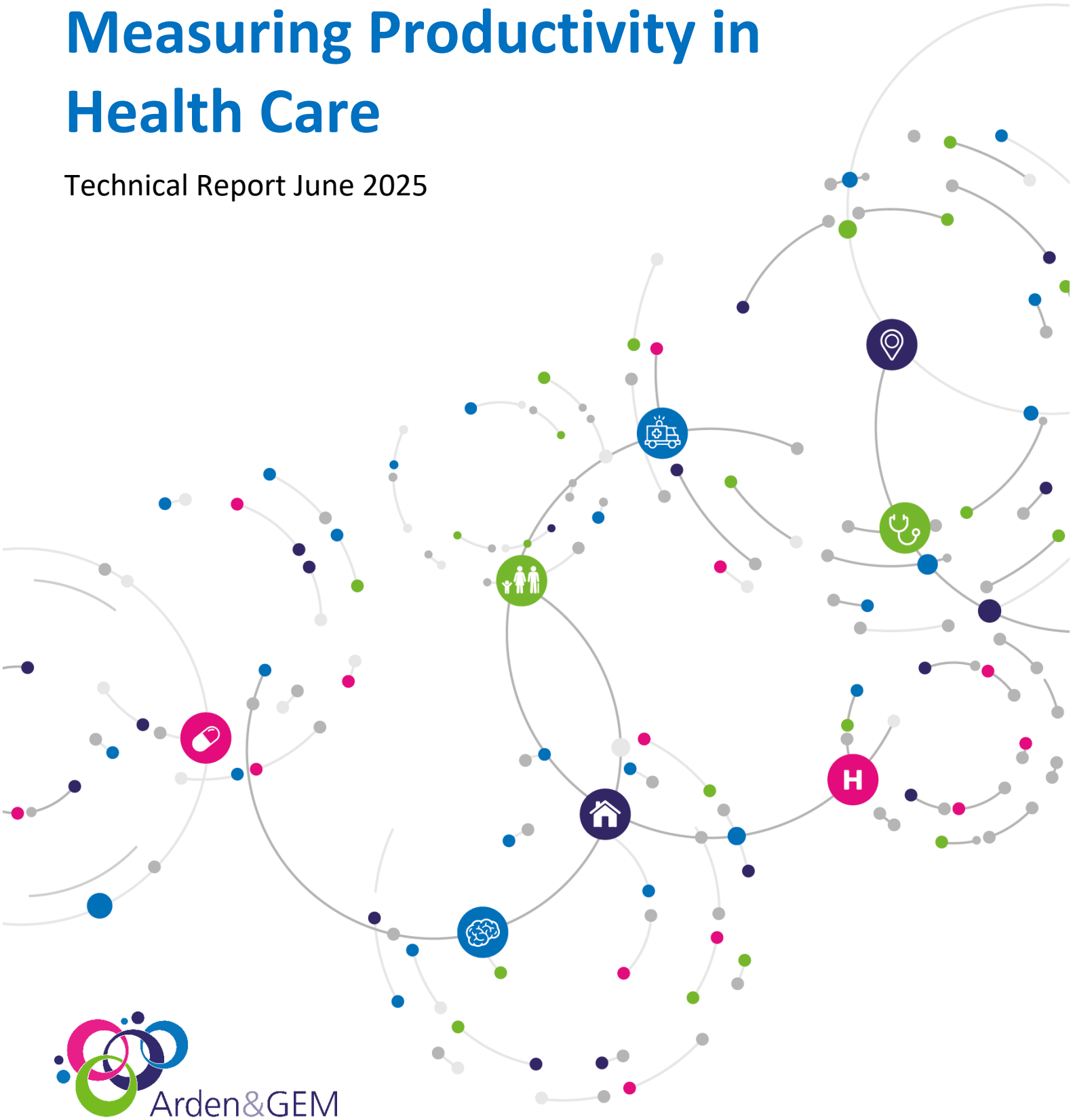
Measuring Productivity in Health Care

Technical Report June 2025



Arden&GEM
Health and social care systems support

Contents

Acknowledgments

This research project to explore productivity measurement in the NHS is supported by **the Health Foundation**, an independent charitable organisation working to build a healthier UK.

NHS Arden & GEM Commissioning Support Unit (CSU) have been commissioned by the Health Foundation to conduct the first phase of the programme, exploring approaches to productivity measurement. Arden & GEM is an NHS organisation that provides a wide range of support services to health and care systems including expertise in areas such as business intelligence, digital transformation, clinical support, procurement, and service redesign. Their services are designed to assist Integrated Care Boards (ICBs), Integrated Care Systems (ICSs), NHS England, local authorities, and various care providers in delivering efficient and effective health care services.

This report was written by project leads **Sheena Dungey** (Senior Health Economist, technical lead) and **Rebecca Gossage-Worrall** (Senior Improvement Consultant, qualitative research lead).
The following AGCSU staff contributed to this evaluation including writing relevant sections of the report: **Mui Wan** (Associate Director; SRO), **Krys Garside** (Project Management), **Heather Humphreys** (Senior Health Economist).

We thank **Katie Fozzard,** senior economist at the Health Foundation, and **Dr Adriana Castelli**, health economist and internationally recognised expert in the field of health care productivity measurement, for the valuable input and feedback throughout the project.

Through the course of the project we have engaged a number of stakeholders, in interviews and/or workshops, our thanks for their time and insights to: Alastair Brodlie (Department of Health and Social Care), Dr Rahul Chodhari (The Royal Free London Foundation Trust), James Cooper (Office for National Statistics), Karl Claxton (University of York), Stephen Gallagher (Mid and South Essex Integrated Care Board), Bruce Hollingsworth (Lancaster University), Rishub Keelara (Organisation for Economic Co-operation and Development), James Lewis (Office for National Statistics), Ben Roberts (Bradford Teaching Hospital NHS Foundation Trust), Lee Rowlands (Manchester University NHS Foundation Trust), Robert Scott (NHS England), Neve Patel (NHS Bedfordshire, Luton and Milton Keynes ICB), Marcus Pratt (NHS Nottingham & Nottinghamshire ICB), Dr. Johannes Wolff, (Guy's and St Thomas' NHS Foundation Trust).

Executive summary

NHS productivity is always of critical importance because it determines the level of care that can be delivered with a given level of staff and other resources. The recent shock to NHS productivity as a result of COVID-19, as well as anticipated future challenges faced by the NHS in meeting demand have led the Health Foundation to make productivity a research priority. Understanding productivity, how it has changed over recent years and what drives it, is critical for the management and sustainability of national health care services. But how do we and how should we measure productivity in such a complex system as the National Health Service?

This report presents findings from a recent commission by The Health Foundation to explore two questions relating to productivity measurement in the NHS:

- What are the strengths and weaknesses of current approaches to measuring productivity, what types of questions can those metrics be used to address, and how they might be improved?
- What are the priority areas for development in measuring productivity, to fill the gaps identified in this research?

Over a seven-month period, NHS Arden & GEM Commissioning Support Unit conducted a comprehensive review to understand and appraise current productivity measurement approaches within the NHS. Central to the project has been the creation of a framework categorising productivity metrics by intended audience and purpose for measurement. We've also considered key drivers of productivity change, including staff experience, joint working across sectors including social care, and the impact of technology, and then extent to which these can be measured.

Here we report our assessment of the extent to which existing productivity measures are comprehensive and provide meaningful insights. This includes how well measures address challenges, such as incorporating patient outcomes, quality, and considering whole pathways of care. It also identifies opportunities to enhance these measures using available data.

Insights have been informed by a rapid literature review and a number of stakeholder discussions including professionals from NHS economics, finance and policy roles at national, system and hospital levels, as well as academia, government departments and external associations.

Key points

No single measure can capture everything that is important about productivity. We identified a diverse array of metrics currently used to measure NHS productivity.

To support navigating this complex policy and analytical space, a framework is presented to classify metrics by their intended audience and purpose for measurement.

The metrics were grouped into six methodological areas:

1. Growth accounting and multi-factor productivity, comparing the system inputs and outputs over time.
2. Workforce productivity, comparing outputs with the number of staff.
3. Clinical and operational measures, which cover a range of metrics often used by health care providers to improve and track performance.
4. Frontier and other scoring approaches that compare resource use across providers.
5. Methods for evaluating impact of interventions of productivity.
6. Methods for measuring health gain from system resource allocation.

A gap analysis highlights the strengths and weaknesses of different metric types.

Common challenges across various audiences were identified and opportunities for improvement discussed, considering these alongside existing development work in this field such as the priorities highlighted in the recent Office for National Statistics Public Sector Productivity Review.

1. Introduction

1.1 Project introduction and objectives

NHS productivity determines the level of care that can be delivered given the size of the workforce and availability of other resources. In recent years, despite increased spending on the NHS, measures of productivity indicate that the amount of care delivered has not kept pace with funding and workforce growth[1], particularly in acute care. Over the longer term, levels of ill health in England are projected to rise, creating additional pressure on NHS services. With low economic growth and constraining tax revenues, improving productivity is vital to ensure the sustainability of the NHS[2].

Looking forward, the Chancellor's 2024 Autumn Budget announced further increases to the budgets of the Department of Health and Social Care (DHSC) and the NHS. It also set out an ambitious productivity improvement target of 2% per year and to support this, an emphasis on improving the way health and care productivity is measured. This target was reaffirmed in the recent (2025) spending review and whilst a welcome budget increase was announced for DHSC, this remains below estimates of that required for restoring services following the covid pandemic, and the target remains a challenging one.

Measuring productivity is critical to decision makers in planning and designing health care services, and to national accountability around public funds. However, whilst productivity can be defined simply as the ratio of outputs to inputs, with such a large, complex system, there are some fundamental challenges to meaningfully measuring how productive the NHS is and why and how this can be impacted. As many have sought to tackle these challenges, there is a wealth of different approaches to measurement including opportunities to learn from other health care systems and other sectors. In light of this, the Health Foundation has commissioned an appraisal of different productivity metrics setting out to explore two questions:

- What are the strengths and weaknesses of current approaches to measuring productivity, what types of questions those metrics can be used to address, and how they might be improved?
- What are the priority areas for development in measuring productivity, to fill the gaps identified in this research?

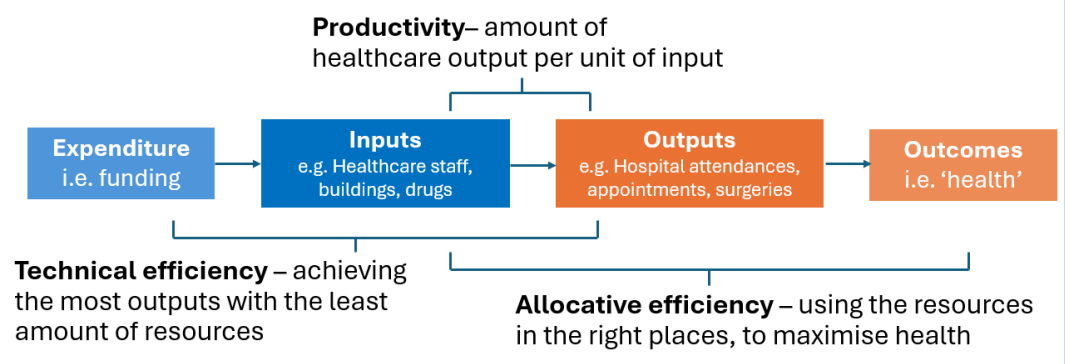
[1] The key ingredients for a sustainable NHS: long-term investment and productivity growth - The Health Foundation

[2] Health in 2040: Projected patterns of illness in England

1.2 A brief background to productivity measurement

A primer on productivity

Broadly, when we talk about productivity, we are referring to 'how much we get out, compared to what we put in' (Coyle, 2023). In the health care example below, this usually means comparing the outputs of the health system, such as the number of GP appointments or surgeries carried out, to the inputs, such as the staff or equipment needed to carry them out.



Economists also often refer to **technical efficiency** and **allocative efficiency**.

Technical efficiency is achieving the maximum amount of output (such as GP appointments or surgeries performed) using the smallest amount of inputs (such as staff and the amount and quality of the facilities and equipment the NHS has).

Allocative efficiency considers how the system is set up and how services are distributed to achieve the greatest benefit to society (or the most 'health). In this way, allocative efficiency is concerned not just with whether existing services and processes are optimised, but whether these are the right services and distributed in the most beneficial way.

Figure 1: adapted from 'Improving public sector efficiency to deliver a smarter state', 2016

1.2.1 Atkinson Review

In 2005, it was recognised that the outputs of the public sector could be better reflected in the national accounts, leading to the Atkinson Review. Atkinson set out a framework to direct and strengthen the measurement of public sector productivity[3], recognising the challenges of measuring these in sectors such as health care, where there is not a traditional 'market' or a set of prices.

The Atkinson principles (set out in full in appendix 1) encourage the use of methods that:

- **Measure the value** of health care outputs such as appointments or surgeries.
- **Reflect the quality of outputs** and their contribution to outcomes.
- Ensure **comprehensive coverage of inputs and outputs** (of services and also in terms of geography)
- Reflect **changes in both prices and service quality** over time.

[3] [ARCHIVED CONTENT] UK Government Web Archive - The National Archives

- Recognise the limitations of measures by **understanding margins of error** and also the intrinsic **limitations of single measures to capture everything that is meaningful**. The framework recognises the need to **triangulate different estimates** and metrics to be able to meaningfully analyse productivity in a complex setting such as health.

Since 2005 the Centre for Health Economics (CHE), University of York, and Office for National Statistics (ONS) have established national measures of productivity, seeking to operationalise the Atkinson principles. These whole-system measures are produced annually, enabling comparisons over time and addressing areas such as how to reflect the differences in complexity of outputs and changes in the quality of care over time, including waiting times and short-term survival rates.

1.2.2 Measurement within the NHS

Within the NHS, in 2015 the RightCare programme was relaunched, focused on value-based commissioning and reducing unwarranted variation, and scaled nationally. It was aligned with the Five Year Forward View and focused on helping Clinical Commissioning Groups (CCGs) reduce unwarranted variation in care and optimising allocative efficiency over system-wide population health pathways (rather than simply increasing activity or reducing cost). Whilst RightCare no longer operates as a named programme, functionally it was absorbed into Integrated Care Systems (ICS) population health approaches and its principles and tools persist.

Whilst RightCare was focussed on the commissioning setting, the (2016) Carter Review[4] (and aligned Getting It Right First-Time programme) focused on identifying, through different metrics, approaches to identifying opportunities to improve hospital productivity. Recent developments have applied approaches to ICS level and leveraged digital technologies and data-driven frameworks, such as the NHS Model Hospital, to enhance benchmarking, real-time monitoring, and efficiency tracking. These initiatives reflect a shift towards embedding quality adjustments, ensuring that productivity measures capture not only efficiency but also improvements in patient outcomes and the value of care.

1.2.3 Recent developments and the ONS review of public sector productivity

More recently, there has been a renewed focus on productivity and its measurement throughout the public sector. Within the NHS, productivity measurement has been a focal point of national planning[5] and the recent Darzi review[6]. More widely, in 2023, the Chancellor of the Exchequer commissioned the ONS to review public sector productivity (PSP)[7], which concluded in March 2025, referencing the changing data and policy landscape and the role of automation and innovation to transform public sector delivery.

The PSP review reiterated the importance of the Atkinson principles and has set out a number of recommendations around appropriate valuing of outputs and quality adjustment. For the health care setting, it recognises ongoing challenges around data quality and coverage, the way that we weigh inputs and outputs according to their value and the particularly thorny challenge of how to incorporate the impact of preventative care into existing measures. It also discusses the importance

[4] Productivity in NHS hospitals - GOV.UK
[5] NHS England » NHS operational planning and contracting guidance
[6] Independent investigation of the NHS in England - GOV.UK
[7] National Statistician's Independent Review of the Measurement of Public Services Productivity – UK Statistics Authority

of being able to drill down into aggregate measures to be able to understand individual elements of such a complex health care service.

This report seeks to review and categorise current health care productivity metrics, including those introduced above and particularly since the Carter Review and the expansion of measurement across different settings, which may be less well established in the literature. We also seek to identify, across the existing approaches, strengths and limitations of metrics and potential areas for development.

1.3 How we approached the project

1.3.1 Rapid literature review

Our objective was to compile a comprehensive list of measures and methods used to evaluate NHS productivity, understanding premise for use and assessing strengths and weaknesses. To inform this, a rapid review of literature was undertaken. Key search terms were developed using synonyms for productivity. Two databases were utilised, PubMed®[8] and Google Scholar, to identify academic and formally published evidence and informal and other grey literature including policy and expert opinion pieces. Details of search terms are included in appendix 2.

Literature prior to 2016 was excluded from the database searches, aligning with the Carter review[4] which marked a shift in focus in NHS productivity measurement to a more operational remit, less established in literature and resources than the macroeconomic models discussed in section 1.2. Key resources providing comprehensive reviews of literature compiled prior to this date were also reviewed (for example the ONS[9] and OECD[10] productivity manuals).

The rapid review also looked at a wide range of grey literature including topics or specific sources recommended by stakeholders during stakeholder interviews (see below). Additionally, we undertook targeted searches on methods used in other health systems and other sectors.

Titles and/or abstracts were screened and articles excluded if they did not describe an approach/method to measuring productivity in the NHS or the approach was very nuanced to a particular health condition or health related process. Key information was extracted from the literature to support this by reviewing abstracts and/or full articles (as detailed in the analysis section) relating to use, strengths and limitations, operational details and feasibility as well as aptitude to deliver insights.

1.3.2 Stakeholder engagement

In addition to the review of the literature, stakeholder interviews were conducted to further understand priorities and challenges for different audiences relating to experiences of productivity measurement. Stakeholders were engaged from settings as outlined in table 1 aiming to encompass a broad perspective.

[8] https://pubmed.ncbi.nlm.nih.gov/
[9] ONS Productivity Manual (last revised 2016): Productivity Handbook - Office for National Statistics
[10] Measuring Productivity - OECD Manual | OECD

Table 1. Stakeholder groups

Stakeholder type	Role (no. interviewed)	Purpose
Technical, health related e.g. finance, economics, analytics, academia	• Academic – health economics (4) • NHS - economics and finance (2) • External - health economics (3)	• Views on and experience of current method/ metric, strengths, limitations, gaps (usability, utility). • Identification of alternate methods/ metrics (and consideration of any non-health approaches) to address current issues or gaps and solutions to future challenges. • Applicability to health of non-health productivity measurement.
Non-health	• Economics and finance (2)	• Views on and experience of methods/ metrics used in other sectors, strengths, limitations, applicability to health of non-health productivity measurement.
Users of outputs of productivity measurement, health related	• NHS national Policy (1) • NHS Commissioner (2) • NHS provider contracting and finance (3) • NHS national Clinical (1)	• To explore utility of current method/metrics, strengths, limitations and gaps. • To explore utility and usability of proposed alternative methods/metrics.

We interviewed 18 stakeholders (either individually or in groups) using a topic guide structured around the topics used for the literature data extraction. We tailored questions to each stakeholder type and their organisation's role in measuring productivity, drivers of productivity or use of the outputs of measurement.

Stakeholders were also invited to take part in one of two workshops, for which the objective was to validate themes identified from the literature review and interviews around areas for developments in productivity measurement. Each workshop had different attendance from the stakeholder groups set out in table 1 (including additional stakeholders who were not also interviewed).

1.3.3 Analysis

To compile a comprehensive list of measures and methods used to measure NHS productivity, understanding what each is commonly used for and assessing strengths and weaknesses, we sought to consider the perspectives of different users and audiences and align where possible to key policy initiatives (such as the operational planning guidance, Long Term Workforce Plan and Plan for digital health and social care).

Data from literature, interviews and group discussions were extracted, tabulated and coded to the following areas:

1. What do we mean by productivity?
2. Why do we measure productivity?
3. How do we measure productivity? Identifying broad method types and analytical variations within a method area (making the measure distinct or novel) including operational details such as reporting mechanism and data sets used.

4. What are the strengths, limitations and aptitude to deliver insights including constraints and development opportunities associated with reporting mechanism and data sets used.

Analysis of themes relating to these areas, informed development of a categorisation of current measures based on use and method type.

A gap analysis was also conducted driven by stakeholder perceptions on shortcomings of existing metrics and promising areas for future development. These were then validated by cross-referencing with the literature to understand any existing innovations within potential gap areas. Stakeholder themes are presented in the next section. The review of different metrics, grouped by method type, is presented in section 3, and their strengths and weaknesses summarised relating these to the gaps analysis and the potential for bridging gap areas.

2. Exploring themes around productivity measurement

2.1 What is health care productivity?

"One of the challenges is about ensuring that there is consistency of use across all of the NHS, I don't think we're there yet. ... different people's backgrounds are associated with different conceptions of productivity." Stakeholder 33

Productivity is typically understood as the ratio of outputs to the inputs used to produce them. In practice, and decidedly in the health care setting, it's a term that means different things to different audiences and fills a busy space in the fields of analysis and economics. Before discussing different ways of measuring productivity, we asked stakeholders what productivity means within their role. Whilst there are many different approaches and metrics for measuring productivity in health care, common themes discussed when setting out 'what health care productivity is' include:

- The need for **clinical resonance**, framing productivity in terms of health gains for patients, avoiding negative connotations such as "doing more for less".
- Relatedly, that it should be **outcome orientated** and focussed on health benefits resulting from NHS expenditure, rather than purely reflecting the number of 'activities' carried out in the NHS.
- That it is an **integrated concept** i.e. a driver of action rather than a goal in itself and that it needs to be embedded within a broader framework to guide decision-making (particularly in supporting the workforce).
- Additionally, the distinction between productivity, quality improvement (QI), and benefit realisation is often unclear. The term 'productivity' is also frequently used interchangeably with efficiency.

"Clinicians are interested in patient outcomes, operational leads are interested in the activity numbers. And then finance is interested in the input numbers. But ideally you bring all three of them together and that's the challenge that I have." Stakeholder 17

Here, we consider a broad perspective to defining productivity, exploring the relationship between definition, use and measurement.

2.2 Why do we measure health care productivity?

Measuring productivity in health care is essential for a range of reasons. Some key categorisations are described below to help set out different motivations for measurement.

1. **Macroeconomic level (system or nationwide perspective):** productivity measures are used as a national indicator of the economy and economy of the health sector as well as for international comparisons. In this context, productivity measures provide accountability and transparency for governments and system planners. They also inform budget setting and the pricing system for hospital reimbursement and resource allocation, support planning supply needed to meet demand and national workforce strategy. Furthermore, they provide insight into whether national policies and investment are working.

2. **Microeconomic level (service, process or organisation perspective)**: hospital and commissioner comparisons of productivity at sector or specialty level can help identify areas for improvement and best practice. Productivity measurement is critical to directing policy and underpins impact assessment when developing new policies and trialling new ways of working. Within the health care delivery setting, productivity measurement extends to quality improvement, programme monitoring and service planning.

3. At both macro- and micro- levels, metrics were reported to be in some cases looking at a **technical** perspective, addressing the question "are we producing more with the same resources?". Their purpose is to assess how efficiently a system or units of delivery (such as health care providers) or individual processes convert resources to outputs (for example, patient treatments or health improvements).

4. In some cases, metrics were more **evaluative**, looking at whether we are achieving value with our resources. This may be evaluating whether individual health care interventions provide relative value for money compared to alternatives or whether, across the system, resources are optimally allocated to maximise health outcomes for a given level of spending. Considering the value of different activities informs decision makers to improve **allocative efficiency** (are we using resources in the best possible way?).

As we present the appraisal of different metrics within the identified broad method areas in section 3, these are discussed in terms of these four concepts i.e. the extent to which they are macro or micro level and whether they look at technical or evaluative productivity. The high-level metric types covered are summarised, according to these concepts, in figure 2 below.

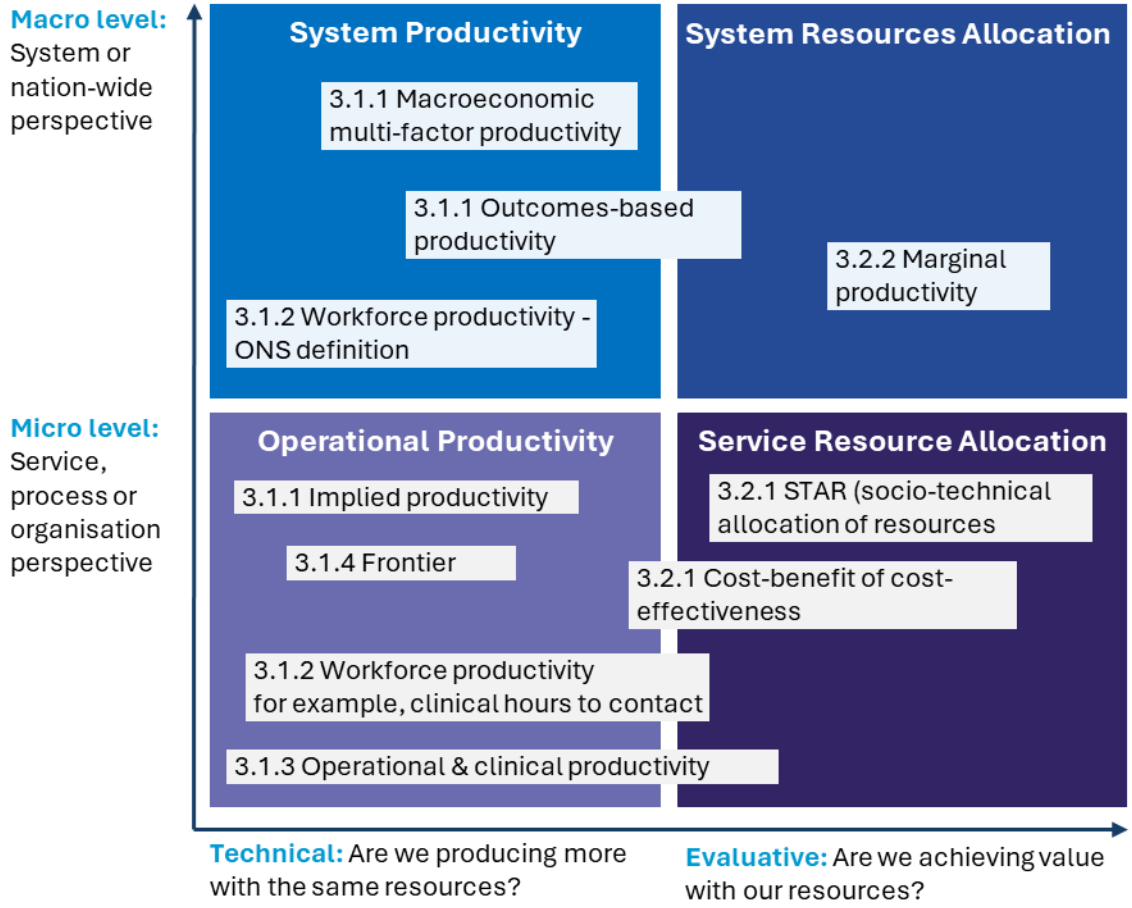


Figure 2: Proposed classification of metrics from rapid review of literature and stakeholder discussions, based on reason for measurement (metrics are denoted with the report section in which they are introduced)

2.3 What are the gaps in current productivity metrics?

Themes were considered relating to stakeholder perceptions on shortcomings of existing metrics and promising areas for future development. Themes were found to be cross-cutting across different stakeholder groups and methodological areas (although some were more relevant to certain uses and approaches). These gap areas have been cross-referenced with the literature to validate and explore any relevant current developments, such as from other countries or settings. Key identified gap areas are presented below and referred to throughout section 3, underpinning discussion of the strengths and weaknesses of different metrics.

"... If it's a new way of doing an admission ... [current] methods will roughly capture that but if we do something completely new we wouldn't have a unit cost, we wouldn't have a data collection at first and there might be a lot going on locally that we're just not picking up." Stakeholder 34

Table 2a: gap areas and focus for development themes, overarching conceptual challenges

Concept	Theme	Stakeholder perceptions	Comments from literature
Coverage	**Comprehensive and integrated data coverage**	It was often discussed that non-acute and (publicly funded) independent sector activity	Review of literature supports the conclusion of limitations beyond

Concept	Theme	Stakeholder perceptions	Comments from literature
		data are poorly represented, limiting system-wide productivity insight. Data coverage was also raised as an issue when discussing how to value outputs or measure their quality (points 2 and 3 below). Social care is also perceived as a gap in coverage, whilst separately funded, social care and NHS services have a dependency on each other viewed as a challenge to improving productivity.	the acute setting and lack of coverage for independent sector. Social care productivity measures are reported in isolation. However, an integrated NHS and social care measure of productivity has not been established. There are promising accounts of the use of linked ICS data sets to offer insights in this area.
Inputs	**Incorporation of input quality and long-term investment**	It was felt that measures should better capture capital quality, workforce capability and future-focused investments, which may otherwise skew short-term productivity interpretation.	Measures of NHS productivity reported in literature do not routinely adjust for longer term investment and input quality, whilst there are some limited examples of the latter, these papers point to lack of data on input effectiveness. Investment as a gap in measurement is increasingly recognised, particularly in conceptual discussions, whilst evaluative techniques do offer a framework for measuring over longer timescales.
Outputs	**Appropriate valuation of health care outputs**	Preventative activities, resilience or those that amount to failure demand are difficult to value, especially within current approaches which value care based on its cost. Often this was raised in the context of working in new ways such as introducing virtual wards or same day emergency care models, with current cost weighting approaches carrying a risk of stifling innovation by underrepresenting value. There is a need to identify an appropriate value weight for health care output which moves away from using unit costs, to a currency that allows us to weigh outputs that are comparable in every clinical	This is echoed in recent developments reported in the literature, particularly work carried out as part of the ONS PSP. Whilst the PSP provides new insights into how preventative activities could be incorporated into the national accounts, it is concluded that this remains a major research challenge which applies to a number of care delivery models.

Concept	Theme	Stakeholder perceptions	Comments from literature
		aspect, including the health outcomes they deliver.	
Quality	**Robust and standardised output quality measures**	Measuring quality has progressed since the Atkinson review, however, many aspects of quality in health care often require subjective or long-term data that can be difficult to quantify and standardise. This was particularly discussed for non-acute care settings.	Whilst there are many instances in the literature around care quality measurement, indicating potential areas for development for both acute and non-acute, there limited instances of these being implemented successfully in the non-acute sectors.

Table 2b: Stakeholder themes, emerging practical challenges

Granularity in input-output attribution	Current methods cannot isolate inputs at specialty or pathway level, hindering understanding of productivity in real service contexts (this was discussed in relation to data development strategies to be able to make more out of current assets).
Fit-for-purpose benchmarking of evolving models of care	It was recognised that whilst there is a nationally set out drive to work differently, integrated and innovative care models challenge conventional productivity metrics and require new approaches to benchmark across varied providers and funding arrangements.

3. Review of metrics currently used to measure productivity in health care

3.1 Technical productivity measures

3.1.1 The growth accounting framework and multi-factor productivity metrics

How it's measured and what it's used for?

Multi-factor productivity (or total factor productivity) measures productivity at the system level. It is defined by the ONS as the growth in the outputs of an industry or economy which cannot be explained by increasing inputs. Thus, multi-factor productivity (in theory) reflects changes in how efficiently inputs are used[9].

Inputs include labour, goods and services and capital consumption (including building costs and reusable equipment) and may be measured directly (measuring volume and price) or indirectly (using total expenditure for a given area). These inputs are combined into an index, to track changes over time.

Atkinson principles determine that the output of health services should be measured reflecting the contribution they make to improved health outcomes[3]. In terms of measurable outputs, the starting point is generally primary and secondary care services (appointments, inpatient stays, outpatient attendances). Unlike industries with revenue-based outputs, public sector outputs generally do not have market prices and hence alternative approaches to valuing outputs and estimating growth are

required. For NHS outputs, the costs of procedures, as collected via the National Cost Collection using different currencies such as attendances, visits and for much of acute care, different procedural Healthcare Resource Groups (HRGs)[11]. These are used as a proxy for prices. The extent to which these can inform improved outcomes is discussed further below.

The ONS[12] and University of York[13] have built on principles from the Atkinson review and the national accounting framework, such as System of National Accounts, to develop national approaches for measuring public sector productivity. These estimates underpin macroeconomic decision making, for example, the HM Treasury (HMT) spending review process[14] for health care budget setting incorporates a productivity goal based on the NHS multi-factor productivity growth trend rate as calculated using methodology developed by the ONS and University of York.

Whilst typically thought of as a macroeconomic concept, NHS England applies a similar framework to calculate Trust level 'implied' productivity[15] using local finance and Secondary Uses Services (SUS) activity data. Differences between the NHS England and ONS and York approaches have recently been highlighted[16]. Furthermore, the ONS have just published a review into public sector productivity (PSP) including health care, revisiting Atkinson principles and exploring how new sources of data could be incorporated to improve measurement. A summary comparing the key MFP metrics and the PSP recommendations by measurement concept, is presented in the pull-out section below.

"I think timeliness is the main thing that goes between advantages and the disadvantages. We've got in year [metrics] which are timely, good for performance management or seeing what's going on right now. But the big disadvantage is they're not taking into account quality at all. They're not taking into account some of the non-acute side of things ... whereas York Uni, ONS, they're much wider scope. But a lot longer timescales... they're very useful for long run trends, understanding the wider economy, how efficient the NHS is and for funding discussions." Stakeholder 34.

Strengths and weakness

Current approaches to multi-factor productivity have several strengths. They draw on comprehensive and systematic NHS activity data, especially in acute services, with more recent developments for community services data, offering high service coverage and a robust foundation for national-level analysis. There have also been promising efforts to adjust outputs for quality[17],[18]. and value such as using cost-weighted measures and exploring links to health outcomes and Quality Adjusted Life Years (QALYs). For example, as part of the ONS public services productivity review, Professor Martin Weale was commissioned to look at novel approaches to valuing national

[11] HRG grouping - NHS England Digital
[12] Public service productivity, healthcare, England - Office for National Statistics
[13] CHE RP Cover
[14] Options for the 2024 Spending Review and beyond | Institute for Fiscal Studies
[15] NHS England » NHS productivity
[16] NHS England » Comparison between the recently published ONS quarterly public service productivity statistics and NHS England productivity statistics
[17] Dawson, D., Gravelle, H., O'Mahony, M., et al. (2005). "Developing New Approaches to Measuring NHS Outputs and Productivity." *Centre for Health Economics (CHE), University of York.*
[18] Castelli, A., Laudicella, M., & Street, A. (2008). "Measuring NHS Output Growth: A Time Series Analysis." *Health Economics, 17(7), 779–800.*

preventative care programme outputs[19] assessing the gains in QALYs due to reduced disease incidence. This is a shift from traditional approaches, where medical treatments typically focus on direct outputs like hospital visits and prevention activities are valued based on their short-term cost, to a perspective that also values long-term preventative health benefits, beyond their short-term costs. This work has led to the recommendation set out as part of the PSP review[7]*:*

"For pre-selected preventative services where high quality data on impact of downstream services can be found, the probability weighted cost of these downstream services can be used as a proxy valuation of the preventative services in the cost weighting methodology".

These developments align with wider shifts toward valuing preventative care and innovations like digital health tools.

Weaknesses:

- **Data coverage:** Non-acute services such as primary, community, and mental health care, whilst included in measures of TFP, suffer from fragmented data and limited nationally agreed currencies, as such they are not represented[20],[21] as well as acute services.
- **How to value outputs:** Outputs are often valued based on cost and complexity, potentially undervaluing lower-cost interventions.
- **How to value input and reflect long term investment**: Investment in infrastructure, prevention, workforce development, or digital systems may reduce apparent productivity in the short term, despite long-term benefits[22].
- **Attributing inputs and outputs at service level:** national metrics cannot currently isolate inputs contributing to outputs at the service or specialty level, reducing their relevance for local decision-making and performance improvement.

"Putting the [MFP] productivity calculation in front of people [at Trusts] raises people's backs. They don't understand it and they don't know what to do with it. So what we need to do is unpack it a lot more and link it to the stuff that they care about." Stakeholder 17

These limitations highlight the need for more nuanced, inclusive and forward-looking approaches to productivity measurement in health care.

Case studies: Looking outside of the NHS: approaches to measuring value

CS1. International case studies for outcomes-based productivity measures

In a 2024 commission looking at national productivity measurement, the Australian government reviewed their traditional approach to measuring MFP, macro-level productivity and introduced an updated measurement incorporating changes in quality of life as related to health care services[23]. The approach mirrors that of U.S economist David Cutler to create a value-based productivity measurement in health care[24], moving beyond simple input-output ratios to assess

[19] Public health measures and the national accounts
[20] Approaches to Measure Efficiency in Primary Care: A Systematic Literature Review
[21] Unravelling-the-mental-health-productivity-puzzle-where-do-we-start
[22] fa3bab05-en.pdf
[23] Advances in measuring healthcare productivity
[24] A Satellite Account for Health in the United States - American Economic Association

health outcomes and economic value. His approach incorporates quality-adjusted life improvements, cost-effectiveness, and long-term health gains.

What are the benefits? A comprehensive measure of quality care that is independent of the specific service setting. It enables consistent comparison over time, even as the quality or nature of care evolves. It also allows different services to be valued based on the outcomes they deliver, rather than the cost of delivery alone. This may encourage the redesign of care around value and outcomes rather than activity.

How applicable is this approach to the NHS? Professor Martin Weale draws some parallels to the Cutler approach in his recent work for the PSP commission. However, there are some limitations to implementing it in the way both Australia and the U.S. have done. For one, the NHS does not so comprehensively link patient health outcomes to spending across care pathways as in both U.S and Australia, these measures are for a limited set of diseases / patient groups so not as comprehensive as ONS / York. There are also differences in funding policy and structures. Cutler's approach relies on willingness-to-pay thresholds which are not fully reflected in NHS measures of productivity. Additionally, whilst the U.S. has a private health care system, Australia's has both public and private components, making productivity analysis more market driven.

CS2. Viewing health care as an investment, comparison with education

Economist Diane Coyle has proposed reframing health care spending not simply as a cost, but as a form of public investment[25], comparable to education. Education is widely recognised as an investment in human capital. Spending on education has been evaluated not just in terms of immediate outputs (e.g. exam results) but through long-term returns[26].

What are the benefits? This approach enables robust public investment over time, supported by frameworks that estimate return on investment (ROI), model lifetime earnings, and factor in societal benefits.

How applicable is this approach to the NHS? In this view, health care contributes directly to the long-term productive capacity of the economy. Investment in services, particularly in public health and preventative care, should be evaluated based on their long-term returns, including improved population health, reduced future treatment costs, and sustained workforce participation. Such an approach could shift policy focus from short-term cost-cutting to long-term value, strengthening the case for upstream and preventative interventions. It also supports arguments for resilience and preparedness. However, estimating ROI for health interventions can be complex, uncertain and difficult to generalise. Current public finance rules and accounting frameworks may not easily accommodate this shift without substantial reforms including to data infrastructure.

[25] Healthcare-as-social-infrastructure.pdf

[26] Corrado, C, M O'Mahony & L Samek. How does education contribute to productivity? An intangible infrastructure approach applied to the UK and the US. ESCoE/IARIW conference paper 2021 https://iariw.org/wpcontent/uploads/2021/08/Education_productivity_paper.pdf

Pullout section: Comparison of key MFP metric publications alongside recommendations from the PSP review by measurement concept

Concept	Office for National Statistics (ONS)	Centre for Health Economics (CHE), University of York	NHS England implied productivity	PSP recommendations I (implemented by the review) N (recommended next steps) R (remaining big issues for research)
Frequency and timeliness of publication	Annual (with quarterly release of 'experimental statistic'). Most recent estimate is for 2022/23).	Annual. The most recent estimate is for 2021/22).	Not currently published to a regular timetable, however, quarterly and annual in-year estimates shared via board (latest for calendar year 2024). Estimates also produced and used internally as part of regular planning cycles.	
Coverage	UK, including hospital and community health services, primary care (inc. dental, ophthalmology), limited coverage non-NHS services commissioned by the NHS. Recent years also include COVID-19 related services.	England, including hospital and community health services, primary care (inc. dental and ophthalmology), limited coverage non-NHS services commissioned by the NHS.	England, published metrics include acute (hospital) settings only, however, developing measures for other settings.	N. Work with devolved governments. N. Data on NHS-funded services contracted from the independent sector.
Inputs	'Direct' measure full-time equivalent (FTE) staff, weighted by salary. Depreciation of assets 'Indirectly' measure other inputs such as equipment and medicines.	'Direct' measure full-time equivalent (FTE) staff, weighted by salary. Depreciation of assets 'Indirectly' measure other inputs such as	Includes staff numbers and associated costs (including bank and agency). Accounts for high-cost drugs.	I. Legal and audit data included. I. Labour inputs weights updated to align with Full-Time Equivalent staff.

Concept	Office for National Statistics (ONS)	Centre for Health Economics (CHE), University of York	NHS England implied productivity	PSP recommendations I (implemented by the review) N (recommended next steps) R (remaining big issues for research)
	Calculated using a Tornquist index and deflators derived from other official ONS statistics.	equipment and medicines. Calculated using a Laspeyres index.		
Outputs	Aggregate units of activity from the National Cost Collection (NCC), weighted by cost. Includes COVID-19 related activity such as vaccinations.	Also uses the National Cost Collection, however, for acute activity York also use Hospital Episode Statistics (HES) patient level data measured using Continuous Inpatient Spells to give a broader activity measure (compared to the Finished Consultant Episodes as per the National Cost collection). Activity weighted by cost. Does not include COVID-19 related activity such as vaccinations.	Secondary Uses Services (SUS) patient level data provides a more timely measure of output than HES or NCC (however, less data quality processing applied) Does not include COVID-19 related activity such as vaccinations.	I. Inclusion of preventative services commissioned to non-NHS providers. I. Incorporated certain screening services. I. Improved measures for primary care. I. Used the same cost-weights where equivalent treatments are provided across different modes of provision. I. Removal of excess bed days activity. N. Quality of National Cost Collection activity data for selected screening services. R. Evaluate the benefits and costs of switching to the available person-level data measuring hospital output. R. Explore how to improve the measurement of preventative services in health care output, including review of potential data sources to ensure consistent application across the range of preventative treatments. R. Investigate whether aspects of the incentives in NHS Payment Scheme can be incorporated in the relative weighting of different services in Health care productivity.

Concept	Office for National Statistics (ONS)	Centre for Health Economics (CHE), University of York	NHS England implied productivity	PSP recommendations I (implemented by the review) N (recommended next steps) R (remaining big issues for research)
Quality adjustment	Adjusts for waiting times, cancellation rates, patient experience and short-term survival rates.	Adjusts for waiting times, short-term survival rates, patient outcomes, life-expectancy, and more recently for avoidable emergency readmissions and hospital acquired infections (Methicillin-resistant Staphylococcus aureus and Clostridioides difficile).	Does not adjust for quality.	N. Development of patient satisfaction surveys. R. Further research to continue to explore the feasibility of improving and expanding the existing health care quality adjustment for waiting times.

3.1.2 Workforce productivity

How it's measured and what it's used for?

Workforce (or labour) productivity metrics may also be constructed according to the growth accounting framework using only workforce as an input. Workforce input may be expressed directly in terms of Full Time Equivalent units[27] grouped by Agenda for Change salary band[28]. Alternatively, workforce input may be measured indirectly from NHS accounts expenditure data on staff, appropriately deflated to ensure changes in prices are adjusted for. There is limited published literature reporting growth accounting type workforce productivity metrics for the NHS. However, some relevant examples include a measure of consultant productivity[29] and metrics reported via the NHS England Model Hospital tool[30].

Considering the perspectives of different stakeholders, workforce productivity is conceptually a much broader term and defined and measured in different ways. Workforce input, for example, can be cost or FTE based (as in the examples above) but also time based and can be further broken down to consider what the time is spent doing? Is it patient-facing or doing administration or training to build up skills?, and is the right person doing the right role? The clinical hours to contact metric (CHtC), for example, measures the ratio of total clinical hours worked to direct interventions with patients for AHP services[31].

In 2023, the NHS Long Term Workforce Plan[32] set out the case for change in relation to how recruitment, retention and reformation of the NHS workforce is managed. It acknowledges the need to grow the workforce to match supply to demand but also to improve productivity by improving the working environment to reduce turnover in the NHS, and to transform working practices.

More generally, workforce productivity is typically discussed in the context of the wider determinants of a productive workforce such as measures relating to wellbeing, experience, skill mix and also indicators of workforce agility to adapt to change[33]. The Global Labor Organization propose a framework emphasising a comprehensive approach to evaluating workforce performance,[29] integrating productivity metrics with additional dimensions such as workforce availability, competence and responsiveness[34]. These approaches aim to move beyond traditional productivity assessments and help organisations and policymakers identify areas for improvement and implement strategies that enhance overall workforce effectiveness.

Additionally, since the pandemic in 2020, the organisational resilience of the NHS and other health care systems, have come under scrutiny and with renewed focus for measurement[35]. Considering motivations for measuring workforce productivity, one important driver is to support optimisation of operational efficiency within a care provider or network of providers. Service managers use digital staff rostering and flexible workforce models, to manage the workforce across provider networks,

[27] Definitions are data elements (including FTE) and provided in the NHS Data Dictionary under National Workforce Data Set
[28] NHS terms and conditions of service (TCS) for Agenda for Change | NHS Employers
[29] www.health.org.uk/publication/year-of-plenty
[30] NHS England » The Model Health System
[31] increasingp-clinical-capacity-of-ahps-using-job-planning-at-ipswich-hospital.pdf
[32] NHS England » NHS Long Term Workforce Plan
[33] Agility: the missing ingredient for NHS productivity | The Health Foundation
[34] Dimensions of health workforce performance: a scoping review
[35] Organizational resilience in healthcare: a review and descriptive narrative synthesis of approaches to resilience measurement and assessment in empirical studies | BMC Health Services Research

monitoring metrics around staff deployment and use of agency staff[36] . These systems reportedly offer opportunities for improving productivity measurement and improvement[37, 38], however, there is more to do to fully realise opportunities and this requires buy in to changing reporting processes.

"So we were doing it as proof of concept and we didn't get past that point because whilst the [Trust] staff [workforce planning and finance], they went 'wow, this is really clever', they were unwilling to really get behind the benefits... We've got away from the practical mathematics of things. To the acceptance, to the ownership of these type of [measurement tools]. So, for me if I'm measuring any level of productivity gain, my observation so far of the NHS is nobody really wants to get behind a number. They don't want to own it." Stakeholder 130

Strengths and weakness

The reforms outlined in the long-term workforce plan infer a need to be able to describe the 'quality' of the workforce including skill mix and level of experience (which may go beyond salary banding). There are examples, particularly looking outside of health, of how determinants of workforce quality can be incorporated into productivity measurement. Some key examples and presented as case studies below. There is additionally a wealth of information relating to workforce captured in the care delivery setting. For example, E-roster, electronic patient records and patient level costing data captured by providers give a detailed picture of workforce, how resources are utilised and deliver care. There are also examples of local data tools and data models which are being used to plan workforce across emerging collaborative models of care.

Weaknesses:

- **How to value input and reflect long term investment**: Whilst there are approaches to incorporating workforce quality aspects into measures, there is a challenge with the long-term nature of investments required to build and sustain quality and resilience.
- **Attributing inputs and outputs at service level:** There are limitations on how service level data can be mapped across local care delivery pathways and aggregated to specialty level.
- **Understanding evolving models of care:** There are reported challenges with uptake and acceptability of new workforce productivity tools.

Incorporating workforce factors which reflect long term capability, sustainability and therefore productivity of the NHS workforce is crucial. Developing metrics will depend on aligning to and promoting within wider NHS data strategy.

Case study: valuing human capital and the Quality Adjusted Labour Input

CS3. Quality Adjusted Labour Input

The ONS produce a Quality Adjusted Labour Input (QALI) index using Labour Force Survey data which captures quarterly earner characteristics, earnings and hours work insights covering a large sample of households. Based on characteristics (such as, qualifications, gender, industry and age), workers are grouped, and marginal productivity of these groups is calculated to produce a

[36] e-rostering-guidance.pdf
[37] General practice workforce model insight tool | NHS England | Workforce, training and education
[38] e-Roster policy: Insights and implications of codifying nurse scheduling - Robert G Drake, 2019

weighted measure of labour input which can be incorporated into growth accounting metrics of productivity to adjust labour input to reflect quality.

What are the benefits? The aim is to give a more accurate measure of labour input, the benefit of an index is that it also accounts for workforce skills changes which affect productivity.

How applicable is this approach to the NHS? While QALI offers a nuanced perspective by accounting for variations in workforce quality, such as education and experience levels, it's not an approach that's currently widely used within the NHS. Understanding how workforce skills-mix varies is highly relevant in the current climate, however, the feasibility of being able to accurately calculate marginal productivity of different groups within the NHS workforce is limited with current available data sets. Arguably there are several critical factors relating to workforce productivity not included in the QALI such as those relating to resilience and it may be that a more holistic measure, such as discussed in CS5, holds more value in this context.

CS4. Valuing human capital based on future productivity

The approach: The World Bank produces a Human Capital Index designed to assess the potential productivity of the next generation of workers based on the education and health outcomes of a country. It was introduced in 2018 as part of the World Bank's Human Capital Project, aiming to highlight the importance of investing in human capital for economic growth and development. It measures how well different countries are fostering human capital based on survival rates, education and health of children populations. The index is scaled from 0 to 1, where 1 means a child born today will achieve their full potential in terms of productivity.

What are the benefits? It helps governments prioritise investments in health and education where the benefits may be realised over much longer timescales than the investment itself.

How applicable is this approach to the NHS? Setting out determinants of future productivity relating to workforce, resilience, other investments would offer valuable insight given both the focus on long-term sustainability of the NHS and aligns with a need to evaluate policies such as those relating to workforce. Dimensions could mirror those currently considered in the framework of workforce productivity such as workforce retention, training and well-being. Any such metric would be dependent on data availability.

3.1.3 Clinical and operational measures

How it's measured and what it's used for?

The concept of operational productivity is set out in the Carter report[39] as productivity within the provider setting, looking at individual resource areas and specialties and using benchmarking to identify unwarranted variation. The scope of operational metrics is vast and given the typical local level application of metrics, there are many iterations, tailored to local productivity challenges. Here we focus on some key metrics, tools and examples of how best practice is promoted in the NHS.

The Model Hospital digital tool[30] was developed in line with Lord Carter's recommendations. It provides trusts with benchmarking data on workforce, clinical services, estates, and procurement,

[39] Productivity in NHS hospitals - GOV.UK

and includes measures of operational and clinical productivity, a term used to describe the relationship between clinical inputs (such as staff time and skills) and outputs (including treatments delivered which may be across a clinical service, and outcomes achieved). A core feature is the use of Weighted Activity Units (WAUs), which standardise activity across different services, enabling comparisons of cost and output between organisations.

The Carter report also launched the clinically led programme Getting It Right First Time (GIRFT)[40], with the aim to reduce unwarranted variations in medical practice and operational processes. The work of GIRFT encompasses data collection and methodological development focussed on measurement linking patient outcomes and hospital efficiency.

Operational productivity initiatives are often paired with quality improvement methods such as Lean and Six Sigma. These approaches, adapted from manufacturing and increasingly applied in health care, aim to improve value by reducing waste, variation, and inefficiency. Lean focuses on streamlining processes and maximising value from the patient's perspective, while Six Sigma uses data-driven methodologies to identify root causes of variation and improve consistency. In the NHS context, these approaches are used to embed a culture of continuous improvement, particularly in areas like elective care pathways, diagnostics, and support services[41].

Strengths and weakness

Initiatives such as GIRFT benefit from clinical leadership and access to patient-level data, which enhances their ability to link operational processes to patient outcomes. The importance of the relationship between NHS productivity and improvement was recently set out by The Health Foundation following the 2024 Darzi review or NHS performance[42], highlighting the need to focus on improvement in care quality to engage clinicians and drive productivity improvement. Tools like the Model Hospital and Digital Maturity Assessment also support peer benchmarking and help identify where systems have the capability to improve productivity, especially as models of care evolve across different providers.

Weaknesses:

- **Quality and output value:** while innovation may improve operational efficiency, national pricing mechanisms don't always keep pace, meaning providers can appear less productive despite delivering care more effectively. Additionally measuring quality is inherently complex when trying to reflect it consistently across diverse services and care pathways[43].
- **Attributing inputs and outputs at service level:** challenges to attributing inputs and outputs at service level constrains the use of productivity metrics for system-wide planning and decision-making across organisational boundaries.
- **Understanding evolving models of care:** current benchmarking methods can struggle to account for local variation, making it difficult to compare providers fairly.

[40] Home - Getting It Right First Time - GIRFT
[41] https://www.england.nhs.uk/improvement-hub/wp-content/uploads/sites/44/2017/11/Quality-Improvement-Theory-and-Practice-in-Healthcare.pdf
[42] How improvement can help NHS productivity - The Health Foundation
[43] The Importance of Surgical Cutting Time as a Key Performance Indicator Alongside Touchtime Utilisation in Operating Theatre Efficiency Optimisation - PubMed

... it's like what are the critical success factors to the most successful productive unit we know? Having a separate cold site, having an independent unit, having access to an EPR, having access to this, it's almost like a metric, isn't it? We have a digital maturity measure, we don't have a productivity maturity measure... Stakeholder 124

As models of care evolve including increased integration across systems, investing in more nuanced, longitudinal, and context-sensitive productivity metrics will be key to ensuring health care is not only efficient, but truly valuable to patients.

Case study: productivity and resilience

CS5. Productivity and resilience: transport case study

In settings such as transport and agriculture, resilience is often assessed using network-based or systems modelling approaches that examine how services respond to and recover from disruption. For example, resilience in transportation networks is defined[44] as the ability to maintain and quickly restore a minimum level of function during disruptions, using metrics such as performance degradation, redundancy, and recovery time

What are the benefits? These methods offer a structured way to quantify the trade-off between efficiency and resilience, a key challenge also faced in health care. Highly efficient systems often run with minimal spare capacity, which can leave them vulnerable to stress. By simulating disruptions (e.g. staff shortages, supply chain issues, or service closures), resilience metrics can highlight weaknesses in the system that might not be visible through traditional productivity metrics. This allows planners to evaluate where investing in redundancy or flexibility could sustain productivity during periods of stress.

How applicable is this approach to the NHS? The NHS, like transport systems, functions as a complex, interdependent network. Resilience-informed productivity measures could better reflect true performance by accounting for sustained service during pressure, not just volume of output in normal conditions. A recent review[45] looked at how resilience is measured in health care and whilst this is an active area for analysis, there is not consensus on how to define and evaluate resilience in this setting. Additionally, a challenge for these methods applied to the NHS is integrating data across organisational boundaries.

3.1.4 Frontier and other approaches to efficiency and productivity scores

How it's measured and what it's used for?

As discussed in the previous section, tools like the Model Hospital systematically analyse differences in productivity by benchmarking providers against their peers. This supports learning from high-performing areas and enables targeted interventions to improve efficiency.

An alternative approach to comparing providers involves calculating productivity scores or cost indices for different decision-making units (DMUs), such as hospitals or GP practices. The most efficient providers define the productivity frontier, serving as a benchmark against which others are

[44] Resilience and efficiency in transportation networks | Science Advances
[45] Organizational resilience in healthcare: a review and descriptive narrative synthesis of approaches to resilience measurement and assessment in empirical studies | BMC Health Services Research | Full Text

compared. Deviations from this frontier highlight inefficiencies, which may stem from non-optimised processes (technical inefficiency) or higher than necessary expenditure on inputs (cost inefficiency). These insights help units understand where improvements can be made and support national efforts to enhance health care efficiency.

Productivity and cost frontiers can be estimated using parametric or non-parametric methods. Stochastic Frontier Analysis (SFA) analysis, a parametric approach, accounts for statistical noise, making it useful when outliers could distort results. Data Envelopment Analysis (DEA), a non-parametric method, evaluates multiple inputs and outputs simultaneously (limited to less than the number of DMUs). Both approaches have been used to provide insights into efficiency variations across providers.

SFA analysis is used for validating and informing efficiency factors in national pricing models. By assessing the efficiency frontier for health care providers, SFA identifies variations in cost and service delivery while accounting for factors beyond providers' control, such as patient demographics and health needs. The method separates inefficiency from random statistical noise, helping determine the achievable best practices for providers operating under optimal conditions[46]

DEA has been applied to compare productivity at commissioning level and for primary care services. By analysing staff levels across different service areas and incorporating geographic factors, one study estimated the additional number of GP appointments that could be achieved if underperforming areas reached the efficiency frontier, providing a basis for resource reallocation[47].

There are accounts that Castelli et al.[48] draw on growth accounting methods to compare productivity differences across NHS hospitals, taking into account their multiple inputs and outputs (an alternative to frontier analysis). They find substantial variation in productivity among hospitals, suggesting scope for productivity improvement.

Strengths and weakness

A key strength of frontier methods like Data Envelopment Analysis (DEA) lies in their ability to manage multiple inputs and outputs simultaneously, offering a valuable way to assess efficiency across diverse health care providers. These methods can help isolate inefficiencies from statistical noise and account for performance differences linked to local models of care. Additionally, combining scores with regression analysis allows for deeper exploration of how external factors, such as geographic variation, deprivation, or demographic characteristics, influence productivity. This layered approach helps provide a more nuanced understanding than a single performance score might offer.

Limitations:

[46] Monitor Research Template

[47] Estimating productivity levels in primary medical services across clinical commissioning groups in England and the impact of the COVID-19 pandemic: a data envelopment analysis | BMC Health Services Research

[48] Examining variations in hospital productivity in the English NHS - PMC

- **Data coverage:** The effectiveness of such analyses is often constrained by inconsistent datasets across providers and commissioners, hampering comparability and national-level insights.
- **Understanding evolving models of care:** Moreover, while frontier methods offer flexibility, they can still struggle to represent the full breadth and complexity of health care outputs and inputs, including quality of care and investment levels. This has led researchers to look to alternative methods for comparing providers[18]. Simplifying performance into a single score can risk overlooking important contextual variables like patient case mix or local structural challenges, potentially leading to misleading conclusions.

3.2 Evaluative productivity measures: What metrics are currently used, and what are their strengths and weaknesses?

3.2.1 Productivity measurement in the context of evaluation and risk stratification

How it's measured and what it's used for?

The HMT Green Book provides guidance on appraising policies, programmes, and projects. "Appraisal is the process of assessing the costs, benefits and risks of alternative ways to meet government objectives" hence Green book approaches are about comparing alternatives to identify best options for available resources[49].

Benefits may include those relating directly to operational productivity, increasing patient treatment with the same workforce, shorter hospital stays reducing bed occupancy cost, better use of capital such as investment in digital tools reducing admin time. It may also be the case that a new treatment or process improves patient health outcomes, for example, technology assisted diagnosis may speed up appropriate care and limit disease progression. This improves the quality of care the patient but also reduces future health and social care demand so has indirect productivity benefits.

Benefits and cost implications can be assessed using economic evaluation approaches such as budget impact modelling, cost-benefit analysis, or cost-effectiveness analysis. These methods help determine whether an intervention provides good value for money and supports decision-making on resource allocation. A discounting rate is applied to costs and benefits with different time spans to be compared on a common "present value" basis to "adjust for social time preference, defined as the value society attaches to present, as opposed to future, consumption."

Relatedly, population health analytics and risk stratification methods provide additional productivity insights by linking resource allocation to health outcomes. By identifying high-risk groups and prioritising preventative interventions, health care systems can improve long-term efficiency, reducing costly reactive treatments and hospital admissions. Evaluation is established as part of the population health analysis cycle[50].

Strengths and weakness

Rather than focusing solely on technical efficiency (doing things right), these approaches also consider whether resources are being directed toward the most impactful interventions (doing the right things). This value-orientated framing enables productivity metrics to be integrated with

[49] The Green Book (2022) - GOV.UK
[50] NHS England » Population health management

strategic decision-making and outcome-focused planning. Incorporating risk stratification into productivity analysis supports a shift from reactive to preventative care, allowing resources to be targeted where they deliver the greatest return in health outcomes.

Evaluation approaches are often integrated into improvement initiatives for local models of care. For example, the STAR[51] (Socio-Technical Allocation of Resources) model, provides a structured way to incorporate both productivity measurement and stakeholder engagement into decision-making. STAR combines value for money analysis with qualitative input from commissioners and health care stakeholders, helping to align resource allocation with local priorities.

Limitations:

- **How to value input and reflect long term investment:** Whilst approaches to managing different time spans for benefits and inputs are established in the Green book methodology, purportedly a shift in how we value evidence of long-term benefits when we invest is required.

.. In the NHS, they do things in very short windows. ... What can you do in two years? ... Where we buy contracts with very, very short periods of time, it's hard to embed it, make it move to BAU status and then to deliver those benefits because it's within a short window, it's very hard to do that. That's just a fact of life. When I was at [private sector IT/software/system provider], our typical outsource contracts were 10 years long. Stakeholder 30

- **Attributing inputs and outputs at service level:** Establishing causal links between inputs and outcomes at the service level remains challenging, particularly in complex systems with numerous confounding factors. Robust evidence of impact is often difficult to achieve in real-world health care environments.
- **Understanding evolving models of care:** As models of care diversify locally, comparative productivity assessment across providers becomes more difficult. Local evaluation findings, while valuable for service improvement, may not generalise easily to national programmes or macroeconomic modelling. These issues limit the consistency and scalability of productivity measurement.

Evaluative methods ultimately answer a different question to the established macroeconomic methods discussed in 2.4.1. However, clearly there would be advantage to aligning the two areas of analysis more closely, for which there is a need to stream-line and standardise evaluation in the NHS such that insights could potentially be used more systematically in productivity measurement.

"We don't have [consistently recognised productivity] impact measures for digitalisation of health systems." Stakeholder 11

3.2.2 Marginal productivity and approaches to system resource allocation

How it's measured and what it's used for?

"So for me, for our purposes, productivity is really about the benefit... what we want to know is what would the health benefits be had we devoted more resources to the NHS" Stakeholder 12

[51] Star guidance document.pdf

Marginal productivity refers to the additional health benefit (e.g. improved patient outcomes, Quality-Adjusted Life Years (QALYs) gained, reduced mortality) produced by an additional unit of health care expenditure or resource input. This concept helps assess whether NHS resources are being used efficiently and where additional funding would have the greatest impact.

Studies have estimated the marginal productivity of NHS expenditure by analysing variations in spending and outcomes across different regions. For instance, Claxton et al[52]. assessed how changes in NHS funding correlate with health outcomes, suggesting that the marginal cost per quality-adjusted life year (QALY) gained is approximately £13,000. The latest study by Claxton et al.[53] investigates the causal association between health care expenditure and mortality rates both for specific disease areas, such as cancer and circulatory disease, and for all-cause mortality.

Given the constraints on NHS funding, decisions about resource allocation arguably should consider marginal productivity. Programme Budgeting and Marginal Analysis[54] is an applied framework that helps NHS decision-makers reallocate resources by comparing the marginal benefits of different spending areas. It involves assessing whether shifting funds from one service to another would improve overall health outcomes. However, the programme budgeting data collection has now ceased.

Evidence suggests that investing in primary care and preventive health services, such as vaccinations, early screening, and chronic disease management, often yields higher marginal productivity compared to hospital-based care. These services can reduce the need for more expensive interventions later, enhancing overall system efficiency. Studies have shown that greater investment in primary care is associated with improvements in health system performance, including increased screening rates[55]. . The marginal cost per QALY literature, as the marginal productivity of health expenditure was previously known as, could also provide a way of uprating values of primary care and preventative care for the macro productivity measures, an approach which has since been considered for NHS reporting.

Strengths and weakness

Methods for estimating the marginal productivity of NHS expenditure, linking spending to outcomes like mortality or QALYs, offer valuable insights into how efficiently resources are being used and where additional investment might deliver the greatest benefit. These approaches can help identify high-value areas for resource allocation. Programme Budgeting and Marginal Analysis (PBMA) further strengthens this by enabling NHS decision-makers to compare the marginal benefits of different services and reallocate funding accordingly.

Limitations:

[52] Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold - PubMed

[53] Martin, S., Claxton, K., Lomas, J. et al. How Responsive is Mortality to Locally Administered Healthcare Expenditure? Estimates for England for 2014/15. Appl Health Econ Health Policy 20, 557–572 (2022). https://link.springer.com/article/10.1007/s40258-022-00723-2

[54] Managing Healthcare Budgets in Times of Austerity: The Role of Program Budgeting and Marginal Analysis - PMC

[55] Does more investment in primary care improve health system performance? - PubMed

- **Data coverage:** The accuracy of marginal productivity estimates can be hindered by data constraints, especially when attempting to link spending to outcomes across specific regions or disease areas. Whilst in some cases it has been possible to robustly establish causal relationships[53], doing so complex and regression-based approaches cannot always confirm that spending directly causes outcome changes. Additionally that programme budgeting data is no longer collected limits this field of analysis.
- **Understanding evolving models of care:** aggregating results at a high level can obscure important local variations, such as differences in population health needs or service delivery models.

Marginal value is a critical area for wider productivity measurement developing because of its key alignment with Atkinson principles which determine that the output of health services should be measured reflecting the contribution they make to improved health outcomes[3].

3.3 Current metric summary: classifying metrics by type and use

Table 2: Summary of metrics discussed according to the proposed metric classification (figure 2)

ID	Metric group	Type	Metric	Classification
1	Multi-factor	Technical	**ONS**	System productivity
2	Multi-factor	Technical	**York**	System productivity
3	Multi-factor	Technical	**NHS E implied productivity**	Operational productivity
4	Multi-factor	Technical / evaluative	**Public health in the national accounts (Martin Weale)**	System productivity
5	Workforce	Technical	**Workforce productivity – ONS definition**	System productivity
6	Workforce	Technical	**Model Hospital workforce metric**	Operational productivity
7	Workforce	Technical	**THF consultant productivity**	System productivity
8	Workforce	Technical / evaluative	**Clinical hours to contact metric (CHtC)**	Operational productivity
9	Workforce	Technical / evaluative	**Workforce performance frameworks (e.g. Global Labor Organization)**	System productivity
10	Operational	Technical	**Model hospital / WAU**	Operational productivity
11	Operational	Technical / evaluative	**GIRFT**	Operational productivity
12	Operational	Technical	**Quality improvement e.g. Lean and Six Sigma**	Operational productivity
13	Frontier and scoring approaches	Technical	**Efficiency / productivity scores derived using Data Envelopment Analysis (DEA)**	Operational productivity
14	Frontier and scoring approaches	Technical	**Efficiency / productivity scores derived using Stochastic Frontier Analysis (SFA)**	Operational productivity

ID	Metric group	Type	Metric	Classification
15	Frontier and scoring approaches	Technical	**Efficiency / productivity scores derived using growth accounting / regression based approach**	Operational productivity
16	Evaluation	Evaluative	**Cost-benefit and cost-effectiveness**	Service resource allocation
17	Evaluation	Evaluative	**STAR (Socio-Technical Allocation of Resources)**	Service resource allocation
18	Marginal productivity	Evaluative	**Marginal productivity**	System resource allocation

4. Conclusions: gaps, challenges and opportunities in productivity measurement

Based on a rapid review of literature and a number of stakeholder discussions, we have explored why and how we measure productivity in a health care setting. We have identified six broad method types and considered variations and different applications within these. We have also assessed the strengths, limitations and aptitude to deliver insights including constraints and development opportunities associated with reporting mechanism and data sets used.

The themes arising from our review of strengths and weaknesses of current approaches, echo Atkinson principles, including the importance of reflecting the quality and the value of outputs and their contribution to outcomes. Also, the need for comprehensive measures recognising the complex settings. Our review demonstrates that there has been considerable development in these areas since Atkinson's work, and more recently, further significant improvements outlined by the ONS Public Sector Productivity Review. But challenges remain, both conceptual (how we think about productivity) and practical (having the right measures to support improvement in real-world settings). Reflecting on these challenges, we have identified broad opportunities for development.

Themes also align with recent recommendations from the ONS Public Sector Productivity Review in the emphasis on developing cost weighting and quality adjustment approaches, the challenge of incorporating preventative care gains and the need to disaggregate macro-level metrics to investigate different components of the large and complex health system. Our findings do indicate some further areas for development, however, particularly expanding on this last point around disaggregation and how this links to the pivotal point in the views of stakeholders, that productivity metrics are most useful when used in a broader analytical framework to identify and understand drivers of productivity. These findings form the basis of recommendations for more multi-faceted measures of productivity, improved reflection of long-term investment, developing system perspective and case study development for preventative care models.

Key recommendations from this project

1. There is a need to adopt **more nuanced and multi-faceted productivity measures**, approaches that can better capture the complexity of health care delivery. Often current metrics are limited in the extent to which they reflect differences in care delivery settings or whether inputs are being managed to ensure sustainable productivity.
    - For example, integrating indicators of **long-term workforce sustainability**, such as staff retention, training investment, and skills mix or those relating to wider resilience of the system, into metrics could strengthen alignment with broader NHS strategy and future service needs.

2. Linked to the idea of sustainability, **long-term investment,** such as in digital systems or infrastructure, pose additional measurement challenges, particularly as their impact may take years to materialise. The challenge of timespan is also true of **measuring outcomes**, which may be improved health over a number of years, rather than short-term procedural outputs. There is recognition of the need to measure these elements more effectively at the macro level and that this could be tackled in different ways, for example:
    - As in point one, by using investment indicators as a proxy for sustainability.
    - There would also be much value in **aligning evaluative and macroeconomic approaches** to ensure that insights from local service evaluations can feed into national productivity frameworks.
    - Strengthening the role of **marginal value** in productivity analysis, consistent with the Atkinson principles, will help ensure that outputs are measured not just in terms of volume, but in terms of the longer-term health improvements they deliver.

3. A shift is evident in the move towards a **wider systems perspective**, one that recognises the interplay between the NHS, social care, and independent sector providers. Currently metrics also often don't offer insight into the way multiple parts of the system are operating to deliver an outcome. As integrated care models evolve, it becomes increasingly important to invest in tools that can assess productivity across different settings and also that recognise pathways of care (and external dependencies such as social care).
    - It's acknowledged that the dynamic nature of health care delivery and the emergence of new models of care across setting, requires **flexible, responsive measures.**
    - Whilst such flexibility will always be difficult to achieve within the scope of top-down macroeconomic metrics, there are important opportunities to strengthen **micro-level measurement**, using locally available data to build a richer, bottom-up understanding of productivity across care settings and pathways.

4. Finally, representing **preventative care** is recognised as a particularly notable shortfall of current productivity approaches. Whilst we currently value prevention using its short-term cost, furthering research on how to do so in the context of productivity forms a key recommendation of the ONS Public Sector Productivity Review. Stakeholders often drew on experiences of preventative care interventions, and it was demonstrated that there would be value in exploring further case studies to engage and develop thinking in this area.

Appendix: Project background and approach – further detail

A1. Atkinson Principles

As summarised by the Office of National Statistics[9]

Principle A: The measurement of government non-market output should, as far as possible, follow a procedure parallel to that adopted in the National Accounts for market output.
Principle B: The output of the government sector should in principle be measured in a way that is adjusted for quality, taking account of the attributable incremental contribution of the service to the outcome.
Principle C: Account should be taken of the complementarity between public and private output, allowing for the increased real value of public services in an economy with rising real GDP.
Principle D: Formal criteria should be set in place for the extension of direct output measurement to new functions of government. Specifically, the conditions for introducing a new directly measured output indicator should be that (i) it covers adequately the full range of services for that functional area, (ii) it makes appropriate allowance for quality change, (iii) the effects of its introduction have been tested service by service, (iv) the context in which it will be published has been fully assessed, in particular the implied productivity estimate, and (v) there should be provision for regular statistical review.
Principle E: Measures should cover the whole of the United Kingdom; where systems for public service delivery and/or data collection differ across the different countries of the United Kingdom, it is necessary to reflect this variation in the choice of indicators.
Principle F: The measurement of inputs should be as comprehensive as possible, and in particular should include capital services; labour inputs should be compiled using both direct and indirect methods, compared and reconciled.
Principle G: Criteria should be established for the quality of pay and price deflators to be applied to the input spending series; they should be sufficiently disaggregated to take account of changes in the mix of inputs; and should reflect full and actual costs.
Principle H: Independent corroborative evidence should be sought on government productivity, as part of a process of 'triangulation', recognising the limitations in reducing productivity to a single number.
Principle I: Explicit reference should be made to the margins of error surrounding National Accounts estimates.

A2. Rapid review of literature

Key search terms were developed using synonyms for productivity. Two databases were utilised to identify academic and formally published evidence, search criteria are provided below. Additionally informal and other grey literature including policy and expert opinion pieces were sourced through targeted online searches and based on recommendations from stakeholders.

The rapid review also looked at a wide range of grey literature including topics or specific sources recommended by stakeholders during stakeholder interviews (see below). Additionally, we undertook targeted searches on methods used in other health systems and other sectors.

Titles and / or abstracts were screened and articles excluded if they didn't describe an approach to measuring productivity in the NHS or the approach was very nuanced to a particular health condition

or health related process. Data was extracted from literature to support this by reviewing abstracts and / or full articles.

Database	Search string	Additional filters
PubMed	(("productivity"[Title/Abstract] OR "efficiency"[Title/Abstract] OR "socio-technical allocation of resource"[Title/Abstract] OR "capital deepening"[Title/Abstract] OR "data envelopment analysis"[Title/Abstract] OR "frontier analysis"[Title/Abstract] OR "cost function"[Title/Abstract]) AND ("NHS"[All Fields] OR ("national health service"[All Fields] AND "UK"[All Fields])) AND ("concept"[Title/Abstract] OR "measur*"[Title/Abstract] OR "identif*"[Title/Abstract] OR "estimat*"[Title/Abstract] OR "defin*"[Title/Abstract] OR "analys*"[Title/Abstract] OR "indicat*"[Title/Abstract]) AND "health*"[Title/Abstract]) AND (2016:3000/12/12[pdat])	Date filter in main string
Google Scholar	"NHS" OR "National Health Service UK"AND ("productivity" OR "efficiency" OR "socio-technical allocation of resource" OR "capital deepening" OR "data envelopment analysis" OR "frontier analysis" OR "cost function")AND (concept OR measure OR identify OR estimate OR define OR analyse OR indicator)AND health	Custom date range: **Since 2016**

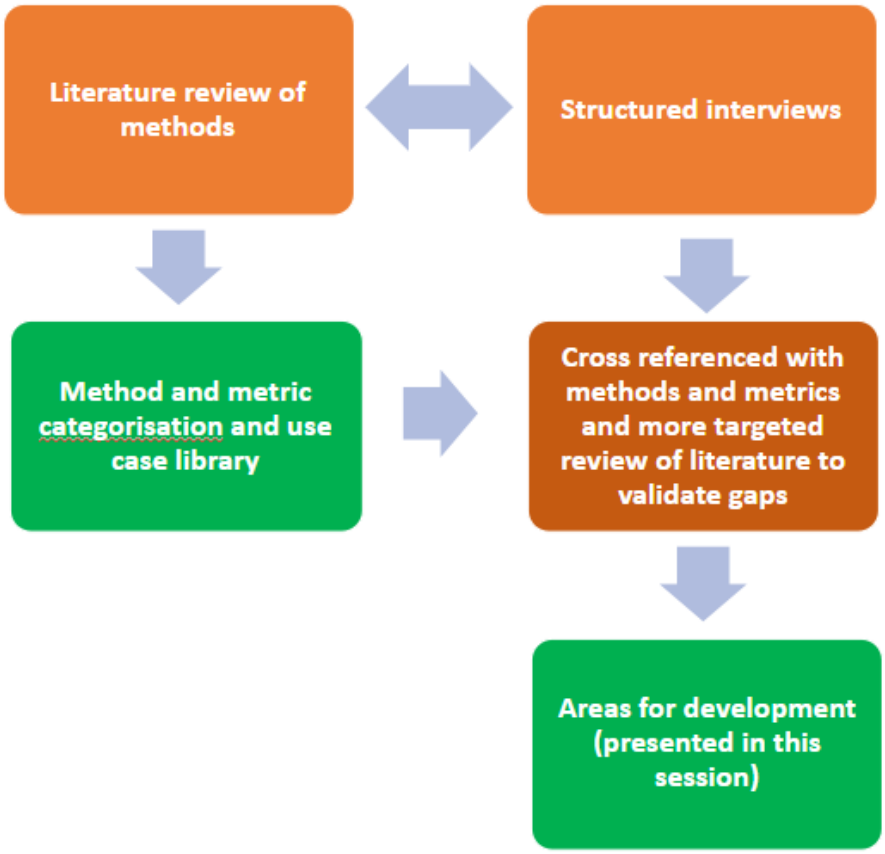
A3. Interviews

Interviews (either individually or in groups) used a topic guide structured around the topics used for the literature data extraction (use, value and limitations, operational details and feasibility as well as aptitude to deliver insights) and tailored to each stakeholder type and their organisation's role in measuring productivity, drivers of productivity or use of the outputs of measurement. Interviews were conducted and recorded using Microsoft Teams®. A summary and transcript were generated and reviewed and key themes were created and organised across and within stakeholder types. On average interviews took approximately one hour.

A4. Workshops

Workshops were conducted and recorded using Microsoft Teams® and used Miro online whiteboards to provide discussion prompts and collate responses. A summary and transcript were generated and reviewed with key themes created within and across the three areas of focus and considered within the frame of stakeholder types.

A5. Analysis process



A6. Information governance

As part of the project initiation, a process map for inviting and consenting stakeholders to participate in an interview (and/or workshop where applicable) was produced along with a consent form explaining the reasons for collecting data and how it will be used during and after the project. It outlined that individual's interview data would be anonymised and combined with other interviewee's data as part of analysis and reporting. We explained that job type/role and sector represented by participants would be collated and presented as part of this report.

A protocol described the work to be undertaken which set out that the project is not research as defined by the Health Research Authority (HRA) and in any case did not require NHS Research Ethics Committee review (based on the HRA Decision Tool assessment of question sets 1-4). Any NHS staff contributing their views and expert opinion was part of their professional role. The protocol therefore did not have regard to the HRA guidance although some content was similar in nature.

Sperate written consent was obtained from stakeholders for the workshops.

Arden&GEM
Health and social care systems support

Get in touch with us at:

www.ardengemcsu.nhs.uk

@ardengem

contact.ardengem@nhs.net