The Strategy Unit.

10-Year Productivity Forecast for the English NHS: An Expert Elicitation Study

October 2025

NHS

Document control

Document Title	10-Year Productivity Forecast for the English NHS: An Expert Elicitation Study
Job No	E004520-0001
Prepared by	Josh Cook, James de Lacy, Jennifer Wood and Mohammed A Mohammed
Checked by	Mohammed A Mohammed
Date	October 2025

Contents

Executive Summary

Background

Healthcare productivity – comparing the outputs of the NHS, such as GP appointments and hospital procedures, to its inputs, such as staff and equipment - is one of the most pressing challenges facing government. With rising demand and limited funding, improving productivity is essential to delivering the government's ambitions within constrained budgets. The COVID-19 pandemic dealt a serious blow to NHS productivity, and while some evidence suggests it may be recovering, it has largely struggled to return to pre-pandemic levels. In response, the government has set an ambitious target of 2% annual productivity growth[1] and introduced a range of policies – from adopting new technologies to workforce reform and service redesign. To support this effort, the Health Foundation launched the NHS Productivity Commission in 2025 to identify practical strategies for raising productivity over the next decade.

Yet, the outlook for NHS productivity remains highly uncertain, particularly trajectory and rate of change. This reflects a lack of knowledge, not just about what will happen, but about how productivity might respond to different interventions and conditions over time. For example, the government's five 'big bets' on technologies such as AI to drive productivity. This poses a serious challenge for long-term NHS planning and achievement of its goals on service standards, as assumptions about productivity directly affect future staffing, funding, and infrastructure needs. Because extrapolating from historical trends alone cannot resolve this uncertainty, the Health Foundation's modelling and the Commission's work draw on expert elicitation – a structured approach to capturing and synthesising expert judgement – to build credible, transparent assumptions about levels of productivity growth.

Objective

We undertook an expert elicitation exercise to elicit probabilistic forecasts for the annual growth rate in non-quality adjusted productivity for the NHS in England over the next 10 years. We also looked to understand the phasing of this growth i.e. whether it would be steady over the 10-year period, or if it might be front or back-loaded.

Method

An in-person 1-day facilitated expert elicitation workshop was undertaken using an evidence-based protocol. The primary outcomes are the pooled subject matter expert (SME) forecasts representing

[1] The 2% NHS productivity target features in the Spending Review 2025 which set departmental budgets for day-to-day spending until 2028-29. The 2% figure was recently reaffirmed as a target over the next 3 years in the 10 Year Health Plan for England: fit for the future.

the % annual productivity growth of the NHS. Each SME was asked to provide their forecast as a prediction interval from a "surprisingly low" (10th percentile – P10) to a "surprisingly high" (90th percentile – P90) value. The prediction intervals were elicited over two rounds. The ONS measure of non-quality adjusted productivity was adopted. SMEs provided their forecasts alongside their rationales using an online app designed specifically for this exercise. The forecast periods were 0-5 years and 5-10 years, from which the experts' 0–10 years prediction were derived. Participating experts were assigned to three teams: Team Frontline (n=6) consisted of senior NHS staff working in healthcare provider organisations; Team Policy (n=4) included senior leaders from NHSE and DHSC; and Team Research (n=4) included senior researchers and academics with domain expertise. After the closure of data collection for this core exercise, a subsequent less formal elicitation was conducted exploring a technology-enabled scenario which is included in the appendix.

Results

Pooled forecasts for NHS productivity are shown in the table below expressed as % compound annual growth rates (CAGR).

Period	All (n=14) Pooled		Frontline (n=6) Pooled		Policy (n=4) Pooled		Research (n=4) Pooled	
	P10 to P90	mean	P10 to P90	mean	P10 to P90	mean	P10 to P90	mean
0-5 years	-1.1% to 2.4%	0.7%	-1.8% to 2.7%	0.5%	-0.4% to 2.5%	0.9%	-0.3% to 1.7%	0.7%
5-10 years	-0.4% to 3.5%	1.4%	-0.6% to 3.8%	1.6%	-0.4% to 2.5%	1.1%	-0.2% to 3.5%	1.5%
0-10 years (derived)	-0.7% to 2.7%	1.0%	-1.1% to 3.2%	1.0%	-0.3% to 2.4%	1.0%	-0.2% to 2.4%	1.1%

While there was significant overlap between the pooled prediction intervals across teams, the Frontline Team's prediction intervals were the widest reflecting the greatest amount of uncertainty.

The means of the pooled forecasts of all experts were considerably higher in the 5-10 years period (1.4%) than the 0-5 years period (0.7%). The prediction intervals were moderately wider in the 5-10 years period indicating greater expert uncertainty.

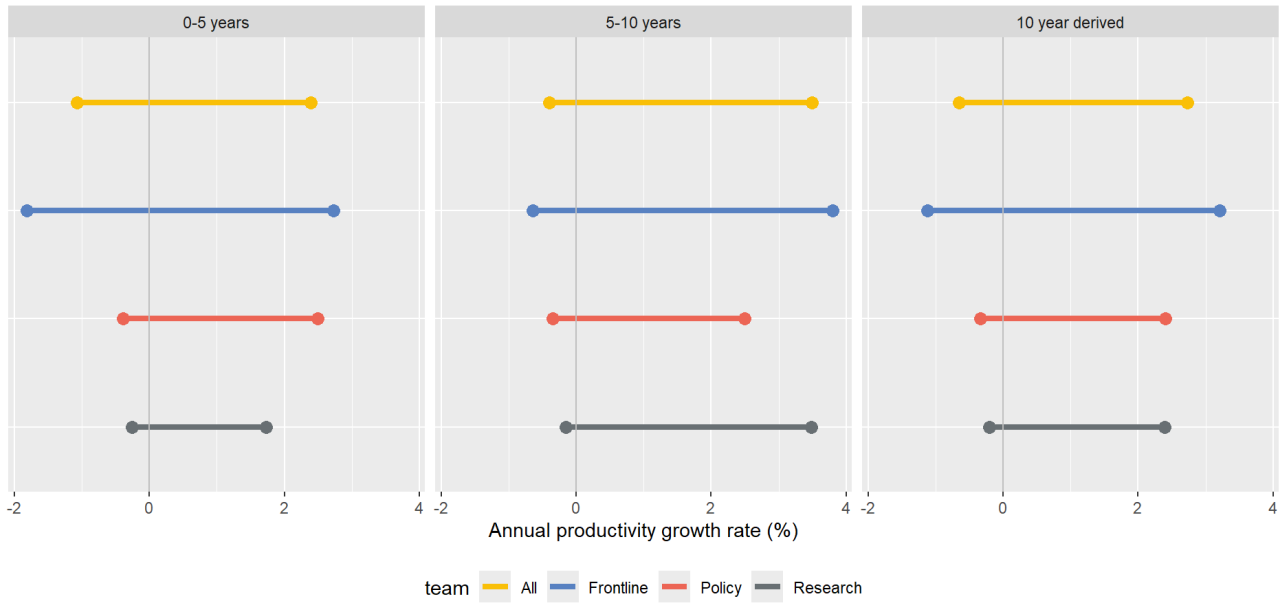


Figure 1 Pooled forecasts by team

Experts provided rationales for their P10 and P90 values describing the barriers and enablers of productivity. Some of the key barriers highlighted in SMEs' P10 rationales were the low morale of staff and high turnover rates; aging population and increasing patient complexity; insufficient capital investment; and weak adoption of technology. In the P90 rationales, some of the key enablers included the successful adoption of new technology; investment in NHS estate; improved data quality; better alignment of financial incentives and greater staff retention.

Conclusion

An expert elicitation exercise has produced pooled forecasts of NHS productivity rates over the next 10 years. At the outset of this exercise the level of expert uncertainty was unknown. The fundamental purpose of the exercise was therefore to quantify the uncertainty of experts and make it explicit using a systematic, evidence-based process. The intention was not to change the uncertainty of the participating experts. As such, the exercise must be judged on the rigour of its method rather than the range of uncertainty it reveals. For the underlying uncertainty to be reduced, policy makers will need to articulate (and demonstrate through robust evaluation) the mechanisms that will generate productivity gains within a complex system such that a greater consensus can emerge on their effects.

Introduction

Healthcare productivity – comparing the outputs of the NHS, such as GP appointments and hospital procedures, to its inputs, such as staff and equipment - is one of the most pressing challenges facing government. With rising demand and limited funding, improving productivity is essential to delivering more care within constrained budgets. The COVID-19 pandemic dealt a serious blow to NHS productivity, and while some evidence suggests it may be recovering, it has largely struggled to return to pre-pandemic levels. In response, the government has set an ambitious target of 2% annual productivity growth and introduced a range of policies – from adopting new technologies to workforce reform and service redesign. To support this effort, the Health Foundation launched the NHS Productivity Commission in 2025 to identify practical strategies for raising productivity over the next decade.

Yet, the outlook for NHS productivity remains highly uncertain, particularly trajectory and rate of change. This reflects a lack of knowledge, not just about what will happen, but about how productivity might respond to different interventions and conditions over time. This includes the expansion of technologies such as Artificial Intelligence (AI), genomics and robotics to drive productivity, but where the potential impacts and time scale are uncertain. This poses a serious challenge for long-term NHS planning and achievement of its goals on service standards, as assumptions about productivity directly affect future staffing, funding, and infrastructure needs. Because extrapolating from historical trends alone cannot resolve this uncertainty, the Health Foundation's modelling and the Commission's work draw on expert elicitation – a structured approach to capturing and synthesising expert judgement – to build credible, transparent assumptions about levels of productivity growth.

Expert elicitation is a widely used method for forecasting when there is uncertainty about a quantity of interest. Our objective was to gather expert judgements on future healthcare productivity by eliciting probabilistic forecasts. The elicitation process is informed by literature on cognitive biases, project planning, and decision analysis as summarised by Hemming et al.[2]:

> Expert judgement can be remarkably useful when data are absent or incomplete. However, experts can also make mistakes. This is often due to a range of cognitive biases such as anchoring, availability, and representativeness, groupthink, overconfidence, and difficulties associated with communicating knowledge in numbers and probabilities. Inappropriate and ill-informed methods for elicitation can amplify these biases. Well designed, structured elicitation protocols can enhance the quality of expert judgments. These protocols treat

[2] Hemming V et al. A practical guide to structured expert elicitation using the IDEA protocol. Methods in Ecology and Evolution. 2018 Jan;9(1):169-80.

each step of the elicitation as a process of formal data acquisition, and incorporate research from mathematics, psychology and decision theory to help reduce the influence of biases and to enhance the transparency, accuracy, and defensibility of the resulting judgements.

Method

The elicitation protocol used in this study is based on the Stanford Research Institute protocol which has five broad steps: motivate, structure, condition, encode and verify, as described in the Handbook of Decision Analysis[3] and the IDEA (Investigate, Discuss, Estimate, Aggregate) protocol as described by Hemming et al[4]. The protocol has been designed to mitigate a variety of cognitive biases and heuristics which can influence expert judgement. The protocol has ethical approval from the University of Bradford research ethics committee.

In collaboration with colleagues from the Health Foundation, we designed and facilitated a 1 day in-person elicitation workshop. Data collection was conducted over two rounds using an online app. The workshop ran from 10:00 to 16:00 with comfort breaks on 2 July 2025.

Participants:

The recruitment of workshop participants aimed to assemble a group of experts who would possess the following characteristics:

• Domain knowledge or expertise or interest in forecasting future NHS productivity rates

• Openness to changing their mind.

• Willingness to follow the elicitation protocol.

• An appetite for making probabilistic forecasts.

It was also intended for the group to represent a diversity of perspectives on healthcare productivity with insights across different healthcare settings. Importantly, we also wanted to ensure that perspective from inside and outside of the NHS would be captured. Senior staff working within the NHS and the Department of Health and Social Care, alongside senior researchers were invited.

Participation was voluntary, confidential and required informed consent with the option to withdraw at any time. The list of participants is provided in the acknowledgement section of this report. In this study participants are referred to as subject matter experts (SMEs). One facilitator made notes during the workshop. SMEs were given the option to make their own notes and share them with the facilitation team if they wanted to. The session was not recorded and was conducted

[3] Parnell GS, Bresnick T, Tani SN, Johnson ER. Handbook of decision analysis. John Wiley & Sons; 2013.
[4] Hemming V et al. A practical guide to structured expert elicitation using the IDEA protocol. Methods in Ecology and Evolution. 2018 Jan;9(1):169-80

under the Chatham House Rule.[5] SMEs were split into three teams to enable us to examine the extent to which there were systematic differences between them. Team Frontline (n=6) consisted of senior NHS staff working in healthcare provider organisations; Team Policy (n=4) included senior leaders from NHSE and DHSC; and Team Research (n=4) included senior researchers and academics with domain expertise.

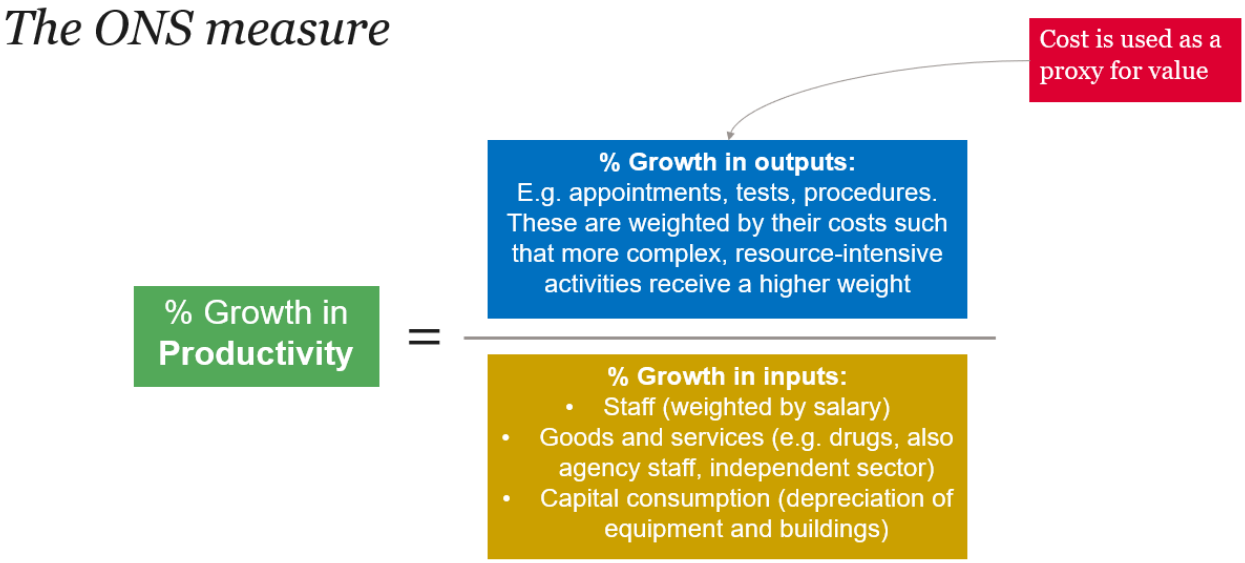
Use of non-quality adjusted productivity measure

The official measure of productivity used at the national level is from the Office of National Statistics (ONS). This measure can be quality-adjusted or non-quality adjusted. In this exercise we used the non-quality adjusted version. A limitation of this measure is that it uses activity (weighted to its cost) to measure NHS outputs. The cost weighting can mean that activity that is equivalent, but delivered in a lower cost setting, appears as a reduction in output. However, a recent review by the National Statistician has made recommendations about developing the measure to better reflect changes to delivery models. We advised SMEs to allow for these specific developments in the measure when making their forecasts.[6]

The decision to use the non-quality adjusted measure was a pragmatic one. The measurement is used in the Health Foundation Real Centre's modelling, is the official accredited statistics and is expected to be equivalent to the target which has been set for NHS England. Furthermore, the quality adjustments currently applied are not comprehensive and are dependent on data availability and collection. As such, we opted against asking experts to predict quality adjusted productivity rates, considering this to be an even more difficult cognitive task. However, we acknowledge the limitations of the non-quality adjusted measurement of productivity in capturing the impact of healthcare on patient outcomes. These challenges are particularly pertinent to healthcare settings outside of hospitals.

[5] The Chatham House Rule is an agreement between participants that allows people to use the information from a discussion but prohibits the disclosure of the speaker's identity.

[6] These developments relate to: the equalisation of weights in acute care (in relation to elective surgery where procedures may be carried out either as an inpatient procedure, a day case procedure or an outpatient procedure); equivalisation of weights in ambulance response (whereby 'a shift in the composition of ambulance activity towards more calls being responded to with advice or care in-place without conveyance will not result in a reduction in output'); the incorporation of certain screening services into the output measure; and improvements to better account for reductions in unnecessary A&E admissions. For further details please see https://uksa.statisticsauthority.gov.uk/publication/national-statisticians-independent-review-of-the-measurement-of-public-services-productivity/pages/11/



Infographic shared with experts

Supporting information

Ahead of the workshop, each SME received a briefing document via email which provided an overview of the aims of the elicitation exercise as well as a description of productivity measurement and summaries of historical trends. Key content from the briefing document was incorporated into a supporting information pack which was provided to SMEs at the start of the workshop. This pack also contained two tables displaying the historical non-quality adjusted productivity trends which are shown below:

Compound annual growth rates (CAGR)	Output (NQA)	Input	Productivity (NQA)
Historical average			
All time (1995/96-2022/23)	4.0%	3.7%	0.3%
Excl. COVID-19 (1995/96-2018/19)	4.3%	3.7%	0.6%
Different time periods			
1995/96-2010/11	4.8%	4.5%	0.3%
2010/11-2018/19	3.4%	2.1%	1.3%
2018/19-2022/23	2.3%	4.2%	-1.8%

Productivity (NQA)		
Historical maximum and minimum* CAGR (1995-96-2022/23)		
Maximum		
5-year	**2.0%**	
10-year	**1.0%**	
Minimum	Excl. COVID-19	Incl. COVID-19
5-year	**-0.1%**	**-5.0%**
10-year	**0.2%**	**-1.7%**

*Here the **maximum** refers to the continuous 5-year and 10-year periods with the highest non-quality adjusted productivity rates (as CAGR) and the **minimum** refers to the continuous 5-year and 10-year periods with the lowest non-quality adjusted productivity rates (as CAGR) between 1995-96 and 2022/23**.**

Anita Charlesworth (Senior Economic Advisor at the Health Foundation and co-chair of the NHS Productivity Commission), and members of the Health Foundation's REAL Centre and Productivity Commission teams were in the room as expert advisors with respect to the historical trend data presented to participants and the quantity of interest. A significant amount of time was spent at the start of the workshop defining productivity and clarifying the ONS non-quality adjusted measure being used in the exercise. The expert advisors outlined the types of productivity gains which were within the scope of the measurement and explained the distinction between productivity and broader terms such as efficiencies. Examples were provided of efficiencies which would not be measured as productivity gains (i.e. reductions in unwarranted care) and therefore should not be factored into the SMEs judgements.

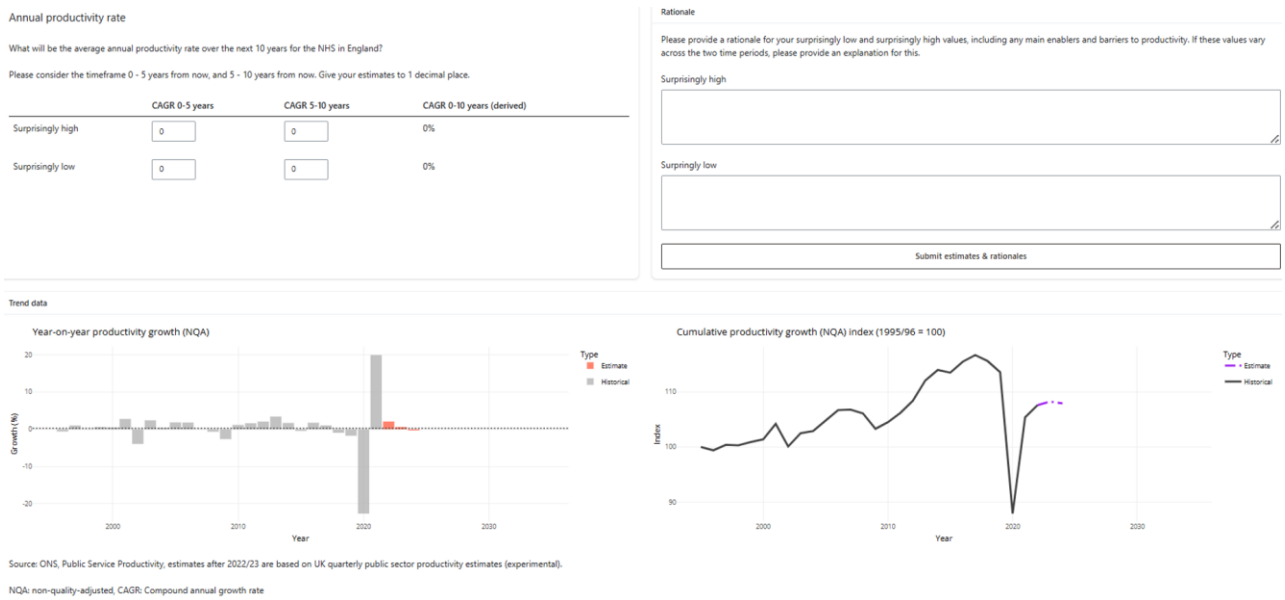
Mini Calibration training

As part of the introduction in the workshop, SMEs completed a mini calibration exercise designed to improve the accuracy of their probabilistic forecasting. SMEs were introduced to the concept of providing prediction intervals reflecting an 80% degree of belief. They answered five calibration questions adapted from Hubbard[7] for which the true value was unknown to the SMEs but known to the facilitators. SMEs were encouraged to reflect on their forecasts in light of the true value.

[7] Hubbard D. *How to measure anything: finding the value of "intangibles" in business*. 2nd ed. John Wiley & Sons; 2010.

Forecasting

The time horizon was 10 years from the baseline year of 2025/26. It was made clear that SMEs were required to **provide annual growth rates (%) not total period growth** via a specially designed online app.



Screenshot of elicitation app

To understand how SMEs anticipated the phasing of productivity growth over the 10-year horizon, they were asked to provide predictions over 0-5 years and 5-10 years. SMEs were informed that their predictions for the two time periods (0-5 years and 5-10 years) would be combined to produce their estimate for the 0-10-years period. SMEs were asked to provide probabilistic prediction intervals with an 80% degree of belief. They were required to use the app to enter the lower and upper bounds of their prediction intervals: their P10 and P90 values. The concept of the P10 to P90 interval was explained by using the qualitative terms "surprisingly low" and "surprisingly high" respectively. When SMEs entered their P10 and P90 values for the 0-5 years and 5-10 years, the app calculated the derived 0-10 years prediction which was visible to SMEs. It was also made clear to SMEs that their prediction intervals would be assumed to describe a Normal distribution[8] (i.e. we assumed that the spread of the uncertainty would be symmetrical taking the classic bell shape - values close to the middle of the interval are most likely. Values become progressively less likely the further away they are from the middle).

[8] In social sciences, Normal distributions are often inferred when the true distribution of a measurement error is not known.

SMEs were asked to provide rationales for their 'surprisingly high' and 'surprisingly low' values describing the key barriers and enablers to productivity. If SMEs values varied across the two time periods, they were asked to include an explanation for this in their rationales.

In round one, SMEs were given 30 mins to enter their predictions and rationales in silence. In round two, the app presented to each respective SME their own prediction intervals for the 0-10 years (derived) period on a dumbbell graph alongside the deidentified data from all other SMEs. They could also view in the app a table containing all SMEs' P10 and P90 values for 0-5 years, 5-10 years and 0-10 years (derived). Presented on a large screen in the room was a dumbbell graph displaying the 0-10 years derived prediction intervals colour-coded by team. Paper printouts showing the prediction intervals alongside rationales were also made available to each participant. Before beginning a 30-minute facilitated discussion, SMEs were given 10 minutes to review this data in silence (although the data remained accessible for the entirety of round two). SMEs then had an opportunity to revise their round 1 responses in silence. Following the closure of round two, the results were pooled.

After the completion of this core exercise, a subsequent less formal elicitation was conducted which explored an NHS technology-enabled scenario. The approach and results for the tech scenario can be viewed in the appendix.

Data processing and analysis: All SMEs were deidentified and assigned a unique code. We assumed that the P10 and P90 values supplied by each SME described a Normal distribution, from which we derived the mean [P10+P90)/2] and standard distribution [(P90 - mean) / qnorm[9] (0.9, mu=0, sd=1)]. SME predictions from round two were pooled by creating mixture distributions[10] using the distr package[11], weighting each SME prediction equally. We present the probability density of these mixture distributions, and extract P10, P90 and mean values, for each team and in total. The GitHub repository containing the R code is here:

https://github.com/The-Strategy-Unit/elicitation_thf_2025_07_results

[9] 'qnorm( )' is a function in R representing the quantile function of the standard Normal distribution.
[10] The pooled (mixture) distribution combines the full parametric distributions of each subject matter expert. For details on the derivation of these mixture distributions, please see https://cran.r-project.org/web/packages/distr/refman/distr.html#UnivarMixingDistribution.
[11] https://cran.r-project.org/web/packages/distr/index.html

Results

Overall results

14 SMEs completed the exercise. There were no missing data from round one to round two. Our pooled results are based on round two.

Pooled forecasts for NHS productivity are shown in the table below expressed as % compound annual growth rates (CAGR).

Period	All (n=14) Pooled		Frontline (n=6) Pooled		Policy (n=4) Pooled		Research (n=4) Pooled	
	P10 to P90	mean	P10 to P90	mean	P10 to P90	mean	P10 to P90	mean
0-5 years	-1.1% to 2.4%	0.7%	-1.8% to 2.7%	0.5%	-0.4% to 2.5%	0.9%	-0.3% to 1.7%	0.7%
5-10 years	-0.4% to 3.5%	1.4%	-0.6% to 3.8%	1.6%	-0.4% to 2.5%	1.1%	-0.2% to 3.5%	1.5%
0-10 years (derived)	-0.7% to 2.7%	1.0%	-1.1% to 3.2%	1.0%	-0.3% to 2.4%	1.0%	-0.2% to 2.4%	1.1%

The means of the pooled forecasts of all experts were considerably higher in the 5-10 years period (1.4%) than the 0-5 years period (0.7%). The prediction intervals were moderately wider in the 5-10 years period indicating greater expert uncertainty. However, this varied across the teams. The pooled prediction intervals of Team Policy were the same across the 0-5 and 5-10 years with the mean being only marginally higher in the latter period (increase of 0.2%). For Team Frontline, the mean of the pooled 5-10 years forecast (1.6%) was more than 3 times greater than the 0–5-year forecast (0.5%) while the width of their pooled prediction intervals were almost identical across those periods. Team Research was the only team whose pooled prediction intervals became wider in the 5-10 years period than the 0-5 years period.

There was a significant amount of overlap between the pooled prediction intervals across teams and the mean values for the 0-10 years (derived) were remarkably similar. However, Team Frontline had a much wider prediction interval than the other teams reflecting a greater degree of uncertainty – the interval between the P10 and the P90 for Team Frontline was 4.3 percentage points compared with Team Policy's 2.7 percentage points and Team Research's 2.6 percentage points.

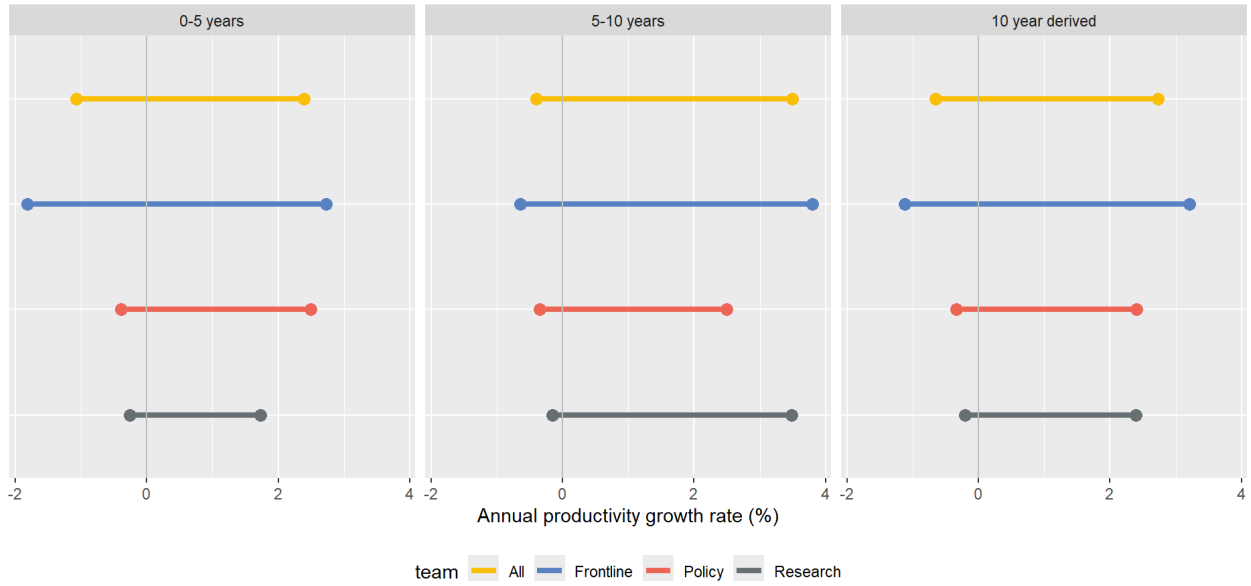


Figure 2 Pooled forecasts by team

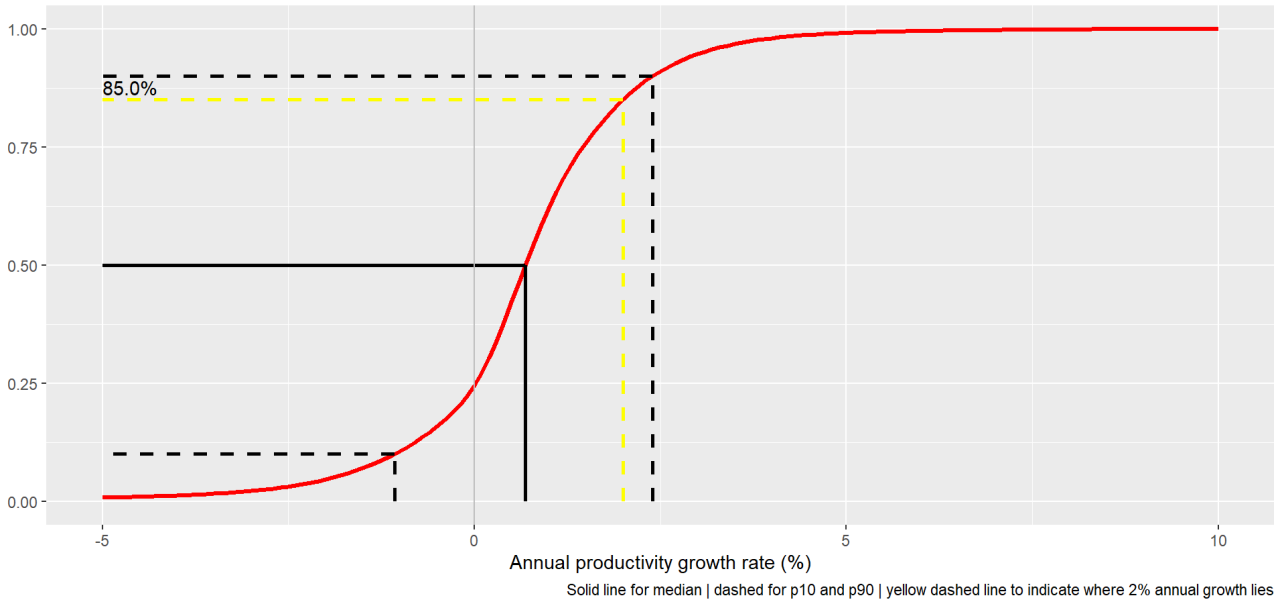


Figure 3 Cumulative distribution function or S-curve showing the pooled forecasts for the 0-5 years period

The S-curve describes the distribution of the pooled forecasts for the 0-5 years period. The Y-axis shows the cumulative probability that the annual productivity growth rate will be less than or equal to a specific value on the X-axis. This cumulative probability ranges from 0 (0% probability) to 1 (100% probability). Plotted on the yellow dashed line is the NHS productivity target of 2%. While this falls between the pooled P10 and P90 values (-1.1% and 2.4%), there is an 85% probability that the productivity rate will be less than or equal to 2% according to the pooled forecasts of all SMEs. The median (P50) value is 0.7%.

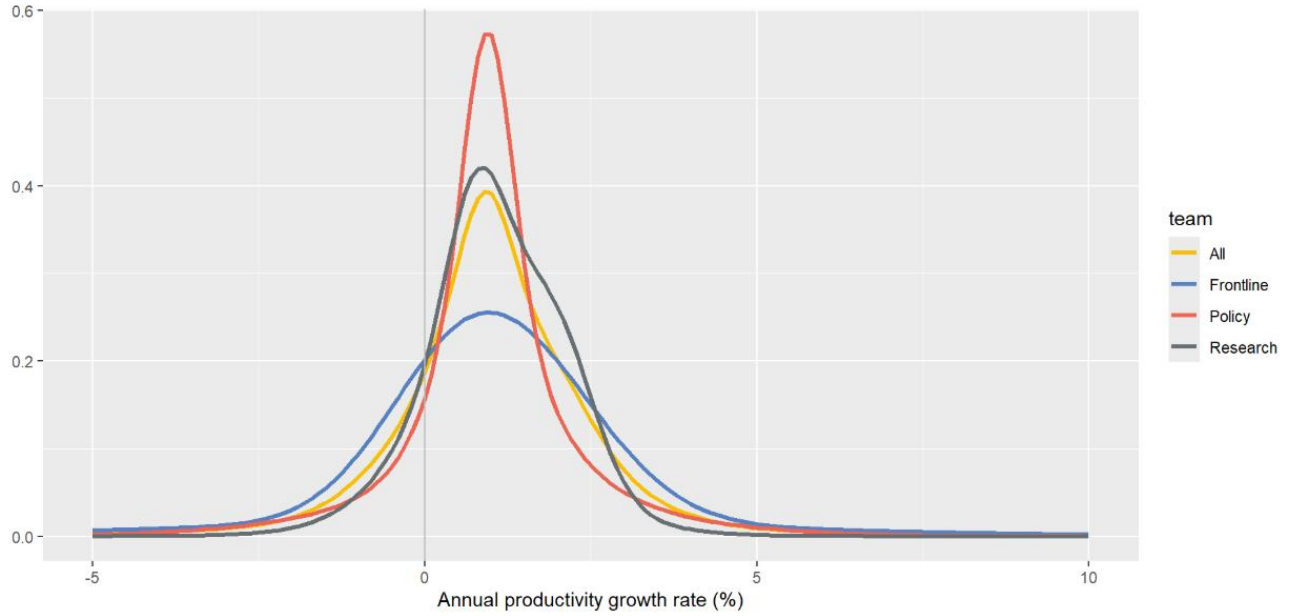


Figure 4 Density plot showing the pooled 0-10 years derived forecasts

Team Policy had the most peaked distribution reflecting the least amount of variation between their SMEs' prediction intervals for 0–10-year derived forecasts. Conversely, the Frontline Team had the greatest amount of variation amongst their SMEs' prediction intervals.

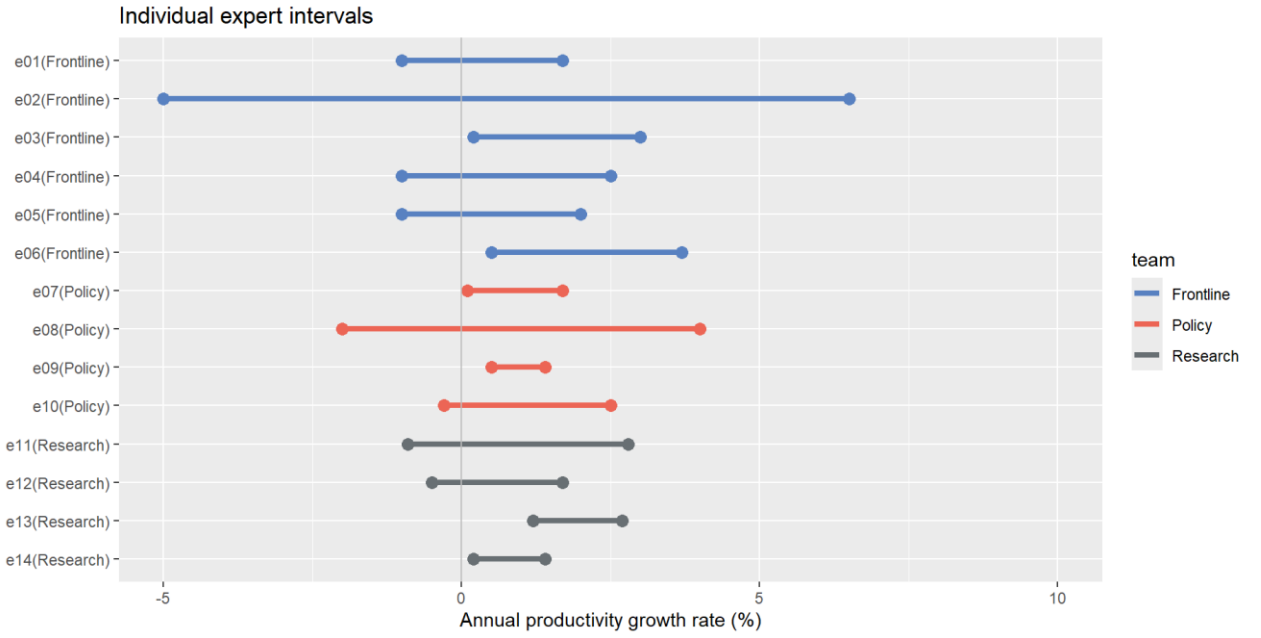


Figure 5 Individual expert forecasts for 0-10 years

There was a considerable amount of variation between individual expert's degree of uncertainty expressed in the width of their prediction intervals. The prediction intervals of e02 (Frontline) indicated the greatest amount of uncertainty while e09 (Policy) reflected the least uncertainty.

The lowest mean value for 0-10 years (derived) forecast was 0.35% (e01 Frontline) while the highest mean was 2.1% (e06 Frontline).

Comparison of round one versus round two results

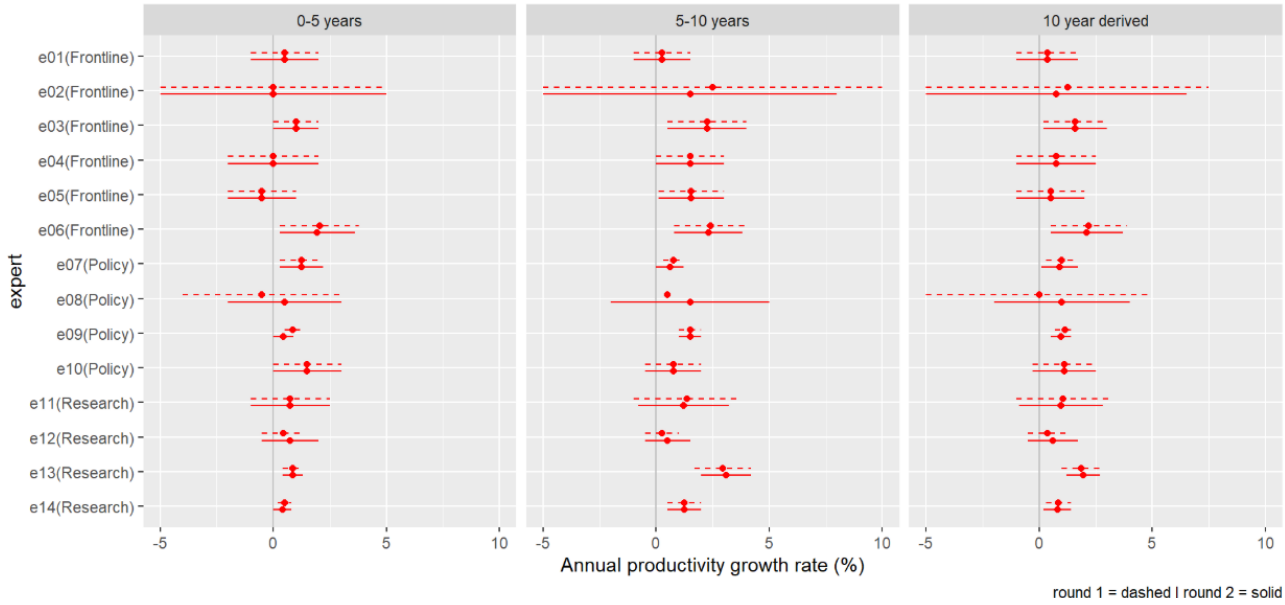


Figure 6 Individual expert P10 to P90 prediction intervals round one to round

In round two, 9 SMEs adjusted their round one forecasts. For 3 of the SMEs, these changes resulted in their 0-10 years derived prediction intervals becoming wider, while for 5 SMEs it resulted in them becoming narrower (in the other instance, the width of the interval stayed the same). 3 SMEs made revisions which increased their mean values for the 0-10 years derived forecast, while 6 SMEs made revisions which lowered their mean values. Some SMEs noted the reasons for adjusting their forecasts in their submitted rationales. Reasons specified by SMEs for lowering their forecasts included low staff moral; more complex patients; staff being worked close to maximum capacity; and overstretch of the system. Of the SMEs who made their forecasts more positive, 2 stated that they had done so because of the facilitated discussion which took place after round one although they did not cite specific factors which explained their revisions. The most significant adjustment was made by e08 (Policy) whose 0-10 years derived prediction interval reduced from a width of 10.0% in round one to 6.0% in round two, with the mean increasing from 0.0% in round one to 1.0% in round two. In their rationale, e08 (Policy) wrote that they had "revised surprisingly low [P10 value] as geopolitical negative unlikely to continue to that degree for ten years."

Rationales

SMEs provided rationales for their P10 and P90 predictions. They were asked to include the main barriers and enablers. A wide range of factors were described by SMEs but in the analysis of the rationale responses some key themes are evident. There were however no apparent factors which SMEs from a particular team were notably more likely to include or omit in their responses. The analysis below is of expert rationales submitted in round two of the exercise.

P10 rationale themes

Supply: NHS workforce sustainability, capacity and capability

Factors relating to the NHS workforce were highlighted in many of the expert rationales. Multiple experts referenced staff burnout, low morale, poor retention and workforce shortages as key barriers to productivity. Some rationales also noted concerns about the possibility of reduced discretionary effort and "quiet quitting", reductions in NHS managers, and further industrial action. Others pointed to a reduction in the supply of overseas staff, which could cause further capacity constraints and an overall less experienced workforce.

> "*Poor staff morale and worsening of goodwill. Issues with retention, less experienced staff, staff shortages making it more difficult to schedule care effectively.*" (e11 Research)
>
> "*Increasing trend away from life-long careers mean overall less experienced staff, less discretionary effort*." (e08 Policy)
>
> "*Industrial Action restarts following pay settlements in 25/26 and beyond which reduces output and increases costs to cover essential care*". (e10 Policy)
>
> "*Insufficient focus on culture, training of workforce - leads to continued poor retention and reduction in foreign recruitment*"(e02 Frontline)
>
> "*Currently high proportion of new staff are recruited internationally and these routes could reduce, both from competition from other countries and also for political reasons*." (e08 Policy)

Demand: demographic changes and increased complexity of patients

The increased healthcare demands of an aging population and an increasing complexity of patients resulting from a greater prevalence of co-morbidities was cited in half of all expert rationales. Some SMEs believed that this would have the potential to create further strain of the healthcare system and while it would require more staff input, the treatment of more complex patients may not result in the equivalent increases in measured output. Additionally, health inequalities and the effects of

the wider determinants of health present add further complexity to patient needs, which could require more intensive interventions.

> "*Overload of the NHS, increasing survival of people with multiple issues, obesity etc is likely to overwhelm the system and this will reduce productivity*." (e01 Frontline)
>
> "*Care of increasingly complex patients not captured in productivity gains*." (e04 Frontline)
>
> "*More complex patients - though picked up in productivity measures - may move the NHS to a breaking point where it becomes hard to do anything properly in the second five years*." (e07 Policy)
>
> "*Wider determinants of health mean that health inequalities worsen.*" (e11 Research)

Investment: Inadequate capital investment

SMEs suggested that delayed capital investments into improving the quality of NHS estate and technologies would be a limiting factor on productivity growth. This focused on the continued practice of sacrificing the longer-term benefits that this investment can bring to productivity and capacity to instead achieve shorter-term survival goals, such as supporting operational funding and propping up the existing estate.

> "*Capital investment failing to keep up with service infrastructure needs*." (e09 Policy)
>
> "*Longer term constraints on capital investment start to bite and the productive capacity of the NHS estate starts to reduce."* (e10 Policy)
>
> "*Potential for doom loop scenarios with capital investment sacrificed to maintain aging estate*." (e08 Policy)
>
> "*If we do not improve estate then we cannot do much more*." (e05 Frontline)
>
> "*Capital budgets continue to be raided.*" (e02 Frontline)

Implementation: Weak adoption of technology

Challenges with the implementation of new technology was frequently highlighted by SMEs. Some rationales included further details about the obstacles to successful implementation. This included possible resistance from staff or the public to adopting new technology, as well as inadequate change management capacity and capability – in relation to the necessary skills and investment

required. In other instances, SMEs suggested that the adoption of technology may not translate into productivity gains or that new investments would be delayed.

"*Current trends such as the massive cut in management capacity paralyses change capacity…There is a lack of focus on implementation and developing the human capital needed to make the most of opportunities such as AI.*" (e02 Frontline)

"*Reluctance of staff to engage with new technologies (e.g. AI interpretation of diagnostic tests) or lack of skills/training for staff to be able to use technologies effectively.*"(e11 Research)

"*Reluctance of public to engage with new technologies*." (e11 Research)

"*Lot of evidence that move to digital can slow staff down.*" (e05 Frontline)

"*New technologies require more staff time than expected (e.g. managing patients on virtual wards, dealing with patient queries about results from at-home monitoring).*" (e11 Research)

Other factors described as barriers to productivity included (but were not limited to) the disruption caused by structural reorganisation; difficulties in moving the point of delivery of care; increased vulnerability to cyberattacks; and political instability.

P90 rationale themes

Unsurprisingly, many of the enablers highlighted by experts were effectively the opposite of the barriers they had described in their P10 rationales.

Widespread adoption of technology

Some of the rationales described a future where technology would be successfully implemented resulting in reductions in the administrative burden on staff and increases in clinical outputs. However, it was noted by several SMEs that a significant proportion of these productivity gains may

"*Technology and skill mix changes allow tasks to be completed at lower pay grades or with less time. (e.g. AI interpretation of diagnostic tests).*" (e11 Research)

"*I feel that the return on investment (not used in a technical way) in capital and technology, including digital and automation, will take time to yield the expected increases in outputs, as Trusts and healthcare providers in general learn to use (steep learning curve in some cases for front-line staff) these new tools, and change established working practices*." (e13 Research)

"*Implementation of technology - particularly ambient AI - massively reduces unproductive activity unequally across sectors.*" (e02 Frontline)

not be experienced until the 5-10 years period. One expert also made explicit their belief the impact of technology would not be equal across the different health sectors.

Faster discharges and delivering care in lower cost settings

In several rationales, productivity gains were envisaged to be delivered through faster discharges from hospital and more care being delivered in outpatient settings or in the community. While in some instances this was the result of technology changes and/or medical advancements, in others it was due to factors such as neighbourhood services becoming established, improved operational processes, cultural shifts and the alignment of financial incentives.

> *"The NHS and public embrace the shift to community, primary care and prevention which enables more activities to switch to lower cost settings which deliver equivalent/improved outcomes." (*e10 Policy*)*
>
> *"Faster discharges and more care delivered at home once neighbourhood services established." (e04 Frontline)*
>
> *"Diversion of some demand via digitally guided self-care." (e04 Frontline)*

Leadership and delegated decision making

Some participants emphasized the role that leadership – particularly clinical leadership – could play in driving productivity gains. This impact extends beyond direct innovation to include indirect benefits such as improved staff retention and wellbeing. Closely related to this was a shared view that devolving decision-making authority and reducing the oversight burden would be critical enablers in realizing these benefits.

> *"Upper limits are only likely to get anywhere near being achieved if there is significant devolution of decision-making authority… move away from current very centralised, hierarchical decision making." (e02 Frontline)*
>
> *"Leadership – particularly clinical leadership – will not only help with implementation and innovation but also retention and wellbeing." (e05 Frontline)*

Some of the other enablers included in rationales were improved NHS estate; better workforce retention; changes to information governance; improved data quality; and regulatory changes bringing about a greater appetite for clinical risk.

Time Profiles

10 out of 14 experts predicted that the annual productivity rate would be higher in the 5-10 years period than the 0-5 years period. While differences in forecasting across the time periods was not made explicit in all rationales, some SMEs indicated that they believed it would take time for productivity gains to be realised following implementation of new technology and capital investment. However, 4 SMEs predicted that productivity would be higher in 0-5 years period. For one SME this was due to increasing patient complexity making productivity gains more difficult in the 5-10 years period. For the other 3 SMEs, a "bounce-back" period will achieve higher growth in productivity in the first 5 years which won't be sustained the following 5 years.

"I think this should be split between an underlying rate and catch up…There is the possibility of high years of one-off productivity gains…high capital investment/measuring new activities/admin reductions which cannot be repeated" (e07 Policy)

"Some recovery from the recent declines and slower progress after this has happened." (e12 Research)

Discussion

This exercise has produced a distribution of pooled forecasts quantifying expert uncertainty related to NHS productivity over the next 10 years. The design of the elicitation provided insight into experts' judgements about the trajectory of productivity growth by gathering predictions for 0-5 years and 5-10 years. Experts also submitted rationales alongside their forecasts describing the key barriers and enablers to productivity. The results can now be used in the Health Foundation Real Centre's workforce modelling and to inform the work of the Productivity Commission.

The pooled results show that experts predicted a modest short-term (0–5 years) productivity growth of 0.7% per year (-1.1% to 2.4%). They anticipated ongoing challenges in funding, supply struggling to meet demand, and the slow adoption and therefore delayed impact of technology. According to the experts' pooled forecasts there is an 85% probability that the annual productivity growth rate will be less or equal to 2% over the next 5 years. This suggests that the 2% NHS productivity target set by the government - which was recently reaffirmed in the 10 Year Plan as a target to be achieved over the next 3 years[12] – is an ambitious one. However, given that the 2% figure is below the pooled P90 value (2.4%), our results indicate that the target is a plausible level of productivity growth.

In the longer-term (5–10 years), experts expected annual productivity rates to rise to 1.4% (-0.4% to 3.5%), driven primarily by the introduction of new care models and AI. However, this depends on successful implementation and adoptions of each of those factors, alongside the necessary investments in people (leadership, culture, skills and capacity) and infrastructure (tech and estates) that are required to enable this.

Strengths and Limitations

Measuring healthcare productivity is complex. Much of the workshop focused on defining productivity and presenting the best available data. We chose to use the ONS measure of non-quality adjusted productivity over the quality-adjusted version, partly due to concerns about how quality adjustments are currently applied, which could complicate an already challenging forecasting task. However, we acknowledge that the non-quality adjusted measure does not fully capture the impact of healthcare on patients and experts also emphasized that increases in this measure do not necessarily reflect greater value in care delivery. This is suggested as a potential area for exploration in a subsequent elicitation exercise.

A key strength of the study was the high level of expert engagement achieved through the facilitated workshop. Each step was treated as formal data collection and supported by multiple

[12] https://www.gov.uk/government/publications/10-year-health-plan-for-england-fit-for-the-future p.16.

learning methods, including synthesised evidence, structured discussions, and interactive tools that combined data, explanations, and pooled results.

The recruitment of participants was undertaken with the intention of capturing a diversity of expert perspectives on future healthcare productivity. SMEs were assigned to three teams (Frontline, Policy and Research) based on their professional expertise although each expert provided their own forecasts. Teams Frontline and Policy consisted of 'inside' experts – each having worked within NHS England or DHSC while Team Research consisted of 'outside' experts. The results from the three teams were comparable.

Conclusions and recommendations

An expert elicitation exercise has produced pooled forecasts of NHS productivity rates over the next 10 years. At the outset of this exercise the level of expert uncertainty was unknown. The fundamental purpose of the exercise was therefore to quantify the uncertainty of experts and make it explicit using a systematic, evidence-based process. The intention was not to change the uncertainty of the participating experts. As such, the exercise must be judged on the rigour of its method rather than the range of uncertainty it reveals. For the underlying uncertainty to be reduced, policy makers will need to articulate (and demonstrate through robust evaluation) the mechanisms that will generate productivity gains within a complex system such that a greater consensus can emerge on their effects.

Given the overlap across the results of the three teams, we recommend that the pooled prediction intervals of all experts be used. These pooled forecasts may now be used to inform the setting of parameters in the Health Foundation Real Centre's workforce modelling.

Having completed an exercise using the non-quality adjusted measure of productivity, a future exercise could seek to elicit expert forecasts of quality adjusted productivity rates.

Acknowledgements

Participants

The authors would like to express their gratitude and thanks to the participants (see below) who took part in the elicitation workshop on the 2 July 2025.

Name	Job Title
Alastair Brodlie	Operational Research Analyst, Department of Health and Social Care
Janos Suto	Deputy Director, Urgent and Emergency Care Directorate Department of Health and Social Care
Rob Unsworth	Deputy Director, Efficiency Strategy, NHS England
Seamus McGirr	Self-employed healthcare management consultant. Former Director of Integrated System and Clinical Analytics, East Lancs ICS
Katharine Halliday	GIRFT lead for diagnostic capacity and former President of The Royal College of Radiologists
Adriana Castelli	Health Economist, Centre for Health Economics, University of York
Martin Weale	Professor of Economics, King's College London
Pete Scolding	Clinical Director of Stewardship, Mid and South Essex Health and Care Partnership
Luisa Pettigrew	Senior Policy Fellow, The Health Foundation
Emily Hopkinson	Senior Policy Manager, Healthcare Financial Management Association
Sara Zella	Head of Public Service Productivity Education, Social Care, and Annual publications, Office for National Statistics
Tanya Pankhurst	Consultant Nephrologist - Medical Chief Medical Officer, University Hospitals Birmingham NHS Foundation Trust
Michael Dimov	Deputy Director of Community Health Service Improvement, NHS England
Johannes Wolff	Associate Director of Costing, Data and Analytics, Guy's and St Thomas' NHS Foundation Trust

We are also grateful to Juney De Aguilar at the Health Foundation for her support delivering the workshop.

The Strategy Unit's Elicitation Project Team

Mohammed A Mohammed, Emeritus Professor & Principal Consultant m.mohammed2@nhs.net

Rhian Davies, Senior Data Scientist rhian.davies25@nhs.net

Josh Cook, Consultant josh.cook@nhs.net

James de Lacy, Senior Consultant james.delacy@nhs.net

Jennifer Wood, Managing Consultant jennifer.wood43@nhs.net

Gabriel Hobro, Senior Analytical Manager, New Hospitals Programme, NHS England
gabriel.hobro@nhs.net

The Health Foundation Project Team

Jack Elliott, Economic Analyst Jack.Elliott@health.org.uk

Katie Fozzard, Senior Economic Analyst Katie.Fozzard@health.org.uk

Stephen Rocks, Head of Secretariat of the NHS Productivity Commission
Stephen.Rocks@health.org.uk

Yi Mu, Economic Analyst Yi.Mu@health.org.uk

Appendices

Technology-enabled scenario

The technology-enabled scenario was a less formal elicitation exercise consisting of one round of data collection. It was conducted during the workshop but after the completion of the core exercise. Comparisons between the results of tech scenario and the core exercise are subject to limitations because of this methodological asymmetry.

Moving from analogue to digital is one of the governments three radical shifts in the 10 Year Plan driven by the potential for technology to transform healthcare. Therefore, we conducted a subsequent exercise to better understand this.

Following the closure of data collection for the core exercise, SMEs were introduced to the technology-enabled scenario. They were asked to imagine a scenario where the government and health service **provide funding**, **secure investment** and **support the adoption of technology**. SMEs were asked to re-estimate their productivity trajectories on this basis.

SMEs were presented with summaries of how technology, data and AI adoption could potentially improve back-office processes and enable more care to be delivered at home and in the community.

SMEs were allowed to ask clarifying questions before being given 20 minutes to independently provide their prediction intervals and rationales in silence. There was no opportunity for SMEs to review peer forecasts or discuss enablers and barriers for this scenario. The tech scenario consisted of one round of data collection. The pooled results are from that single round. Due to unforeseen circumstances, one SME had to leave the workshop after completing round two of the core exercise but prior to beginning the tech-enabled NHS (tech) scenario.

Tech-scenario Results

The table below shows the pooled forecasts for the technology-enabled scenario.

Period	All (n=13) Pooled		Frontline (n=5) Pooled		Policy (n=4) Pooled		Research (n=4) Pooled	
	P10 to P90	mean	P10 to P90	mean	P10 to P90	mean	P10 to P90	mean
0-5 years	-0.4% to 3.7%	1.4%	-2.0% to 4.3%	1.3%	0.4% to 3.9%	1.8%	0.1% to 2.5%	1.2%
5-10 years	0.2% to 6.9%	3.0%	-0.5% to 10.2%	4.1%	0.3% to 5.4%	2.7%	0.4% to 3.8%	2.0%
0-10 years (derived)	0.0% to 5.0%	2.2%	-1.0% to 7.0%	2.6%	0.4% to 4.4%	2.3%	0.2% to 3.1%	1.6%

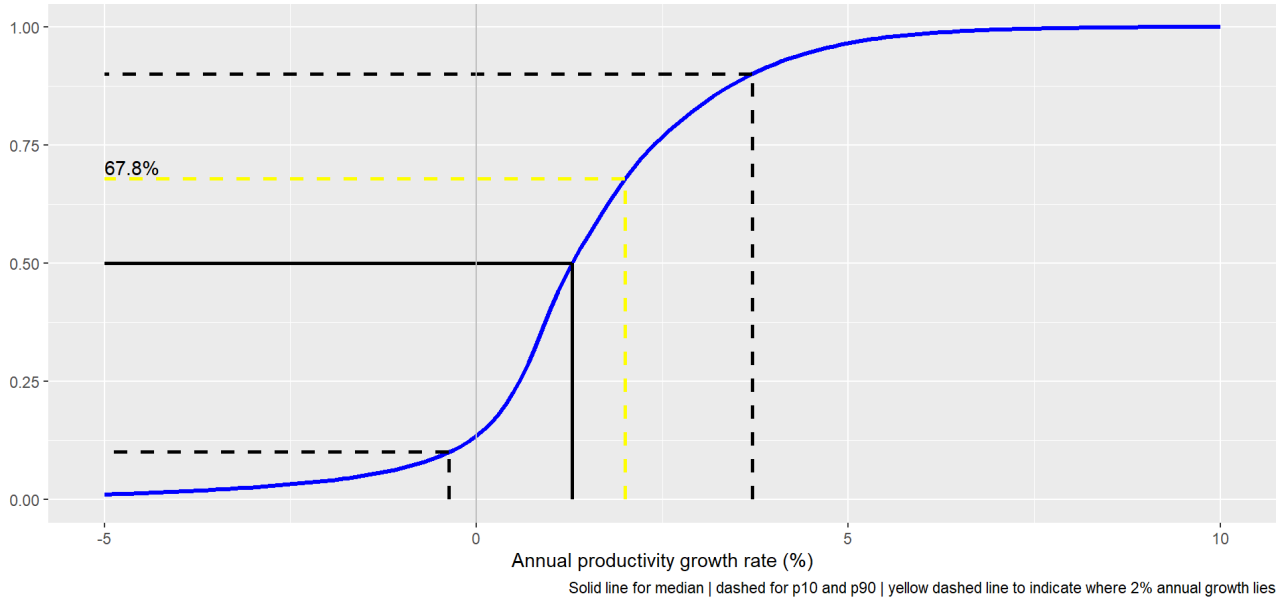


Figure 7 Cumulative distribution function or S-curve showing the pooled forecasts for the 0-5 years period for the tech scenario

The median (P50) value of all experts pooled 0-5 years forecasts for the tech scenario was 1.3%. Plotted on the yellow dashed line is the NHS productivity target of 2%. This falls between the pooled P10 and P90 values (-0.4% and 3.7%). There is approximately a 68% probability that the productivity rate will be less than or equal to 2% according to the pooled forecasts of all SMEs.

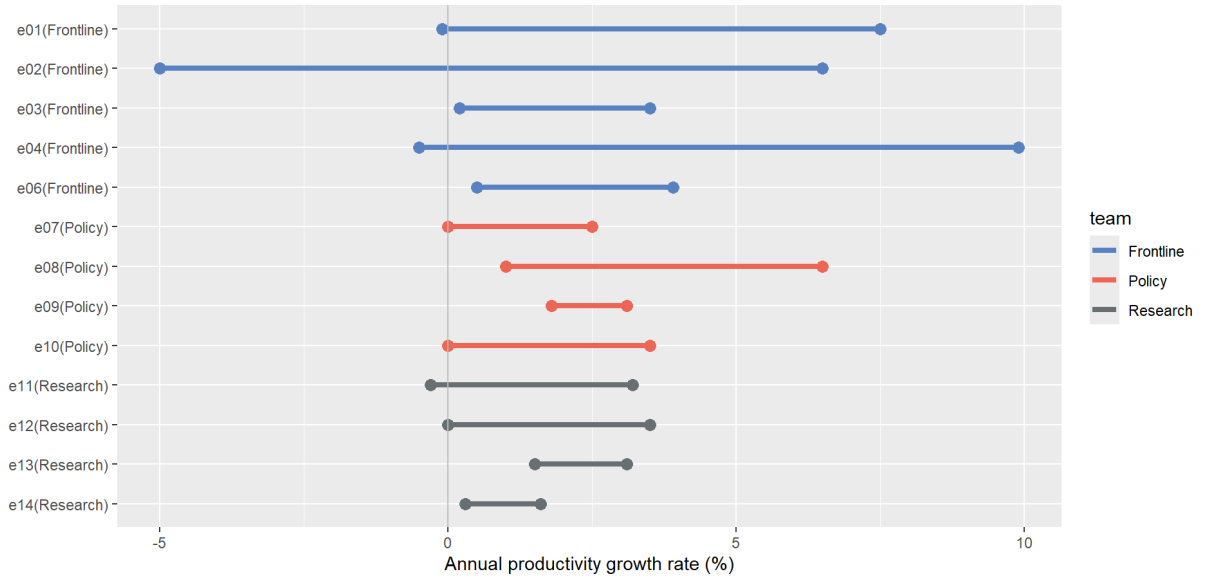


Figure 8 Individual expert forecasts 0-10 years for the tech scenario

The only prediction intervals for 0-10 years which did not overlap were those of e09 (Policy) (1.8% to 3.1%) and e14 (Research) (0.3% to 1.6%). SME e14 (Research) predicted the lowest P90 value (1.6%) while SME e04 (Frontline) predicted the highest (9.9%).

Tech-scenario rationales

P10 Rationales

Cost of investing in new technology negates productivity gains

One SME highlighted the risk of NHS providers having to run old systems and services at the same time as those being introduced with new technology. Here the cost of 'double running' would diminish the productivity gains brought about by new technology. Others suggested that the cost of new technology and innovation could overshadow the savings from reducing staff costs, or that poor decisions may be made about which technology to invest in due to a lack of quality evidence or because of financial incentives.

> *"Even though new tech is delivered and rolled out on time and to budget, existing processes and systems are not decommissioned quickly enough and trusts end up double running systems and services."* E10 Policy
>
> *"The cost of new technology may outweigh the capital and revenue savings from retiring old systems. Investment in innovation and increased output may outpace any financial benefits from reducing staff costs."* E06 Frontline
>
> *"Weak evidence base and/or financial incentives that push people to spend money quickly meaning we do not invest in the most effective digital technologies. (In the absence of evidence, commercial interests drive choices.)"* E11 Research

P90 rationales

The 'surprisingly high' (or P90) forecasts in the tech scenario varied significantly in their level of optimism. This was reflected in the rationales produced by SMEs which revealed marked differences in the extent to which widespread adoption of technology could impact future healthcare productivity.

Tech transforming healthcare delivery

For some SMEs, the conditions of the tech-enabled scenario could have a truly transformative impact on the way care is delivered in the NHS. Technology would enable dramatic changes in diagnostics, the volume of remote care delivered and substantial changes to the workforce.

> *"A true tech revolution - if enabled by cultural change - could see a rapid change in workforce. This has happened in other industries, with swathes of low paid jobs being replaced by technological alternatives."* (e08 Policy)
>
> *"0-10yr - widespread tech which has enabled self-care, remote care, and automated everything that was automatable in healthcare!" (e*04 Frontline)
>
> *"Point of care testing with instant results reduces contacts needed and deals with things faster."* (e01 Frontline)
>
> "*Admin time in the NHS reduces dramatically resulting in less admin workforce with no effect on output*" (e10 Policy)

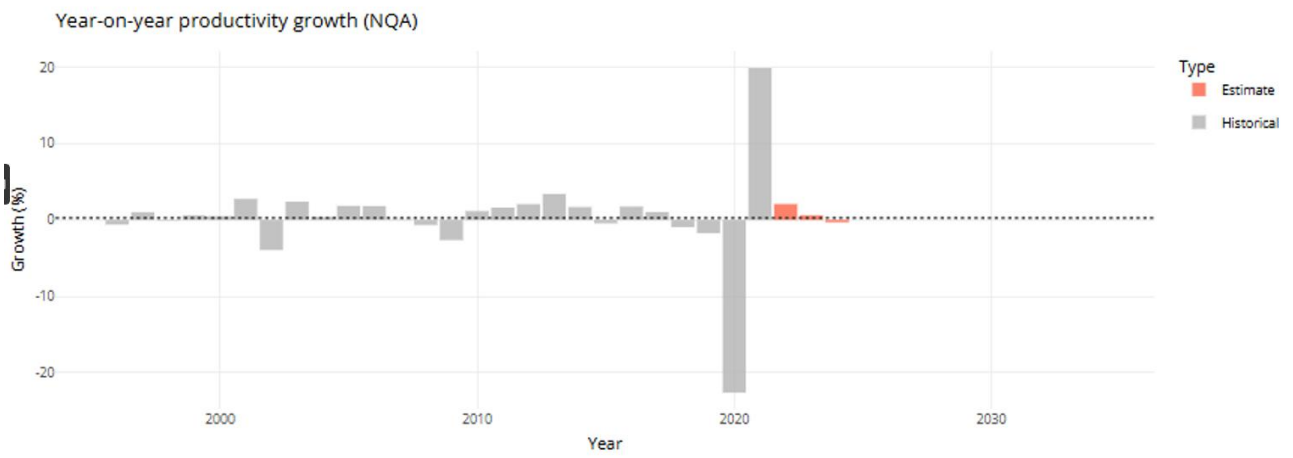
Different interpretations of the scenario

Some of the rationales suggest that there may have been differences in the way in which the conditions of the scenario were understood by SMEs. This may explain the variation in the degree of optimism. For instance, SME e06 (Frontline) made explicit that they were not exploring the possibility of AI replacing diagnostic and clinical tasks. However, other SMEs clearly were factoring this into their judgements. SME e09 (Policy) stated that their forecasted productivity values for the tech scenario were based on the delivery of healthcare activity by AI and that being measured as equivalent to activity delivered by humans.

> *"One scenario not yet explored is the potential for AI to replace diagnostic or even clinical tasks. While still in its early stages, if the tech industry's predictions hold true, such capabilities could emerge well within the 10-year horizon we're considering. If we reach a point where a Dr ChatGPT prompt counts as a legitimate output, the implications for productivity could be very significant*." *(e06 Frontline).*
>
> *"ASSUMPTION: AI delivered activity is captured and counted as equivalent to human activity." (e09 Policy).*

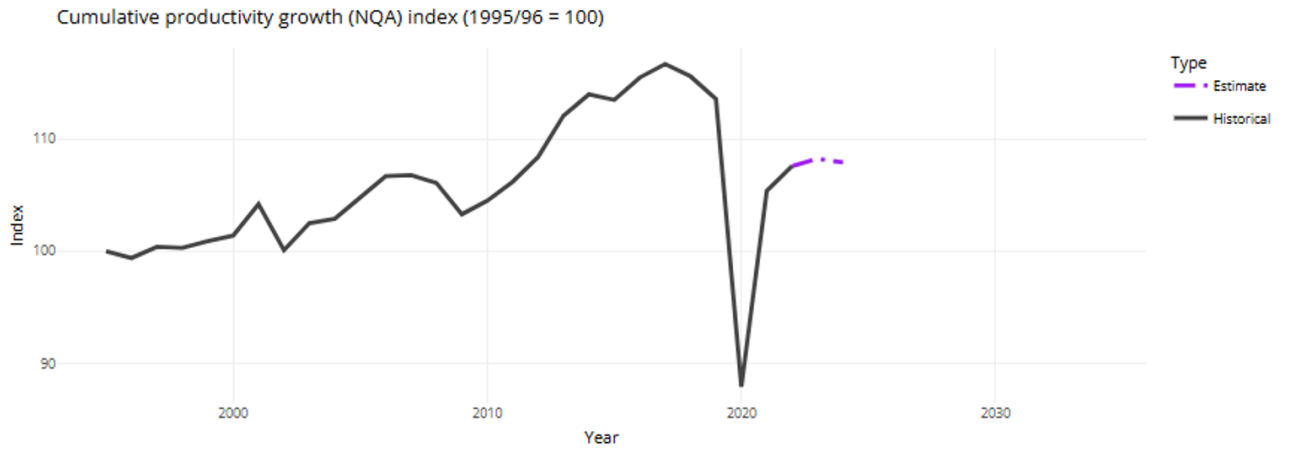
Non-quality adjusted productivity trend data presented to SMEs in the elicitation app

Trend data



Source: ONS, Public Service Productivity, estimates after 2022/23 are based on UK quarterly public sector productivity estimates (experimental).

NQA: non-quality-adjusted, CAGR: Compound annual growth rate



The Strategy Unit
Email: strategy.unit@nhs.net
Web: www.strategyunitwm.nhs.uk