources to parts of the supply chain where patients’ needs can be met most cost-effectively.

If, in future, providers are to take on such budgets, much of the science currently being applied to the resource allocation formula for commissioners can be used to calculate the level of capitation for integrated providers. For example, person-based resource allocation (PBRA), originally developed to support practice-based commissioning and developed further for clinical commissioning group allocations, could be developed to calculate risk-adjusted capitations for integrated providers.

In our view, if applied in this way, PBRA has greater potential than previous approaches to change the incentive structures of individual clinicians. It should make the opportunity costs of clinical decisions clearer; it could also support specialisation and cost-effective service redesign. For example, providers with expertise in diabetes care would have an incentive to increase the number of patients they treated if they could deliver good-quality care at lower than the capitated payment rate for these patients. It should also help to disincentivise providers from not taking on or ‘dumping’ more costly patients, since they would be rewarded more accurately for taking on these patients, rather than receiving average costs only.

The use of PBRA in this way would change the resource allocation process significantly, bringing financial and clinical incentives together for integrated providers for the first time. But while this could have major benefits, there are also significant risks since, in our view, PBRA – far more than any other technical development post-RAWP – has the potential to really begin to affect the incentives and actions of clinicians. While that could be for good, as set out above, it would need to be monitored very carefully to ensure that there are no negative effects; in particular, it could lead to issues with skimping on quality, unless quality is well measured and observed.

But if resource allocation is used to support more integrated provider organisations, this then raises a question about what becomes of clinical commissioning groups (CCGs) – at least in their current role as commissioners. What is their role and function, and on what basis should resources be allocated to them if more of the funding and risk is passed on to provider organisations? We could foresee a situation where commissioners could either receive budgets calculated on a similar basis to that used for the integrated providers, and they use their expertise to commission from the pool of providers, or, if they maintain
significant additional functions, their allocations could resolve into a simpler formula, dependent on the relative size and risk pool of their other responsibilities.

Whatever the method of allocation, in this scenario, commissioners will need a much stronger focus on measuring and maintaining quality to counteract the sharper financial incentives created by the development of more precise payment methods, such as PBRA.

**Supporting mandated services through bottom-up costing**

When RAWP reported in the mid-1970s, clinicians and managers had very little clinical information on costs, effectiveness and cost-effectiveness to aid them in allocating resources. Since then there has been an explosion in clinical effectiveness and cost information, and the development and widespread use of tools such as programme budgeting to inform provision and commissioning decisions.

In this context, it is possible to think of a future where clinical commissioning groups in particular will increasingly argue for a more specific and clinically meaningful approach to resource allocation. Furthermore, the National Institute for Health and Clinical Excellence (NICE) continues to identify a growing set of services which its guidance recommends should be funded, and the NHS Constitution is strengthening the rights of patients to access these services. Over time, these incremental changes could develop into a more explicitly defined set of NHS benefits, though the challenges of doing this are not insignificant (Rumbold et al 2012).

Were such an approach adopted by design or default there may be an even stronger case for commissioners’ budget allocations to be based on delivering these mandated services.

There are a number of ways that such a budget could be derived bottom-up. For example, programme budgets could be allocated to national bodies (eg, the National Cancer Network) or local clinical commissioning groups, allowing them to allocate resources linked to need and demonstrated outcomes. NICE quality standards could be integral to such an approach, and Appleby and Harrison (2007) have set out how such a system might work.

It is not impossible to foresee the increasing information available on the costs of delivering care pathways, and ‘years of care’ for client groups with multiple co-morbidities to be used as the basis for determining allocations, linked to a better direct assessment of population needs for services akin to the Welsh focus on direct needs assessment, perhaps collected as one part of Joint Strategic Needs Assessments.

Depending on how it was rolled out, large proportions of the NHS budget could, as a result, be carved out for specific services – either nationally, regionally or locally – based around certain types of patient or clinical services; these could increasingly be seen as challenging clinical commissioning groups for the lead role as strategic commissioners of locality-based population health care. Unless done well it could also present many of the same risks of fragmentation and creating silos around diseases that we have raised elsewhere (Appleby et al 2012) and here as a core criticism of the current reforms to resource allocation.

So, a more clinically led and integrated NHS, supported by a bottom-up and evidence-based resource allocation system, is possible. We have already seen a significant increase in the NHS Commissioning Board’s specialised commissioning budget announced for 2013/14. But any more radical moves in this direction could have significant implications for localism, the role of clinical commissioning groups in future and, unless undertaken intelligently, could reinforce fragmentation.
Public health resource allocation

Our view is that public health resource allocation needs to fit the vision of the public health system it is intended to support. That vision, in reality, and as set out in Healthy Lives, Healthy People (Department of Health 2010), is an uncomfortable but perhaps unavoidable blend of two opposing approaches. The public health reforms are a halfway house between a ‘traditional’ approach to public health – the prescribed delivery of particular services – and a much more radical outcomes-based approach, reflected in the development of the Public Health Outcomes Framework (PHOF). How resources are allocated should reflect this compromise, and adapt if and when it changes.

A public health system based on tight definitions of prescribed or mandated services would require local authorities to receive a budget based on a bottom-up assessment of costs, service by service, taking into account the size of local populations and their needs. In contrast, a public health system that is geared around good performance on the PHOF implies a much looser definition of services to be delivered, since outcomes in local areas will be driven by very different needs, requiring local flexibility on how to spend the money. This approach is more suitable to a formula-based approach to allocation.

Since the government’s new vision for the public health system is clearly a blend of these approaches, it makes sense that its approach to resource allocation should also be a blend. In fact (as set out earlier) the decision has been made to go with a formula alone based on SMR <75. But in practice, the future system will retain significant elements of mandated services, such as sexual health, and in large metropolitan areas such as London, such provision will account for a significant proportion of budgets.

Public health resource allocation therefore needs to take this into account, since the danger of relying on a formula-based approach alone is that it cannot help determine the overall scale of resources needed. When local authorities have to provide specific services, this could lead to systematic under-funding because the overall amount allocated to public health is simply not adequate, even if the share of the budget received by a certain authority can be justified. Table 2, below, shows a more practical, blended approach to public health resource allocation, with a significant mandated element.

Moving beyond the current decision, we therefore welcome ACRA’s proposals to the Secretary of State to investigate other approaches, including elements for bottom-up costing where appropriate (see ACRA 2012) and the House of Commons Communities

Table 2 A blended approach to public health resource allocation

<table>
<thead>
<tr>
<th>Functions</th>
<th>Key characteristics</th>
<th>Implications for allocation method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health protection*</td>
<td>■ Large fixed costs. Relatively clear definition of services, with high risks to population health if not in place</td>
<td>■ Bottom-up costing of key services that need to be in place</td>
</tr>
<tr>
<td>Wider determinants</td>
<td>■ Great variety of drivers and interventions which differ by area. No clear definition of services that need to be in place</td>
<td>■ Formula based on outcomes to be achieved. Should include within-area as well as between-area focus</td>
</tr>
<tr>
<td>Health improvement</td>
<td>■ Prescribed health improvement services (eg, sexual health)</td>
<td>■ Bottom-up needs-based, where possible based on cost-effective service configuration</td>
</tr>
<tr>
<td></td>
<td>■ Non-prescribed services. Needs differ according to population demographics, and different areas will innovate and take different approaches</td>
<td>■ Formula based on outcomes to be achieved. Should include within-area as well as between-area focus</td>
</tr>
</tbody>
</table>

* Healthy Lives, Healthy People (Department of Health 2010) signalled that Public Health England will be taking its own approach to health protection allocations.
and Local Government Committee’s recognition that this is an important issue (House of Commons Communities and Local Government Committee 2013), and we urge the Secretary of State to commission them to do so. Given the shift in public health responsibilities and funding to local authorities, there is also an opportunity to consider these alongside other allocations to local authorities (as explored above).

Conclusion

Resource allocation in England has a long and distinguished history. In times of plenty this process goes largely unnoticed and unremarked, as everyone gets a slice of the growth in health resources. That is set to change as growth in health funding stalls and as more stakeholders, including clinical commissioning groups and local authorities, will be affected by health resource allocation decisions.

The basis on which these decisions are made therefore needs to be – and be seen to be – sound and robust, as well as fully transparent.

It is clear that resource allocation will become more fragmented as the new system is implemented and as the number of commissioners proliferates; this poses substantial risks to a system that needs to integrate around whole care pathways and populations. There is as yet no sign that the Department of Health is addressing these risks.

Transparency and clarity will be critical if the NHS Commissioning Board is to convince clinical commissioning groups, let alone local authorities and local populations, that its allocations are fair. The formula approach’s complexity makes it difficult to understand for all but a few expert statisticians. Furthermore, it is not clear that the complexity involved is actually making material differences compared with a simpler approach.

One option would be to build on the current formula-based approach. The NHS Commissioning Board could introduce more transparency and more opportunities for consultation. It is also critical to look again at how needs are measured and the speed of pace of change, as well as whether resource allocation decisions should support broader health policies rather than sitting in isolation.

But the fundamental review is an opportunity for a more radical rethink of the resource allocation process given the changes wrought by the health and social care reforms. It is time to assess whether resource allocation should be increasingly based on outcomes, more integrated with other public service budgets, or more explicitly linked to clinical standards and mandated services, as well as considering the implications of changes in the purchaser-provider relationship as a result of the development of integrated care organisations. While none of the approaches we explored would be straightforward, we believe they need to be appraised as it is vital to ensure that the resource allocation process is aligned with the broader system within which it operates.

Above all, the fundamental review must address this question: is resource allocation simply a mechanism for moving centrally collected funds to local areas, or is it a tool for achieving wider policy goals? If it is the former, this could suggest a return to a much simpler approach since the complexity does not seem to be making much difference to final target allocations. But if it is accepted that resource allocation is, or should be, a policy tool then a much more fundamental review is called for. And, as Bevan (2008) noted: ‘It would seem to be beneficial to enable ACRA to take time out to explore possible directions in the longer term that are not subject to two constraints: the imperative of producing changes within a year, and inadequacies in existing data’.

The fundamental review needs to allow ACRA to take a longer-term approach, with the authority to commission research and substantial data collection to support the testing
of more innovative approaches and pilots as the new system evolves. While ideally, that process should inform the 2014/15 clinical commissioning group allocations, as proposed, it should not be restricted to that.

We would also suggest that the decision to base public health allocations solely on a formula approach may need to be revisited, and we would encourage Public Health England to partner with the NHS Commissioning Board and include public health allocations in the fundamental review.

References


Appleby J (2011). ‘An NHS ice age may have only just begun’. Health Service Journal, 1 June.


About the authors

David Buck is a Senior Fellow at The King’s Fund, specialising in public health and inequalities. Before joining the Fund in 2011, David worked at the Department of Health as head of health inequalities. He managed the previous government’s PSA target on health inequalities and the independent Marmot Review of inequalities in health and helped to shape the coalition government’s policies on health inequalities. While in the Department he worked on many policy areas including diabetes, long-term conditions, the pharmaceutical industry, childhood obesity, and choice and competition. Before working in the Department of Health, David worked at Guy’s Hospital, King’s College London and the Centre for Health Economics in York, where his focus was on the economics of public health, and behaviours and incentives.

Anna Dixon is Director of Policy at The King’s Fund. She has conducted research and published widely on health care funding and policy. She has given lectures on a range of topics including UK health system reform and patient choice. She was previously a lecturer in European Health Policy at the London School of Economics and was awarded the Commonwealth Fund Harkness Fellowship in Health Care Policy in 2005/6. Anna has also worked in the Strategy Unit at the Department of Health, where she focused on a range of issues including choice, global health and public health.

Acknowledgements

The authors would like to thank Gwyn Bevan, John Appleby and two anonymous reviewers for reviewing the text and Steve Casson for research support while working at The King’s Fund as a public health trainee.