MAKING THE MOST OF THE MONEY

Efficiency and the long-term plan
## CONTENTS

- Foreword 4
- Key points 5
- Introduction 7
  1. The efficiency challenge 10
  2. Cost reductions 17
  3. Productivity improvements 22
  4. System efficiencies 25
  5. Support 32
  6. Solutions and recommendations 37
- Conclusion 39
- Annex 40
- References 41
The forthcoming funding increase and long-term plan for the NHS present a valuable opportunity to restore the service to sustainable good performance and financial health. However, in exchange for the funding increase, ambitious and stretching improvements will be required of the health service. Whether the new money is enough to fund these improvements will significantly depend on the extent to which the NHS can free up resources by reducing waste and becoming more efficient and productive. This is a challenge for individual organisations, but increasingly requires system partners – providers, clinical commissioning groups (CCGs) and local authorities – to work more effectively together to ensure they deliver the best possible value to the public for every collective pound.

The discussion on provider sector efficiency has often been conducted as a one way, top down, debate with NHS national system leaders setting the overall level of efficiency ask each year and outlining the areas where they believe efficiencies can be made. But it is front line NHS trust leaders who have to own that task and deliver the required savings. If they do not feel ownership of the task, or if they feel the nationally set task is undeliverable, then the potential for greater efficiency can be easily lost.

As the membership organisation for the NHS provider sector, we therefore wanted to give trust leaders a voice in the efficiency debate, particularly as national system leaders finalise the new NHS long-term plan. We have canvassed opinion from trust leaders via an online survey, and conducted a series of qualitative interviews with trusts and national policy makers. We asked trust leaders about their confidence in delivering a stretching efficiency task, where the opportunities lie, what barriers exist and what support they need to continue to improve their productivity in a financially constrained environment.

We hope that their response is a timely and helpful addition to the national debate and would like to thank all of our contributors.

Chris Hopson
Chief Executive, NHS Providers
The new NHS funding settlement and the forthcoming long-term plan offer an opportunity to address systemic operational and financial pressures in the provider sector and transform the way healthcare is delivered. How far trusts can drive efficiency will be a crucial factor in determining how much the new money will result in an improved offer to the public.

Trusts have consistently delivered impressive efficiency savings in each year since the beginning of austerity in 2010/11. Providers delivered almost £7bn of recurrent efficiency savings in the past three years alone (NHS Improvement, 2015, 2016, 2017, 2018), on top of the £20bn ‘Nicholson challenge’ delivered from 2010 to 2015. The NHS already performs well on length of stay, administration spend and drugs compared with other health systems internationally. These are significant achievements which should be celebrated.

However, trust leaders across ambulance, community, mental health and hospitals are clear that they have now largely exhausted the ‘easily realisable’ savings from within their own organisations. They are increasingly reliant on non-recurrent savings or on more ambitious, longer term transformation plans which require central support and upfront investment, as well as productive partnerships in local systems.

There are three main categories of efficiency: cost reduction schemes, productivity improvements and system efficiencies. There should be an acknowledgement that, for each of these categories, the scale of the opportunity and time required to realise efficiencies will vary significantly and are dependent on local context. Trust leaders are also clear that some efficiency activity leads to improvements in care quality, as opposed to the release of cashable savings. This needs to be taken into account more effectively in national assumptions on efficiency.

Since 2010/11, the assumption has been that the rate of savings will accelerate if trusts are given a stretching efficiency requirement through a real terms deflating of payments to providers. To some extent the provider sector has responded well to this and achieved unprecedented levels of efficiencies. However, the evidence since 2014/15 demonstrates that an over ambitious efficiency requirement simply leads to a larger provider sector deficit. Using an ambitious efficiency requirement to make the NHS budget balance is inherently risky.

Trusts believe that better system working, integrating services and addressing workforce challenges offer the greatest opportunity to improve quality of care and patient experience, as well as reduce costs and use resources more efficiently. Trusts are committed to pursuing these approaches, but need adequate time to implement and deliver on ambitious local system-wide efficiency programmes. Trusts believe this work could be accelerated if there was a better balance between, on the one hand, support for longer-term transformation and, on the other, frameworks that encourage short-termism, such as the cost improvement programme (CIP) and control total regimes.
While trust leaders can see efficiency opportunities, both within their organisations and, to a greater extent, at a system level, they are less confident that they can continue to make efficiencies at the current rate. They are therefore looking to the long-term plan to set a stretching, yet deliverable, efficiency requirement of the sector. This requirement should be based on realistic assumptions and be accompanied by a framework setting out the potential areas of savings and the support trusts can expect from peers and national bodies.

Trusts welcome the support offered by NHS Improvement and embrace the value of programmes such as Lord Carter’s efficiency and productivity work, *Getting it right first time* (GIRFT) and Model Hospital. They are, however, more sceptical about some of the bold extrapolations these programmes have made about the size and speed of efficiency savings that can be realised. Trusts told us that they would like more support, rather than punitive action, from the central bodies.

Overall, trusts recognise their responsibility to plan for and deliver a stretching efficiency ask within their trusts and local systems. Trust leaders believe that a stretching efficiency requirement is more likely to be delivered if:

- national system leaders, trusts and local systems work in partnership to agree a realistic efficiency requirement and timeframes for delivery, rather than have these centrally imposed, top down, from the national system level
- national system leaders do more to help trusts and local systems identify and share learning from across the country about where efficiencies can be delivered
- national system leaders are realistic about the extra support that trusts and local systems will need to realise the transformational savings and local system focused savings that must now form the focus of efficiency and productivity activity
- any efficiency delivery assumptions explicitly match the capacity and capability available on the frontline given the ambitious list of other priorities trusts are expected to deliver.
INTRODUCTION

The NHS has been given a real terms funding increase of £20.5bn over the next five years. In exchange, ambitious and stretching improvements will be required which demonstrate a clear improvement on this investment to the government and the public. Whether the money is enough to fund the expectations set out in the forthcoming long-term plan will depend largely on whether the NHS can free up resources by cutting waste and becoming more productive.

The improved funding package for the NHS, which was announced by the prime minister in June, offered 3.4% average real terms growth over five years. This is significantly more generous than any settlement offered to any other public service and recent NHS funding increases. However, it is below the long run average NHS funding growth and will barely allow existing models of care to keep pace with rising demand (Institute for Fiscal Studies, 2018).

The NHS is currently struggling on most major finance and performance indicators:

- the 95% for A&E waiting times has not been met since July 2015 (NHS Providers, 2018)
- in June 2018 the 18-week waiting list is likely to have been around 4.3 million (NHS Providers, 2018)
- the target for 92% of patients to wait no more than 18 weeks from referral to treatment has not been met since early 2016 (House of Commons Library, 2018)
- the provider sector has recorded a deficit for the last five years while CCGs have recorded deficits in two of the last three years
- the provider sector finished almost £1bn in deficit in 2017/18 (NHS Improvement, 2018)
- the underlying deficit now stands at £4.3bn (NHS Providers, 2018).

Trusts are confident that the long-term plan can set out a vision for a transformed and sustainable NHS – but, to be deliverable and have credibility, it must give trusts a realistic and achievable task on efficiency. We must learn the lessons from past policy initiatives including the Five year forward view where realising expected efficiencies has proved problematic.

The Five year forward view won widespread support with its vision of a more prevention-focused health service. However, delivering this vision has proved difficult, in part because of the plan’s over-optimistic assumptions of how much the rate of efficiency could increase. For the budget to balance, the government’s injection of £8bn real terms growth over five years depended on the NHS delivering an unprecedented 2-3% rate of efficiency compared with a long run average of 0.8% (NHS England, 2014). Service redesign on the scale required, and the timeframe for releasing efficiencies from new models of care, turned out to be more complex than envisaged.

In addition, since funding began to be constrained in 2010/11, there has been a national assumption that deflating the tariff, and other payment mechanisms, would cause an acceleration in the annual rate of efficiency savings achieved by the provider sector and that doing this consistently over a number of years would generate a sustained rate of efficiency above the long run average achieved by the service.
To some degree this has proved successful as in the three years up to 2017/18 trusts delivered £6.9bn efficiency savings – £2.3-2.4bn of sustainable, recurrent efficiency savings each year (NHS Improvement, 2015, 2016, 2017, 2018). However, in part as a consequence of the tariff deflator, in 2017-18 102 trusts (44%) finished the year in deficit, with financial deficits now the norm in the acute sector. In addition, trusts have become increasingly reliant on unsustainable one-off, or ‘non-recurrent’ savings such as technical accounting adjustments, land sales, vacancy freezes and delaying essential maintenance works to deliver the efficiencies required.

The debate around NHS efficiency is often based around a single, headline efficiency requirement for the whole NHS, which trusts believe is over-ambitious yet national leaders feel is appropriate given that waste exists in the system.

As the NHS develops its new long-term plan, we believe it is important to have a more sophisticated debate on efficiency. It is frontline NHS trust leaders who have to own the efficiency task and deliver the required savings. If they do not feel ownership of the task, or if they feel the nationally set task is undeliverable, then the potential for greater efficiency can be easily lost.

As the membership organisation for the NHS provider sector, we have used our unique access to trusts to explore the efficiency issue in depth – using a quantitative survey and a series of qualitative interviews. This document sets out the views we have gathered from trusts. It aims to share the trust level perspective to better inform the policy debate in this area.

The survey was completed by 157 provider leaders, representing 54% of trusts and foundation trusts. This work, carried out between July and September this year, forms the basis for this report and its recommendations. In our research we sought to find out:

● how trusts have risen to the efficiency challenge over the past three years
● how trusts feel about their ability to deliver an increasingly stretching efficiency requirement
● where savings can still be made
● what national leaders can do to help trusts and local systems become more efficient.

This report examines these issues over six chapters. The first examines current performance on efficiency and productivity and how confident trusts are that it can be sustained. The following three chapters explore the opportunities offered by three main categories of efficiency providers cited: reducing costs, making existing services more productive and restructuring systems to make them more efficient. The final two chapters look at the approach taken by national leaders, and provide a set of recommendations for the future.
More than half (54%) of trusts in England responded to our survey on NHS efficiencies.

**VIEWS ON NHS EFFICIENCY SAVINGS**

- **Provider confidence levels over the efficiency ask of the sector**

<table>
<thead>
<tr>
<th></th>
<th>Very confident</th>
<th>Confident</th>
<th>Neutral</th>
<th>Not confident</th>
<th>Not at all confident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Your trust can increase its rate of efficiency savings in the next five years?</td>
<td>10%</td>
<td>15%</td>
<td>45%</td>
<td>20%</td>
<td>10%</td>
</tr>
<tr>
<td>Your trust will be set a reasonable efficiency target in the next five years?</td>
<td>15%</td>
<td>22%</td>
<td>41%</td>
<td>23%</td>
<td>11%</td>
</tr>
<tr>
<td>Your trust will begin 2019/20 in a better financial position than it began this year?</td>
<td>10%</td>
<td>16%</td>
<td>34%</td>
<td>24%</td>
<td>14%</td>
</tr>
</tbody>
</table>

- **59%** disagreed that they had delivered efficiency savings at a rate they could safely continue.
- **54%** disagreed that saving expectations have been realistic.

**WHERE CAN THE SAVINGS BE MADE AND WHAT ARE THE MAIN BARRIERS?**

**Top 5 initiatives which offer the best opportunity for large efficiency savings**

1. Reducing unwarranted clinical variation
2. Better workforce planning
3. New models of care
4. More productive use of facilities infrastructure
5. Changing skill mix

**Top 5 barriers to making efficiency gains**

- Short termism encouraged by current financial framework: 69%
- Operational pressure: 66%
- Lack of transformation funding: 46%
- Lack of access to capital: 39%
- Lack of system working: 27%

**KEY RECOMMENDATIONS**

- Realistic delivery plans for each area, agreed by local trusts
- Support for sustainable savings schemes rather than single year efficiencies
- More focus and resource from NHSI to help trusts make improvements
This chapter examines the starting point for any future efficiency challenge. It describes whether trust leaders believe there is waste in the system, and how confident they are that they can maintain or accelerate their current performance on efficiency. It also describes the main current measures for reporting efficiency – CIPs and implied provider productivity. It explores the rise in non-recurrent savings, and the difference between efficiencies that release cash and those that do not. It also briefly reviews the current NHS provider sector efficiency approach.

Trust leaders’ confidence in realising efficiencies

Trusts acknowledge that there is waste in the NHS (figure 1) but they are strikingly lacking in confidence (figure 2) that they can eliminate it or continue to tackle it at the rate delivered to date.

Our survey suggests the reasons for this potential paradox are that trust leaders believe:

- there is no such thing as a system without waste and the NHS already compares favourably with other systems internationally
- the most easily realisable savings have largely been made and continuing to identify further efficiency gains is becoming increasingly difficult
- the current degree of day-to-day operational pressure in trusts is limiting management bandwidth for other initiatives, including realising efficiency savings
- trusts do not have access to the resource required to deliver the complex transformation needed to improve efficiency for the medium to long term
- there is a danger of too great a national focus on short-term savings being generated within single organisations, potentially diverting attention away from system transformation that trusts believe would now yield the largest, long-term, efficiency benefit.

Figure 1 demonstrates that trust leaders can still see waste in the NHS – however the comments attached to this part of the survey show that they also believe the NHS compares well with other systems internationally. This is consistent with a 2017 study published by the Institute for Fiscal Studies which found that the NHS performed well on length of stay, administration spend and drugs costs (Institute for Fiscal Studies, 2018).

Nearly three quarters of trust leaders surveyed agreed or strongly agreed that the NHS wastes too much money through inefficiency, while around half agreed or strongly agreed that their own trust is too inefficient. The fact that more respondents acknowledged waste in the NHS overall than in their own organisation suggests that trusts believe that reforming local system working now represents a greater opportunity for efficiency gains than cost reductions or productivity gains within a single organisation. This is explored in more detail in chapters 2, 3 and 4.
**Figure 1**

To what extent do you agree that the NHS and your trust wastes too much money through inefficiency?

<table>
<thead>
<tr>
<th></th>
<th>Your trust (n=157)</th>
<th>The NHS (n=157)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly agree</td>
<td>48%</td>
<td>12%</td>
</tr>
<tr>
<td>Agree</td>
<td>22%</td>
<td>59%</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>21%</td>
<td>17%</td>
</tr>
<tr>
<td>Disagree</td>
<td>11%</td>
<td>11%</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

"International studies show the NHS is one of the most efficient health systems in the world. I think improvements can be made but it is overstated."

Operations director, combined acute and community trust

"There is room for further efficiency, however there is also a tremendous amount of work done as a result of the goodwill of staff. If this were included the NHS would be highly efficient on any benchmark."

Chief executive, mental health and community trust

"There is always room for improvement and much scope for increased efficiency comes from the systems that we [are] operating, the regulatory burden we are under and the need to maintain a wide range of performance metrics rather than a focused view. There is scope to release time to care but requires a brave look at competition, contracting, performance measurement procurement and regulation, as well as a rigorous focus on value and productivity within teams."

Chief executive, combined mental health and community trust

"I think the inefficiency is strongly linked to capacity and any business would be inefficient if it is running at over 100% occupancy so efficiencies cannot be addressed without addressing urgent and emergency care."

Finance director, acute trust

However, figure two demonstrates that 61% of respondents were ‘not confident’ or ‘not at all confident’ that their trust can increase its rate of efficiency savings in the next five years.

Only 15% are ‘confident’ that their trust will enter the next financial year in a better position than it began this year – 48% were ‘not confident’. Meanwhile, 59% ‘disagree’ or ‘strongly disagree’ that they have delivered efficiency savings at a rate they can safely continue. Our survey also found that almost two thirds were ‘not confident’ or ‘not at all confident’ that their trust will be set a reasonable efficiency target in the next five years.
In interviews, finance directors from all sectors consistently told us that they think the easy wins have been exhausted, and the savings available to them now are harder to realise than they were three or four years ago. They described a task that gets incrementally more difficult each year.

Although trusts have so far sustained their performance on delivering efficiency savings, past achievements do not guarantee future success. It is also concerning that respondents from some solvent, high-performing trusts felt that 2018/19 would be the first year where meeting efficiency targets does not look achievable.

In their comments, survey respondents suggested that their low confidence levels stemmed from unrealistic expectations set by the national bodies:

“Although there are efficiencies to be gained, the requirements centrally are far in excess of this.”
Divisional director, acute specialist trust

“The overall ask will not be reasonable due to a mismatch between resources and expectations/demand.”
Finance director, acute trust

Trusts also told us that their existing CIP requirements would be unsustainable in the future, as they had exhausted many of their options:

“The trust has delivered 4% efficiency annually for the last seven years. The ability to deliver this recurrently in 2019/20 is not going to be possible”
Chief executive, mental health/learning disability trust

“We are one of the few trusts to make financial targets consistently for the last six years, but now running out of ideas of how we can take 5% every year out.”
Chief operating officer, combined acute and community trust
“We are being increasingly driven from the top, with block contracts for acute work, multiple regulators pulling in different directions, and demand not being met in the community. This will make delivery of efficiency savings required very difficult.”

Strategy/Transformation director, acute trust

Respondents also cited the current uncertain context within which they are working as a major challenge to their trust making further efficiency savings. This context included workforce pressures, projected increase in demand, increasing acuity, changes to payment mechanisms and uncertainty over Brexit.

Trusts did not feel that the NHS would have the capacity to continue making efficiency improvements at the current rate because the resource available to make improvements – management time and the investment of staff time in new ways of working – is finite, and is increasingly committed to efforts to improve performance and quality of care in the face of a relentless growth in demand.

One trust leader told us that staff were exhausted by change, which is exacerbating issues with retention, and will in turn increase costs and reduce performance overall. Another said a plan to implement a new electronic patient record system would probably divert attention from schemes that could save money quickly, even though it would improve how the trust operated in the long term. These are just two examples of the frontline operational trade offs trust leaders must make in order to balance efficiency requirements with other imperatives.

Taken together, this demonstrates the scale of the task facing national leaders as they draw up the long-term plan. There is a clear risk that the difference between the money available and the understandable aspiration for improvements will result in an efficiency requirement that the service does not believe can be delivered.

**CIPs, efficiency and productivity – what do they all mean?**

The debate around efficiency in the NHS is clouded by a lack of clarity in the terminology commonly used. NHS Improvement’s quarterly reports include details on two measures: implied provider productivity and CIPs.

The implied provider productivity measure is calculated by reviewing the annual change in provider costs, adjusted for inflation, and then comparing these with the change in provider outputs. It is a pure productivity measure, which does not take quality of care or operational performance into account. It therefore does not measure increases in value, which may be increased by an improvement in quality for no extra cost. For 2017/18, the implied productivity measure was a 1.2% gain.
CIPs are individual trusts’ efficiency targets, which are reported to NHS Improvement and aggregated up to give a national figure. The efficiency targets are set at the beginning of the financial year based on the projected income for each trust, set against expected costs. Trusts must reduce costs, contain expected cost growth or demonstrate productivity improvements to calculate their CIP. For 2017/18, trusts overall made CIP savings of 3.7%, against a plan of 4.3%. This is an average figure – as the comments quoted above make clear, many trusts’ targets will be 5% or more.

There are problems with using CIP as a metric for comparing efficiency between organisations. While CIPs are reported centrally, there is no consistency from one trust to another about how they are calculated or what is included in them. For example, a trust may deliver an increase in productivity for a particular service for low marginal cost, which will reduce the average unit cost of each episode of care. While this does not release any cash, some trusts include the reduced unit costs in their CIPs, while others do not. CIPs also include non-recurrent savings, which may release funding in-year but do not represent genuine sustainable efficiency. Many trusts have also used income gains, for example by increasing their elective surgery activity, to contribute to their CIP savings targets.

Non-recurrent savings

The percentage of trust CIPs being delivered via one-off, or non-recurrent, savings has been steadily rising in recent years. In 2015/16 non-recurrent schemes represented 23% of total savings, but by 2017/18 this had risen to 26% (NHS Improvement, 2018). There are two main drivers for this: the growing gulf between the efficiency gains trusts can make and what they are being asked to deliver and the control totals regime potentially penalising trusts that do not hit their surplus/deficit targets. Trust leaders tell us that the latter factor has rewarded them for making short-term and what, in some cases, they describe as “desperate” savings.

Non-recurrent savings can encompass any measure that improves a trust’s in-year financial position, but which do not fundamentally change the cost of running the organisation and cannot necessarily be repeated in subsequent years. These might involve land sales, holding vacancies open and delaying maintenance work. They may also include technical adjustments, such as revaluing estates or changing the way annual leave is accounted for.
Figure 3
How do you expect the proportion of non-recurrent savings made by your trust in 2018-19 to compare to last year?
(n = 51)

<table>
<thead>
<tr>
<th>Category</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher than last year</td>
<td>39%</td>
</tr>
<tr>
<td>Same as last year</td>
<td>37%</td>
</tr>
<tr>
<td>Lower than last year</td>
<td>24%</td>
</tr>
</tbody>
</table>

Figure 3 demonstrates that, of finance directors who responded to our survey, 39% said they would have to rely more on non-recurrent savings to hit their cost improvement plan targets this year. Just 24% expected a lower rate of non recurrent savings compared to last year.

However, this rising reliance on one-off measures cannot be sustained indefinitely. As the finance director of one high performing acute trust put it: “We are rapidly running out of family silver to sell”.

Cash releasing savings

We asked finance directors roughly what percentage of efficiency savings they could consistently make every year which would release cash or resource that could be reinvested in services.

There was a range of responses to this question. Some were optimistic in their response and suggested that 100% of these efficiency savings could fund additional services, whereas 45% of respondents felt that these efficiency savings would release no additional resource to be spent on services. The average (mean) percentage given by finance director respondents was 29% and the median was 10%.

Although waste exists in the system, trust leaders have told us that efficiencies cannot continue to be delivered at their current rate in the way they have been in the past (figure 2). There is an increasing reliance on non-recurrent savings and many CIPs do not release cash. For example, a trust may make excellent productivity gains, and be able to treat more people per day. These gains may count towards a trust’s CIP, but will not reduce its costs or release cash, because it will still be employing the same number of staff and running the same facilities.
Similarly, a trust may move its back office staff into cheaper accommodation offsite. This may allow an office area to be converted into a ward and therefore increase inpatient capacity more cheaply than if it had to be built from scratch. Again, such an intervention, while valuable, does not release resource to be spent elsewhere in the way that closing the back office function entirely and ceasing to employ its staff would have done. This is especially true for services where there is unmet need. Mental health trusts particularly tell us when they make their services run more efficiently, it can only ever enable them to reach more people – not take costs out or free up resource to be used elsewhere.

It can therefore simultaneously be true that the NHS is performing impressively on improving efficiency, and that this is not freeing up much resource to accommodate new demand or improve the service offer.

Current approach to provider sector efficiency

The NHS planning process sets an annual provider sector efficiency assumption. This is then reflected in the tariff which determines the prices acute providers are paid for different treatments and procedures and the underlying uplift to the block contracts prevalent in the mental health and community provider sectors.

The danger of a national efficiency assumption historically is that it has been treated as the “balancing item” against the overall financial allocation for the service. Given that annual NHS funding increases since 2010 have mostly been significantly lower than the increases in demand and costs, this has led to very stretching efficiency targets for the provider sector.

Different efficiencies will be available to different trusts – responsibility for identifying and delivering these lies with the trust board. One of the limitations of a blanket, tariff-driven approach is that it does not take into account vital factors such as the differences between local systems, whether a trust is in difficulty or not, whether there are structural problems in its local system that are beyond its control, whether it needs investment, or how easily it can make the savings required to break even or record a surplus.

NHS Improvement’s operational productivity team helpfully aims to identify where inefficiency sits in the system and provides benchmarking data to the provider sector. This is done through national programmes, such as Lord Carter’s operational productivity work, GIRFT and Model Hospital. These initiatives arm trusts with information and data, however, this by itself does not guarantee that inefficiency can be eliminated.

The next three chapters address the main approaches that may be taken to generate further savings. Efficiency initiatives can release more than one type of productivity gain however we have aimed to organise the material against the primary driver of efficiency: cost reductions, productivity improvement and system change.
This chapter focuses on where the opportunities are to reduce costs within the NHS as distinct from improving productivity or redesign that will result in a more cost-effective system. It addresses three major areas where efforts have been concentrated in cost reduction – reducing transactional spend, typically via better contracting, reducing the costs of agency staff and procurement. It also explores national factors such as the prices of drugs and devices, which influence the NHS’s cost base but which trusts cannot control individually.

When asked where the greatest opportunities lie for further efficiencies, trust leaders say the biggest savings are not necessarily those that will be realised quickly. Figure 4 shows that trust leaders believe that cutting agency spend and making technical accounting adjustments offer the best immediate opportunity. It also shows that the largest efficiency saving opportunities tend to require complex, long-term transformation, which may require agreement across the clinical community or work across several institutions in a local system. For example, two opportunities identified by trust leaders for efficiencies at scale were addressing workforce challenges through better workforce planning and changing the skill mix.

**Figure 4**

*Which specific initiatives offer the best opportunity for efficiency gains for your trust in the next five years?*

(n = 157)

**Quick savings.**

<table>
<thead>
<tr>
<th>Initiative</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reducing agency costs</td>
<td>53%</td>
</tr>
<tr>
<td>Technical/accounting measures</td>
<td>34%</td>
</tr>
<tr>
<td>Reducing unwarranted clinical variation</td>
<td>31%</td>
</tr>
<tr>
<td>Improving pathways within your organisation</td>
<td>31%</td>
</tr>
<tr>
<td>Changing skill mix</td>
<td>27%</td>
</tr>
<tr>
<td>Back office consolidation</td>
<td>25%</td>
</tr>
<tr>
<td>Better workforce planning</td>
<td>23%</td>
</tr>
<tr>
<td>More productive use of facilities/infrastructure</td>
<td>23%</td>
</tr>
<tr>
<td>Networking services with other trusts</td>
<td>18%</td>
</tr>
<tr>
<td>New models of care</td>
<td>13%</td>
</tr>
<tr>
<td>Economies of scale via trust mergers/takeovers</td>
<td>7%</td>
</tr>
<tr>
<td>Outsourcing</td>
<td>6%</td>
</tr>
</tbody>
</table>
Large savings available.

Reducing unwarranted clinical variation: 55%
Better workforce planning: 55%
New models of care: 46%
More productive use of facilities/infrastructure: 38%
Changing skill mix: 38%
Improving pathways within your organisation: 34%
Networking services with other trusts: 31%
Reducing agency costs: 31%
Economies of scale via trust mergers/takeovers: 30%
Back office consolidation: 18%
Outsourcing: 8%
Technical/accounting measures: 5%

Reducing transactional spend

Removing overhead costs could involve transforming a trust’s use of property and estates – for example moving staff onto sites it owns and out of leased accommodation. But it could also include commercial savings. An ambulance trust told us they saved up to £4m by better managing their contracts with firms they lease their fleet from, for example by challenging unnecessary repair works.

Trusts have also outsourced transactional services such as payroll, catering or outpatient pharmacy, to firms that can run these services more cheaply at scale. Interventions such as these do not necessarily make a trust run more efficiently – but they do make it more cost-effective.

Trusts also told us that saving money through better contract management is not “free” – it can only be achieved where trusts continue to invest in commercial skills. Plus, saving money on contracts may require more intensive management, such as monitoring and challenging contractual performance, which can lead to additional costs.
Agency spend

The nationally coordinated drive to reduce the costs of agency nursing shifts has been an unqualified success. Trusts have managed to reduce expenditure on agency staff from £3.7bn in 2015/16 to £2.4bn in 2017/18. Respondents to our survey rated it highly as an effective contribution to the efficiency drive from NHS Improvement (see chapter 5), while one interviewee said it was a “game changer”.

The approach taken by NHS Improvement, which set caps on the amount that could be paid, and a clear set of rules for breaching the cap when safety required it, shows how national bodies can act effectively on behalf of all providers. Trusts acting alone could not have achieved similar reductions, as the market in agency labour made them competitors, driving prices up.

The reduction in agency spend also illustrates an important nuance in the efficiency debate about reducing cost, increasing value and releasing cash. In 2017/18 for example, while trusts reduced expenditure on agency staff by over £500m, overall spending on temporary staff remained broadly flat because they ended up spending more than expected on bank staff. This represented better value, as trusts get more shifts per pound from bank spending than they do from agency and most likely improved quality of care by maintaining staffing levels.

However, it also raises an important point – while an intervention may appear to reduce costs when viewed in isolation, it may have knock-on effects elsewhere in the system. In this case trusts argue that while the agency cap had the potential to release cash, they still spent this money on temporary staffing, rather than being able to use it in other places within the trust. Therefore these approaches are most certainly worth doing, but may not release money that can be spent, for example, on pump-priming transformation.

Procurement

The move to centralise procurement under the Department of Health and Social Care’s (DHSC) ‘future operating model’ scheme is intended to have a similar impact to the agency cap, by introducing national negotiation to yield savings that trusts acting alone cannot.

However, the approach is not without risk. Firstly, any reductions in prices paid will have to cover the substantial increase in administrative costs before they yield cashable savings. Trusts will be expected to fund almost £0.5bn over the next two years to pay for the new centralised procurement service. Secondly, the analogy with agency spending is not perfect. With agency staff, competition between trusts was driving up prices – this does not apply to routine consumables and services. Finally, procurement is already undertaken at scale through regional hubs – the DHSC plans for the proportion of regional procurement to fall from 40% to 10%, and the national spend to rise from 40% to 80%. However, it is not clear that moving from regional to national purchasing will lead to significantly lower prices.
One finance director in a trust providing some extremely specialised services noted that their organisation was already effectively conducting national purchasing, where it was buying devices for services that few other trusts provided. In cases such as these, providers will have long-established relationships with suppliers, and expert knowledge about which products support the best outcomes. In such cases, it is reasonable to doubt whether centralising procurement will result in either better prices or better results for patients.

**Technology**

More than one interviewee told us IT procurement constitutes an opportunity for a more coordinated approach to release savings. One acute finance director said trusts on their own in his region had been unable to coordinate purchasing computers, tablets and smartphones because they had different clinical cultures. Another community trust told us that Microsoft licences, which used to be paid for centrally, now cost their organisation £300,000 a year. These could potentially be purchased more effectively at scale.

The opportunities that better use of technology offers for system working are detailed in chapter 4.

**National factors**

There are significant risks and opportunities impacting on costs that are beyond the control of individual trusts and sometimes the leadership of the national bodies.

This year marked the beginning of the end of eight years of pay restraint in the NHS. This is necessary and welcome, and staff have borne the burden of austerity for too long. However, the new pay deal will increase trusts’ cost base in 2018-19 – and by more than the £800m made available by the Treasury to cover the pay rise. Holding down pay rises at rates below inflation since 2010 has artificially made the NHS appear more productive: outputs have risen, while cost inputs have risen more slowly (National Audit Office, 2018). While paying staff more is the right thing to do, it will have the effect of making the NHS appear to be losing ground on some measures of efficiency as input costs begin to rise more quickly with no related increase in outputs.

Another important national factor beyond trusts’ control is the price of drugs. The total drugs spend in the NHS was around £17.4bn last year (The King’s Fund, 2018), meaning even just a 6% shift could cost the NHS around £1bn, equivalent to 35% of the total planned recurrent CIPs. It is clearly vital that national bodies use the collective bargaining power of the NHS to get the best deal possible for medicines. But it is also dangerous to assume that they are not already achieving good value. It should be remembered that the NHS remains exposed to changes to exchange rates, for example, or inflationary pressures specific to the pharmaceuticals industry, which could have a major impact on the NHS’ cost base.
The NHS cost base is complex. Some of it can be controlled by trusts or national leaders, and some of it cannot. However, the introduction of the agency cap policy suggests that where national co-ordination is appropriate, it can have significant, positive impact.
The second main type of intervention in the efficiency agenda is productivity – finding ways of doing more work with the same resource. This is sometimes known as technical efficiency. Figure 4 demonstrates that trusts believe reducing unwarranted clinical variation offers the most potential to make better use of existing resources and improve quality in the next five years.

This will typically involve streamlining working practices within a department, specialty or ward, to reduce wasted time, effort and resource. This may lead to improvements in quality, and increased activity for providers – for instance, an increase in the number of operations a trust performs in a week. However, under an activity-based payment model this will cost commissioners money.

This chapter explores two major initiatives to improve productivity with a focus on how clinical staff work: NHS Improvement’s GIRFT programme and ‘lean’ working. It also looks at the difficulty that trusts can face in improving productivity when they are under major operational strain.

**GIRFT**

The appetite among trust leaders for reducing unwarranted clinical variation is encouraging for proponents of GIRFT, which seeks to identify unwarranted variation in cost and patient outcomes, find out why it exists, and eliminate it. GIRFT is divided into workstreams by clinical specialty and directed by leading medics in their field.

The clinically-led nature of GIRFT has proved helpful for trusts that have in the past struggled to engage senior clinicians. Acute finance directors often say their consultants see attempts to introduce more efficient working practices as being an attack on their clinical judgement and autonomy. However, GIRFT is seen as being more successful in opening these conversations because it is clinically-led and uses clinical outcome data, rather than being manager-led and based on financial data.

One acute finance director told us: “GIRFT is the glue between performance, workforce, quality and finance”.

However, trusts that are positive about GIRFT tend to be reluctant to put a figure on the added productivity gains it may bring, and whether that can be generalised across the whole NHS. To attempt to do so may be “simplistic”, in the words of one acute finance director.

Trust leaders also see GIRFT as a long-term approach. One interviewee said. “GIRFT is a cultural change programme. It will take five or ten years to do properly – why would you attach a savings figure to that?”

While GIRFT is attracting attention as a centrally run standardisation programme, trusts have been making inroads in this direction on their own. Though GIRFT is not yet focused on
community care, we heard of one community trust running a programme under its own dedicated transformation team working with clinicians to change their interactions with patients, with the aim of using technology more and standardising processes.

Finally, we should also recognise that the GIRFT approach will not necessarily improve productivity in the same way in each specialism it tackles. GIRFT began in orthopaedic surgery, which typically has high volumes of relatively straightforward procedures which are easy to cost. While variation will exist and should be reduced in specialisms where there is more long-term condition management and less surgery, it will be more difficult to come up with evidence-based targets.

‘Lean’ working

In 2015, five trusts, with the backing of national leaders, brought the Virginia Mason Institute to England to work with them to improve cost effectiveness, safety and quality. Those trusts’ gradual adoption of ‘lean’ working practices is bearing fruit and attracting attention. One trust, which had several years of implementing lean working, listed a series of small improvements they had made internally that had made their organisation more productive.

These included:

- Putting all items needed for steroid injections on a standard trolley. Previously, the items had all been in different places. The intervention cut the time it took to set up an injection from 85 seconds to five seconds.
- Improving processes for managing cases of diarrhoea which have cut diagnosis times from two days to six hours, reduced the time spent by nurses gathering supplies for personal care from 7.5 minutes to 1.5 minutes, cut the time taken to implement a treatment plan after diagnosis from 29 hours to 30 minutes and brought down the time needed for patients being put into an isolated room from over 20 hours to four hours.
- Using a computer on wheels during ward rounds for elderly care reduced “non-value added time” spent with patients from 19 minutes to just under 12 and eliminated defects in reporting.
- Cutting set up times for ultrasound guided injections from 13 minutes to seven minutes per patient.

As with GIRFT, trusts are confident that such improvements have eliminated waste and resulted in better care. One trust told us: “Our focus has always been on quality, safety, experience and outcomes. What this work does is add an additional strand which focuses on the relentless pursuit of reducing waste while putting the patient first and reducing the burden of work on staff.

“We rarely mention reducing costs or increasing productivity or efficiency because this seems to be just the wrong motivator for staff. We simply don’t have the resources to track all of the financial benefits and if we did it would only be proper to track the costs too.
We know this is the right thing to do, we know there are cost benefits which help our financial performance but we aren’t looking to track everything to the nth degree as that process in itself could be considered a waste."

National leaders should bear in mind that it is difficult to objectively say how much any given intervention “saved”, and that giving staff the space and opportunity to change processes to improve patient experience may be more valuable than trying to put a number on it. It will also be essential to share the learning and positive experience of the ‘lean’ approach from these trusts, widely across the sector.

**Operational pressure and capacity**

One of the greatest challenges facing trusts is operational pressure – across acute, mental health, community and ambulance services.

This issue manifests itself quite clearly in the lack of inpatient capacity. A hospital that had 95% of its beds full may appear more productive than if it was running at 85% occupancy. In fact, as occupancy rates tip over 90%, hospitals become less efficient. This is because as inpatient beds become scarce, elective operations are cancelled – increasing the risk of complication, and duplicating the process of booking a patient in for their operation. Furthermore, patients that are admitted may be placed in any bed that is available, rather than on the most suitable ward. The finance director of a large acute trust told us this results in “safari ward rounds”, where staff have to manage patients dispersed across the hospital instead of on one site. This cuts down patient contact time.

In our report *The NHS funding settlement: recovering lost ground* (NHS Providers, 2018) earlier this year, we argued that hospitals with insufficient inpatient capacity will be less able to hit the four-hour A&E standard because they cannot achieve the necessary levels of flow.

Investing in surplus capacity will be needed to ensure hospitals can maintain basic effective processes and will also improve efficiency beyond the acute sector. Ambulance trusts currently lose time queuing at hospitals unable to take patients because they have no beds to admit them to. Increasing hospital capacity could improve use of resource across the whole system.

Increasing capacity will be necessary to improve productivity in the short term. However, trusts are also clear that this is not a permanent fix, and that without system reform, the same problems will arise when the demand increases to fill any extra capacity that can be created. A long-term answer will involve designing a more efficient local system with greater investment prevention and in primary care, community services and social care.
Respondents felt that, over the next five years, the biggest gains could come from making their local systems work more efficiently, with 61% of survey respondents (NHS Providers, 2018) ranking this as the top contributor to future savings.

This chapter details the types of improvements that could be made when trusts act collectively with other partners in their systems. These are sometimes known as ‘allocative’ efficiencies. Often, the financial gains may not be shared equally between those partners – a key barrier to these savings being realised is that it takes time for relationships between providers to develop to the point where there is enough trust between them to share financial risk and reward.

Respondents gave a number of examples where opportunities for future efficiency gains at a system level included:

- integration of out of hospital care providers (GPs, community and social care)
- pooling budgets (to get better value from collective resource)
- reduced bureaucracy and regulatory burdens across the system
- new workforce roles and staffing models
- further improvements in clinical pathways
- consolidating commissioning functions
- effective working with partner organisations
- prevention and community investment
- digitisation and online technology
- further estate rationalisation projects
- avoidance of unnecessary hospital admissions.

Trusts tell us the progress on system efficiencies could be accelerated through the development of system level incentives. To date there has been a focus on trust-specific control totals and trust-specific CIPs. There is currently no equivalently firm mechanism for holding systems to account on their ability to collectively use resource better and some trusts tell us that current incentives do just the opposite.

Holding providers and commissioners to account separately for their own efficiency targets (in the case of CCGs, via the quality, innovation, productivity and prevention (QIPP) programme) can sometimes thwart system working. Two mental health trusts told us that CCG QIPP savings led to annual cuts in their core funding, which reduced their inpatient capacity, in turn escalating costs as patients had to be sent to other, more expensive, providers out of area. Adding this operational and financial pressure made progress on transformation all but impossible. Expanding community mental health services was the only way they felt acute beds could be sustainably cut – but this would present double running costs that trusts were unable to fund on their own. When service redesign efficiencies were achieved, they tended to count towards CCGs’ QIPP requirements, rather than the trust’s CIP – despite the trust being instrumental in realising them.
Growing and shrinking

As an illustration of how trusts’ individual requirements can be counter-productive from a system point of view, several acute trusts have told us they have made efficiency gains through increasing activity levels. One trust leader told us: “It is easier to grow your way out of a problem than cut your way out of it”. The gains can materialise in two main ways: through doing more procedures without increasing overheads, giving a notional decrease in average costs per procedure that can count towards a CIP, or by increasing profitable work which is paid for by activity. One finance director told us it was much easier to engage staff in a discussion about how to do a lot more work for little extra resource than it is to engage staff in a discussion about cuts and removing headcount.

However, this cannot be the answer for the entire sector. At a time of financial constraint and endemic provider side deficits, not all trusts can grow their way out of difficulty at once, nor would it be of benefit to patients as we move towards a new model of delivering care which is carried out in the community and in people’s homes.

Other trusts have told us they have had to shrink their way out of financial problems – one community provider saved £500,000 by withdrawing from contracts which it was losing money on. However, while this is possible in the community sector, where services are tendered routinely and commissioners may have several bidders to choose from, it is not necessarily an option for acute care.

Trusts have also told us that the pressure they are under to meet their control totals has forced them to challenge commissioners more on coding and costing to ensure they receive all the income they can. Again, while this makes sense for individual organisations, it is perverse at a system level, and demonstrates the tension between the national policies of developing sustainability and transformation partnerships and individual control totals for trusts.

Reforming the accountabilities of trusts and commissioners to better reward providers for long-term system efficiencies could produce a shift away from short-term, transactional savings towards transformation. This would also deliver a more efficient system and more joined up, integrated care for patients.
Integrated working

Providers hope that a move towards integrated system working could remove wasted effort, such as arguments with CCGs about coding and costing, and, particularly in the community sector, bidding for tenders. One provider said: “Taking out the purchaser/provider split is the one structural change that could make a difference.”

There is general agreement among providers that closer working between hospital care, community care, social care and GP services offers the greatest opportunity to improve patient experience and ensure people are treated in the most appropriate setting. However, there are many ways to achieve this. Some believe that bringing acute and GP services into a single organisation, under a single integrated management structure, is needed, while others emphasise the importance of improving working relationships and IT.

Genuine system working almost always requires frontline professionals to begin working more closely together than they have in the past – often in mixed teams and across organisational boundaries. For example, an ambulance trust told us that at present, once paramedics have reached a patient who has fallen at home, there are currently few options other than conveying them to hospital. This begins a pathway that too often leads to delayed discharges while the patient waits for home assessments after their inpatient treatment has finished. A better solution may be for community teams to work with ambulance staff during the initial call out to begin treatment at home, and immediately assess and begin to improve their housing.

Integration does not only mean collaboration with neighbouring trusts, it also means working with other agencies. This could involve mental health trusts working with local housing associations to ensure security of accommodation for patients, as losing this can exacerbate mental ill health. Or it could involve working jointly with ambulance and police services to triage patients who are picked up on the street needing care.

Although there is a widespread belief that much greater integration will improve patient care and reduce costs, we were struck by two things. One was that it was a belief, rather than a plan with a clearly evidenced savings target, the other was a simultaneous acknowledgement that demand for care is not going to go down.

One high-performing integrated acute and community trust finance director told us: “I don’t think that, whatever we do in the rest of the system we will slow growth in demand for acute services. I cannot envisage a set of circumstances where there will be a sustainable reduction. We see growth in demand consistent with an ageing population.” People with experience running new care model vanguards are also cautious about the financial savings available. One finance director said: “The teams are running better. There are fewer hospital admissions but cost savings have not been realised.”

The assumption that better care will necessarily cost less because it is the right thing to do will be tested in the coming years as more systems become more collaborative. A 2018
long-read from The King’s Fund stated that “integrated care and population health should not be expected to save money but have the potential to enable resources to be used more effectively.” While there is a potential improvement in value, that will not necessarily lead to cash releasing savings (The King’s Fund, 2018).

Pathway redesign enabled by technology

Our survey found that trust leaders believe technology and digital innovation should be one of the top three priorities for the forthcoming NHS long-term plan, alongside workforce and improved out of hospital care.

There are trusts that still barely use an electronic patient record and are overly reliant on paper notes. However, this is not the norm and these organisations are working to tackle the issue – although the impact shifting to digital often has on ordinary trust operations should not be overestimated.

More commonly, interviewees highlighted the improvements technology could bring through automating back office functions and reducing stationery costs. Another repeated theme in both the survey and our interviews was how investment in the right technology could support much better system working across organisations.

One acute trust in the south has established a virtual fracture clinic, enabling new patients to be assessed ahead of their appointments – this has led to the cancellation of more than three quarters of new patient appointments in favour of GP-led self management. This has freed up consultant time to use for urgent referrals from elsewhere in the trust.

We also heard of plans to enhance links between paramedics and consultants, to improve treatment when the patient is first reached and in the ambulance, and to ensure only patients who need acute care are brought to hospital.

Other trusts were keen to improve the interoperability of their patient record systems, so a single record to be amended by GPs or consultants in any of the trusts a patient visits during their treatment. We also heard from one trust that was working closely with a neighbour on digital collaboration through the Global Digital Exemplars programme, this had enabled them to share imaging more effectively, which had reduced duplication when patients pass from one organisation to another.
Trust mergers, groups and scale

While it is easy to assume that there are too many providers and that merging them into larger organisations will take out duplication and so reduce costs, it should be noted that the vast majority of NHS spend goes on frontline services rather than the back office – even the trusts with the highest back office costs only spend 11% of turnover on administrative functions. (Department of Health, 2016). Removing the costs of running two boards and two sets of back office support functions may be desirable, but a merger will not on its own affect the cost of providing services to patients, which accounts for the overwhelming majority of the spend (NHS Improvement, 2017).

Similarly, while a move to a group model may improve consistency and deliver some economies of scale, the primary focus of much of this work is on eliminated unwarranted variation in quality and making better use of the collective workforce.

The finance director of a relatively small, but successful, integrated acute and community provider told us that his organisation was viable, but because it focused on efficiency at the level of individual units, such as wards or operating theatres, rather than focusing on questions of organisational form. He argued that whether a hospital is efficient or not depends on how efficient its individual wards were, and it could not become more efficient simply by increasing the number of wards or theatres it runs, for example by taking over another trust. Their argument was that energy is better focused on making each unit more efficient than on organisational mergers. His trust, he said, had efficient wards partly because they incorporated very few individual single rooms – although this did have an impact on patient experience and, sometimes, infection control.

Outpatients

It has become increasingly clear during 2018 that there is appetite among national policy makers for reform of traditional consultant-led outpatient services. NHS England chief executive Simon Stevens described it as an “obsolescent” model for managing long-term conditions, while NHS Improvement chief executive Ian Dalton signalled an intention to begin a “major transformation” in outpatient services.

Next year could see changes to the payment system for outpatient services to incentivise a move away from face to face meetings between patients and consultants.

There is appetite among trusts to adopt more cost efficient ways of providing outpatient services. One survey respondent told us that they would transform outpatients if they had enough project management capacity. Another said they wanted to see redesigning outpatients as a priority in the forthcoming five and ten-year plans for the NHS.

We also spoke to a finance director in a specialist provider who was working with commissioners to change the outpatient model, even though their organisation had done
well financially from episodic payments for the service in its current form. The finance director acknowledged that the management of long-term conditions should involve fewer trips to hospital. However, they also pointed out that the consultant-led model, while expensive, was popular with patients who placed great value on the continuity of care and access to expertise it offered.

The Nuffield Trust (HSJ, 2018) has recently described three types of mistake in previous attempts to reform outpatient services – attempts to reduce referral rates did not tackle fixed costs on the provider side, moving the services to less expensive facilities simply reallocates the same overheads to other services and policy has relied too much on blanket assumptions, rather than seeing outpatients as many services. Most relevantly for the efficiency debate, the Nuffield Trust has cautioned against assuming that improving services will necessarily release money to be redeployed. While there is “undoubtedly scope for major change in outpatients, which will make the experience better for patients and staff”, it is “much more doubtful that this will produce major cash savings”. They describe a set of local pathway improvements example specialist doctors running teaching and discussion sessions with GPs, and setting up virtual clinics.

If outpatient reform is expected to save money, national leaders need to be realistic about how long this will take to deliver and be careful about assuming a single approach is uniformly applicable.

Pathology

It is 10 years since Lord Carter completed his first independent review on pathology, which recommended scale provision by standalone providers such as joint ventures or outsource companies. The review estimated potential savings of between £250m and £500m a year. It stated: “We have studied the evidence before us and listened carefully to what we have been told. This is not another ‘false dawn’: we believe our proposals will enable pathology services to meet the needs and demands of health care provision for the future by strengthening the delivery of high quality, efficient, effective and responsive pathology services. Our confidence stems from the clear consensus we have found about the barriers to progress and the steps needed to overcome them.”

However, reform was not rapid. As with the Five year forward view, some pilots were launched but detailed guidance and a model implementation plan, describing the route from the existing service configuration to an optimal one, was lacking. Procurements for pathology networks were slow to come to market, partly due to the complexity of the service redesign and tender processes involved.

One finance director we interviewed told us his trust scaled up its pathology facilities so they could bid for new work, only to find fewer contracts than expected came to market, leaving it with a lab that is not running at full capacity even today. This, along with other examples of joint ventures set up that did not deliver the expected benefits, demonstrates the distance
between savings estimated and savings delivered, and shows that aspiration does not always guarantee success.

Between Lord Carter’s 2008 review of pathology and his wider 2016 review of operational productivity in hospitals, there was some consolidation of pathology services, and some adoption of outsource provision. In the 2016 review Lord Carter again found that some of the most efficient and high quality pathology services were provided across several trusts at scale, and estimated £200m could be saved if best practice was adopted across the service. In 2017, NHS Improvement published recommendations for pathology networks covering the entire country, along with savings targets for each network. Trusts were required to begin setting up the networks or detail how they would bring about equivalent savings.

Why has it taken so long to establish widespread scale working in pathology? There was an initial focus on organisational form and an emphasis on commercial tender processes which proved to be distracting and time-consuming. The complexity of this approach – and the risks it introduced which crystallised in the failure of The Pathology Partnership – confirmed suspicions that the initiative was not worth the effort or the risk.

The collaboration necessary for partnership working on pathology is still not well established in many areas, after many years of national policy encouraging competition between trusts. Interviewees repeatedly told us that, as with new models of care, collaboration on pathology can only work where relationships at all organisational levels are mature. Trust takes time to develop – especially where providers have separate clinical cultures. Providers are responsibly bearing in mind the impact of changes to pathology on other services. They will not rush into change without considering the overall effect on local services.

The lesson for national and local leaders is that there must be an acknowledgement that change takes time and that potential benefits are not the same as cash in the bank: they are always balanced by risks. Evidence that change is necessary does not negate the need for full engagement with clinical staff. A detailed implementation plan setting out how change will be managed and where savings can be made is essential. Also, experience tells us that savings tend not to materialise as fast as projections initially suggest.
A consistent theme across our survey and interviews was that, although most leaders know where efficiency gains can be made, they need support beyond their own organisation to be able to realise them. This chapter addresses the gap between what trusts need to be able to make more progress on efficiency and their current capacity.

Figure 5 demonstrates that only a fifth of respondents agreed that they had received enough support from central bodies (20%), and had the resources they needed to make progress on efficiency (19%). The majority disagreed with both these statements.

Yet some of NHS Improvement’s most high profile activity driving operational productivity has been well received. Respondents told us that the most useful of NHS Improvement’s contributions over the past three years had been data sharing and trust benchmarking exercises (such as the Model Hospital), with 77% suggesting this had been useful. More than two thirds (69%) of respondents had also found the coordination of national cost saving programmes (such as agency caps) useful in supporting efficiency savings at their trust.
From the survey and our interviews, the overarching message has emerged that sharing data on efficiency is helpful but trusts need more support to deliver.

**Benchmarking**

Under the Model Hospital, NHS Improvement shows trusts how they measure against their peers on a range of cost, productivity and efficiency measures. These include spend per unit of activity within service lines, spending on routine consumables, staffing costs, medicines use, operating theatre productivity, estates costs, and so on.

One the whole, the tool largely applies to the acute sector. However, we consistently heard from community and mental health trusts that they wanted access to better benchmarking information. These trusts feel they would benefit from more comparison data to identify which areas of their organisation they should be targeting for savings. There is one note of caution from the mental health trusts, their overheads are lower than the acute sector and benchmarking on patient outcomes rather than inputs would provide much more value.

Generally, acute trusts were positive about the Model Hospital that has supported internal scrutiny and helped boards identify where gains could be made. However, trusts are also realistic that benchmarking is mainly useful in prompting questions. “It gives you a clue for where you might look – but it doesn’t provide an answer” was one typical comment. “If one out of ten findings turn out to be worth acting on, then it’s a good thing”, was another.
However, some variation can be explained by the differences in how trusts cost bases are accounted for. The Model Hospital does not yet account for the fact that overhead costs are apportioned to different departments in different ways from one trust to another. The tool will not be wholly reliable unless this data is recorded comparably by all providers.

Likewise, in the ambulance sector, trusts record activity data differently. For example, some trusts will include hoax calls or repeat calls in some activity indicators, while others will not. This makes it difficult to compare trusts on the proportion of calls that lead to patients being treated. The implementation of the new ambulance quality indicators have gone some way in addressing this, but there is a period of bedding before data is truly comparable.

Another important caveat is that, while unnecessary high costs or low productivity should be identified and tackled, there will always be some variation between providers – any savings projections associated with reducing variation must be based on a realistic projection for the difference between the worst and the best, and not assume that everywhere will be as good as the current top performers.

Providers with experience of the Model Hospital were keen to point out that the data it uses is not perfect, and sometimes throws up misleading results. For example, one acute trust explained its workforce costs skewed unfavourably because it employs some staff on behalf of other organisations – and then had to defend itself when questioned by NHS Improvement.

It should also be noted that not all variation is “unwarranted”. An example that two trusts gave was the price of food. Both spent more than the average on food for their patients – both had had this highlighted to them via the Model Hospital. Both acknowledged the higher spend, and said they would continue to spend more on food because they believed it complimented the medical and nursing elements of patient care and supported patients’ recovery.

The support offer

While benchmarking is seen as broadly a good thing, trusts would like more support to act on the data. Specifically, trusts felt that NHS Improvement was putting a lot of effort and resource into analysing questionable data, however, they would also welcome more support to make improvements as trusts do not always have the capacity to make large scale change happen quickly. Within the cap set out in Lord Carter’s 2016 report (trusts should limit administrative spending to 7% of turnover), trusts may be able to sustain normal operation but are unlikely to be able to undertake large change programmes.

“Trusts know what they have to do but they do not have the capacity or resources,” one finance director told us.
Too often, we heard, NHS Improvement feels like an analytics function armed with regulatory powers to ensure compliance. We have consistently found that, in future, trusts want NHS Improvement to focus on helping trusts to improve. (NHS Providers, 2018) This would require a change in behaviour at different levels within NHS Improvement, however the organisation’s collaboration with NHS England presents a key opportunity as it repositions its offer around improvement and support, and develops new regional teams.

This national support should initially come in the form of including a framework within the long-term plan setting out where trusts could realise efficiencies and working with trusts to agree a stretching, yet deliverable, efficiency requirement. More ongoing and direct support could take the form of practical support for people trying to bring about improvements. Another suggestion was to identify top performers and introduce a structured programme for them to spread best practice, along the lines of the Global Digital Exemplars, where organisations that have independently made progress are matched with trusts in need of development.

Trusts would also like to spend less time producing data returns for central bodies. The more these processes can be automated, the more time trust staff can devote to making change happen within their organisations.

**Barriers to improving efficiency**

**Figure 7**

*What are the biggest barriers to making efficiency gains?*

(n = 157)
Figure 7 shows that more than two thirds of respondents cited short-termism encouraged by the current financial regime (69%) and operational pressures (66%) as barriers to making efficiency gains. Almost half (46%) also felt the lack of transformation funding was a barrier and 39% pointed to a lack of capital investment.

The control totals regime was introduced in 2016-17 to establish central grip over financial performance. NHS Improvement has used the provider sustainability fund (formerly the sustainability and transformation fund) to enforce compliance with stretching year end targets, with the aim of bringing the trust sector back into a sustainable position.

While this has not happened, the regime could be credited with a slowdown in the deterioration of the overall sector position after the rapid decline seen in 2015-16.

However, trusts tell us it is failing to incentivise system efficiencies, which they say present the greatest opportunity, and instead encourages potentially unhelpful, organisation-centric behaviour such as growing profitable activity, stopping unprofitable services and arguing with commissioners over contracts. It also does nothing to encourage trusts to invest in schemes that may pay off over many years, but not immediately.

This, combined with a lack of dedicated money to fund service change, and operational pressure that diverts all management, clinical and physical capacity into hitting performance standards, has created a situation where many trusts will find it impossible to make progress on transformation.
In order to maximise the efficiency of the provider sector, our survey suggests the following solutions and recommendations should be explored at national, regional and local levels.

- The forthcoming long-term plan offers an opportunity to identify realistic and deliverable improvements that can be made and chart a credible path to realising them. It should not assume that placing providers under ever greater financial pressure will automatically drive efficiency improvements at a rate faster than has been achieved in the past.

- Any new efficiency requirement should be based on realistic assumptions and collaboratively agreed with providers. This ask should take the starting position of the trust into account and be accompanied by co-produced realistic delivery plans at organisational, system and national levels.

- NHS staff at all levels should engage in the efficiency agenda, from joint ownership of efficiency across all board roles, to empowering frontline staff to make improvements to their own working practices.

- Distinctions should be drawn between cost reduction schemes, productivity improvements and system efficiencies, and the centre should share nationally compiled learning about where trusts could make savings under each of these different areas.

- There must be acceptance at a national level that some variation between organisations, and between local systems, is natural, explicable and justifiable. The national bodies and long-term plan should only describe variation as “unwarranted” where there is clear evidence that it can and should be removed, and there is a clear means of doing so.

- Accountability must shift away from single-year efficiency programmes and control totals to provide incentives for both commissioners and providers to prioritise more sustainable efficiency schemes with the potential to transform services and release greater productivity gains over a longer time period.

- Trusts and local partners should embrace programmes such as Model Hospital, GIRFT and lean working practices that contribute to efficient ways of working. However, they should recognise that these programmes are driven by a focus on quality of care not cost base and support more efficient use of existing time and resource.

- Community and mental health trusts should have access to relevant benchmarking data of similar quality to the acute sector as soon as possible.

- Trusts and NHS Improvement should not rely on organisational reconfigurations like mergers or takeovers to make unsustainable services sustainable, and while new care models will result in more efficient ways of working and better quality of care for patients, the centre should be cautious about the extent to which they produce cash releasing savings.

- To reduce the reporting burden on trusts, NHS Improvement should automate as many information collection processes as possible and put more effort and resource into supporting trusts to make improvements, even if this means reducing the size of its analytics function.
NHS Improvement should consider identifying top performers on efficiency and introduce a structured programme for them to spread best practice through peer-to-peer support, along the lines of the Global Digital Exemplars.

As poor facilities, particularly IT, are a source of frustration for trusts and a block on increased efficiency, it is vital that capital is made available to the frontline and that the DHSC clarifies the rules around access to capital as soon as possible.
Opportunities for further efficiencies exist. Trusts are clear that GIRFT, ‘lean’ working, better commercial practices, integration, and smarter use of technology can all play a part in ensuring the taxpayer pound is spent as wisely as possible. Better system working and addressing workforce challenges have also been identified by trust leaders as ways to achieve more significant efficiencies over the longer term.

However, our survey results demonstrate a low level of confidence that trusts can realise more savings, despite showing an overwhelming acknowledgement that there is still inefficiency in the wider NHS system.

Overall, trusts recognise their responsibility to plan for and then deliver a stretching efficiency ask within their trusts and local systems. Trust leaders believe that a stretching efficiency requirement is more likely to be delivered if:

- national system leaders, trusts and local systems work in partnership to agree a realistic efficiency requirement and timeframes for delivery rather than have these centrally imposed, top down, from the national system level
- national system leaders do more to help trusts and local systems identify and share learning from across the country about where efficiencies can be delivered
- national system leaders are realistic about the extra support that trusts and local systems will need to realise the transformational savings and local system focused savings that must now form the focus of efficiency and productivity activity
- any efficiency delivery assumptions explicitly match the capacity and capability available on the frontline given the ambitious list of other priorities trusts are expected to deliver.

NHS Providers also recognises the role we can play in supporting trusts and we remain committed to working with the national bodies to help shape constructive offers of support from the centre, and to promote what works well as we have done with the GIRFT programme. We fully recognise our role in helping trusts share and showcase good practice.

The additional investment in the NHS and the forthcoming ten-year plan represent a rare opportunity to reset the operational pressure and endemic financial problems that are currently blighting the NHS. Increased efficiency is possible, but it should not be used as a financial balancing item reconciling a hugely ambitious improvement trajectory with a modest funding rise. Trusts must have access to the right mix of support and be given time and headroom to pursue long-term system efficiencies. If lessons are learned from the past, and the new settlement comes with a realistic efficiency target and a clear and robust delivery plan, it can succeed.
ANNEX 1
Methodology

The findings in the report are drawn from a survey of NHS trusts and structured in-depth interviews.

The mixed methods approach included an online survey conducted by NHS Providers that was open for three weeks during August/September 2018. The survey was open to various executive director posts including chief executives, finance directors, HR directors, medical directors, nursing directors and chief operating officers.

There were 157 responses to the survey, from 122 NHS trusts. This constitutes 54% of the provider sector, and responses covered all types of NHS trust (acute, ambulance, community, mental health and learning disability, specialist and combined).

To explore the survey findings in more detail, we carried out interviews with 10 executive directors from a range of different types of trusts. In addition, we also interviewed eight professionals and experts working in areas related to NHS finance policy, drawn from various health bodies and think tanks.

To allow individuals to speak openly about the issue of efficiency in the NHS, the interviews were anonymous. Where we have included direct quotes we have given the job title and type of trust, as long as it did not identify the individual.
References


Suggested citation

NHS Providers (October 2018),
Making the most of the money: efficiency and the long-term plan.

Interactive version

This report is also available in a digital format via:
NHS Providers is the membership organisation and trade association for the NHS hospital, mental health, community and ambulance services that treat patients and service users in the NHS. We help those NHS foundation trusts and trusts to deliver high-quality, patient-focused care by enabling them to learn from each other, acting as their public voice and helping shape the system in which they operate.

NHS Providers has all trusts in membership, collectively accounting for £84bn of annual expenditure and employing more than one million staff.