HFMA introductory guide to NHS finance

Chapter 16: How providers of NHS services are paid

Chapter 16. How providers of NHS services are paid

Overview

This chapter sets out how funding is transferred from commissioners to providers of healthcare services. It considers the evolution of the payment mechanism, the legal framework surrounding it and the way that payments can be used to influence behaviour.

Primary care funding flows are covered in chapter 6.

16.1 What is a payment mechanism?

A payment mechanism is a system of financial flows that is used for commissioners to reimburse providers of NHS healthcare in England. It does not affect the financial allocations that commissioners receive from NHS England. The payment mechanism can be used to incentivise provider and healthcare system behaviour, depending upon national policy.

The payment mechanism sets out the elements that contribute to the payment amount, such as activity undertaken or quality of care, as well as any penalties for under achievement. The NHS standard contract described in chapter 15 plays an important role in supporting the application of the payment mechanism, through setting out the conditions that apply in a particular context.

The payment mechanism currently in use is called the NHS payment scheme and was introduced in April 2023. We will look at it in more detail later in the chapter.

16.2 Key components

Unit of activity

For the payment mechanism to work, it is important to decide what is being paid for – the unit of healthcare activity (sometimes referred to as the currency). Different parts of the NHS use different units of activity to reflect the way that patients are treated and care delivered.

The healthcare resource group (HRG) is the unit of activity used for admitted patient care (covering a spell of care from admission to discharge), procedures undertaken in outpatients and accident and emergency attendances. HRGs group services that are clinically similar and require similar resources for treatment and care. They recognise the different costs associated with elective and non-elective procedures, treating patients of different ages, those with multiple co-morbidities (related illnesses) or where there are additional complications. Overall, there are around 4,000 different HRGs.

The unit of activity for outpatient attendances is the attendance itself, split between first and follow-up attendances, the broad medical area (defined by a treatment function code) and whether the attendance relates to a single professional or a multi-professional team.

For adult mental health and learning disability services a unit of activity known as the 'care cluster' was introduced in 2012. It describes the common needs of a group of patients/ service users over a period. Each patient is assessed based on their symptoms and needs and assigned to one of 21

clusters. However, clustering has not been widely adopted and NHS England is working to develop a new model for mental health units of activity.

NHS community health services do not have defined units of activity. Development work is ongoing in five areas: children and young people with disabilities; single episodes of care; long-term conditions; frailty; last year of life.[305]

For ambulance services, four units of activity are used:

- urgent and emergency care calls answered
- hear and treat/ refer
- see and treat/ refer
- see, treat and convey.

Units of activity do not necessarily have a payment directly attached but they provide an agreed way to measure and count patient activity. They therefore underpin all payment approaches where it is necessary to understand activity in order to fund a service at the correct level.

Types of payment

Payments from commissioners to providers fall into three broad categories:

- activity-based payments – a payment is made for each individual unit of activity delivered
- block payments – a lump sum is paid for all activity delivered
- blended payments – a combination of the previous two approaches with fixed and variable elements.

16.3 Evolution of the payment mechanism 2000-2022

In this section we will examine the history of the payment mechanism to understand how it has evolved in line with national policy.

Payment by results

The first version of the payment mechanism - payment by results (PbR), was a key element of reform set out in the 2000 NHS Plan[306]. The government wanted to be sure that the planned increases in NHS resources would be used to deliver a higher volume and quality of clinical services and enable patients to choose where they were treated.

PbR fundamentally changed the way funds moved between commissioners and providers of acute secondary care services. Unlike the previous system of block contracts, these providers were paid for each patient seen or treated at a national rate set by the Department of Health and Social Care (DHSC). Therefore, if more patients were treated than originally planned, the provider would receive more income but the commissioner would face a financial pressure. The opposite was true if fewer patients were treated.

[305] NHS England, *Annex B: Guidance on currencies,* March 2023
[306] NHS England, *The NHS plan*, July 2000

The impact of payment by results

PbR allowed money to follow the patient for the first time. It helped to encourage patient choice and reduce waiting lists in the areas it covered. But it increased risk and uncertainty for NHS organisations and was only implemented in the acute sector.

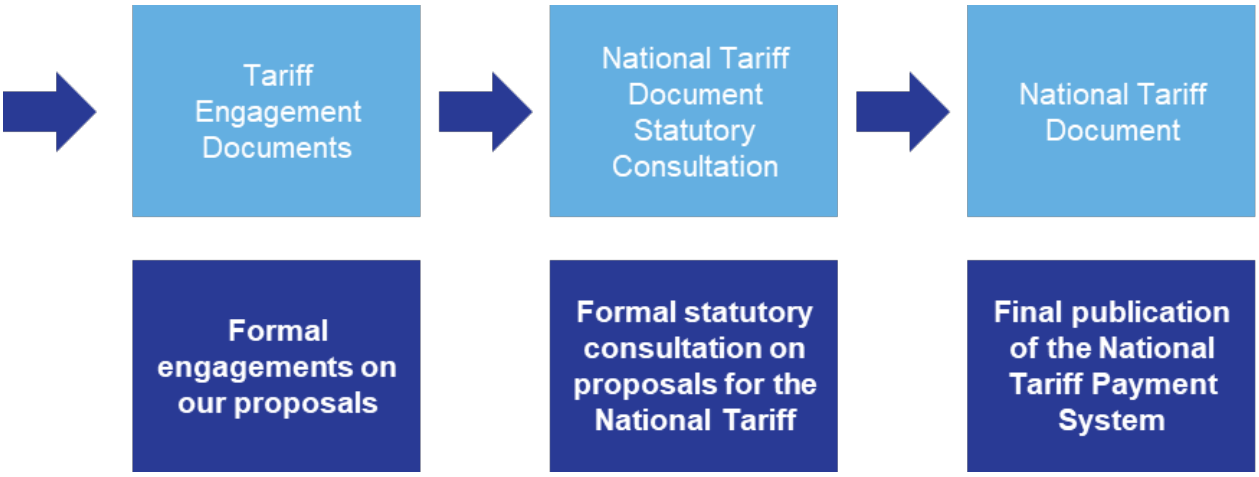
National tariff

The Health and Social Care Act 2012[307] transferred responsibility for the NHS payment mechanism to NHS England and Monitor (the regulator of foundation trusts at the time).

Introduced in 2014 and in use until 2023, the national tariff[308] was a legal framework that covered prices for treatments and procedures, the methodology for setting them and the underpinning rules. Under the national tariff, payments made to providers of secondary acute care for NHS patients were linked to the services actually provided using national prices and HRGs as the units of activity.

The 2012 Act set out the process for producing an annual national tariff document, including a period of statutory consultation, as set out below.

Development of the national tariff



The national tariff primarily covered elective (including outpatients) and non-elective acute care (including urgent and emergency care), and ambulance services. It did not cover all healthcare services but for those services included within it, the rules had to be followed. However, there was provision within the tariff for local variation to prices and units of activity, through an agreed process. This flexibility was included to enable service transformation. Attempts were made to extend the use of activity-based payments to the mental health sector using care clusters as the unit of activity. But in practice most providers of mental health services used the permitted flexibility to remain on block contracts.

Influencing behaviour

The national tariff was used to incentivise behaviour of provider organisations and encourage improvements in quality. The commissioning for quality and innovation (CQUIN) scheme linked a small proportion of contract values to the delivery of specific quality initiatives. Best practice tariffs rewarded providers with higher payments if they adopted guidelines recommended by the National Institute for Health and Care Excellence (NICE).

---

[307] UK Government, Health and Social Care Act 2012
[308] NHS England, *National tariff,* 2023

The national tariff also set out the national efficiency requirement that applied to all secondary healthcare services. Billing, payment and activity reporting processes were set out in the national tariff document, along with payment for NHS funded services provided by local authorities. All licensed healthcare providers were required to comply with the national tariff and provide information to support its development.

The impact of the national tariff

Like PbR before it, the national tariff helped to improve efficiency and quality by encouraging NHS providers of acute services to compete. As national prices were fixed, competition could only be based on quality.

However, as an activity-based payment system the national tariff created a significant number of transactions, reconciliations, queries and disputes between NHS organisations. Continued pressure on the NHS from a growing and ageing population created a change in the emphasis of national policy away from competition towards collaboration and system working, and by 2020 the national tariff had become a barrier to the required reforms.

Covid-19 arrangements

The covid-19 pandemic outbreak in 2020 completely disrupted the normal flow of activity within the NHS, forcing the cancellation of elective care, the repurposing of facilities and the redeployment of staff. The enforced drop in planned activity would have created significant deficits if national tariff payments had been maintained.

Consequently, all NHS organisations were immediately moved to a simplified block payment system allowing the NHS to focus on the needs of the pandemic. It freed up finance staff to support clinical colleagues and allowed organisations to work together with fewer barriers to cooperation.

While it was accepted that the simplified financial regime was not sustainable and may be causing wastage, many welcomed the reduction in transactional processing and the removal of financial barriers between organisations. These factors influenced the next change to the payment mechanism.

16.4 The current payment mechanism (2023 onwards)

NHS payment scheme

To support the move to system working, the Health and Care Act 2022[309] set out changes to the legislation surrounding the payment mechanism. In April 2023 the national tariff was replaced by the NHS payment scheme (NHSPS)[310]. The legal requirements for consultation and publication of documents remain in place.

The NHSPS[311] sets out four payment approaches:

- contracts with NHS providers over £0.5m per annum use the aligned payment and incentive (API) approach. This is a form of blended payment – the fixed element funds an agreed level of activity; for providers of acute services the variable element covers all elective activity on an activity-based payment basis.
- contracts with NHS providers under £0.5m per annum use the low volume activity (LVA) approach. This is a block payment with a nationally set value.

---

[309] UK Government, Health and Care Act 2022
[310] NHS England, *NHS Payment Scheme*, 2023
[311] NHS England, *2023/25 NHS Payment Scheme*, updated December 2023

- contracts with non-NHS providers where there is a nationally set price use an activity-based payment approach.
- a local payment approach can be used if none of the above approaches apply.

The first NHSPS covers two years – 2023/24 and 2024/25, with a process to update prices in 2024/25. Although they have a reduced role compared to the national tariff, prices are still published within the NHSPS. There are two categories:

- unit prices are used for activity-based payments and relevant variable elements of API
- guide prices to benchmark information and support local payment approaches.

Published prices are based on previous years' prices with adjustments made for cost uplifts – for example, pay growth, drug costs and other inflation, and efficiency, a downward adjustment to encourage providers to improve their use of resources. Historically, prices were set based on the average reported costs from all NHS providers, but this data hasn't been used to update prices since 2018/19 to avoid distortions caused by the Covid-19 pandemic.

The NHSPS sets out five principles that must be followed when developing a local payment approach or proposing any departure from the API approach:

- the payment approach must be in the best interests of patients
- the approach must promote transparency and good data quality to improve accountability and encourage the sharing of best practice
- the provider and commissioner(s) must engage constructively with each other when trying to agree payment approaches
- the provider and commissioner(s) should consider how the payment approach could contribute to reducing health inequalities
- the provider and commissioner(s) should consider how the payment approach contributes to delivering operational planning guidance objectives.

The NHSPS came into effect on 1 April 2024 and applies through to 2024/25, however some amendments for 2024/25 have been proposed and consulted on[312].

Many of the changes are in support of the delegation of specialised services to ICBs. NHSE consider that the impact of the proposed changes would be relatively minor at organisation level.

The NHSPS remains a legislative requirement for all NHS secondary care services. It was built on the foundations of the previous payment mechanisms and the learning from the impact of the covid-19 pandemic. It effectively moves nearly all payments between NHS commissioners and providers onto an API basis – a blended payment that seeks to find the right compromise between certainty, flexibility, incentivisation and administrative burden, while following the identified principles and encouraging collaboration.

[312] NHS England, *Consultation: Proposed amendments to the 2023/25 NHS Payment Scheme,* December 2023

Key learning points

- A payment mechanism is the way that funding is transferred from a commissioner to a provider.
- The NHS has used several different mechanisms, each of which has its own advantages and disadvantages.
- The Health and Social Care Act 2012 required the NHS to publish a national tariff that sets out prices for most acute sector activity.
- The Health and Care Act 2022 replaced the national tariff with the NHS payment scheme (NHSPS) to support the move to collaborative system working.

Additional HFMA resources

The HFMA maintains a directory of resources which provides links to other HFMA outputs such as briefings and webinars across a range of subjects. It also highlights online learning courses that are available. The directory of resources can be found here.