

Contracting for health outcomes

from concept through theory to implementation

May 2024

Prepared by:

Luigi Siciliani, Department of Economics and Related Studies, University of York

Steven Wyatt, The Strategy Unit, NHS Midlands & Lancashire CSU

Matthew Revell, The Royal Orthopaedic Hospital NHS Foundation Trust

Peter Spilsbury, The Strategy Unit, NHS Midlands & Lancashire CSU

Document control

Document Title	Contracting for health outcomes
Job No	935
Prepared by	Steven Wyatt
Checked by	Fraser Batty
Date	May 2024

Contents

Foreword	3
1. Introduction	5
2. Elective knee replacement surgery	7
3. A short history of NHS contracting models for hospitals	9
4. Contracting as a means to improve outcomes	11
5. Fairness in outcomes-based contracts using casemix-adjustment	13
6. Contract theory	15
7. Imagining an outcomes-based contract for knee replacements surgery	16
8. Calibrating an outcomes-based contract: the Commissioner’s role.....	23
9. How should providers respond to an outcomes-based contract?.....	34
10. Looking forward	38
Appendix – Technical description of the scheme	41

Foreword

Payment by Results, the mechanism that determines hospital Trust income, is an important and enduring feature of health policy in England. But its critics claim it's a misnomer: that it rewards hospitals according to the number of patients that they treat, rather than the results or outcomes they deliver. Policy-makers have attempted to address this challenge by adjusting payments to take account of some aspects of service quality, but these adjustments are modest and marginal.

In recent years the policy has been renamed as the National Tariff, but the question remains: why isn't hospital income tied more closely to patient outcomes? There's no shortage of support for the idea, from domains as diverse as politics, academia, and patient groups. The fact is that outcomes-based contracts are very difficult to design and implement. So much so that the rhetoric has only rarely turned into reality.

But there's good news. Analysts and researchers have been steadily chipping away at the technical barriers to implementation. Outcome measurement, the raw material of outcomes-based contracts, has matured through the roll-out of national clinical audits and the PROMs programme. Casemix-adjustment methods are increasingly used to ensure that Trusts are compared fairly and incentives to cherry pick 'easy' patients are moderated. And more recently, academics from the University of York have shown how to calibrate the incentives in outcomes-based contracts. This is the last significant piece of the technical jigsaw, and a robust, real-world implementation of an outcomes-based contract now seems feasible.

As doubts about feasibility recede, another, more fundamental question takes centre stage. Should we pay providers based on outcomes? This is not a technical question, but a question of principles, priorities, and fit with the policy and fiscal context. Two challenges dominate health policy today: (1) how to enable healthcare services to cooperate and integrate to improve service quality and patient experience, and (2) how to increase service capacity, efficiency and equity as the population and the health system recover from the effects of the COVID-19 pandemic. Outcomes-based contracts could certainly feature in the mix of policy responses to these challenges, freeing up services to innovate, and ensuring that providers maintain focus on outcomes rather than on throughput. But developing and implementing outcomes-based contracts will require considerable effort from NHS England, providers and commissioners, and there are many other issues competing for these agencies' attention.

If we set aside the option of a wholesale adoption of outcomes-based contracts, which would be premature and unhelpfully disruptive at this point in time, then there are perhaps two rational stances that policy-makers could adopt. The first would be to explicitly postpone decisions on the development of outcome-based contracts for say, three years, until the context is more settled. This would quell the call for outcomes-

based contracts, and give providers and commissioners some certainty, but without ruling out the possibility in the future. The second would be to trial an outcomes-based contract for a selected service, evaluating it fully and robustly to inform future decisions about wider roll-out.

Elective knee-replacement surgery is a prime candidate for trialling outcomes-based contracting. In this paper, we show how an outcomes-based contract for this service might be designed.

Acknowledgments

This study was funded in part by the National Institute for Health and Care Research (NIHR) Policy Research Programme, conducted through the NIHR Policy Research Unit in Economics of Health Systems and Interface with Social Care, PR-PRU-1217-20301. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

1. Introduction

In 2010, the Commonwealth Fund published a comparative analysis of health system performance across seven countries: Australia, Canada, Germany, Netherlands, New Zealand, the United Kingdom and the United States. Whilst the UK health system performed well in terms of the quality, accessibility, efficiency and equity of care, its health outcomes were comparatively poor. The analysis raised an important question. Why, if the NHS provides high quality, efficient and increasingly accessible care, does this not consistently translate into good population health outcomes?

This publication was one of several factors that generated a renewed interest in healthcare outcome measurement and improvement in the UK. The clinical community established a number of national clinical audit programmes. NHS Digital launched a bold programme to capture patient reported outcome measures for patients receiving specific surgical interventions. And the Government set out how it would use national policy to this end.

One potential policy intervention that was frequently proposed, was the use of contractual mechanisms to incentivise healthcare providers to deliver better health outcomes. At the time, health service commentators had grown weary of the improperly named 'Payment by Results' policy which rewarded acute hospital providers based on the volume and types of activity they carried out. They claimed that a true 'Payment by Results' policy would reward providers for restoring function, slowing disease progression, and preventing premature deaths, rather than for carrying out a surgical intervention of unknown quality. The view gained traction, and many believed it would only be a matter of time before such contracts were commonplace. But this transformation in contracting models failed to materialise. Whilst contracts increasingly featured quality bonuses, these almost exclusively rewarded providers based on input or process measures. Indeed, we are aware of only one NHS scheme that genuinely reimbursed providers based on outcomes, and this arrangement ceased in 2011, falling victim to the NHS reorganisation.

1

The reality is that outcomes-based contracts are far more difficult to design and implement than their proponents had imagined. Robust data on outcomes is difficult to collect and marshal, and it's taken time for academics to understand how to apply their theories and methods to the question of designing outcomes-based contracts in healthcare. But 13 years on from the Commonwealth Fund report, the NHS might be ready to put the theory of outcomes-based contracts to the test. We set out in this report

¹ McLeod H, Blisset D, Wyatt S, Mohammed MA, Effect of Pay-For-Outcomes and Encouraging New Providers on National Health Service Smoking Cessation Services in England: A Cluster Controlled Study, PLoS One 10(4): e0123349.

how such a contract might function for the provision of elective knee replacement surgery, summarising the key concepts and illustrating the fiscal consequences of such an arrangement for commissioners and providers. We go on to explore the residual barriers to implementation and how these might be overcome.

This paper builds heavily on the recent work of Luigi Siciliani, Professor of Economics, and his colleagues at The University of York.²

² Siciliani L, Gaughan J, Gutacker N, Gravelle H, Chalkey M, Paying for Health Gains, CHE Research paper 183. July 2021
https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP183_Paying_Health_Gains.pdf

2. Elective knee replacement surgery

Elective knee replacement is a common form of planned surgery where a damaged knee joint is replaced with an artificial one. A primary knee replacement describes surgery carried out to replace a natural knee joint. An artificial knee joint can last many years, but in some circumstances may need to be replaced. This is known as revision surgery.

Osteoarthritis and inflammatory arthritis of the knee are the most common conditions that lead to knee replacement surgery. The conditions cause joint pain, stiffness and can limit mobility. The risk of developing these conditions increases with age and it is more common in women than in men. Congenital and childhood conditions, family history, trauma, infection, and body weight can increase the risk. Osteoarthritis of the knee is a common condition. It is estimated that 18% of the UK population aged over 45 years have sought treatment for the condition.³ Rheumatoid arthritis, a form of inflammatory arthritis is less common, affecting about 1% of the population. Both conditions are typically chronic and progressive.

116 thousand knee replacements were carried out in England, Wales and Northern Ireland in 2019, 87% of which were funded by the National Health Service (NHS). The number of knee replacements increased by 26% between 2013 and 2019 before reducing sharply with the emergence of COVID-19 in 2020. Rates of surgery recovered somewhat in 2021 but remained below pre-pandemic levels.⁴ Pre-pandemic trends can be explained with respect to three factors: (1) increases in the prevalence of osteoarthritis and inflammatory arthritis, (2) innovations in surgical practices have lowered risks and have made surgery viable for an increasing number of people, and (3) supply capacity and NHS funding have increased. Despite these increases the supply of knee replacements remains highly inequitable.⁵ Patients from deprived areas, who might benefit from a knee replacement are considerably less likely to receive surgery than their counterparts living in other areas. Waiting times for surgery improved considerably between 2000 and 2014 but have subsequently regressed, a problem exacerbated by the pandemic.⁶

Surgery is usually considered when pain or poor knee function substantially limits day to day activities and when less invasive interventions, such as painkillers and physiotherapy, are no longer adequate. Knee replacement surgery is considered to be highly

³ Arthritis Research UK, Osteoarthritis in general practice. Data and perspectives. Arthritis Research UK. 2013 <https://www.versusarthritis.org>

⁴ National Joint Registry, NJR reports centre. <https://reports.njrcentre.org.uk/>

⁵ Judge A, Welton N, Sandhu J, Ben-Shlomo Y, Equity in access to total joint replacement of the hip and knee in England: cross sectional study BMJ 2010;341:c4092 <https://www.bmj.com/content/341/bmj.c4092>

⁶ NHS England, Referral to Treatment (RTT) Waiting Times. <https://www.england.nhs.uk/statistics/statistical-work-areas/rtt-waiting-times/>

preference-sensitive since the decision about whether to proceed with surgery is informed not only by objective clinical assessment of the benefits of treatment, but also by the importance of these benefits to the patient's lived experience, their attitudes towards pain, surgical risk, recovery time and treatment alternatives. The shared decision-making paradigm ensures that both clinical and patient perspectives are taken into account and have equal prominence.

The Oxford Knee Score (OKS) is a validated measurement scale that is widely used to assess knee function and the need for knee replacement surgery. The scale comprises 12 questions and delivers a score from 0 to 48, with higher scores indicating better knee function.

Since April 2009, the NHS has systematically collected pre-operative and post-operative Oxford Knee Scores for patients receiving an NHS-funded elective knee replacement. In 2019/20 the average (mean) pre-operative Oxford Knee Score was 19.2 and the average post-operative score was 36.3, indicating a substantial improvement in knee function and pain. The average gain in knee function after surgery, indicated by the difference between the mean pre- and post-operative scores has increased modestly but steadily from 14.6 in 2009/10 to 17.1 in 2019/20.⁷

Studies that compare the costs of knee-replacement surgery with the improvements it generates in quality of life conclude that is a highly cost-effective treatment.⁸ Wider societal benefits are also likely to accrue in the form of increased productivity and reduced dependency on care.

⁷ NHS Digital. Patient Reported Outcome Measures <https://digital.nhs.uk/data-and-information/publications/statistical/patient-reported-outcome-measures-proms>

⁸ Dakin H, Gray A, Fitzpatrick R, et al Rationing of total knee replacement: a cost-effectiveness analysis on a large trial data set. *BMJ Open* 2012;2:e000332

3. A short history of NHS contracting models for hospitals

For most of the history of the NHS, hospitals were allocated budgets that were set to ensure adequate provision of acute hospital care to a local population. Over time, as the counting of hospital activity became more sophisticated, these budgets were accompanied with expectations about the numbers and types of clinical activity that a hospital should deliver. During the first decade of the new millennium, and in response to increasing waiting times for planned surgery, a new payment mechanism was introduced. Payment by Results, now known as the National Tariff, assigned specific payments (tariffs) to groupings of clinical activity, Healthcare Resource Groups (HRGs). The tariff was based on the national average cost, but the tariff paid to a particular provider was adjusted to take account of local market forces (i.e., unavoidable cost differences). This new payment system provided incentives for providers to increase supply.⁹ Like most service providers, hospitals can make economies of scale which means the unit cost of provision falls as the scale of provision rises. Since the tariff was based on an average unit cost, the additional (marginal) income associated with increasing service provision outweighed the provider's extra (marginal) costs. The policy was backed by considerable additional funding and led to an increase in service provision and a reduction in waiting time. Tariffs, market forces factors and HRGs were adjusted annually. Over time, HRGs became increasingly stratified, taking account of hospital activity, patient characteristics, the mode of delivery (emergency, elective, day case etc) and staffing skill-mix. In the most recent version of the National Tariff there are 21 distinct HRGs and associated tariffs for elective knee procedures.

The global economic downturn that began in 2008, led to a tightening of NHS budgets and tariff increases were limited for several years, shifting the focus of the National Tariff from increasing supply to improving service efficiency. As tariff growth was constrained, some specialist providers argued that the average unit cost approach failed to account for the casemix and complexity of patients they typically handled. Specialist provider top-ups were introduced to ensure the sustainability of these providers. These included modest top-ups for specialist orthopaedic hospital Trusts.

Growing concerns about the quality and variability of service provision led to a further innovation in the National Tariff in 2010/11. Best Practice Tariffs (BPTs) sought to incentivise providers to improve service quality and reduce unwarranted variation in

⁹ Farrar S, Yi D, Sutton M, Chalkey M, Sussex J, Scott A Has payment by results affected the way that English hospitals provide care? Difference-in-differences analysis *BMJ* 2009;339:b3047 <https://www.bmj.com/content/339/bmj.b3047>

clinical practice, by paying enhanced tariffs to providers that could demonstrate that they had adopted specific evidence-based practices and complied with enhanced reporting. BPTs for knee replacements were introduced in 2014/15 to incentivise providers to participate in the National Joint Registry and Patient Reported Outcomes Measurement programmes, and to use a particular form of surgery (cemented or hybrid prostheses) in at least 80% of patients aged 70 or over. In addition, providers were assessed in terms of the average health gain their patients reported following knee replacements. Specifically, the health gain reported could not fall significantly below the national average. Providers that failed to meet all of these criteria were paid at a tariff 10% below the standard rate.

During the initial phase of the SARS-CoV-2 pandemic, NHS hospitals were instructed to cancel most forms of elective surgery, to free up capacity for COVID-19 patients. The National Tariff arrangements were suspended, and hospitals were instead paid on a block contract basis.

The most recent developments in contracting for hospital services are designed to support a smooth transition back to the National Tariff, by combining a fixed payment and a variable element based on actual activity and performance. Where implemented, the effect of these arrangements is to moderate the provider incentive to increase supply, whilst enabling commissioners to incentivise other aspects of service provision, such as quality and outcomes. The detail of these arrangements is subject to considerable local determination.

Waiting times for elective care have once again, risen to the top of the political agenda for health service provision, carrying with it, support for policies such as the National Tariff, that incentivise providers to increase activity.

4. Contracting as a means to improve outcomes

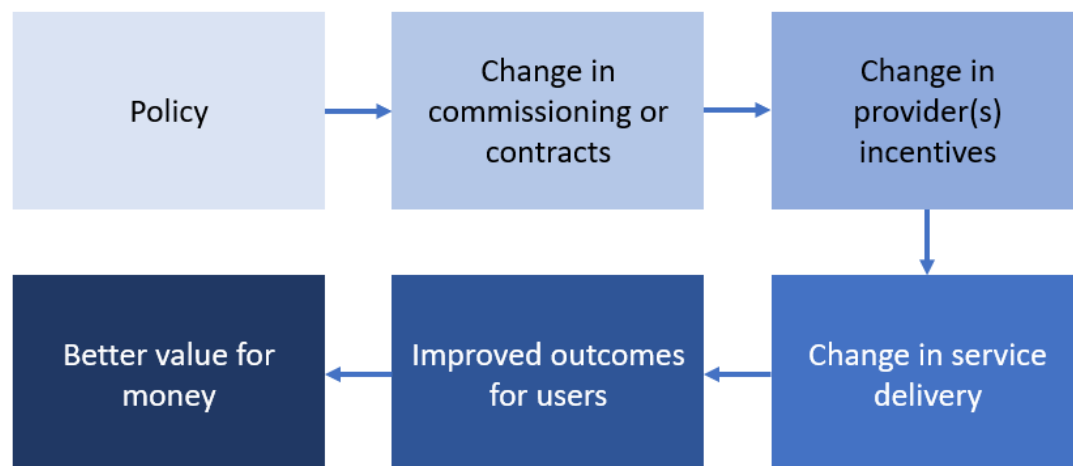
Improving service outcomes is always desirable from a patient perspective. And if these improvements can be delivered at an acceptable cost, then funders are also likely to be supportive. But it does not necessarily follow that we should use contracts to achieve this objective. Can we be sure that outcomes-based contracts will lead to improved outcomes? What might be the unintended effects? And are there alternative, less resource intensive, or more reliable ways of achieving the same objective?

Contracting is one of many means by which quality and outcomes might be improved, alongside staff training, recruitment, continuous improvement or PDSA cycles, peer review, performance management etc. Each of these will consume resources and management time which are always limited and stretched. Decisions about whether to adopt outcomes-based contracts should be taken in light of these alternatives.

If outcomes-based contracts for knee-replacement surgery are adopted within the NHS, then evaluations can be used to determine their efficacy. In the meantime, we must rely on other types of information that provide some imperfect insight: theory and empirical evidence from similar contractual interventions.

Battye (2015), describes the theory of change for a policy that seeks to improve outcomes through contractual mechanisms.¹⁰

Figure 1: Theory of change, adapted from Battye (2015)



¹⁰ Battye F, payment by Results in the UK: Progress to date and future directions for evaluation, Evaluation 2015, Vol 2(2) 189-203.

In the same paper, the author suggests that the high levels of political commitment to these contractual mechanisms, arise primarily from an appeal to values, rather than the evidence. The evidence base has however, developed considerably in recent years. A 2021 Cochrane Review¹¹ concluded that:

"The evidence base on the impacts of P4P (pay for performance) schemes has grown considerably, with study quality gradually increasing. P4P schemes may have mixed effects on outcomes of interest, and there is high heterogeneity in the types of schemes implemented and evaluations conducted."

The authors go on to say that:

"P4P is not a uniform intervention, but rather a range of approaches. Its effects depend on the interaction of several variables, including the design of the intervention, the amount of additional funding, ancillary components and contextual factors."

Battye also highlights risks or unintended consequences of these contractual mechanisms. The approach assumes that patient outcomes are a function of service quality. But these outcomes are also influenced by other factors such as patient casemix and outcome recording practices. Providers might therefore seek to increase their income by cherry-picking certain patients or by improving the quality and completeness of outcomes data. Furthermore, a provider might adopt practices that increase its outcomes, but at the expense of others, by poaching for example the best surgeons from their peers. Commissioners will need to take a view on the legitimacy of these practices, and design means of curtailing or moderating undesirable provider behaviour. However designed, transaction costs for outcomes-based contracts are likely to be higher than for simpler contractual models.

In conclusion, both theory and evidence suggest that the success of an outcomes-based contract is highly contingent on its design and implementation.

¹¹ Diaconu K, Falconer J, Verbel A, Fretheim A, Witter S. Paying for performance to improve the delivery of health interventions in low- and middle-income countries. Cochrane Database of Systematic Reviews 2021, Issue 5. Art. No.: CD007899

5. Fairness in outcomes-based contracts using casemix-adjustment

Concerns about fairness feature heavily in debates about whether to adopt outcomes-based contracts. These concerns take two forms: fairness for providers and fairness for patients. The first focuses on whether providers will be reimbursed fairly for the work that they have done and the patients they have treated. It may be more difficult for a provider to achieve a certain level of average health gain if it tends to work with patients with more complex needs or in areas of higher deprivation. Providers want reassurance that contract payments will take account of relevant differences in the mix of patients that they treat.

The concern from a patient perspective is that providers might only treat those patients who are likely to see the greatest gain in health. This practice, often referred to as 'cherry-picking' or 'cream-skimming', could exacerbate existing inequalities, reduce access to effective treatments and discourage innovation.

Fortunately, a method, known as casemix-adjustment, can address both concerns. The method adjusts the average health-gain that a provider achieves according to the mix of patients that it treats. There are several ways to achieve case-mix adjustment, but the most common approaches are based on well-established statistical techniques such as standardisation and regression.

Casemix-adjustment is a pragmatic, rather than a perfect solution. Analysts must make judgements about which factors to adjust for, and this in turn is constrained by the data that is available.¹² The approach tends to be more effective, and less prone to methodological problems, when used to adjust outcomes for a specific service, such as knee-replacement surgery, than for hospital services en masse. It also tends to emphasise case complexity in terms of co-morbid risks and demographic factors. It is more difficult to account for the surgical view of case complexity. A primary knee replacement may involve a spectrum of procedures from a resurfacing of the joint to a stemmed locked hinge prosthesis depending on a number of surgical factors that will largely be assessed clinically and radiologically as well as intra-operatively.

NHS Digital have published average casemix-adjusted health gain scores for providers that carry out elective knee replacement surgery for several years. The methods are updated and improved periodically in response to feedback. The most recent version takes account of patients' age, sex, ethnicity, deprivation of area of residence, disability

¹² Factors that affect patient outcomes but are not observed or measured, cannot be directly adjusted for, and so may bias any results.

status, living arrangements, PROMs form completion arrangements (self or by proxy), selected comorbidities, pre-operative Oxford Knee Score, pre-operative symptom period, and procedure type (primary vs revision surgery).

Use of these casemix-adjusted health gain measures within outcome-based contracts should in theory lead to somewhat fairer comparisons between providers and moderate provider incentives to 'cherry pick' patients.

6. Contract theory

How might health systems use the flexibilities set out in the latest National Tariff guidance to maximise the benefits that accrue from the delivery of elective knee replacement surgery? What exactly should be incentivised and how big should the incentives be? Whilst it may be fair to assume that both provider and commissioner want to improve health outcomes for patients receiving knee replacement surgery, it is important to acknowledge that these two organisations have subtly different imperatives, challenges, information, and competing priorities. If incentives are poorly designed and scaled, then the desired improvements might not be realised, scarce resources might be wasted, and organisations may be financially destabilised.

Contract theory is an area of academic research within the broader discipline of micro-economics that is specifically concerned with the design of contractual arrangements.

Contract theory explores how individuals and organisations respond to contractual arrangements in the presence of incentives and risks. It can be used to ensure that incentive payments are sufficient to change behaviour and improve outcomes, and to ensure that a provider can just cover the costs associated with delivering the desired outcomes.

Contract theory was first described in the 1960s and the discipline was extended and improved in the subsequent decades. The theory has been used to address issues in a wide range of settings including CEO pay levels, public utilities, used car sales, and telecommunications. Within the health sector, contract theory has been used to investigate the properties of different payment systems for secondary and primary care providers.

In his recent paper (Siciliani et al 2021)¹³, Professor Luigi Siciliani and colleagues describe how the PROMs data collection arrangements and contract theory might be used to design and scale financial incentives to improve health outcomes from elective joint replacement surgery. In the next chapter we outline how such a contract might operate, how the key figures might be set, before moving on to explore how these incentives may alter commissioner and provider costs and income. We aim to illustrate that such contracting arrangements are, at last, feasible and within reach of competent commissioners and providers.

¹³ Siciliani L, Gaughan J, Gutacker N, Gravelle H, Chalkey M, Paying for Health Gains, CHE Research paper 183. July 2021
https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP183_Paying_Health_Gains.pdf

7. Imagining an outcomes-based contract for knee replacements surgery

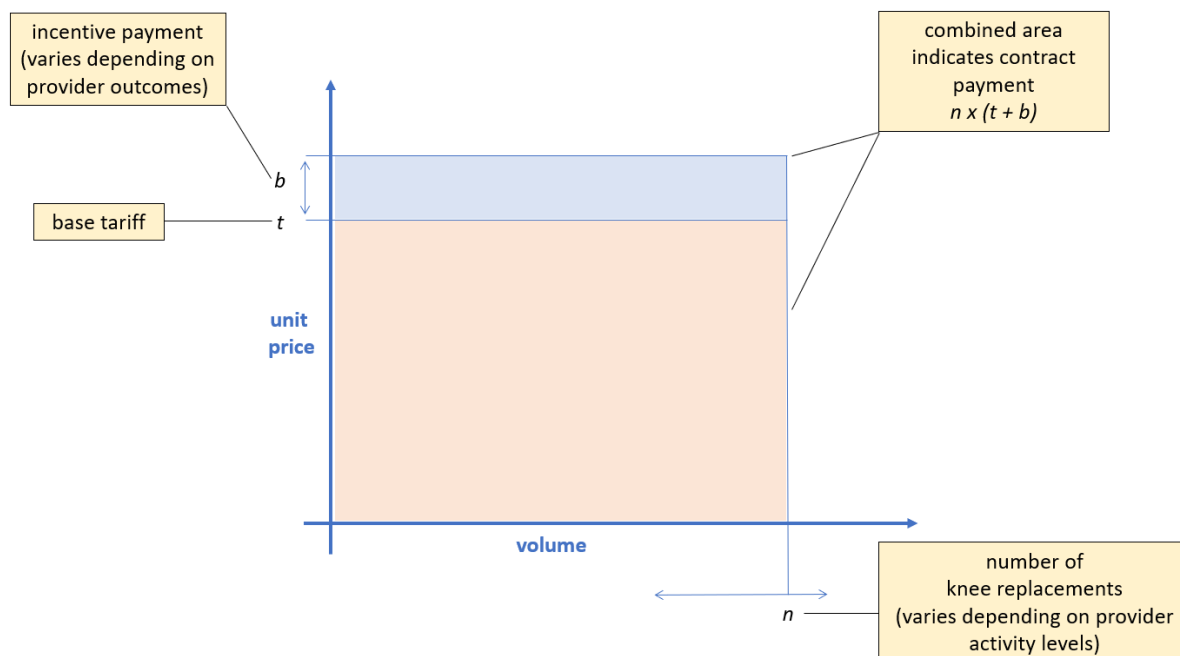
We set out in this chapter how an outcomes-based contract for elective knee replacement surgery might operate. We illustrate how funding flows would have been affected in 2019/20 if such a contract had been in operation and had been calibrated in a particular way.

7.1 Contract form

Our proposition is that a provider payment from the commissioner should be based on three quantities:

- (1) the volume of knee replacements carried out (n),
- (2) a base tariff for each knee replacement surgery carried out (t) and
- (3) a bonus or incentive payment (b) for each knee replacement surgery carried out that varies by provider according to the average gain in knee function experienced by the provider's knee replacement patients.

Figure 2: An outcomes-based contract



The commissioners might want to set the value of the base tariff (t) and the incentive payment (b) so that:

1. the total commissioning spend does not increase substantially, but the contracting arrangement supports the provider to seek improvements in outcomes
2. the incentive payment is sufficient to cover a provider's reasonable costs, associated with improvement programmes
3. the incentive reflects the complexity of cases that the provider treats and does not inadvertently introduce bias in the patient selection process.

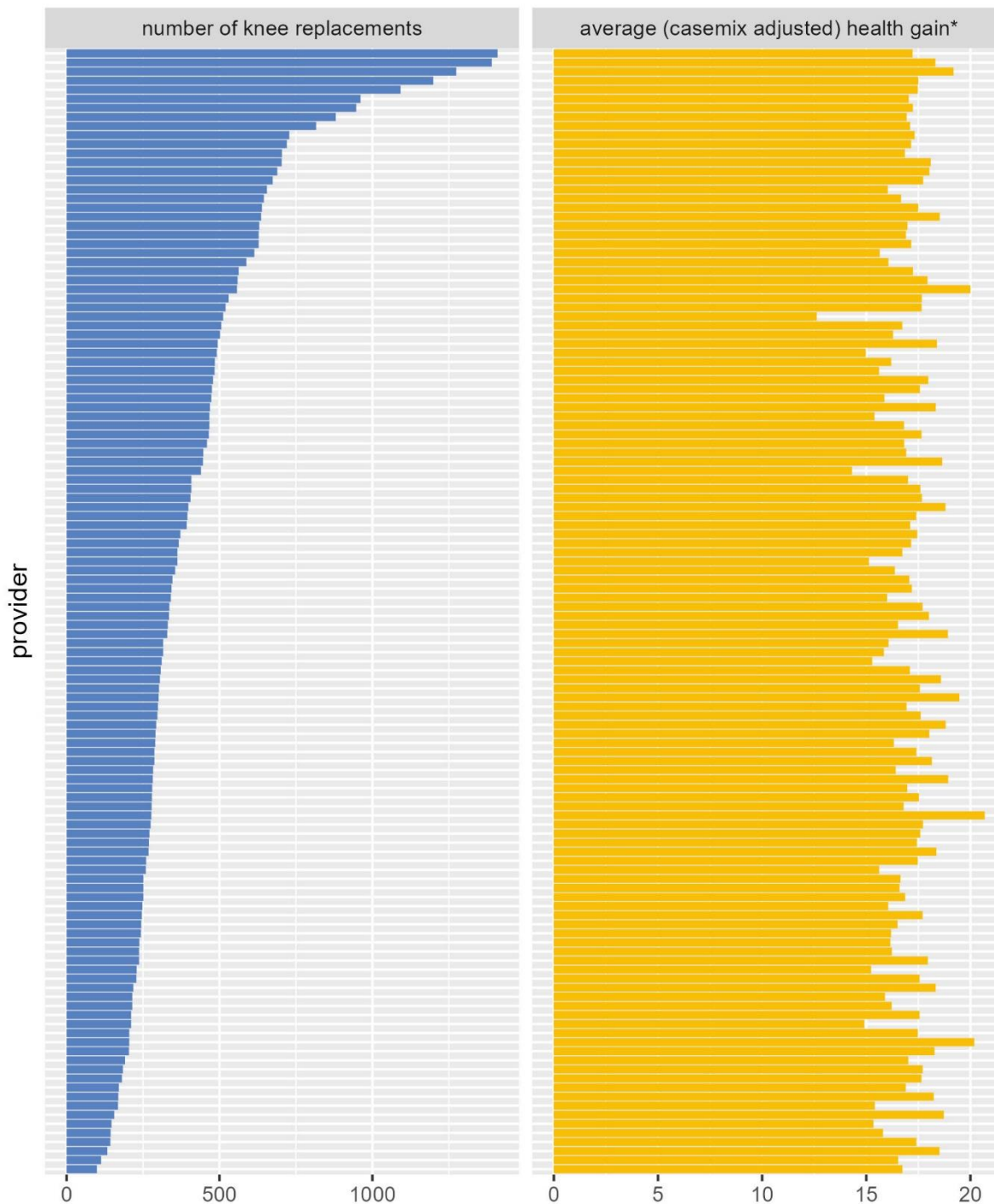
7.2 How might an outcomes-based contract change provider income & commissioner expenditure?

In chart 1 below we show the number of elective knee replacements that providers carried out in 2019/20 and the average gain in knee function that patients reported following surgery, having adjusted for the casemix of the patients treated.

Chart 1 - Provider activity volumes and outcomes

NHS-funded knee replacements, volume & health gain

124 providers in England | 2019-20



** measured using the Oxford Knee Score | casemix adjusted using NHSD method*

We set out below how such a contract might have influenced commissioner expenditure and provider income in 2019/20 based on credible values for the base tariff (t) and the

unit incentive payment (b). For this example, we set the base tariff to 86% of the current tariff, and the bonus payment at £200 per patient for every unit of average casemix adjusted gain in Oxford Knee Score above 12. These figures are used purely to illustrate the impact of such a contract. In the next chapter we set out how these values might be set in line with alternative commissioner objectives and our best knowledge of provider costs.

Whilst our hypothetical outcomes-based contract is designed to incentivise providers to improve health outcomes, its impact will depend on the extent to which providers respond to this incentive. In table 1 below we show how commissioner costs (provider income) at a national level, would change under three scenarios:

- 1) the outcomes-based contract produces the same health outcomes as the existing national tariff best practice tariffs (no health improvement).
- 2) the outcome-based contract has a uniformly positive impact on health outcomes for all providers, equivalent to a one-point increase in average Oxford Knee Score gains.
- 3) the outcome-based contract increases Oxford Knee Score gains from 0 to 2 points across providers. Providers with the lowest casemix-adjusted Oxford Knee Score gains at baseline achieve the greatest increases.

Table 1: Impact on total Commissioner Costs (England)

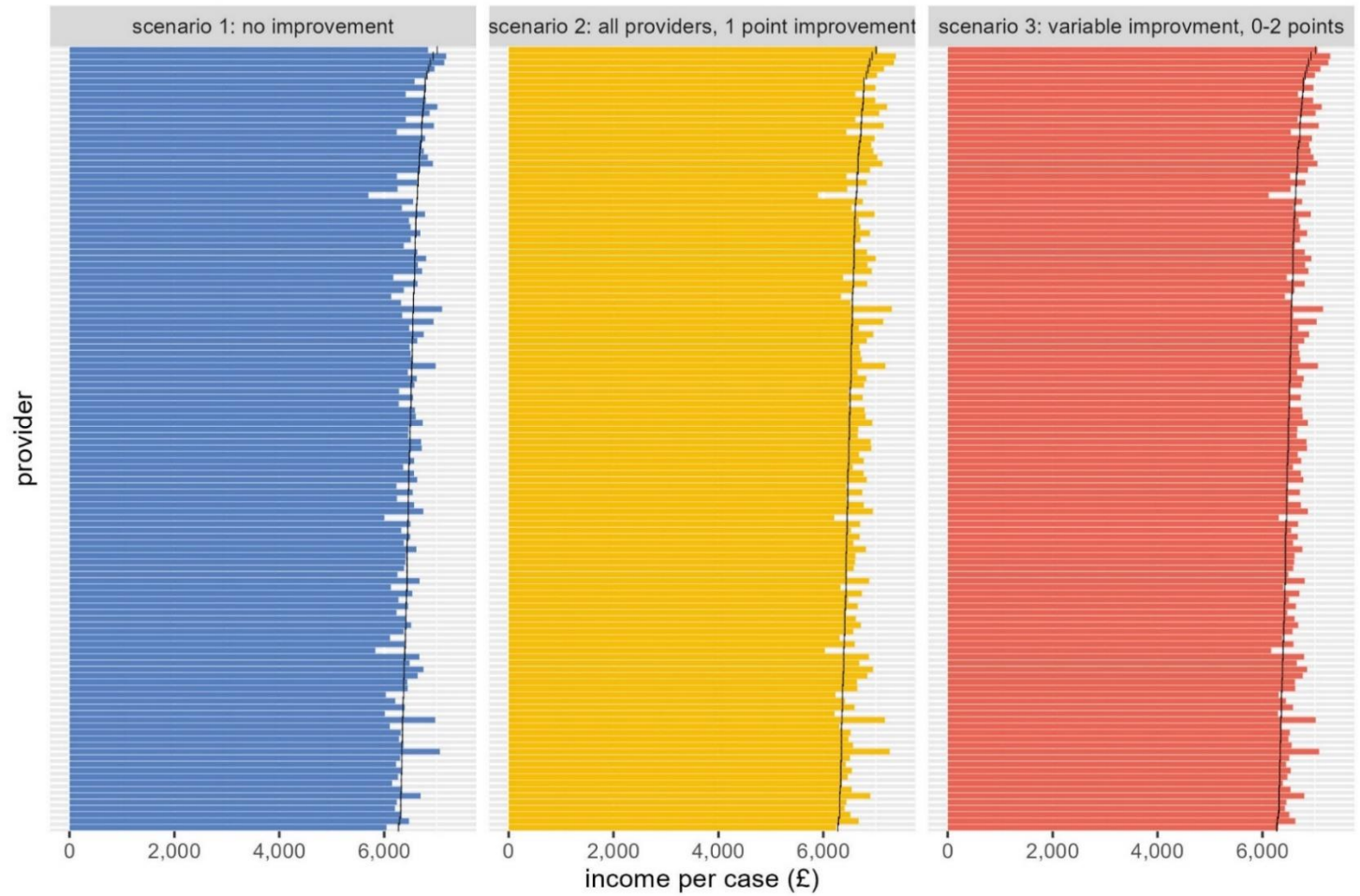
Scenario	Impact of outcomes-based contract on average (casemix adjusted) gain in Oxford Knee Score	% change in all England commissioner costs relative current contractual model
1	No improvement in health gain	+0.0%
2	1-point improvement for all providers	+3.1%
3	0 to 2-point variable improvement	+2.7%

These are the effects at a national level. The effect on individual providers will vary according to the outcomes they deliver. Chart 2 shows the level of income for each provider under these three scenarios.

Chart 2 - Potential Impact of the contract on provider income

Provider income under outcome-based contract

Three health gain scenarios



black line indicates income in under current contract

In chart 3 we present this same information as changes in provider income relative to the income received under the current arrangements. In scenario 1, when the outcomes-based contract leads to no change in health gain, the net effect on commissioner costs, aggregated at a national level, is zero. But 48% of providers, those delivering lower levels of case mix-adjusted health gain, would see income reduce, whilst those delivering higher levels of health gain would see income increase. These changes in income range from -14.1% to +11.4%. In total the contract would reallocate 1.2% of the total commissioning budget from providers with poor outcomes to those with better outcomes.

In scenario 2, when the contract leads to all providers increasing health gain by one point on the Oxford Knee Score scale, total commissioner costs increase by 3.1%. 80% of providers would see an increase in income and provider income changes range from -11.1% to +14.6%.

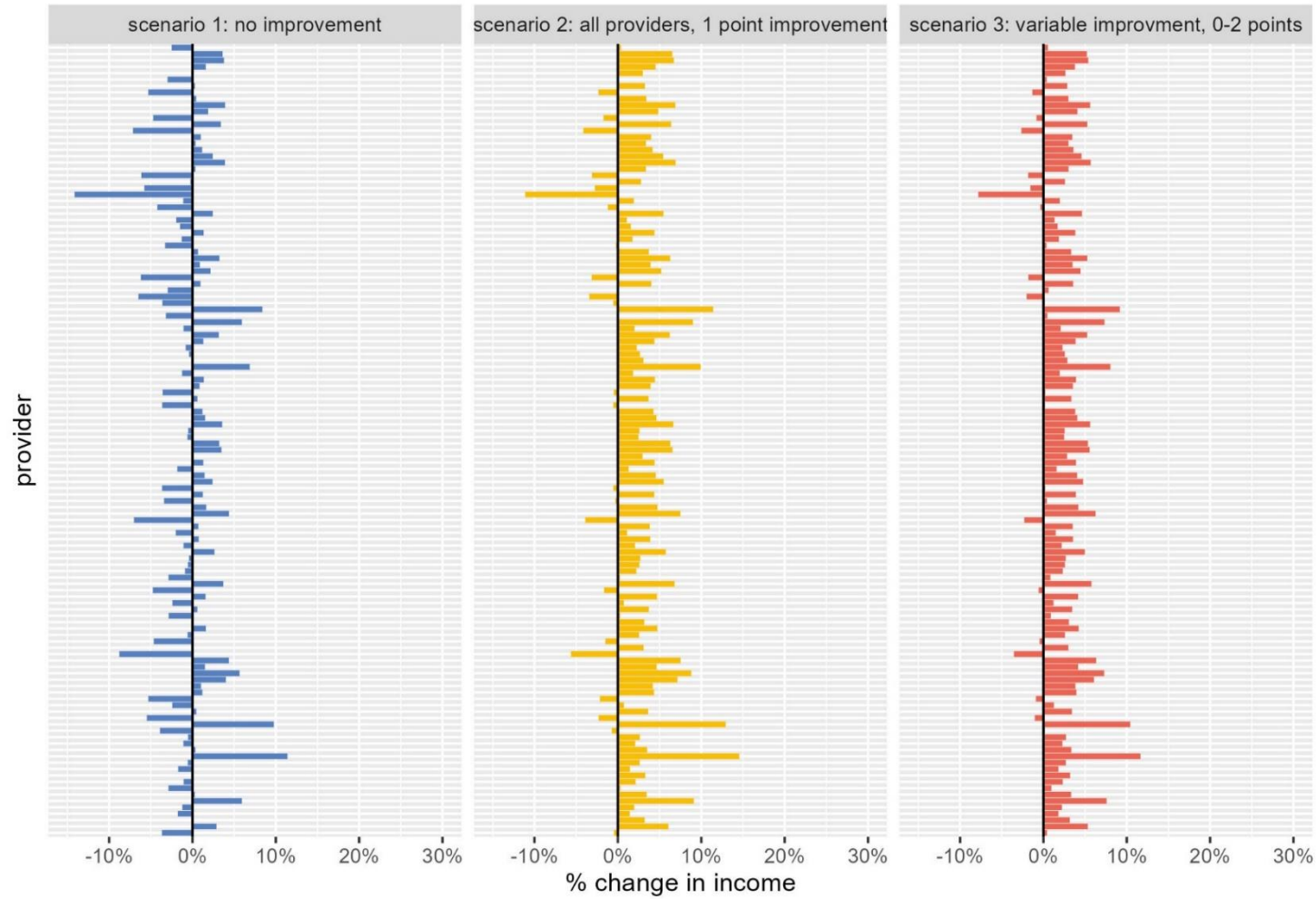
In scenario 3, we assume that those providers with the lowest casemix-adjusted health gain before the outcomes-based contract was introduced, (i.e., those with the greatest potential for improvement) see the largest increase (2 points) whereas those with the highest health gain at baseline see no improvement.¹⁴ This leads to an increase in national commissioner costs of 2.7%. 84% of providers would see an increase in income with changes in provider income varying from -8.1% to 11.4%.

This example, and our selected scenarios, shows how an outcomes-based contract would redistribute funds given specific values for the base tariff (t) and the incentive payment (b). In the next chapter we show how contract theory can be used to set these values rationally and the factors that should be considered when setting t and b for a specific outcomes-based contract.

¹⁴ This may be driven by the contractual incentives, but also by regression to the mean.

Chart 3 - Impact of the contract on income, by provider

Change in provider income under outcome-based contract Three health gain scenarios



8. Calibrating an outcomes-based contract: the Commissioner's role

In the example above, we have shown how provider income and commissioner expenditure for knee replacement surgery might have been different if a specific outcomes-based contract had been in place. But these income and expenditure levels are dependent on the bonus payment and the base tariff values. If we had set these at different levels, then economic theory suggests that this would have induced different provider behaviour and generated different payments.

In this chapter, we set out how a commissioner might calibrate an outcomes-based contracting scheme for elective knee-replacement surgery. We start by setting out what data the commissioner will require and what judgements the commissioner must make. We then show, using a set of flow diagrams, how this information combines with provider performance data to determine provider payments. We provide these diagrams for completeness and acknowledge that these can appear somewhat complicated. It is not necessary however, for a reader to assimilate the details of this process until a decision is taken to implement the approach. The key insight here is that the commissioner must obtain some data and make some judgements in order to calibrate the contract payments in line with sound contract theory. Some relatively simple arithmetic is then used to work out the details.

8.1 Designing the scheme

How should a commissioner set the level of bonus payments and base tariff, whilst demonstrating good stewardship of public resources? There are six things the commissioner needs to decide or establish. Two are judgements that the commissioner needs to make: the target average health gain and the threshold above which a bonus payment is made. And four are data items that the commissioner must source: the baseline average health gain, and the incremental costs of providers delivering improvements in health gain.

At the point of implementation, two further pieces of information are required to determine the payment that a specific provider will receive under the scheme: the providers' average case mix-adjusted health gain and the tariff paid to providers before the outcomes-based contract was introduced.

We set out these judgements and data items below. For the purposes of this paper, the 'average health gain' that we are interested in is the improvement in knee function, as

measured by the Oxford Knee Score that patients experience following surgery, averaged over all patients treated, having adjusted for the patient casemix.

Box 2: Data and judgements required to calibrate and implement the scheme

Judgements required for scheme design

J1 Target health gain: What level of average health gain does the commissioner want to incentivise providers to achieve?

J2 Bonus threshold: What is the minimal level of average health gain above which the commissioner is prepared to reward providers for delivering?

Data required for scheme design

D1 Baseline health gain: What is the average health gain delivered across all providers now?

D2 Incremental cost at baseline: What is the average incremental cost to providers of delivering one additional unit in average health gain from the baseline (D1)?

D3 Incremental cost at target: What is the average incremental cost to providers of delivering one additional unit in average health gain to the target level (J1)?

D4 Incremental cost from baseline to target: What is the average incremental cost of providers increasing health gain from the baseline (D1) to the target (J1)?

Data required for scheme implementation

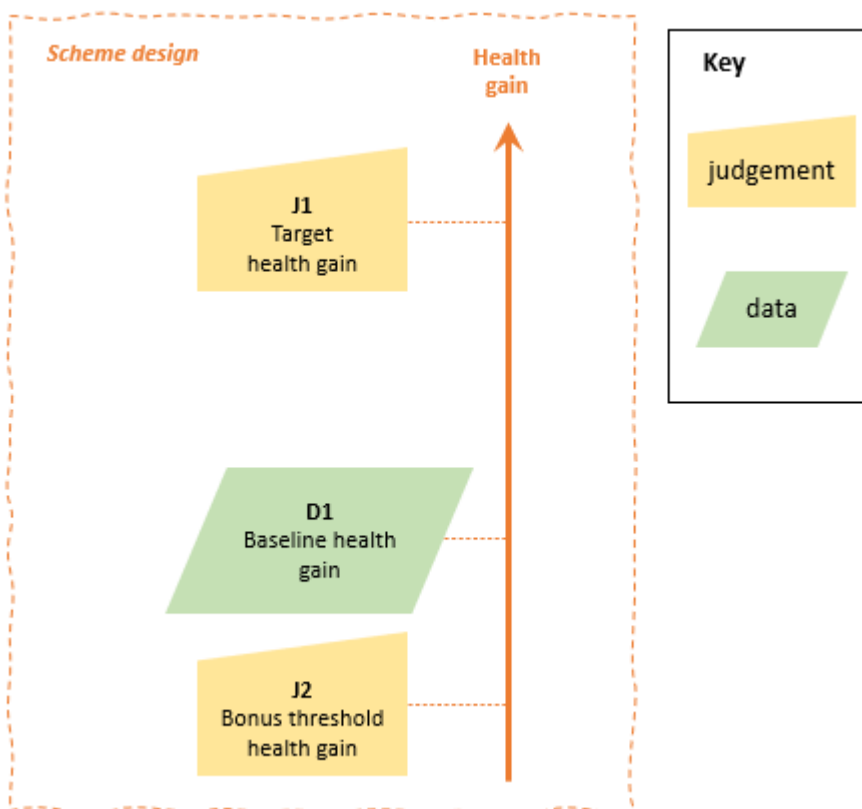
D5 Provider health gain: What level of average health gain did the provider deliver under the new scheme?

D6 Current tariff: The levels of payment that providers received in the absence of an outcomes-based contract. Note that this will need to be adjusted to take-account of inflation and anticipated efficiency gains as with all other tariffs.

Baseline, target, and bonus threshold health gains (J1, J2 & D1)

The two judgements, J1 and J2 and the data item D1, all operate on the same scale, health gain.

Figure 3: Baseline, target, and bonus threshold health gains



D1, the baseline health gain is straightforward to source. It is calculated and published each year by NHS Digital. The provisional data for 2020-21 suggests that the average, casemix-adjusted gain in Oxford Knee Score for patients receiving an elective knee replacement in England was 16.7 pts.¹⁵ This is the difference between the average pre-operative and post-operative scores having adjusted for casemix.

J1, the target health gain, articulates the commissioner's objective for the contract. The maximum possible gain is the difference between the average pre-operative Oxford Knee Score and 48, the maximum post-operative score. Delivering improved knee function requires expenditure, and the scheme must recognise this and reimburse providers accordingly. Incremental costs (the cost of delivering an additional point of health gain on

¹⁵ The Oxford Knee Score takes values from 0 to 48.

the Oxford Knee Score scale), are likely to increase as the target health gain rises. It will cost more on average for a provider to increase health gain from say 20 to 21 points than it would to increase it from 17 to 18 points. So affordability, and competing calls on resource, will lead commissioners to set a target below the maximum possible health gain.

J2, is the level of health gain above which, providers are paid a bonus. The simplest way to set this value is at the lowest casemix-adjusted health gain score that a provider delivered in recent years. That way all providers are likely to be engaged in the scheme and have a realistic chance of increasing their income by attempting to improve the health gain they deliver. If, however, commissioners believe that this minimum level represents poor practice, then they may choose to set the threshold for bonus payment at a higher level. Whatever level this is set at, the commissioner has a further opportunity to manage their expenditure, and ensure provider costs are covered, by adjusting the base-tariff. We explore this issue below. In the diagram above, we have shown J2 below D1 (the baseline health gain). This would normally be the case, but J2 can take any positive value below the target, J1.

It is important to draw this distinction between the bonus threshold health gain (J2) which is averaged over all patients, and an acceptable level of health gain for an individual patient. A patient with intractable night pain, for example, may report a significant improvement in their quality-of-life following surgery, with only a modest improvement in their Oxford Knee Score.

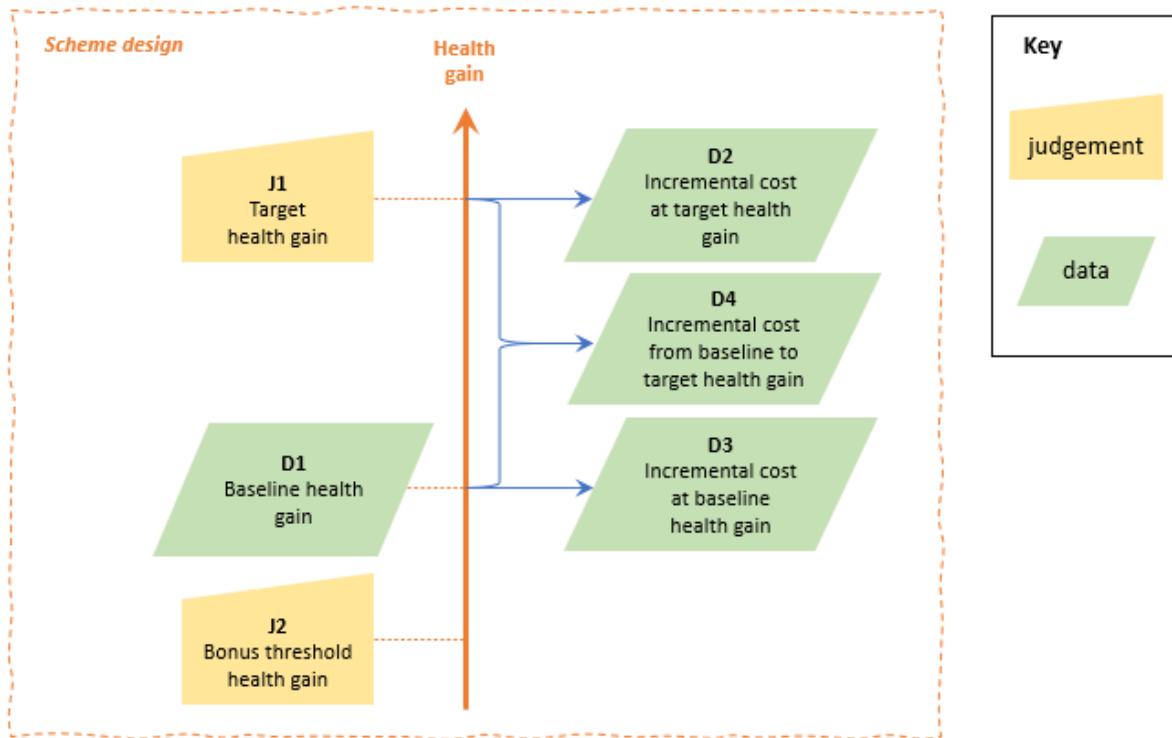
Average incremental costs (D2, D3 and D4)

Having set the target health gain (J1) and sourced the baseline health gain (D1), three further data items must be estimated. These all take the form of incremental costs, a provider's additional costs associated with increasing health gain. The three values are: D2, the incremental cost of increasing health gain by one unit to the target (J1); D3, the incremental cost of increasing health gain by one unit from the baseline (D1); and D4, the cumulative costs of increasing health gain from the baseline to the target. In all cases, these incremental costs should be averaged over all providers, and patients.

These are more difficult quantities to estimate but might be informed from two sources. Econometric methods might be used to estimate these values from historical data on provider costs and outcomes. Alternatively, or ideally in addition, commissioners might elicit expert views to estimate these quantities. This might be done informally or using an established method such as Delphi. Commissioners should not seek perfection here. Estimates, derived from empirical studies and expert opinion should be sufficient. In

future years, these values can be calibrated against the provider response and to incorporate new research.

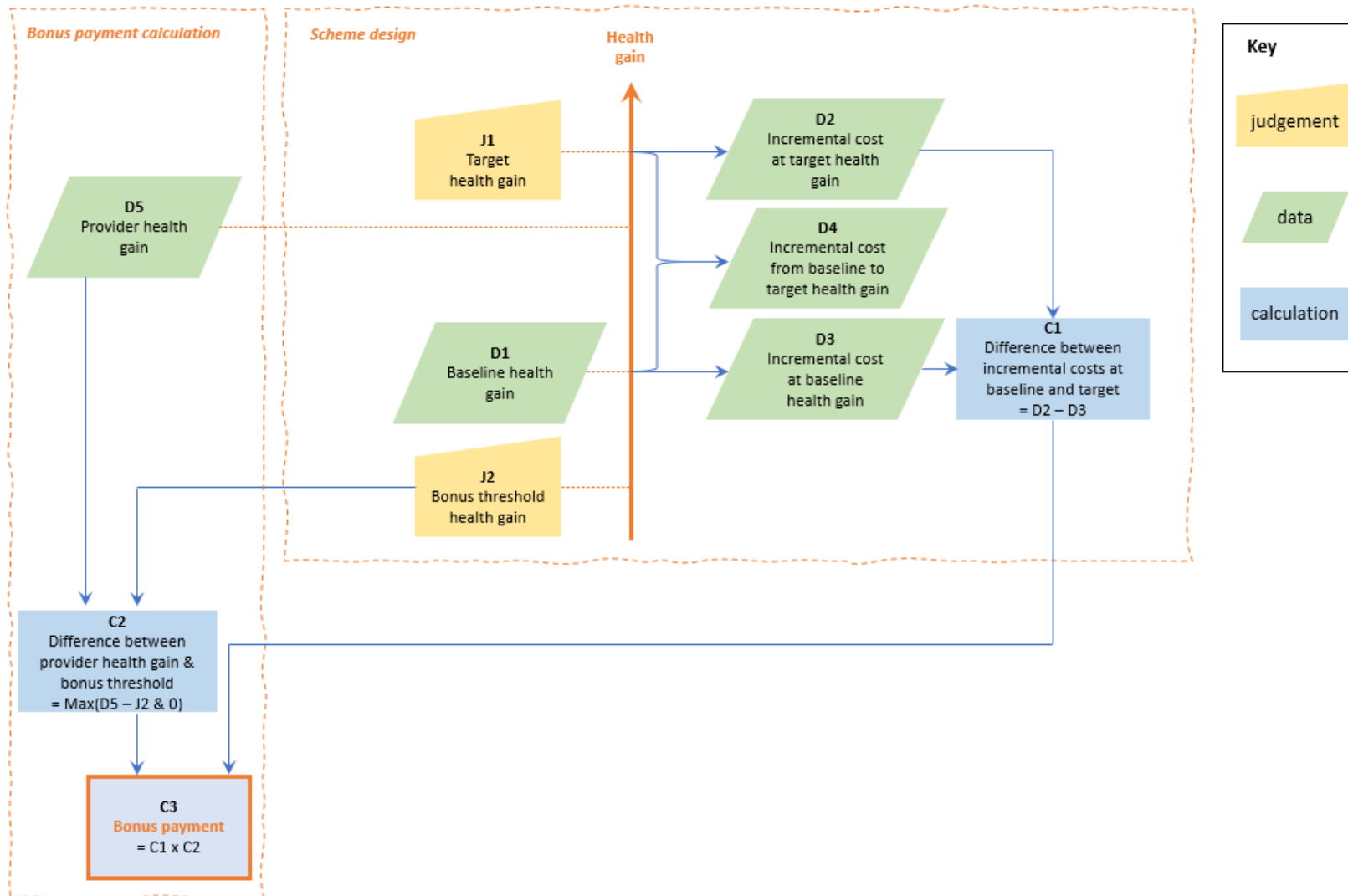
Figure 4: Incremental costs of health gain improvement



8.2 Setting the bonus payment level

Once these values have been obtained, contract theory can help us calculate the bonus payment for a provider given the health gain it delivers (D5). If the provider fails to deliver a health gain greater than the bonus threshold (J2) then no bonus payment is made. Otherwise, the bonus payment the provider receives (C3) is equal to the difference between the provider's health gain and the bonus threshold (C1), multiplied by the difference in the incremental cost of a unit of health improvement at target health gain and the incremental cost of a unit of health improvement at the baseline health gain (C2).

Figure 5: Calculating the bonus payment



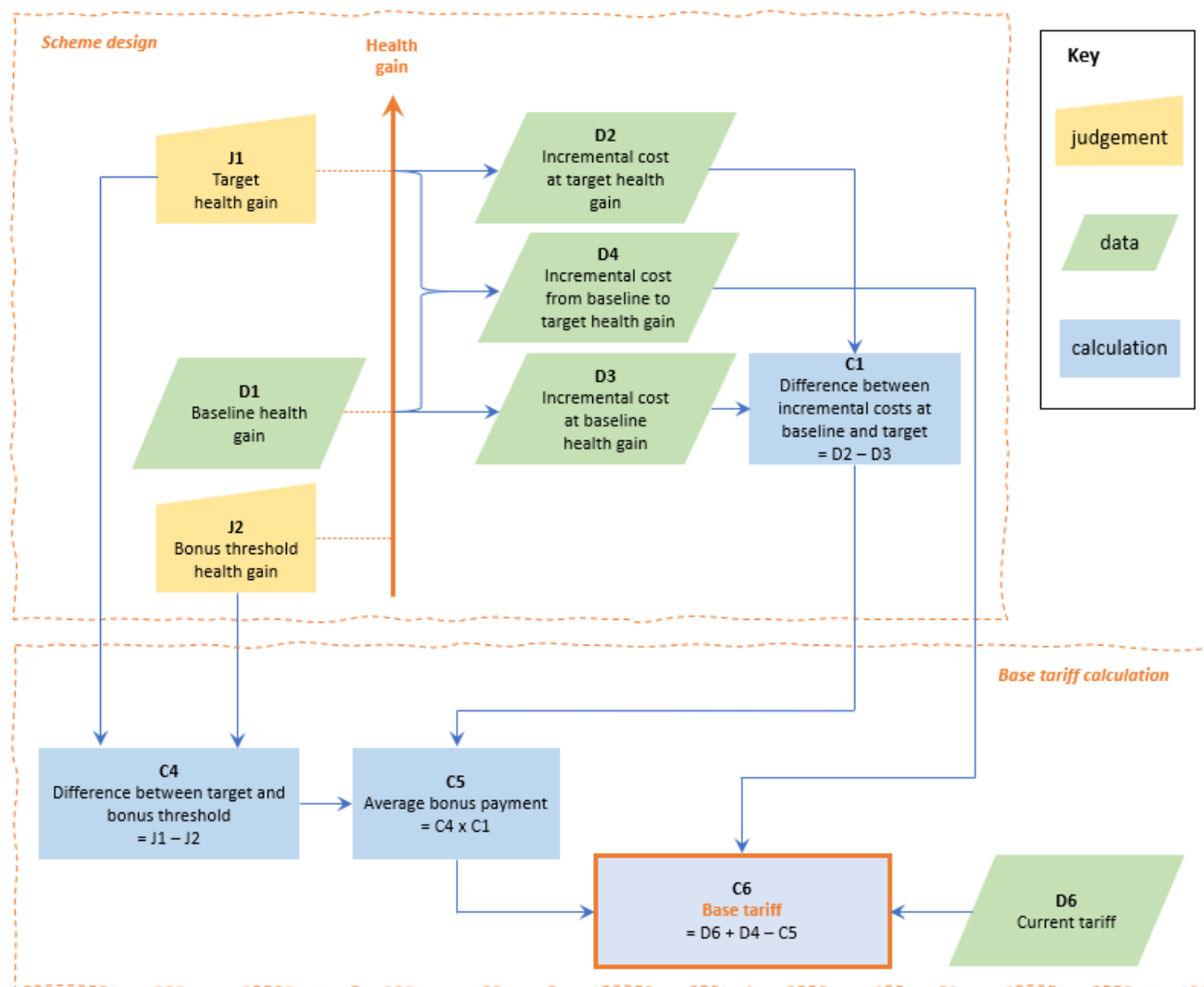
8.3 Adjusting the base-tariff

Next, we need to calculate the base tariff that providers will be paid, irrespective of the health gain they deliver.

Historically, national tariffs have been set to ensure that, on average, the income that providers receive is equal to their expenditure. This isn't to say that all providers break-even; some will make a surplus, and others will make a loss. But break-even is the average experience of providers.

In an outcomes-based contract, the payment a provider receives per patient is made up of two parts, a base tariff and a bonus based on the average health gain delivered. If commissioners want to ensure that on average this new scheme is sufficient to cover provider costs on average, then the base-tariff must be adjusted from its current level to take account of the anticipated bonus payments and the additional costs associated with delivering the improved outcomes (D4). The average cost across providers following the bonus payment and taking into account the health improvement is equal to the base tariff (D6) plus the additional costs associated with delivering improved outcomes (D4). The anticipated bonus payment is the payment that a provider would receive if it increased health gain to target (C5). The adjusted base tariff (C6) is therefore the average cost (D6+D4) minus the average bonus payment (C5).

Figure 6: Adjusting the base tariff



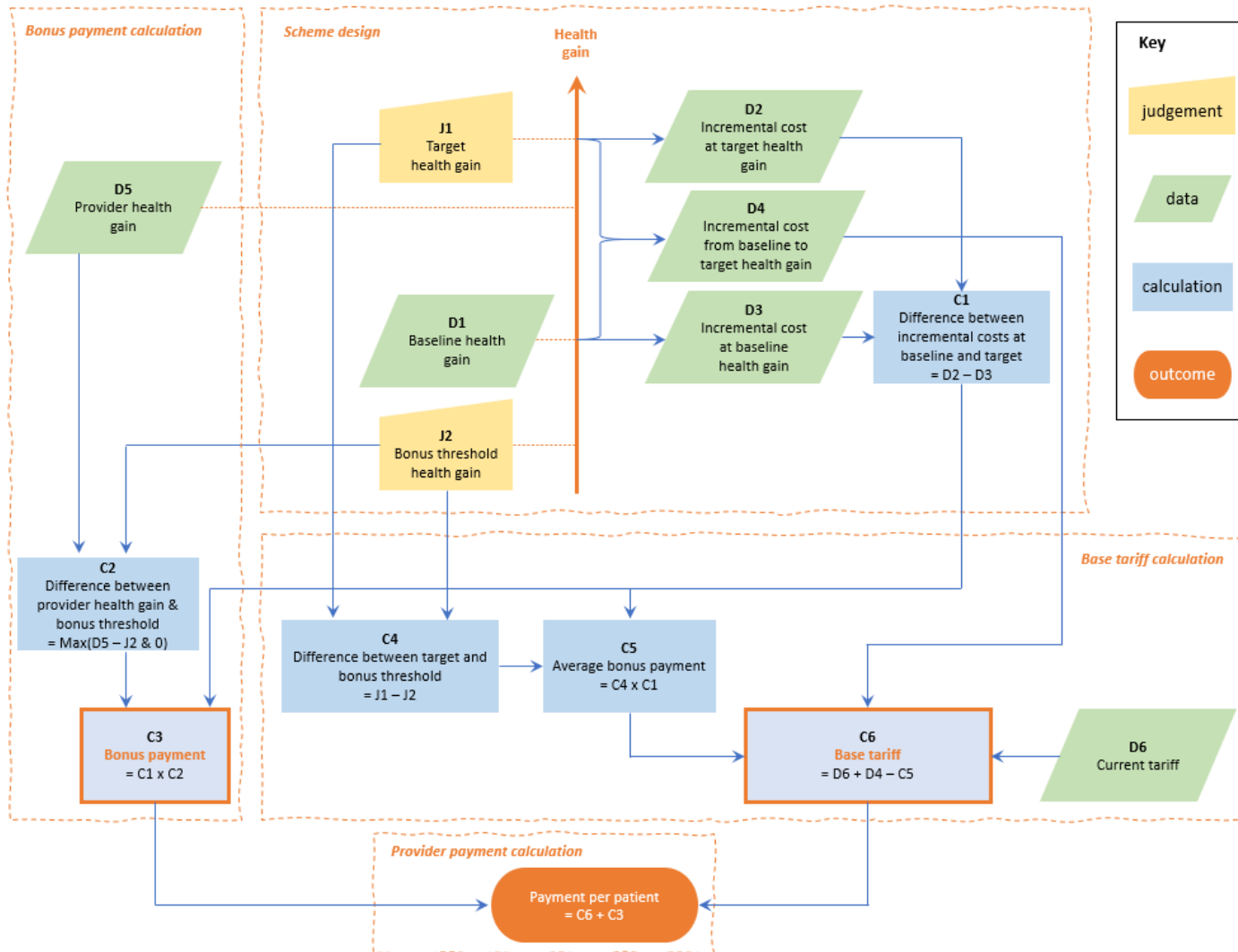
8.4 The total payment to providers

Having calculated the base tariff and the bonus payment, the total payment for a provider, for each knee replacement surgery carried out, is the sum of the base tariff (C6) and the bonus payment (C3). Note that the base tariff is the same for all providers, whereas the bonus payment depends on the average casemix-adjusted health gain that the provider delivers.

The work of Professor Siciliani confirms that this is the optimal design for the bonus payment and base tariff, in the sense that this is the most cost-efficient way for the commissioner to achieve its target health gain whilst fairly reimbursing providers.

A more mathematical description of these arrangements can be found in the appendix.

Figure 7: Combining the base tariff and bonus payment



9. How should providers respond to an outcomes-based contract?

Healthcare providers are constantly seeking to improve outcomes for their patients. There is rarely a shortage of ideas about how this could be achieved. We set out some example schemes in box 1 below.

Box 2: Examples of interventions to improve outcomes from elective knee replacement surgery

a) Assertive patient follow-up using outcomes data to guide

This service improvement initiative will seek to increase the coverage of post-operative follow-up, by employing a small team of telephony staff to contact patients after surgery to identify early signs of post-operative complications or slow recovery and to prompt completion of post-operative PROMs surveys.

b) Enhanced post-operative group physiotherapy

This service improvement initiative seeks to improve patient outcomes by offering two additional group physiotherapy sessions to patients after surgery. This would improve contact between the service and patient post-operatively, increase patient confidence, and accelerate recovery of function.

c) Targeted, enhanced post-operative physiotherapy

This initiative will also seek to improve patient outcomes by increasing the level of post-operative physiotherapy, but under this scheme the physiotherapy will be offered on a 1:1 basis for those identified pre- or post-operatively to be at risk of a poorer outcome.

Deciding which ideas to pursue is less straight-forward. These decisions are constrained by funding availability and by management capacity. Outcomes-based contracts do not change provider objectives but alter the context within which these decisions are made, by providing an income stream to offset the costs of service improvement. But these funds are only realised if improvements lead to better health outcomes, and it is not possible for providers to know with certainty the extent to which a proposed service improvement will lead to improved health outcomes.

Under a given outcomes-based contract, decisions to invest in a service improvement will be influenced by:

- The expected cost of implementation of the service improvement, both initial set-up costs and ongoing, additional costs
- The increase in income associated with an expected increase in average, casemix adjusted knee function that will result from the service change
- The patient health benefits arising from such improvement.

A provider may be more inclined to adopt a service improvement initiative when the income associated with improved health outcomes equals or exceeds the expected costs of delivery. And the bigger and more certain the gap between expenditure and income for a given service improvement, the more likely it is that a provider will adopt it.

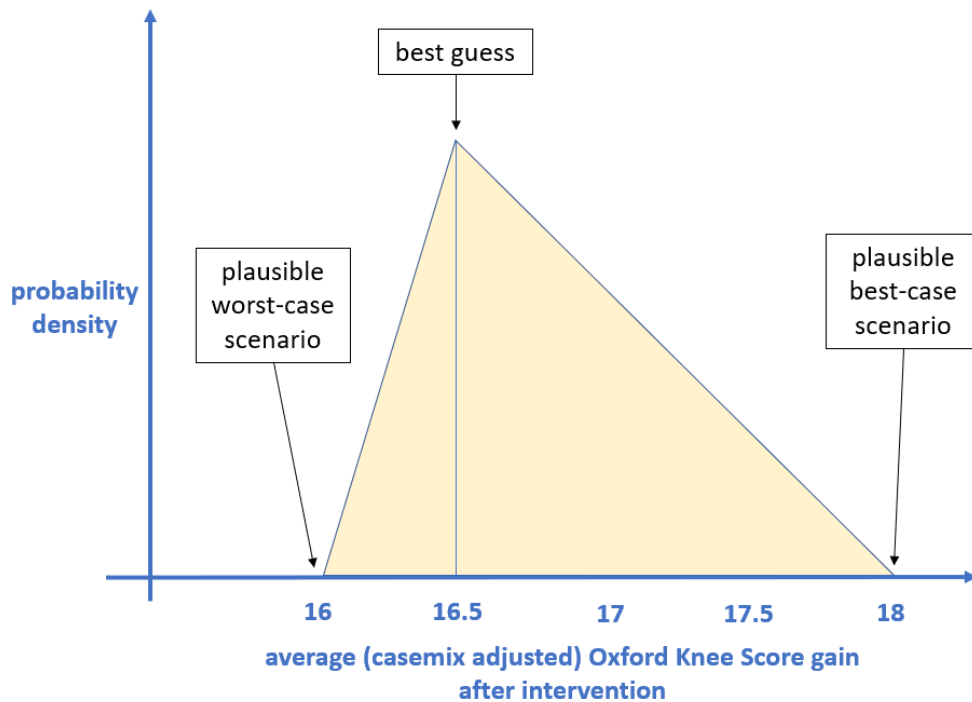
We will illustrate this decision-making process with an example. Let's assume that a Trust is considering implementing an improvement scheme that it hopes will improve outcomes and increase post-operative Oxford Knee Scores. The Trust has costed the scheme and estimates it will cost £70,000 to implement. It expects to deliver 800 elective knee replacements next year. In previous years, patients at the Trust have gained on average, 16 points on the Oxford Knee Scale as a result following their knee replacement surgery. This is a little below average of all providers, which was 17 points last year. The Trust operates under an outcomes-based contract that includes a bonus payment of £200 per patient per point of a provider's average, casemix adjusted Oxford Knee Score gain, above 12 points.

The improvement scheme is novel and so the Trust is unsure whether and by how much the scheme might improve post-operative Oxford Knee Scores.¹⁶ A group of knee surgeons at the Trust are asked to estimate the impact of the scheme by offering three values: (1) their best guess, (2) a plausible worst-case scenario, and (3) a plausible best-case scenario. They agree that the plausible worst-case scenario is that the scheme will generate no improvement, so the Trust average, casemix-adjusted Oxford Knee Score gain would remain at 16 points. The surgeons believe that plausible best-case scenario is that the average, casemix-adjusted Oxford Knee Score gain will improve by 2 points, from 16 to 18 points. This would put the Trust in the upper-quartile by performance. But the surgeons believe that the most likely scenario is that the Scheme will result in a 0.5-point improvement, moving the trust from 16 to 16.5 points, close the national average.

¹⁶ If the intervention has been tried elsewhere, then there may exist, some published literature, estimating its impact on knee function. This evidence can inform the Trust's decision, but some judgement will always be required, since the Trust is unlikely to be identical to the setting where it was trailed and evaluated.

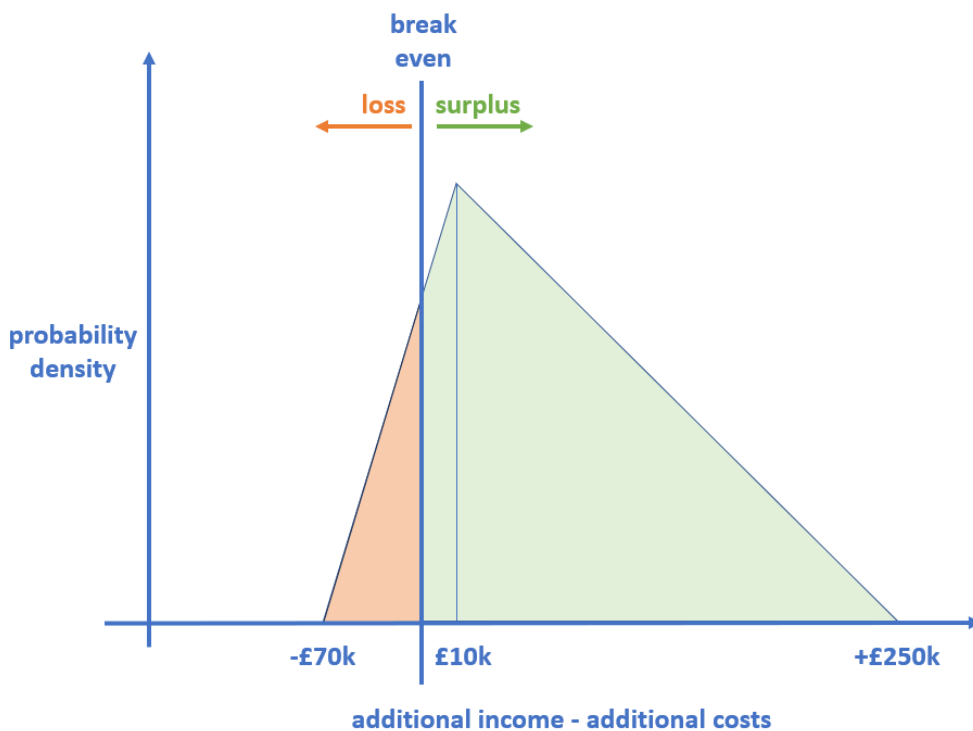
We can think of the surgeon's estimates as a way of describing the probability of the improvement in Oxford-Knee-Scores gains, when the intervention is implemented.

Figure 8: Probability distribution based on Surgeon's estimates of the likely impact of the improvement scheme



If the plausible worst-case scenario comes to pass, the Trust will see no return on its £70,000 investment. If the plausible best-case scenario was realised, then the Trust's outcome-based bonus would increase by £320,000 [= 200 (bonus payment) * 800 (patients) * 2 (improvement in OKS health gain)]. This would net off against the cost of the intervention, leaving the provider £250,000 better off. If the surgeons' best guess proves to be correct, then the Trust's outcome-based bonus would increase by £80,000. Having deducted the cost of the intervention, that would leave the trust with an additional £10,000.

Figure 9: Probability distribution of the cost consequences of the scheme



With a bit of (fairly simple) statistics, it's possible to derive some additional information. For example, we can estimate there is a 63% chance that the intervention will leave the Trust better off. And the expected return is £63.3k.

This analysis cannot remove the uncertainty in the Trusts decision-making process, but it can inform it.

In this example we have assumed that the degree of improvement in post-operative knee scores is the only source of uncertainty affecting the Trust's decision. In reality, the Trust might also be unsure about volume of knee replacement that they will carry out next year, or the cost of the intervention. These uncertainties can be captured in a similar way to inform the Trusts decisions about whether to approve the proposed improvement scheme.

10. Looking forward

To date there has been a gap between the rhetoric and the reality of outcomes-based contracts in the NHS. The work of Professor Siciliani and his colleagues at the University of York, shows how contract theory might help the NHS bridge the gap and design a credible outcomes-based contract for elective hip and knee replacement surgery. If implemented, this could become a test-case, providing an opportunity for the NHS to refine its approach before wider adoption is considered.

Some of the challenges inherent in outcomes-based contracting for hip and knee-replacement surgery, will only become clear through real-world testing. But others can be surmised and mitigated in advance, drawing on lessons learnt elsewhere from attempts to increase quality through contractual mechanisms. We set some of these out below along with potential solutions.

Telling the story of outcomes-based contracts

The story of outcomes-based contracts can be told in one of two ways. From one perspective, outcomes-based contracts are designed to enable providers to seek quality improvements for their patients. It assumes that providers want service improvement and have credible ideas about how this might be achieved but are held back by the cost-consequence under the current contractual model. Improvement schemes cost money, and providers should be recompensed. In this version of the story providers, commissioners and patients are partners seeking the same objective. The contract is simply a mechanism to fairly redistribute funds to allow these objectives to be achieved.

The second version of the story focuses on rewards for outcomes, and penalties for (quality) failure. This story pits providers and commissioners against each other, with the providers trying to maximise their income and the commissioners' squeezing providers for financial value.

How we choose to tell the story has consequences, creating an environment within which the policy is adopted. Both versions of the story are compatible with the concept of outcomes-based contracts, but the first is more likely to lead to positive working relationships between providers and commissioners. It is also more in keeping with national policy and its emphasis on collaboration over competition.

Supporting commissioners to calibrate outcomes-based contracts

As we saw in chapter 8, calibrating outcomes-based contracts is not straightforward. Different choices can create very different incentives and induce different behaviours. Competent analysts will be required to guide and support commissioners. The NHS should continue to support empirical research estimating the incremental cost of increasing health gains and

facilitate the elicitation of expert opinion that commissioners can draw on to appropriately design outcomes-based contracts.

Supporting providers to make rational investment decisions

Outcomes-based contracts will only lead to health outcome improvements if providers design and implement improvement schemes. Deciding which schemes to invest in can be challenging, but as we have seen in chapter 7, there are approaches that can be used to ensure these decisions are taken rationally based on a provider's expectation of costs and outcomes. Before rolling out outcomes-based contracts, training and associated tools should be made available to providers to maximise the chance that they will act rationally (in the narrow economic sense) in response to the incentives created by the contract.

Managing the transition

As with any change in contracting arrangement, the move to an outcomes-based contract will lead to some cost redistribution between commissioners and providers. Sudden changes in expenditure and income can create instability, and so coordinated transitional arrangements at a regional and national level will be needed to moderate these effects in the short term.

The impact of incentives on data quality

If outcomes data are used to determine provider payment levels, then we might expect an improvement in collection rates and the rules that govern the collection arrangements to be subject to greater scrutiny. In the medium-term this will lead to improvements in the quality, consistency and completeness of this data. These changes might also lead to some increase in commissioner costs. This should be seen as a secondary benefit of outcomes-based contracts, rather than as a cost associated with providers 'gaming' the collection system.

Curating patient outcome data

Outcomes based contracts for elective hip and knees replacement are now viable, only because patient reported outcome measures are routinely recorded before and after surgery. The wider adoption of outcomes-based contracts relies on the collection of appropriately designed outcomes data for other interventions and forms of care. The responsibility for developing and publishing outcomes appears to have shifted in the recent years, from NHS Digital to the National Clinical Audits overseen by the Healthcare Quality Improvement Partnership. A

summary of the outcome measures collected via these audits and their relevance and readiness for use in outcomes-based contracts, would help guide plans for a wider roll-out.

Making these outcome scores available to patients, via existing electronic portals, would fit well with other initiatives to empower patients, such as early supported discharge, remote monitoring, and patient-initiated follow-up.

Appendix – Technical description of the scheme

Recall that the payment for each surgery is $t+b$, where t is the base-tariff. The bonus payment that we consider is such that it rewards financially each additional point of improvement in the Oxford Knee Score scale above a minimum threshold. Define H is the (average risk-adjusted) health gain of the patients treated by the provider, and \tilde{H} as a minimum health gain threshold above which payment is received, and p as the payment for unit of health improvement. Then, the bonus payment is

$$b = p(H - \tilde{H})$$

As an example, suppose that average health gain in a given hospital is 15 points on the Oxford Knee Score scale, and that the minimum threshold is 11 points, while the payment for each additional point is £200. Then the bonus payment for each patient is £800, as $b = p(H - \tilde{H}) = £200(15 - 11) = £800$.

How should the commissioner set the payment per unit of improvement?

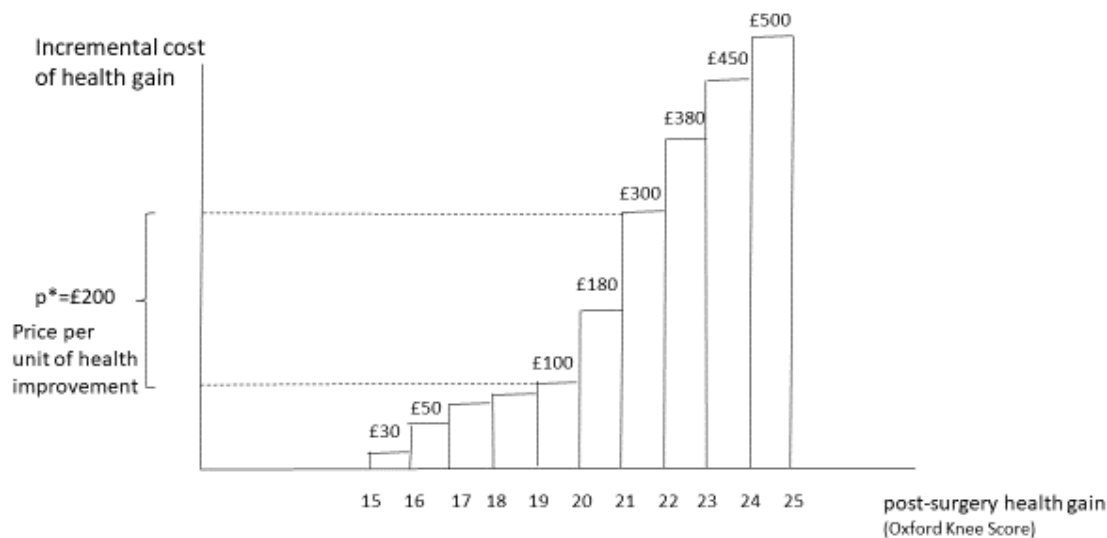
In the first step, the commissioner has to identify the target average health gain that is trying to incentivise with the introduction of the bonus. Let us refer to this target level of health gain as \hat{H} . If the current average health gain by the provider is H_0 , the commissioner would like to design a scheme that induces the provider to improve from H_0 to the target average health gain equal to \hat{H} . For example, if the provider is starting an average health gain equal to 19 and the target is 22, then the incentive payment should be sufficiently strong to induce the provider to improve the average health gain by three points on the Oxford Knee Score scale.

But how can the commissioner calculate the payment per unit of average health gain improvement that will induce the provider to improve from H_0 to \hat{H} ? Contract theory suggests that this payment should be set as the difference between the incremental cost (of one unit of health improvement) measured at the level of target health \hat{H} and the incremental cost (of one unit of health improvement) measured at the level of health pre-policy intervention H_0 . If we define $c(h)$ as the cost of improving health gain by one point then the optimally-designed payment per unit of improvement, denoted with p^* is equal to (Siciliani et al. 2021):

$$p^* = c(\hat{H}) - c(H_0)$$

Figure 1 below illustrates. Suppose that before the policy intervention the provider generates an (average) health gain of 19 points using the Oxford Knee Score. The diagram illustrates the cost of increasing the Oxford Knee Score from 15 to 16 points, from 16 to 17, 18 to 19 and so forth up to a maximum of 25 points. A key plausible assumption is that it is increasingly difficult for the provider to keep increasing the health gain. For example, the figure suggests that the cost of increasing the health gain from 15 to 16 points is £30, the cost from 16 to 17 points costs is £50, and so on, and from 24 to 25 points (the highest achievable health gain) is £500.

Figure 1. Provider incremental cost of increasing patient health gain



Consider the example where the current average health gain is 20 points, and the commissioner target is 22 points on the Oxford Knee Score scale. Then, the payment per unit of improvement should be $p^* = c(22) - c(20) = £300 - £100 = £200$, namely the difference between the cost of improving the health gain from 21 to 22 points and the cost of improving the health gain from 19 to 20 points.

For a less ambitious target of 21 points, the payment per unit of improvement should be $p^* = c(21) - c(20) = £180 - £100 = £80$, while for a more ambitious target of 23 points, the payment per unit of improvement should be $p^* = c(23) - c(20) = £380 - £100 = £280$. The intuition behind the bonus design is that the funder has to compensate the provider for the incremental

cost at the target level of health which is above the incremental cost at the current (pre-bonus) level of health.

Adjusting the base tariff

Once the incentive bonus b has been determined, the commissioner can re-compute the base HRG tariff to ensure the provider covers the costs. This approach is in line with several Best Practice Tariffs that introduce bonus schemes to incentivise process measures of quality while simultaneously reducing the base tariff (examples include BPTs for stroke or hip fracture).

Suppose that the average cost of treating a knee replacement with an average health gain of 22 points on the OKS scale is £6,000, and that the incentive scheme successfully increases the average health gain to 22 points. Suppose also that the minimum health gain threshold above which payment is received is 17 points, so that the bonus payment is $b = p(H - \tilde{H}) = £200(22 - 17) = £1,000$. Then the base HRG tariff can be computed as the difference between the average cost and the bonus payment with $t = £6,000 - £1,000 = £5,000$. The sum of the bonus payment and the base tariff cover provider costs.

The
Strategy
Unit.

The Strategy Unit

Tel: 0121 612 1538

Email: strategy.unit@nhs.net

Web: www.strategyunitwm.nhs.uk

Twitter: @strategy_unit



Midlands and Lancashire
Commissioning Support Unit