NPL
National Physical Laboratory

NPL REPORT IEA 20

# EVALUATION FRAMEWORK FOR THE NATIONAL MEASUREMENT SYSTEM: A FRAMEWORK FOR ASSESSING THE IMPACT OF NMS LABORATORIES

DR MIKE KING, SOLOMON OLAKOJO, GABI GORRINGE

October 2023

# Evaluation Framework for the National Measurement System: A Framework for Assessing the Impact of NMS Laboratories

Mike King, Solomon Olakojo, Gabi Gorringe
Analysis and Evaluation, Partnerships Directorate

## SUMMARY

The document presents the Evaluation Framework for the National Measurement System (NMS) and the UK Measurement Strategy. The framework provides the over-arching approach for the evaluation of the NMS to guide the monitoring and evaluation activities to be undertaken by organisations that comprise the NMS. It presents a description of the role of the NMS to set the context for the evaluation and provides details of the primary evaluation methodology and the data collection and analysis methods that will be used to provide evaluation evidence. Each laboratory within the NMS will develop its own evaluation plan to guide its monitoring and evaluation work. Such plans will be based around the common NMS evaluation framework but should be in-line with the scale and content of their specific activities.

ISSN: 2633-4194

DOI: https://doi.org/10.47120/npl.IEA20

National Physical Laboratory
Hampton Road, Teddington, Middlesex, TW11 0LW



Approved on behalf of NPLML by
David Skelton, Strategic Programme Leader.

## CONTENTS

## 1. INTRODUCTION

The reliability and acceptability of the UK's physical, chemical, and biological measurements is maintained by a series of high-level measurement laboratories funded by the Department for Science, Innovation and Technology (DSIT). This document is an evaluation framework for assessing whether this programme of funding - knowns as the National Measurement System or the NMS - is generating outcomes as anticipated and the extent to which its impacts constitute value for money. As this programme is ultimately funded by UK taxpayers, the costs and benefits will be assessed from the perspective of the whole United Kingdom, rather than just the direct customers of these high-level measurement laboratories.

This document provides a framework and general principles for designing and implementing evaluation activities. These principles will align with guidance in HM Treasury's Magenta Book, as well as the overarching evaluation strategy of the DSIT. This document seeks to take the guidance supplied by these documents and tailor it to the specific nature of the NMS.

Lastly, we look to review this evaluation framework for the NMS every five years; which allow sufficient time for any changes to the depth and scope of our evaluation activities to take effect.

## 2. THE NATIONAL MEASUREMENT SYSTEM

Every time you use your GPS, put petrol in your car, or receive a medical diagnosis, you are putting your trust in measurements that are underpinned by a system that ensures they are both reliable and internationally recognised. The UK National Measurement System (NMS) consists of a core infrastructure of measurement laboratories and a wider community of service providers that ensure you can have confidence in the measurements you make or are made on your behalf. Our reliance on the NMS is often overlooked, but, like many of our infrastructures such as roads or water, it would soon be noticed if it did not work. Hence, the fact that this infrastructure is invisible to many is indicative of its success. Nonetheless, our economy, our quality of life, and occasionally our very lives depend on the robust and reliable measurements that this infrastructure enables.

The UK, like all developed nations, has a national measurement infrastructure that ensures a robust system of measurement and forms an essential component of being part of a global economy. At its core, the NMS ensures that the measurements made in the UK are consistent with the global common system of measurement units: the International System of Units – the SI (Système international d'unités).

The common SI system of units underpins much of the daily use of measurement in the UK. However, there remain areas of measurement such as in chemistry, biology, and food science with somewhat more tenuous links to the SI system of units. Primary measurements underpinning these areas are established and disseminated through reference methods and/or materials. These areas are often where new measurement knowledge needs to be developed to tackle emerging needs or challenges.

Internationally, each country has one National Measurement Institute (NMI), whose role is to take the lead in international representation and to underpin delivery of a measurement infrastructure consistent with the SI system. In most countries, there are one or more Designated Institutes (DI), that support the NMI by delivering specific measurement

capabilities and are recognised internationally as the lead measurement organisation in that country for a particular quantity.

**The NMS Laboratories**

In the UK, the National Physical Laboratory (NPL) is the UK's NMI and works in partnership with six designated institutes:

- NML (National Measurement Laboratory at LGC) – designated for chemical and biometrology
- NEL (National Engineering Laboratory) – designated for fluid flow metrology
- OPSS – responsible for legal metrology
- NGML (National Gear Metrology Laboratory) – designated for gears metrology
- NIBSC (National Institute for Biological Standards and Control) – designated for bioactivity metrology
- NIAAM (National Institute of Airborne Acoustic Metrology) - designated for airborne acoustics and based at the University of Salford.

Such is the importance of measurement to the economy and society that organisations such as NPL and NML (formerly LGC) have been in existence for more than 100 years.

## KEY ROLE OF THE NMS

The UK Measurement Strategy (2022)[1] is based on broad themes that have been used to construct a logic model for the NMS. These themes are based on mechanisms highlighted within the UK Measurement Strategy and summaries of the economic evidence for the NMS compiled during successive government spending reviews. To a first approximation these themes can be put into two broad categories:

- Enabling technological change and innovation:
  - Performing applied research to develop instrumentation, reference data, and measurement protocols.
  - Helping organisations to innovate by providing direct support through collaborative R&D projects.
- Maintaining confidence in measurement:
  - Helping organisations to move closer to the technological frontier by supplying consultancy and training services.
  - Underpinning the national quality infrastructure by providing the technical basis for trade and standards.
  - Supporting local authorities to apply weights and measures legislation

The sections below describe these themes in more detail and provides examples of how the NMS is used in practice. Further examples are provided in an evidence document that accompanied the previous UK Measurement Strategy (2017).

1 The UK Measurement Strategy and evidence document can be found at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1125285/UK_measurement_strategy_2022.pdf

## ENABLING TECHNOLOGICAL CHANGE AND INNOVATION

### Scientific Research and Technical Infrastructure

The universities perform curiosity driven research and sometimes this leads to radically new technologies that revolutionise large parts of the economy and displace the established ways of working. Examples of these technologies might be artificial intelligence, engineering biology, and quantum computing. However, the development of such technologies requires measurement tools, reference data, and protocols to aid the efficiency of R&D activities by improving the reproducibility of results. This technical infrastructure helps to unlock significant private investment by reducing uncertainty about the viability and performance of the technology.

The government funds the NMS laboratories to perform the research needed to develop and extend this technical infrastructure; and the capabilities this generates form the basis of leading-edge products and services supplied by the NMS laboratories. Typically, the users will be firms operating near the technological frontier and, for a time, the NMS laboratories offer a bespoke service used by a small number of innovative firms.

However, once demand for such a service begins to grow, the role of the NMS laboratories starts to change. Rather than supply the service to end-users (e.g., high tech firms), the NMS laboratories work with the commercial calibration labs who can provide the service much more efficiently and cheaply by using a factory-like system to exploit economies of scale.

***Providing Confidence in measurement***

*Flow meters are the cash registers of the energy industry, raising billions for the treasury every year in fuel duty. To ensure accurate measurement, these meters are calibrated every year at a cost of tens of thousands of pounds. For over 40 years, NEL has led the research and development of ISO 5167, the key standard for differential pressure (DP) meters covering design and use which, when used correctly, means that these meter types do not require calibration, reducing costs for industry while ensuring accurate performance. This standard has been exclusively developed as a result of NEL research for the past 40 years, including a recent update to all parts in 2022.*

***Measurement supporting emerging technologies***

*NPL's Quantum Metrology Institute aims to provide leading-edge measurements for verifying the performance of new quantum technologies.*

*A recent application concerns Quantum Key Distribution (QKD). Key exchange is at the heart of cryptography for secure communication. This distribution of keys using photons over an optical network can provide an extra layer of data security which is inherently future-proofed. Toshiba Research Europe, BT, ADVA Optical Networking and NPL performed the first demonstration of QKD and data encryption over a single field-installed fibre, with NPL providing traceable calibration of the QKD optical system. This showcased the technology and highlighted the importance of traceable measures for system verification.*

Finally, the research performed by the NMS laboratories sometimes generates patents and know-how connected with instrumentation. Granting access to this intellectual property through licenses and collaboration agreements stimulates a second round of research activity in the public and private sectors. Moreover, once any commercial opportunities have been exploited, the knowledge is also made available to external researchers through the publication of journal articles.

**Knowledge Transfer and Innovation**

The NMS laboratories support innovation by encouraging the diffusion of knowledge and best practice. This knowledge transfer activity takes three main forms: Firstly, the NMS laboratories produce good practice guides that are accessible to measurement users and provide practical guidance for engineers or technicians working in the public or private sector. The most important of these guides describes how to work out the uncertainty associated with a measurement process and function much like a technical standard. The ability to estimate uncertainties is important for ensuring that a production process is operating within the required tolerances and can produce parts to a certain specification. If a company is unaware that its production process is outside the required tolerance, then much of its output will be defective and the defective parts will be scrapped or reworked, causing delay and increased costs.

The NMS laboratories offer a range of training and educational services. Firstly, the laboratories work with universities to develop PhDs that attract interest and funding from industry. Furthermore, PhD students with an interest in industrial research are more likely to pursue industrial careers and thereby to directly transmit leading edge thinking to their employers. In addition, the NMS laboratories offer training for technicians through a mixture of online resources and classroom learning (Olakojo and King, 2023). Training ranges from bite-sized introductory modules (often provided free of charge) to more substantial hands-on (paid for) courses, apprenticeships, and industrial PhDs. Moreover, some of the courses are accredited by established exam boards and so lead to qualifications, which enhances the career prospects of those that pass. Therefore, the NMS laboratories work to increase the pool of skilled people and thereby make it easier for employers to recruit staff while boosting the wages of individual learners.

Finally, the NMS laboratories are based across the UK and the larger labs are establishing regional hubs to ensure regional access to NMS expertise. These hubs raise awareness of what the NMS laboratories can offer and gather information about the measurement needs of local organisations. Furthermore, some of the NMS laboratories have staff dedicated to visiting manufacturers to offer advice about how to update their measurement and quality control processes in support of improved productivity. These visits are, typically, short and combine an audit with specific practical advice. These activities are directed at aiding the diffusion of innovations rather than innovation itself. Nonetheless, it's through diffusion and uptake of innovations that the economic and social benefit is realised.

In addition to the three main forms of knowledge transfer discussed above, experts from the NMS laboratories undertake and participate in workshops, seminars, and conferences to disseminate the outputs of their research to stakeholders and potential users in business, academia, and public sector communities (Olakojo and King, 2023).

## UNDERPINNING TRADE AND REGULATION

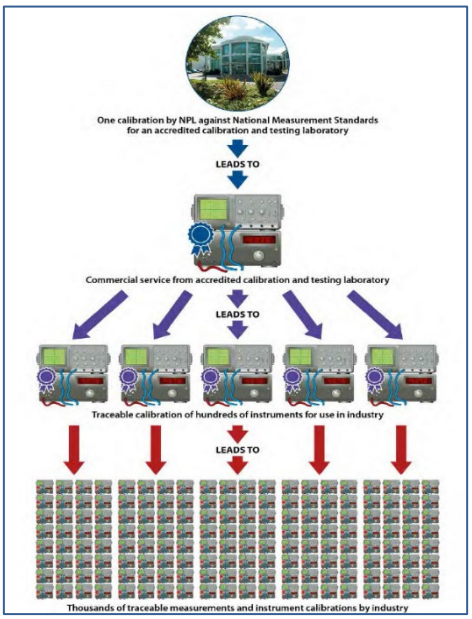
**Maintaining confidence in the nation's technical infrastructure**

The NMS laboratories offer a range of products and services that were developed through earlier research projects. These products and services draw on primary standards - often embodied in a piece of laboratory apparatus - that require regular maintenance and updating. Such maintenance activity is largely focussed on updating accreditation, software developments and performing the key comparisons and proficiency testing needed to ensure measurements performed in the NMS laboratories are comparable to measurements

performed in equivalent institutions in other countries. Such key comparisons are at the heart of international mutual recognition arrangements (MRAs) that ensure that a traceable measurement performed in the UK is acceptable in another country and does not need to be duplicated. These arrangements, along with product certification (e.g., CE marking), are thought to significantly lower non-tariff barriers and so support international trade.

The NMS laboratories offer products and services to public and private users. Customers send their instruments to the NMS laboratories for calibration against a primary standard to ensure the bias and precision of their instrument is known, enabling them to use it appropriately. The main products supplied by NMS laboratories are reference materials and reference artefacts and traditionally the main service is the calibration of measuring instruments. A reference material or reference artefact will have known characteristics (such as chemical composition or geometrical properties) and customers use this to calibrate their own instruments.

In terms of calibration services and reference materials or artefacts, the key users of the NMS laboratories are commercial calibration companies. By taking traceability from the NMS laboratories, these calibration companies disseminate reliable calibrations to their customers. In short, these companies are the next link in a chain that links measurements done by the public and private sectors to the primary standards maintained by the NMS laboratories. Along with a system of protocols, accreditation, and international agreements, this chain of interlinked calibrations ensures that a measurement can be performed once but accepted everywhere. Domestic and foreign confidence in this system avoids duplicating tests and circumvents disputes, which reduces the cost of verifying that goods conform to specification, as well as, aiding consistent regulation and taxation.




This system of traceable measurements tends not to cover emergent measurement needs. In particular, new technologies, new markets, and regulatory reforms mean that the nation's measurement infrastructure needs extending and updating. A major step towards this goal is the development of new technical standards by industry representatives sitting on the committees of BSI (the UK's national standards body). Scientists from the NMS laboratories support these committees by drafting the measurement protocols that give documentary standards traction over products and processes. In particular, agreeing these protocols is a prerequisite for the accreditation and conformance testing that enables the certification of goods and services and in so doing grows markets through commoditisation.

***Accurate measurements for manufacturing***

*Rolls-Royce purchases many calibrations a year from NPL to ensure that the measurements it makes are traceable back to national standards and therefore robust and reliable. The company uses over 200,000 measuring instruments in the production of its engines and the entire life cycle of the company's products and services is underpinned by measurement.*

*Rolls Royce also collaborates with NPL in a wide range of measurement disciplines to access innovations in measurement methods and technologies to support continuous improvement in product and business performance.*

Confidence in measurement data, assured via the calibration chain, also supports effective and efficient implementation of public policy and regulation. The NMS underpins, for example: accurate air and water quality data; the measurements that ensure delivery of safe and effective radiotherapy doses to cancer patients; the measurements to protect workers from hazards such as radiation (King & Renedo, 2019).

**Helping local authorities to enforce weights and measures legislation**

Weights and measures legislation (legal metrology) and the work of trading standards officers exists to ensure fairness and maintain the confidence of buyers. The measuring instruments governed by weights and measures regulation range from supermarket scales to railway weighbridges, beer glasses to fuel dispensers, and taximeters to utility meters. Where goods are sold on the basis of volume, length or weight, sellers are required to use approved instruments and to calibrate them appropriately. In cases where goods are prepacked and sold on the basis of an average quantity, the distribution of quantities across a large number of packets must conform to rules governing the tolerable negative error. The Office for Product Safety and Standards (OPSS) provides certification for new measuring instruments (type approval) and has ultimate responsibility for many of the standards applied by local authorities during their enforcement activity. Moreover, the expectation is that improvements in how these responsibilities are discharged enable both more effective regulation and more cost-efficient compliance by end-users performing day-to-day measurements. Local authorities also provide traceability to local businesses using their standards traceable to the primary standards at the NMS laboratories.

The benefits of legal metrology can be seen by considering what might happen if the rules were relaxed. Firstly, suppose that a small independent petrol station buys some nominal volume of fuel from a large supplier each week, but the new rules permit significant variation in amount of fuel delivered. For simplicity, imagine that the retailer receives too much on some occasions and too little on others so that over the course of a year the gains and losses cancel out. Nonetheless, if the retailer's profit margin is small and liquidity is a concern, then this weekly variation in the volume of fuel will cause them economic harm. Weights and measures legislation helps to keep this harm to a minimum by putting strict limits on the level of variation that's allowed. (Remember that as the supplier makes many sales each week the gains and losses will cancel out ensuring that there is very little variation at the aggregate level. In short, the large supplier is not affected by the variation in individual deliveries, but their customers are affected.)

***Preventing counterfeits***

*LGC produced the first of a set of three reference materials (RMs) certified for carbon isotope ratios traceable to the SI unit of mass, the kilogram, through calibration standards of known purity. Isotopic analysis forms a part of the authentication process for confirming the origin of food and drink products with Protected Denomination of Origin (PDO) status. It can also be used to disrupt the enormous global market in counterfeit pharmaceuticals which was estimated in 2012 to be worth close to $200 billion, with over 800 counterfeit product lines in circulation worldwide.*

Finally, consider what might happen if local authorities were to significantly reduce their spending on enforcement activity. In this situation, a minority of sellers will be tempted to offer short measures and to systematically overcharge customers. The concern is that honest sellers will be undercut by the rogue sellers. The danger is

that honest sellers can't compete and are forced out of the market. In short, without adequate enforcement, the 'good' sellers are driven out of the market by the 'bad' sellers.

## 3. INTRODUCTION TO PROGRAMME EVALUATION

**What is Evaluation?**

As this document is an evaluation framework for the NMS, it's helpful to begin by describing what we mean by evaluation. There are essentially two types of relevant evaluation activity:

- An impact evaluation seeks to understand the difference that the programme made. That is, the difference between what is observed to have happened with the public support and what would have happened anyway without this public support. In short, impact analysis is about working out how much of what we observe (e.g., total R&D spending by a certain set of companies) is truly attributable to providing public support. Lastly, impact evaluation will include a mixture of quantitative and qualitative methods. The quantitative analysis will focus on finding statistical evidence that the programme has had a causal effect on some outcome of interest; and the qualitative analysis is concerned with why and how this effect was achieved.
- An economic evaluation is concerned with comparing the benefits attributable to the programme to the public and private expenditure attributable to the programme. These benefits can be economic or social and may not be easy to monetarise even if the outputs are countable. Typically, this type of cost-benefit analysis builds on the findings of an impact analysis and takes expenditure figures from a process evaluation.

**Why Perform Evaluation?**

Evaluation work is necessary to ensure accountability, as well as to enable future funding to be allocated based on evidence of what worked well in the past. Strategic decision making about how to design a programme and allocate resources should be informed about how it is functioning, whether it is delivering the intended outputs, and whether those outputs are effective at generating impacts. In short, an understanding of the insights generated by evaluation work is needed to continuously improve the operation of the programme. More specifically, it's important to know whether activities are generating the expected outputs and whether the mechanisms for converting these outputs into the desired outcomes are operating as hypothesised during the earlier design phase of the programme.

It is recommended that those in charge of government programmes follow the established ROAMEF cycle, where the letters in this acronym stand for: rationale; objectives; appraisal; monitoring; evaluation; and feedback. When applied at a large scale this stylised policy cycle is highly appropriate, but the process is likely to be rather less linear in the case of smaller schemes. Nonetheless, the ROAMEF cycle remains the ideal and shows the role played by evaluation activities.

### Concepts and Terminology

**Attribution**

The complexity and variety of the NMS's activities make the attribution of impact difficult. Because most outcomes are the result of multiple activities involving inputs by the users themselves, as well as inputs from other organisations, such as Innovate UK. Furthermore,

outcomes can also be affected by a range of external factors, such as whether a company wins a grant. All this makes it difficult to attribute a particular outcome to one specific activity. Nonetheless, evaluation should try to consider all influences and seek to identify those that originate from the NMS laboratories.

**Additionality**

To find to the net additional impact of the NMS services, the observed outcomes need to be adjusted for:

- **Deadweight** - the outcome that would have happened anyway without the support. For example, a firm may have successfully developed a new product without support from the NMS.
- **Displacement** - when positive outcomes promoted by providing support to users are offset by a negative outcome elsewhere. For example, perhaps the NMS employing people with specialist knowledge of metrology makes it harder for industry to recruit such people and inadvertently puts an upward pressure on the wages of those they do manage to recruit.
- **Substitution** - when the effects of an intervention on a particular individual, group or area are only realised at the expense of other individuals, groups or areas. For example, if support makes a given firm more efficient and productive its share of the market could grow at the expense of firms who did not receive support. That is, the supported firm could expand by taking on more employees, but the unsupported firm could contract and so employ less people.
- **Spillovers** - when NMS's activities generate benefits for firms that don't directly use its services and aren't even aware of its role. For example, suppose that one of the NMS laboratories manages to extend the range of one of its calibration services, which enables the calibration labs who take traceability from them to pass on this extension of the range to their customers. Initially, only a few calibration labs might offer calibrations in this range which could allow them to charge a premium to firms who benefit from this new capability. However, as more calibration labs update the service they offer, this price premium will be eroded by competition and the full benefit of the new service will be passed on to the customers of the calibration labs.

## 4. THE PURPOSE OF THIS FRAMEWORK

This document sets up an evaluation framework for the National Measurement System (NMS) to provide a common structure around which each laboratory will take responsibility for the development of its own evaluation plans. This framework aims to promote overall coherence by having a common theoretical framework and guidelines on developing a set of indicators along with how to approach evaluation work in general. Each laboratory will be responsible for the production, validation, and publication of its own evaluation work.

As these laboratories work in different fields of science and focus on different sectors of the economy, the objective is to avoid being too prescriptive to allow each laboratory sufficient freedom to develop a plan that's tailored to its activities. Moreover, the scale of each laboratory's evaluation work should be proportionate to the scale of the funding it receives. In particular, as NPL is the single largest laboratory in the UK measurement system (representing 80% of the NMS investment) its evaluation plan will be considerably more extensive than that

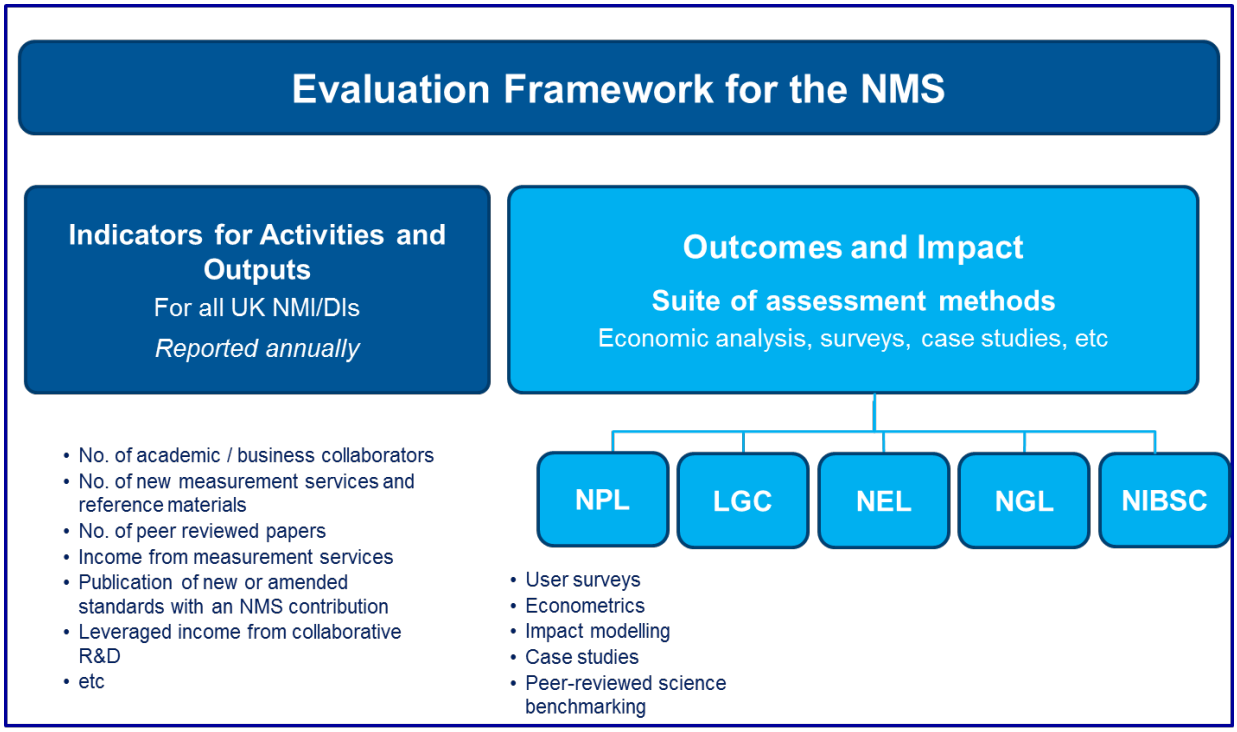
of the other laboratories. Also, the extremely small levels of NMS funding that go to NGML and NIBSC make it unreasonable to expect much (if any) evaluation activity by these laboratories.

There will also be some joint studies looking at the whole of the NMS as opposed to specific laboratories. Providing data sharing proves possible, this includes a customer survey and an econometric study, which are discussed in more detail in later sections of this document. These studies would seek to provide an assessment of NMS programme but would not seek to evaluate individual laboratories.

In the future an independent consultant-led evaluation of the NMS laboratories could be commissioned by DSIT It is expected that this overarching evaluation would draw on all the available evidence including studies and analysis provided by each laboratory. In the intervening period the NMS indicators will be reported annually and outputs of specific studies within the framework will be reported as they are completed.

An overview of the framework is presented in Figure 1 and described in Section 7.

Figure 1: Stylised Evaluation Framework




## 5. LOGIC MODEL

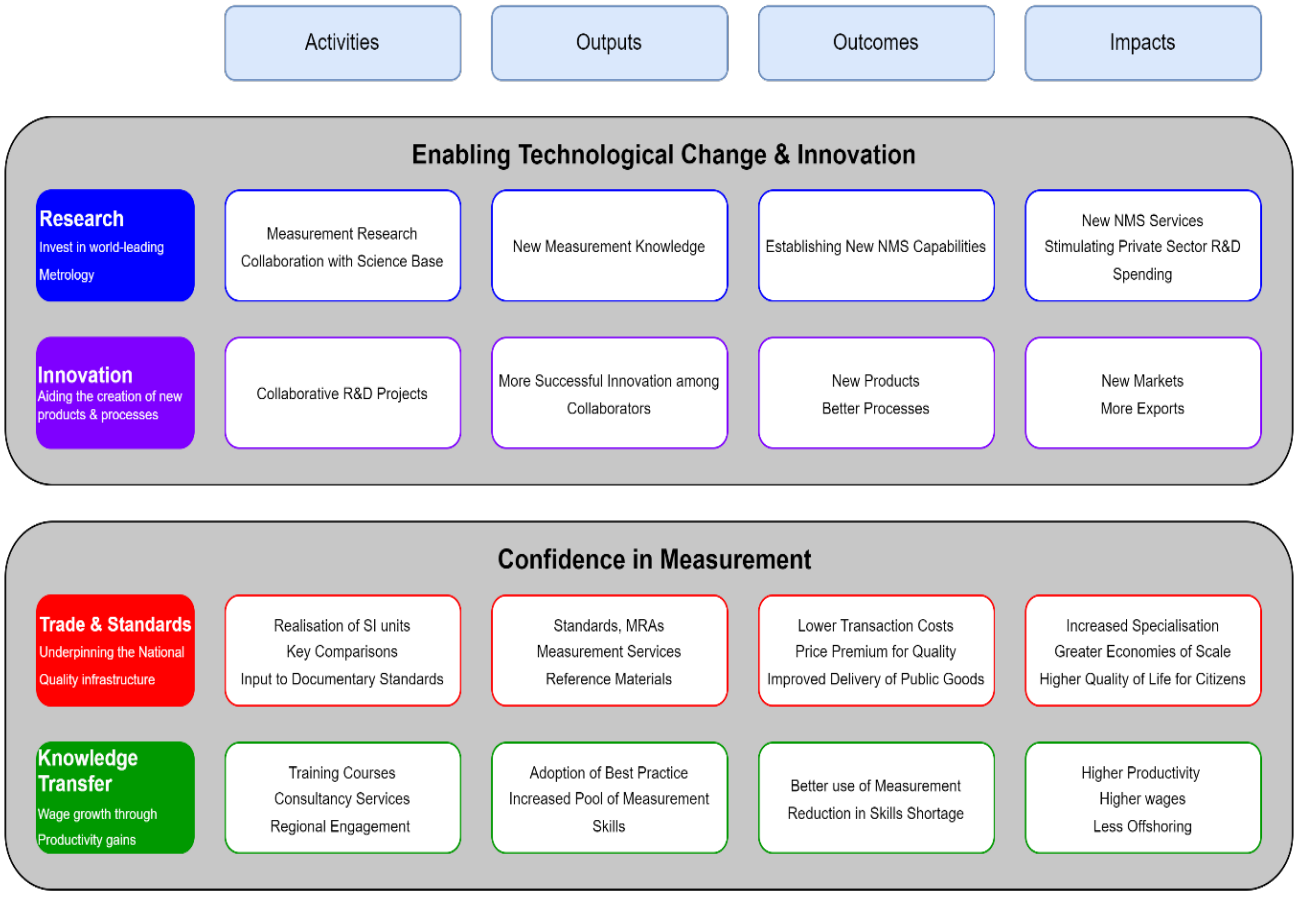
This section of the document sets up a logic model for the NMS based around themes in the UK Measurement Strategy. A logic model is a stylised way of describing step-by-step how the programme is intended to operate and the anticipated pathways to impact. It is a diagram setting out how inputs lead to output and how these outputs lead to outcomes and, finally, how these outcomes bring about economic and social impacts. Such a diagram shows links in the causal chain and identifies data that need to be collected to understand whether the programme is functioning as expected.

The logic model that follows (see Figure 2) provides a stylised representation of how the activities of the NMS laboratories, along with other parts of the national quality infrastructure, leads to economic and social impacts. This logic model has been designed around four impact mechanisms:

1. Investing in world-leading research into the field of metrology alongside other areas of measurement science
2. Accelerating the innovation activities of other organisations through engaging with them in deep collaborative R&D projects
3. Underpinning the national quality infrastructure through delivering unique measurement services that provide the technical basis for trade and standards within the UK
4. Helping organisations get closer to the technological frontier by fostering the adoption of best practice through a range of active knowledge transfer activities

Finally, a detailed theory of change to accompany this logic model is presented in Annex A.

Figure 2: Logic Model for the NMS




## 6. EVALUATION CHALLENGES

There is an extensive body of academic literature to support the analysis of science and innovation policies (Wang et al, 2022; BEIS 2021; Friedrichs, 2018). This literature can help analysts working for the programme to evaluate the effect of the NMS laboratories on the innovation outcomes and performance of organisations that make regular use of the services

and support on offer. However, this body of literature cannot help us to evaluate the infrastructural role played by the NMS laboratories or quantify the indirect benefits that come from traceable measurement.

A small number of notable academics in the US and UK made significant contributions to the understanding of 'infra-technologies' in the early 2000s but this work was more theoretical than empirical in nature. This would suggest that while it's possible have a good understanding of why the NMS laboratories are required, it remains very difficult to estimate the size of the loss that would occur if this infrastructure was dismantled.

## THE CHALLENGES OF ASSESSING IMPACT ON DIRECT USERS

The impact on direct customers of using leading edge products and services from the NMS laboratories can be assessed using econometric methods. Access to these products and services helps customers to adapt and innovate, which improves their performance, and this can be detected using statistical techniques. This subsection discusses both challenges encountered when applying these methods, as well as potential solutions. Furthermore, by complementing such studies with interviews and case studies it should be possible to understand why and how these impacts are created.

**Endogeneity**: Firms differ in terms of the ability and culture of their management, which means that some firms would be more inventive than others regardless of whether supported. Furthermore, the more inventive firms will probably have more need for the leading-edge services of the NMS laboratories than the less inventive firms. If so, customers would have been more inventive than non-customers even if they had been denied access to the NMS laboratories; and this self-selection would confound a simple statistical analysis.

Proposed solution: The collection of longitudinal data will enable evaluators to track the performance of individual firms over time and to look for changes in performance following the use of NMS services. Moreover, providing the Companies House Reference Numbers of customers are recorded, experts from the Office of National Statistics (ONS) will be able to link information from invoices and other administrative records (e.g., year of payment) to official datasets, such as the Business Structure Database which provides basic data on about two million UK businesses. By this means, it is possible to assemble large longitudinal datasets without the need for regular surveys. (Companies House Reference Numbers can be used to link payment data to the Inter Departmental Business Register and thus to all other ONS datasets.)

**Unique Primes**: The NMS laboratories work closely with a collection of very large multinational companies (primes) that dominate high technology sectors, such as aerospace, pharmaceuticals, and the extraction and petrochemicals. It's difficult to find a convincing control group for the primes because all such primes use the laboratories and in any case each prime is unique. That is, for a given prime who uses one or more of the NMS laboratories on a regular basis, it's difficult to find a similar looking counterpart who does not use these laboratories. This makes it difficult to apply cross sectional or matching methods when evaluating impact on the primes.

Proposed solution: It is possible to collect detailed annual data on an outcome of interest, as well as how intensively they interact with the NMS laboratories. (That is, although a prime will use one or more of the labs every year there is variation in the number and scale of payments.) Next, if there's a strong association between outcomes in consecutive years (e.g., levels of employment), then the outcome for a given year can be predicted based on the outcome in earlier years. Dynamic panel data methods (e.g., Arellano & Bond, 1982; 1991) use these forecasts to construct a counterfactual that can be subtracted from the observed outcome to find the additionality.

**Overlapping interventions:** Multiple organisations and interventions contribute to the same outcome, which makes it hard to isolate the impact of the NMS laboratories from other parts of the system. For example, the introduction of a new product could be the result of a grant from Innovate-UK and support from one of the NMS laboratories. (Similarly, the introduction of the NPL's Quantum programme is related to wider government investment in the development of quantum technologies by companies and universities.)

Proposed solution: The Business Support Database will be used to find firms that received public support from institutions, such as, UK Trade & Investment (UKTI). Similarly, the Gateway to Research database will be used to find firms that have been partners in projects with research council funding or that have received grants from Innovate UK for their innovation activities. The names and addresses of the firms in these databases can be used to find their Companies House Reference Numbers; and these company IDs should make it possible to link information from the support/innovation databases to more general firm-level databases, such as, the Business Structure Database. Hence, it's possible to track many forms of public support received customers of the NMS laboratories, as well as the support received by similar looking non-customers.

**Time lags and the skewed distribution of benefits**: It may take several years before the benefit of working with one of the NMS laboratories shows up in a firm's performance. For example, in the short term, the introduction of a new process by a firm may lead to disruption that causes its productivity to drop. Moreover, it could be that although innovation is generally beneficial, a small number of firms receive very large benefits while most firms receive small benefits. It follows that unless these fortunate firms are included in the sample, we will underestimate the average benefit.

## THE CHALLENGES OF ASSESSING IMPACT ON INDIRECT BENEFICIARIES

The benefit to individuals and organisations from having reliable and accepted measurements, underpinned by primary standards, generally can't be evaluated by empirical or statistical methods. This subsection highlights the challenges encountered when attempting to apply econometric methods to assess such benefits.

Proposed solution: To avoid underestimating the average benefit from an instance of support, it is necessary to ensure that the datasets used in our analysis are as comprehensive as possible. Furthermore, to avoid dropping the Primes from the dataset, the analysis will be based on dynamic panel-data methods rather than matching methods with a high attrition rate (see the discussion above). Finally, the construction and use of large longitudinal datasets should make it possible to detect delayed effects, such as, turnover and employment growth, that occur three or four years after a certain innovation activity took place.

It follows that a more theoretical approach is needed, and this will be mainly based on inferring impacts through the development of models that draw on evidence from interviews, surveys, and case studies.

**No Experiments:** It is sometimes suggested that such programmes should be evaluated by performing experiments. While this type of approach works well elsewhere it is generally not going to be appropriate for evaluating the NMS. For example, it's not practical to stop performing certain key comparisons to see if trade is damaged by a failure to maintain mutual recognition agreements.

Proposed solution: There are rare instances where the NMS laboratories withdraw from certain areas of metrology, such as, when NPL scaled back its capabilities in the field of acoustic measurements. These occurrences could be used as the basis of event studies or case studies focussed on how customers of discontinued services adapted to the change. Interviews with past users could be conducted to understand how the change effected their activities and whether they were able to access alternative services from national measurement institutes in other countries. For example, could they access similar services from PTB in Germany or VSL in the Netherlands.

**Cannot track indirect benefits:** Sometimes an organisation or individual can benefit without directly using one of the NMS laboratories. In particular, firms that take traceability from commercial calibration labs will benefit from the maintenance of primary standards by the NMS laboratories. This indirect benefit coming from secondary or tertiary calibrations is often referred to as the "fan-out". Moreover, firms that engage in manufacturing or production benefit from the traceable calibration services and mutual recognition arrangements (MRAs). These indirect benefits are impossible to track and make statistical methods unfeasible. However, in certain situations it might be possible to estimate the benefits through event studies and case studies.

Proposed solution: Part of the NMS customer survey could be focussed on the UKAS accredited calibration laboratories whose services are directly traceable to one of the NMS laboratories. In particular, the calibration laboratories could be asked about how many of their customers are using services that can be traced back to standards maintained by the NMS laboratories.

It may also be possible to get a broader picture of measurement activity across the economy by resurveying respondents to the UK Innovation Survey, which covers the whole business population and is carried out by the Office of National Statistics (ONS) every two years on behalf of DSIT. It's likely that this would be done in collaboration with other parts of the UK Quality Infrastructure (BRE, UKAS, BSI).

**Very few big changes:** The NMS laboratories have existed in a similar form for at least thirty years and in some cases much longer. Furthermore, most of the capabilities have evolved gradually over many years.

Proposed solution: Occasionally, there are major events such as the COVID crisis that cause major disruption to the operation and activities of the NMS laboratories. From an analytical standpoint, such events create an opportunity to assess the criticality of services or capabilities. What was the effect of suspending access to certain services? Were there some capabilities that couldn't be mothballed because attempts to do so had immediate adverse effects on critical parts of the economy or national infrastructure?

Furthermore, there are very few examples where the laboratories have shut down a capability or withdrawn a popular calibration service. This high degree of continuity makes it difficult to perform studies that use a major policy change that separates the time before from the time after.

**Quality of Life Benefits**: The NMS laboratories often work for government departments (e.g., DEFRA) or public sector organisations (e.g., NHS hospitals). Hence, the NMS laboratories indirectly contribute to the delivery of non-market goods, such as, health and the environment. The support going to public sector organisations often helps to sustain some of the quality-of-life benefits for which these organisations are responsible. Even when the impact on quality-of-life is direct it can be hard to quantify such benefits in monetary terms. Generally, the effect of NMS support on how public sector organisations deliver their quality-of-life benefits to citizens cannot be assessed statistically and so must be estimated using models and case studies.

Proposed solution: It may be possible to infer the scale of benefits through the use of models that show how improvements in how something is measured translates into an improvement in the quality or delivery of a non-market good (King & Renedo, 2019). There has been some past success with a case study looking at how improvements in radiation dosimetry translate into improvements in the delivery of radiotherapy (Font et al, 2015). It's hoped that something broadly similar could be done to find the benefits of emissions monitoring services on reductions in greenhouse gas emissions.

## 7. THE PROPOSED APPROACH

To address these challenges the overarching approach to the evaluation is Contribution Analysis – a mixed methods approach that seeks to test and provide evidence for the theory of change (i.e., how the NMS creates benefits for the UK) as described in section 2 and summarised in the logic model. A range of data and quantitative and qualitative analysis methods have been selected to provide evidence of NMS activities, outputs, outcomes, and impacts and the links between them. Each laboratory within the NMS will select and implement evaluation methods in line with the scale and content of their NMS activities.

### CONTRIBUTION ANALYSIS

Contribution analysis is a theory-based approach that requires a clearly articulated theory of change as embodied in a logic model. In particular, its focus on mechanisms helps to generate research questions and hypotheses to be tested. This approach to evaluation seeks to identify each of the various elements in the underlying logic model and examines the links between each element. Contribution analysis is conducted through a series of steps which are progressed through with appropriate iteration.

- Set out the expected attribution of impact to be considered
- Develop a theory of change, based on the programme's logic model
- Gather existing evidence on the theory of change
- Assemble and assess the contribution story and any challenges to it
- Seek out additional evidence, including through the evaluation methods set out above
- Revise and refine the contribution story

Hence, contribution analysis provides a coherent framework around which quantitative and qualitative evaluation methods can be structured. This approach provides the overarching framework for understanding, systematically testing and refining the assumed connections between the provision of support and the anticipated impacts. The aim is to understand and quantify the contribution that the NMS laboratories have on businesses and the science base. The focus is not only on understanding whether a policy has worked, but why, and under what conditions a change has been observed.

## MAPPING METHODS TO THE LOGIC MODEL

The evaluation is a mixed-methods approach, collecting both quantitative and qualitative data against a range of activity, output, outcome, and impact categories. The methods are listed in Table 1 and their role in assessing the different elements of the logic model is presented in Table 2. An overview of each method is provided in each of the sections that follow.

Table 1: Methods

| **Qualitative methods** | **Quantitative methods** |
|---|---|
| Scientific Peer-Review<br>Qualitative Case Studies<br>Stakeholder Interviews | Indicators<br>Econometrics<br>Economic Modelling<br>Quantitative Case Studies<br>Event Studies<br>Surveys |

Table 2: Mapping of methods to the logic model

| Themes | Activities | Outputs | Outcomes | Impact |
|---|---|---|---|---|
| **Investing in a world-leading metrology. (Research)** | Indicators | Indicators<br>Case studies<br>Peer Review | Indicators<br>Case studies<br>Interviews | |
| **Underpinning the national quality infrastructure. (Trade & Standards)** | Indicators | Indicators<br>Case studies | Modelling<br>Event Studies<br>Case Studies<br>Survey<br>Interviews | Modelling |
| **Aiding the creation of new products and processes. (Innovation)** | Indicators | Indicators<br>Survey<br>Case studies | Econometrics<br>Indicators<br>Case Studies<br>Survey<br>Interviews | |
| **Wage growth through productivity gains. (Knowledge Transfer)** | | Indicators | Indicators<br>Event Studies<br>Case Studies<br>Survey<br>Interviews | Econometrics |

## QUALITATIVE APPROACHES

Qualitative methods can be used to explain 'why' and 'how' a given impact was generated. That is, case studies are useful for testing the hypothesised theory of change embodied in the logic model. Case studies provide a narrative that's often missing from quantitative studies.

**Case Studies**: There will be essentially two types of case study:

- Case studies of major impact that is partly or wholly attributable to an output from an NMS laboratory to understand how this impact occurred and the role played by the laboratory. This type of case study will require looking at why something done twenty or thirty years ago led to changes and impact that are visible today. Furthermore, there will be a focus on identifying elements outside the laboratory that influence the effectiveness of its output, as well as whether the output had any unanticipated consequences (Cormin & Mestieri, 2013). Lastly, it would be helpful to identify organisations or groups that did not engage directly with the NMS laboratory but were nonetheless impacted.
- Case studies of specific outputs for the different elements (i.e., the rows) of the logic model from the more recent past (five or ten years ago). These will work both forwards and backwards in time to identify the activities that generated this output and investigate whether it led to any clear outcomes. In some cases, the intended outcome from a given output will not have been realised and so it would be useful to identify any external factors that have blocked or delayed the expected outcome. Lastly, we will examine whether the applicability of the theory of change embodied in our logic model depends on the size, sector, or region of an intervention's intended beneficiaries.

**Interviews**: In addition to the case studies, in-depth interviews with key stakeholders will be used to look at wider issues than those typically found in a case study. Interviews will be selected to cover the breadth of NMS activities, and beneficiaries:

- Internal interviews with key scientists from within the NMS laboratories. Such interviews might focus on how outputs from the NMS are strengthened by the R&D they perform. It would also be useful to conduct interviews with the key scientists responsible for maintaining core capabilities.
- Interview the principal users of the measurement services supplied by the NMS laboratories, primarily, the large calibration labs. Ask what they think might happen if traceability to measurement standards held by the NMS laboratories were withdrawn. The focus of this work would be on quantifying the cost and disruption caused by having to go overseas for high-level calibration services.
- Interview researchers in fields like life-science or data-science to understand whether greater awareness of metrological principles would benefit these fields. We will also conduct interviews with researchers in medical schools and NHS staff to better understand how products and services from the NMS laboratories affects the efficient delivery of public services and thereby patients' quality of life.

Most of the impact from recent R&D activity will occur in the future, and so their ultimate impact can't yet be evaluated. Consequently, we will focus on activity-based indicators for recent R&D projects together with an assessment of scientific quality based on a peer review process that

consciously imitates the REF (Research Excellence Framework), which is being used by the research councils to assess university departments.

**Peer Review**: Qualitative methods will also be used to assess the quality of the scientific outputs of the NMS. NPL's scientific work makes up the largest portion (80%) of the NMS and outputs are regularly review by external experts. The International Science Review is an expert and peer review commissioned by NPL's Science and Technology Advisory Council (STAC) as an independent assessment of the quality and relevance of the science output of NPL, NML and NEL. It should be noted that 'science', in the context of this review, covers all areas of scientific activity, including all STEM subjects, measurement services, knowledge transfer, training and consultancy.

The Review Board uses the reports from the Expert Reviewers, other documentary evidence, interviews with senior scientists and managers, personal knowledge, judgement, and technical expertise, to draw a collective conclusion about the overall performance of NPL. The output of the review is a report, prepared by the Review Board independently of NPL and the STAC, which concludes on the overall performance of the organisations, and presents detailed observations of individual research areas. In both cases, the report presents recommendations which the Review Board feel would be suitable to address their observations and conclusions. The International Science Reviews were conducted about every four years.

## QUANTITATIVE APPROACHES

As described above, the character of the NMS makes it challenging to evaluate. Therefore, the methodologies presented include both established methods and, importantly, a range of experimental new methods. The aim is to deliver robust evidence, based on the latest thinking, and develop models that can be applied in future evaluations.

**Indicators**: Data from the laboratories' management information systems will be used to construct indicators for the activities and outputs of the NMS laboratories. Indicators have been designed to meet two criteria: to align with the logic model and be based on reliable data. Furthermore, the amount of information supplied by a given laboratory should be commensurate with the funding it receives. Nonetheless, if the laboratories can supply standardised information on activities and outputs, then it should be possible to construct aggregate indicators for all NMS laboratories. These indicators will be reported annually to provide evidence that the NMS laboratories are conducting activities and producing outputs in line with the logic model and are therefore on the route to delivering the expected outcomes and impact. The indicators are presented against the logic model in Table 3.
The indicators that feature in Table 3 are limited to activities and outputs. Whist this provides important information to monitor how the programme is functioning it provides no information about the programme's ultimate outcomes or impacts. Annex B outlines our approach for the development of new indicators to track outcomes and impacts related to the socioeconomic challenges emphasised in the UK Measurement Strategy: '*prosperity*', '*health*', '*environment*', and '*security*'. The development of these new metrics will make it possible to tack the strategic alignment of the NMS programme with government's priorities and objectives.

Table 3: Mapping of methods to the logic model

| Themes | Activities | Outputs | Outcomes |
|---|---|---|---|
| **Investing in a world-leading metrology. (Research)** | 1.1 No. of academic collaborators | 1.2 No. of peer reviewed papers | 1.3 Citations |
| **Underpinning the national quality infrastructure. (Trade and Standards)** | 2.1 Number of active measurement services and reference materials sold | 2.2 Income from measurement services and reference materials sales<br>2.3 Publication of new or amended standards with an NMS contribution | |
| **Aiding the creation of new products and processes (Innovation)** | 3.1 Number of business collaborations<br>3.2 The number of active new measurement services and reference materials | 3.3 Leveraged income from collaborative R&D and consultancy, instruments | |
| **Wage growth through productivity gains (Knowledge Transfer)** | | 4.1 Number of people accessing measurement training through web resources<br>4.2 Participation in face-to-face training | |

**Surveys**: The NMS has significant experience with undertaking customer surveys going back as far as 2005 but these are not longitudinal as a random sample was selected each time. The total database of both users and customers contains around 7,000 companies and, typically, a sample of 400 - 600 telephone interviews is conducted by an independent market research company (King and Tellett, 2020). The response rates have been high, with the majority of customers prepared to participate. In total four waves of the survey have been undertaken with many of the questions being repeated in a similar manner so that there is now a well-established methodology and questionnaire.

A new beneficiary survey will be used to obtain information about the experiences and perceptions of users and customers. Such surveys will enable collection of standardised information about their experience of using the products and services offered by the NMS laboratories, as well as whether they believe this contributed to the introduction of new products and processes. Customers and users who say the NMS laboratories contributed to the success of their innovation activities will be asked to describe why and how the support they received contributed to these innovation outcomes. Finally, those reporting positive outcomes will be asked whether support from NMS laboratories led to additional benefits in other areas of their business (or other parts of their enterprise group) or whether the benefits were transmitted to other independent business, such as through subsequent collaboration.

Finally, in addition to the NMS customer survey (King & Tellett, 2020), we anticipate that there will also be a survey of organisations who use reference materials or get their instruments calibrated. The intention is to use this survey to assess the scale of the fan-out along the calibration chain. Lastly, there may also be a survey of people and companies that have used the NMS laboratories' training services.

**Modelling**: Models will be used to estimate impact in those cases where the length and complexity of the impact pathway means that it's unrealistic to pursue an empirical evaluation. That is, modelling is useful in cases where the distance between an NMS laboratory and the ultimate beneficiary is large and there is a wide range of factors that would confound an empirical evaluation. A model-based approach is appropriate when quantifying the effect of a new protocol or standard on the costs and benefits of a certain type of measurement activity. These models will be based around archetypes within our archive of impact case studies and parameter estimates will be based on interviews with the beneficiaries and trade associations. The intention is to develop studies that blend accepted models from engineering (or medicine) with economic ideas about the value of information and risk aversion. Studies that blend engineering, information theory and economics have been used before by academics and consultants and so this is not an entirely new approach.

**Event Studies**: We will also seek to use industry and stock market data to perform event studies based on the introduction of a particular standard or protocol. The basic idea is to find the abnormal return attributable to the introduction of a new standard or regulation by adjusting for the return that stems from the price fluctuation of the market. This approach has not been tried before in the context of the NMS but is an established method for assessing things like regulatory interventions. In particular, it might be a good way to assess the impact of MRAs (mutual recognition agreements) or the introduction of a new standard.

**Econometrics**: It is important to have objective statistical evidence for the causal effect of the NMS laboratories on innovation and economic outcomes. Furthermore, the application of statistical techniques requires large samples sizes and thus means that the analysis needs to be conducted at a high level, preferably by aggregating across multiple laboratories.

Economists from NPL and Innovate-UK worked with Frontier Economics and Belmana Ltd to quantitatively assess the effect of direct support - receiving grants from Innovate-UK and/or using the NMS laboratories - on the survival and growth of firms over the next four and five years, respectively (Frontier Economics, 2016; Belmana, 2020). These studies were the two attempts to make use of firm level data to assess the connection between direct support from state funded agencies or laboratories and a firm's future economic performance.

It is proposed that we build on the previous studies but expand it in several ways:

- The previous analyses were based on administrative data over a 5-year and 7-year period, respectively (2008 to 2012; and 2009 to 2015) on grants awarded by Innovate UK and income for paid NMS services. However, it would now be possible to extend the sample to cover a longer period, which should result in a usable sample of around 2,000 observations for the NMS case. Repeating these analyses with this sample would increase the precision of the estimates and may make it possible to find statistically significant effects on turnover growth.
- This new study would combine data from the Business Structure Database with the much wider range of information on firms' performance and activities available from FAME and ORBIS. In particular, these platforms provide data on overseas sales (exports), patenting activity, and R&D spending. This would make it possible to assess the effect of direct support on a range of outcomes, as well as providing a much richer set of control variables than can be found in the Business Structure Database.

In keeping with the approach taken by Frontier Economics (2016) and Belmana (2020), this study will combine Propensity Score Matching (PSM) with Difference-in-Differences (DiD). The matching approach involves constructing a control group by pairing a firm that received direct

support each year with similar firms that didn't receive such support that year (and did not go on to receive such support in later years). This matching is done on fixed characteristics like age, size-class, sector, and region. With DiD, once an appropriate control group has been defined, measurements of the outcome of interest are taken before and after the support was received. The primary assumption is that any unobservable differences between the supported group and the control group (e.g., managerial ability and ambition) are roughly constant over about ten years. The second semi-testable assumption is that in the years before some of the firms received support, outcomes for the treated group track the outcomes for the control group. If these assumptions hold, then it's reasonable to claim that any observed difference in the rates at which this outcome measure grows or declines for the two groups is caused by the support received by the treated group.

## PUBLIC SECTOR EQUALITY DUTY

The analysis now being conducted to support Public Sector Equality Duty (PSED) involves a comprehensive examination of how publicly funded organisations and institutions are meeting their legal obligations to promote equality, eliminate discrimination, and foster good relations among different groups. In accordance with the Equality Act of 2010, the NMS has a responsibility to pay due regard to the need to eliminate discrimination, advance equality of opportunity and foster good relations between people of different protected characteristics. This covers the need to consider all individuals when delivering services, as well as in relation to employees of the laboratories funded by NMS.

Evaluating PSED for NMS will encompass an assessment of the extent to which PSED has been effectively embedded in NMS programme design and delivery. This will focus on ensuring that PSED is not merely a compliance exercise but rather has become an integral component of decision-making processes. It will involve examining whether PSED considerations are systematically integrated into all phases of NMS programme development, from the initial conceptualisation and design through to implementation. This will also involve assessing the clarity and specificity of equality objectives, the presence of diversity in stakeholder engagement, the influence of PSED on policy design, and the impact on NMS service delivery. It will equally include a review of data collection and the monitoring mechanisms that exist to track progress. Ultimately, the objective is to ensure that equality and diversity considerations are woven into NMS planning, decision making and delivery, fostering an inclusive and equitable public sector.

It is important to note that each of the laboratories funded by NMS (NPL, NEL, and NML) are individually responsible for evidencing their commitment to fulfilling PSED. Each laboratory is responsible for providing evidence of the degree to which they are satisfying PSED and identifying areas for improvement. Approaches of each laboratory for evidencing PSED are highlighted as follows:

### NEL's Approach

NEL currently monitor gender balance, gender pay disparities, age demographics and representation of women in management positions as part of its commitment to PSED. Recognising the need to enhance understanding of NEL workforce's diversity and the need to better tailor strategies to address potential disparities and create a more inclusive and

equitable working environment, a comprehensive D&I survey will be conducted in first quarter of 2024. The survey will align perfectly with the PSED's requirements to advance equality, eliminate discrimination, foster good relations among different groups.

With this survey, the objective is to gather insights into a broader spectrum of characteristics, such as ethnicity, disability, sexual orientation, religion and more, to ensure a more comprehensive understanding of the NEL workforce's unique composition and needs. To ensure buy-in from the staff, there will be engagement in outreach and communication to encourage all employees to participate in the survey. Additionally, the survey results will be analysed to identify areas where inequality exists, and subsequently, targeted action plan to address these disparities will be developed. This approach will provide a more nuanced perspective on diversity and inclusion challenges and underscores NEL's commitment to embedding the principles of equality and diversity into the very core of its culture, in fulfilment of the PSED.

### NML's Approach

The PSED evaluation plan within the NML/GC focuses on its current practices and future endeavours to promote equality and diversity. Presently, the primary focus is on monitoring gender data and maintaining gender balance across various organisational levels. There have been positive results from these efforts, with equal pay being consistently maintained between genders and no significant pay gap disparities observed. Moving forward, while the importance of gathering additional data on protected characteristics is recognised, there is commitment to ensuring that any such data collection is voluntary and respects individual privacy. This approach aligns with the goal of fostering an inclusive and respectful work environment.

In a broader context, LGC has proactively embraced diversity, equality, and inclusion (DEI) initiatives. A comprehensive Diversity, Equality, and Inclusion Strategy & Roadmap, implemented during the fiscal year 2020-21 (FY21-FY24), provides guides on actions and commitments. This strategy encompasses expanded data and reporting mechanisms to enhance understanding of the organisation, sponsorship, support, and investment in employee resource groups (ERGs) and Networks, and regular updates and reviews to ensure future plans are aligned with DEI objectives. Furthermore, an independent review of recruitment and promotions processes has been conducted, validating, and refining approach, including considerations for DEI factors such as inclusive language in job descriptions, gender-balanced shortlists for leadership roles, and providing unconscious bias training for hiring managers. Currently, two ERGs/Networks are supported: the Women's Network and PRYSM (LGBTQIA+). Additionally, various HR policies, including an Equality & Diversity policy and a Family Leave Policy that includes enhanced maternity/adoption pay have been implemented, to support and reinforce LGC's commitment to fulfilling the obligations of the PSED.

### NPL's Approach

Monitoring and evaluating the demographics of NPL's staff is crucial to ensuring that it plays its part in nurturing an equal, diverse, and inclusive society. NPL has demonstrated its continued commitment to this challenge by professionalising these activities through the hiring of a Diversity and Inclusion Human Resources Business Partner. NPL currently track its Gender Pay Gap (NPL, 2021). Using these statistics, NPL can benchmark its performance to

the broader sector, such as, professional, scientific, and technical activities within UK economy, and identify relevant interventions to address any negative impact. Internal benchmarking will also help to determine if those with protected characteristics are having less or more opportunities in a systemic way.

Below is a detailed framework that NPL will be using to conduct its PSED analysis:
To investigate equality within NPL, the analysis will mirror the three objectives of PSED by focusing on elimination of unlawful discrimination, the advancement of equality of opportunity and, fostering good relations between people who share a protected characteristic and those who do not, particularly, amongst people within NPL's workforce.

Given the nature of its objectives and data availability, PSED analysis should take a "mixed methods" approach, which combines quantitative and qualitative analysis. The qualitative techniques required to implement the analysis are highlighted as follows:

- *Data Collection and Analysis:* In terms of sequencing, data on key equality indicators, such as ethnicity, gender, disability, and age, will be collected to understand the demographics of staff. Subsequently, the collected data will be analysed to identify disparities, gaps, and areas where inequalities may exist in key outcomes, such as promotion, pay zones, and inclusion, within NPL's workforce using combination of techniques including descriptive statistics analysis, significance testing, contingency table analysis, Herfindahl-Hirschman Index (HHI) to measure inclusion (that is, how concentrated is a particular protected characteristic in the work force) and Gini index which is a measure of top level (in)equality. There will also be analysis involving multiple years of data, specifically trend analysis statistical testing to see whether there are changes in the underlying parameters between periods.
- *Case Studies*: Specific case studies or examples that illustrate how NPL has either successfully addressed equality issues or encountered challenges will also be considered. This will be relevant, specifically, in addressing the PSED objective of fostering good relations, and in assessing any programme delivery impacts.
- *Engagement and Consultation*: It would be worth to evaluate how NPL engages with stakeholders, including diverse groups, to gather input and feedback on staff satisfaction and internal policies. This consultation is important in identifying and addressing equality issues.
- *Existing Equality Objectives*: Examining existing NPL's equality objectives to assess how they align with the specific duties of the PSED is also important to gauge its level of awareness as an institution funded with public resources. The objective is to analyse the progress made toward achieving these objectives, including any impact assessments conducted.

Alongside these qualitative techniques there will also be data collection and techniques for quantitative analysis: Firstly, data on key equality indicators, such as, ethnicity, gender, disability, and age, will be collected to understand the demographics of staff. Next, the collected data will be analysed to identify disparities, gaps, and areas where inequalities may exist in key outcomes, such as, promotion, pay zones, and inclusion, within NPL's workforce.

This quantitative analysis will be conducted using combination of descriptive statistics and the following techniques:

- Significance testing using contingency tables, and multinomial logistic regression.
- The Herfindahl-Hirschman Index (HHI)[2] to measure inclusion.
- The Gini index to measure (in)equality.

There will also be analysis involving multiple years of data, specifically trend, as well as associated statistical testing, to see whether there are changes in the underlying parameters between periods.

*Equality Impact Assessment*: The focus will be on assessing the implementation of PSED by NPL given the analysis conducted, whilst also examining NPL's effort at addressing any identified inequalities and the impact of its mitigation actions. The centre of attention would be on NPL's efforts to comply with the PSED regarding the promotion of equality, elimination of discrimination, and the fostering of good relations.

*Monitoring and Reporting*: To monitor any positive, as well as, adverse or negative impacts, identified through the equality impact analysis, a similar analysis will be conducted for NPL every year and reviewed every two years as part of measures to monitor fulfilment of its public sector equality duty.

## PROCESS EVALUATION

**The National Measurement System (NMS)**: The NMS utilises its funding to maintain and develop measurement capability and expertise under the Challenge Areas as defined within the UK Measurement Strategy[3]:

- Health ('the health and wellbeing of a growing population')
- Environment ('managing and reducing our environmental impact')
- Prosperity ('increasing prosperity and supporting innovation')
- Security and Resilience

In addition, the NMS provides crucial measurement research, capabilities and infrastructure to support, for example: the definition and maintenance of measurement traceability through international comparisons; input to government policy, regulations and standards (the National Quality Infrastructure); and provision of expert advice to the public and private sectors. The different Challenge Areas will have their own unique areas of research and type of end user, and intensity of activity type (e.g., measurement services or research). Therefore, it is probable that the NMS process evaluation will require bespoke research questions for each to account for the specifics of each Challenge, whilst also providing general findings at programme level.

**Process Evaluation**: HMT's Magenta Book defines process evaluation as an examination of '*activities involved in an intervention's implementations and the pathways by which the policy was delivered*'[4]. The framework recognises that there may be variation in these components and suggests covering questions that inquire into mechanisms that are working well, or less well, and why in terms of programme delivery. These can include the level of resourcing, the influence of external factors, and the extent to which the programme has reached its intended users. Process evaluations may use a range of methodologies, according to the scale and nature of the programme, covering '*subjective issues (such as perceptions of how*

2 The HHI index gives a measure of how concentrated a particular protected characteristic is in the workforce.

3 BEIS, "UK Measurement Strategy for the National Measurement System," Department for Business, Energy & Industrial Strategy, 2022.

4 HMT, "The Magenta Book: Central Government guidance on evaluation," HM Treasury, 2020, p.15.

*well a policy has operated) and objective issues (the factual details of how an intervention has operated, typically using administrative data, where available).*'[5]

**Process Evaluation for the NMS**: Undertaking a dedicated process evaluation for the NMS will not only serve as a key element of the ongoing evaluation of the NMS, but also identify potential lessons to support the continuous improvement the programme's delivery. In this sense, the process evaluation for the NMS will serve to provide feedback at strategic level, which is a core component of the ROAMEF[6] cycle by, for example, utilising established NMS stakeholder mechanisms (e.g., Programme Expert Groups [PEGs]).

It is proposed that the NMS follows the framework set out by The Magenta Book, with each laboratory commissioning their own process evaluation from an external consultant with expertise in assessing programmes of a similar size and complexity. Each laboratory's evaluation lead and/or programme manager will work in close collaboration with their external consultant, from planning through to delivery of the outputs, to maximise the potential insights from a process evaluation.

Building from the Magenta Book's framework, areas for investigation through a process evaluation for the NMS may include:

**Strategy**:
- How was the rationale and design of the programme defined at conception stage?
- What steps were taken to link the NMS to the Challenge Areas and core components of the UK Measurement Strategy?

**Delivery**:
- To what extent did governance, management, monitoring and coordination processes generate the intended results?
- Can any areas be identified for learning, in terms of success factors or barriers?
- What mechanisms are in place for continuous improvement (e.g., feedback during delivery informing process refinements)?

**Stakeholder Engagement and Benefits**:
- What is the route to benefits uptake (e.g., by the external users) from the outputs of the NMS?
- What can be learned from stakeholders in terms of benefits realisation and dissemination, and what role do they play?
- Are formal mechanisms established to drive stakeholder engagement and input?

Where appropriate, due consideration for Equality, Diversity, and Inclusion (as part of the Public Sector Equality Duty) should be planned for and applied across all stages examined through the NMS process evaluation.

**Method**: As previously outlined, the NMS is a complex programme in terms of its breadth of activities, users, collaborators, and strategic alignments. Therefore, it is anticipated that a process evaluation for the NMS will adopt a mixed methods approach, utilising qualitative methodology, (e.g., stakeholder interviews to obtain a depth of insights on programme delivery) and quantitative data for factual details on performance (e.g., from laboratories'

5 HMT, "The Magenta Book: Central Government guidance on evaluation," HM Treasury, 2020, p.15.
6 Process for policy development, from rationale through monitoring and evaluation; see:
HMT, "The Green Book: Central Government Guidance on Appraisal and Evaluation," HM Treasury, 2022, pp. 14-15.

existing administrative financial, monitoring and FTE [headcount] data). To optimise a process evaluation with the NMS' complexities, the appointed external consultants will be collaborated with, and empowered by, the commissioning laboratories.

Evaluation leads from the NMS laboratories will be engaged with the external consultants throughout the planning, delivery, and output generation of a process evaluation, supporting the consultants in terms of the definition of challenge-specific research questions and the identification of programme data sources and stakeholders.

We look to review the evaluation framework for the NMS every 5 years. This is to allow sufficient time for evaluation quality.

## REFERENCES


Arellano, M., & S. Bond. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58: 277–297.

Arellano, M., and S. Bond. (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics* 18: 47–82.

BEIS (2022). *UK Measurement Strategy: for the National Measurement System*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1125285/UK_measurement_strategy_2022.pdf

BEIS (2021). UK Innovation Strategy: Leading the future by creating it. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1009577/uk-innovation-strategy.pdf

BEIS (2017). *UK Measurement Strategy: Confidence in Investment, Trade, and Innovation*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/605955/uk-measurement-strategy-accessible.pdf

Belmana. (2020). *Public Support for Innovation and Business Outcomes.* London.

Friedrichs, S. (2018), Trend-analysis of science, technology and innovation policies for BNCTs. *OECD Science, Technology and Industry Working Papers*, No. 2018/08, OECD Publishing, Paris, https://doi.org/10.1787/1566a6ce-en.

Font, A., de Hoogh, K., Leal-Sanchez, M., Ashworth, D. C., Brown, R. J., Hansell, A. L., & Fuller, G. W. (2015). Using metal ratios to detect emissions from municipal waste incinerators in ambient air pollution data. *Atmospheric environment, 113*, 177-186.

Frontier Economics. (2016). *The Impact of Public Support for Innovation on Firm Outcomes.* London: BEIS. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/604841/innovation-public-support-impact-report-2017.pdf

King, M., & Renedo, E. (2019). *Estimating the impact of improved dose control on clinical outcomes in radiotherapy.* Teddington: NPL.

King, M., & Tellett, G. (2020). *A Customer Survey for Three of the Core Labs in the National Measurement System.* Teddington: NPL.

NPL. (2020). *Quantifying the Scale and Make-Up of Measurement Activity in the UK (Forthcoming monograph).* Teddington: NPL.

NPL (2021). *Gender pay gap report*. Available at: https://www.npl.co.uk/about-us/corporate-information/gender-pay-report 2021_final#:~:text=By%20the%202021%20snapshot%2C%20the,%25)%20but%20the%20median%20increased.&text=To%20calculate%20quartiles%2C%20we%20have,this%20list%20into%20four%20quartiles

Olakojo, S., & King, M. (2023). The impact of NPL and the National Measurement System. Available at: https://www.npl.co.uk/about-us/who-we-work-with/government/nms-infographic.aspx

Wang, J., Ma, X., Zhao, Y., Zhao, J., Heydari, M. (2022). Impact of scientific and technological innovation policies on innovation efficiency of high-technology industrial parks – A dual analysis with linear regression and QCA. *International Journal of Innovation Studies*, 6(3), Pages: 169-182.

## ANNEX A: THEORY OF CHANGE FOR THE NMS

This section sets out the theory of change that explains the casual links between elements of the previous logic model (i.e., activities, outputs, outcomes, and impacts).

### Research

The NMS enables a national approach to measurement capabilities that deliver world-leading measurement science in partnership with scientific institutions, global NMIs and industry. The NMS performs the research needed to develop and extend the technical infrastructure (measurement tools, reference data, protocols); the NMS capabilities generated form the basis of leading-edge products and services. The technical infrastructure is leveraged by industry and academia, increasing R&D activity, generating better measurement within the research base, and increasing external investment.

The NMS provides measurement and expertise to undertake pre-standards research, supporting the generation of new and improved standards. Furthermore, this research extends the application and dissemination of physical reference materials and standards (SI units), informing redefinition and digitalisation activities. Research activities undertaken in-house by the NMS increase and expand its capabilities for knowledge dissemination to industry and academia, for example generation of new and improved measurement methodologies which are then published in good practice guides or standards.

NMS metrology research increases the stock of knowledge in key areas such as quantum and biosciences. This knowledge stock can be broken down into two elements: the first being “proof-of-concept research” which establishes technology platforms for a basis for developing products; the second is a technical infrastructure which includes the analytical tools and standards needed for measuring and classifying the components and assessing the adequacy of new technology. The long-term impact of these forms of knowledge includes an increase in the productivity and intensity of R&D in the UK.

The route to the NMS delivering this long-term impact starts with the NMS conducting primary research in key areas through the funds received via the government’s contribution as the NMS and other sources. These research capabilities are used to produce explicit knowledge which is articulated and disseminated through peer-reviewed publication; the NMS published just under 600 peer-reviewed papers and collaborated with 677 academic institutions in 2020.

Research performed by the NMS laboratories sometimes generates patents and know-how, and increased capability. Granting access to this intellectual property through licenses and collaboration agreements stimulates a second round of research activity in the public and private sectors. The increase in capability further widens the NMS knowledge assets for exploitation in private and public sectors; in 2020, 102 patents cited NPL publications.

An example of the theory of change explained above is the NMS metrology research programme on 3D imaging of organic electronics. The collaboration with more than 40 instrument manufacturers, industries, and academics led to 23 publications and provided vital information enabling the method to be reliable, reproducible, and trusted. NPL’s research unequivocally demonstrated that argon cluster sputtering is superior to C60 sputtering (an alternative technique), informing critical research and development decisions. The result gave instrument manufacturers (Kratos, Thermo, ION-TOF, PHI and Ionoptika) confidence to develop argon cluster ion sources, involving significant R&D investment.

The NMS labs engage with a broad community of stakeholders to know the acceptance of their scientific outputs:

The NMS Customer Survey (King & Tellett, 2020) found that two thirds (68%) of academic collaborators said they work with the NMS labs to better inform their strategy and research direction. Moreover, a third of these users (34%) said the impact created by working with the NMS labs could not have occurred without their support.

The NMS brings in international scientific experts to evaluate the quality of its research and science, benchmarking NMS science against peer groups, assessing economic and social impact, and so on. The latest International Science Review in 2016 concluded that all the NMS science areas perform at an internationally competitive level, and that more than half of them are considered to be internationally leading (for more detail see here).

The NMS forms part of the critical national infrastructure, via SI traceability and supporting entities such as the NHS and Defra; the knowledge stock its research generates is an enabling factor for the UK to be safe, green and more prosperous. This subsection is a clear indication of the strength of the NMS research activity, and its wider impact within the measurement infrastructure and industry. The goal is to begin extending the NMS to whole emerging and evolving areas such as quantum technologies, industry 4.0, advances in life sciences and net zero.

**Innovation**

The NMS seeks to get better connected to end users to increase awareness, access, and uptake of best practice measurement. The goal is to increase the productivity of private spending on R&D and innovation. This will contribute to bring up investment in R&D closer to the 2.4% target set by Government.

Measurement plays an essential role in innovation because advanced tools and measurement knowledge are required to design and incorporate new or improved features into technologies; companies may not have in-house knowledge, nor the funds to invest in capital for R&D activities. The NMS recognises that there are some of its research capabilities are tacit in nature and can only be shared through close interactions e.g., consultancies and collaborations. These activities include programmes where the NMS collaborates with end-users, e.g., Innovate UK, Analysis for Innovators (A4I), Measurement for Recovery (M4R) and Sharing in Growth (SiG) programmes. These programmes increase the number of end-users the NMS works with, stimulating regional and SME engagement.

The adoption of measurement capabilities by companies provides them with information that is essential in controlling complex systems, testing the validity product claims, and overcoming innovation-impeding accuracy problems. Innovation programmes support the 2.4% agenda by leveraging the national NMS assets and making them more accessible to industry, and industry perceive the initiatives to be a cost effective, timely and de-risking mechanism through which to undertake R&D. The outcomes for companies on these programmes often occur during a project, enduring for many years. Outcomes reported by companies include new products, IP, improved and standardised processes, and additional investment secured; indirect outcomes include knowledge transfer, new jobs and studentships, and increased confidence in measurement data.

Innovation can also be initiated as a response to a challenge or regulation. The NMS supports companies innovating emerging green technologies, driven by a response to support net zero and meet regulation, in areas such as battery technology and energy storage, and new measurement techniques and sensors to monitor emissions. Industry utilises the NMS' capabilities to improve efficiency in their production processes, benefitting from a reduction in

waste and emissions; one company has cited an estimated £0.25M per month cost of quality saving attributable to a 3-month project with NPL[7].

The creation of new products and efficient process encourages investments into these firms, as investors view these firms as profitable ventures which would guarantee a future flow of returns. Additionally, if firms become generally innovative within a sector, its competitors, as well as the firm, must continually invest in product research and development (R&D) to maintain their competitive advantage, thus leading to long term market competitiveness.

AgPlus Diagnostics is an example of a near-to-market NMS user. Agplus developed a portable diagnostic platform that could quickly identify a wide range of illnesses from a blood or saliva sample, allowing point-of-care diagnosis within minutes. However, during a scale-up of their manufacturing process, they experienced high failure rates in production, limiting further exploitation of the technology. AgPlus undertook an A4I project with the NMS to develop and refine the manufacturing of their key cartridge components. NPL found the printing thickness was unsuitable and optimised AgPlus's cartridge component. Following this optimisation, AgPlus was able to mass-produce the platform.

Standards are of particular importance in innovation, due to the high level of uncertainty associated with R&D and novel characteristics. As per the Innovation Strategy, standards help to accelerate innovation by providing the agreed method to demonstrate against performance criteria. Furthermore, they stimulate investment by bringing confidence to decision making, such as de-risking investment, for innovative products. The NMS supports companies with the certification of innovative new products through the outputs of collaboration and measurement services, enabling claims to be demonstrated on new products.

The case study of Johnstons of Elgin (JoE) provides an example of innovation leading to standards and stimulated R&D investment. JoE undertook an A4I project in collaboration with NPL to create new standards for the quality finish on cashmere. As a result of their project, JoE have reported cost of quality savings in the region of £900K per annum and have set unprecedented quality standards in the industry. In this their first formal R&D project, JoE have since committed to 22 R&D work streams over 3 years.

Support to innovation is probably the better understood route to impact and has been extensively analysed to econometrically quantify the benefits of supporting innovative business. A summary of the econometric evidence can be found in the evidence section below. The last wave of the NMS Customer Survey (King & Tellett, 2020) specifically asked users about innovation benefits. The survey estimates that:

80% of supported businesses introduced a new or improved product or process. Furthermore, 19% of those that made a change attributed the innovation to the support from the NMS labs.

Users believe that, without the support of the NMS labs, their annual sales of new and improved products would have decreased by at least £539 million.

The NMS enables innovation and accelerates the development, commercialisation, and take-up of new technologies. The NMS pre-standards research underpins the innovation ecosystem (as evidenced in the Innovation Strategy), which will enable the UK to act with agility in generating new products, processes, and technologies required for the green industrial revolution. The NMS continues to drive the prosperity of customers supported via

7 A4I Round 3, company wishes to remain anonymous

collaborative R&D, increasing their number of patents held, products released, and number of highly skilled green jobs.

**Trade and Standards**

The NMS supports the UK's Critical National Infrastructure (CNI) and is depended on by the UK for its provision of traceability to the SI units, measurement assets and knowledge. Under the first Covid lockdown in 2020, NMS critical services were fully maintained (e.g. NPL Time, services supporting cancer treatment – see Annex 6 for more detail). However, some services were brought online sooner than planned, driven by customer demand. These customers included the NHS and public bodies, requiring the NMS for continuity of their critical activities. This highlights that the services the NMS provides are not just supportive of the CNI, but also deemed to be critical and essential by its customers.

National standards come from achieving global agreement with other countries on how measurements should be made, expressed, and used, and avoid a duplication of effort. International comparisons are carried out only by those NMIs with the highest level of skills in the measurement involved to deliver the reference value for each key quantity. NPL is the UK's National Metrology Institute (NMI) – the designated organisation representing the UK's measurement internationally and influencing the development of international standards. NPL and NML at LGC have coordinated 86 comparison exercises combined since 1998, placing the UK at the forefront of international measurement along with other leading countries such as the US, South Korea or Germany.

The NMS champions measurement across Government to ensure good policy and regulation. Public institutions implement regulations and policies to manage complex systems across markets, and to manage and minimise negative externalities (e.g., pollution). Regulations can also internalise the social cost of the goods firms produce (e.g., the Carbon Emissions Tax). The process of making good regulations is iterative and steered by information from monitoring the results from current regulations, i.e., whether the current regulation has had impact on subject. The systems for monitoring these regulations benefit from the established standards, reference materials, and capabilities of the NMS. This ensures that methodologies used for enforcing or monitoring these regulations are fair, traceable, and consistent.

The NMS can assist firms to identify areas where they may be incurring additional cost under regulatory requirements. NPL conducted a project with National Grid Gas to undertake real-time methane monitoring for leak detection. NPL's system detected individual leaks up to 16 tonnes of methane per year from different locations across the site. Scaling this saving to cover NGG's 23 sites would provide a saving of 230 tonnes of methane (an effective CO2 saving of 6400 tonnes per year). The commercial saving of gas lost equated to £47k per year (£384k per year using the non-traded value for carbon).

In 2020, the NMS contributed to the amendment and publication of over 100 standards. With representatives on hundreds of technical standards committees, the NMS provides valuable technical input to the development of new documentary standards. This representation ensures that measurement methods – as codified in technical standards - are underpinned by sound measurement practice. Standards and regulation are referenced as a source of

information at a higher rate than publications[8], and as such the technical input the NMS provides to the process is vital for dissemination of sound measurement knowledge. Arguably, the contributions of the NMS are essential as very few organisations have a substantial amount of expertise in the field of metrology. The availability of documentary standards and reference materials allows firms to have access to resources which they can use to reduce both the technological and market risks.

Standards can provide the means to demonstrate regulatory compliance, helping to minimise cost and disparity by providing defined measurement requirements. NPL worked directly with the Financial Conduct Authority (FCA) in the run up to the implementation of MiFID II, EU legislation which establishes how financial markets must operate. The FCA have stated the support was 'instrumental in the development of the Regulatory Technical Standards on clock synchronisation (RTS 25), which were essential to ensure a consistent approach to recording precisely when market transactions happen'.

The NMS enables measurement users to have confidence in their data, reducing risk and allowing sound decisions to be made across the policy, standards, and regulatory environment. The NMS labs generate around £15m in sales of measurement services each year and provide circa 400 measurement services and reference materials. NMS calibration services underpin the use of national standards; commercial calibration laboratories customers in turn calibrate the instruments of their customers and diffuse measurement accuracy through a chain of linked calibrations. The NMS Customer Survey (King & Tellett, 2020) estimates that the NMS' reach extends through the 'fan-out' of calibration laboratories to around 74,000 organisations.

NPL manages a series of networks around the UK which measure air pollution, and regularly checks and calibrates measurement equipment in the field. NPL further ensures that measurements are traceable to national standards so their results can be relied on to make important decisions. NPL are entrusted by Defra, the Environment Agency, and the National Environment Research Council (NERC), as well as Local Authorities to lead air quality monitoring networks and provide measurements for use in climate change models. These activities contribute to the UK's goal of achieving net zero greenhouse gas emissions.

The NMS plays an essential role in the development and assessment of evidence-based policy. For instance, a recent analysis by NPL scientists aimed to fingerprint emissions from six municipal waste incinerators and then test if these 'fingerprints' could be found in ambient air samples (Font et al, 2015). The analysis finds no evidence of incinerator emissions in ambient metal concentrations around four UK waste incinerators. The conclusions of this paper were used in Parliamentary briefing notes on waste incineration facilities and incineration of industrial and commercial waste.

The NMS develops, supports and steers generation of standards, regulation and legislation through maintaining and developing its measurement infrastructure. The NMS helps reduce the cost of compliance and enable effective global trade for firms and provides robust and accurate data in response to policy and challenges - to improve quality of life and reach the targets of net zero.

8 UK Innovation Survey 2019: Statistical Annex, 2016-18 data https://www.gov.uk/government/statistics/uk-innovation-survey-2019-main-report

**Knowledge Transfer**

The NMS aims to develop talent to build the UK's capability in measurement across all sectors. Improving metrology skills in the workforce reduces user's measurement costs and can lead to increased productivity and wages. To maintain a talent pipeline, it is essential to offer a range of technical and academic routes into careers as well as opportunities to upskill along the way. Upskilling serves critical requirements in industry (e.g., via Laser Safety training) and strengthens the knowledge ecosystem (e.g., via Climate Data and Earth Observation courses to meet the challenges of Net Zero).

The NMS' provision of services empowers the workforce with the knowledge required to understand and work within standards requirements. For example, training for Aerospace Standards AS 13003 and AS 9145 provides companies with the skills to meet industry-specific requirements. The NMS provides UKAS accredited courses to users, both directly and via agreement with external delivery parties. These accredited courses work directly to address immediate skills requirements and serve as 'points' towards a longer-term UKAS qualification.

As new technologies arise, so does the demand for labour increase. 6.3% of all employments in the UK involve taking measurements and from 2010-2017 the number of employees in measurement-related occupations has risen 2% more than overall employment in the UK (NPL, 2020). The NMS recognises this growing demand in measurement skills and engages in various activities to help to address it. The labs continually develop and offer training materials, apprenticeship programmes, secondments, and PhD programmes. In some cases, these materials and programmes are developed in partnership with industry and academic institutions. The NMS, via training providers, conducted just over 1200 face-to-face training to organisations in 2019, and 428 in 2020[9]. Moreover, in 2020 the NMS provided free access to online metrology courses during the UK lockdown between April – June. This resulted in over 9000 additional sign-ups and 10,000 courses being completed, with the total accessing NMS online training resources at over 34,000 for 2020.

The Postgraduate Institute for Measurement Science (PGI) was founded in October 2015 by a strategic collaboration between NPL, the University of Strathclyde and the University of Surrey. The PGI hosts approximately 200 postgraduate researchers from across the UK (a growth of 82% in 5 years), they are co-supervised by NPL staff in collaboration with over 30 UK universities and industrial partners. The PGI provides students with industry-relevant experience and enables a smoother transition into an industry role, improving the knowledge and skills transfer from university to industry. Students at the PGI support NMS research working across a variety of sectors and often on projects that directly collaborate with industry. Of the 140 graduates who have completed their research since 2015, the largest proportion have moved into industrial roles (43%), indicating their industry readiness, with a number making a rapid move into leadership positions within 3 years of graduation. PGI's alumni are also established in academic and other research roles (29%) at prestigious organisations worldwide, with a further 19% becoming full-time researchers at NPL. These PhD students are arbiters of the industrially relevant knowledge they have acquired through the programme and spread this knowledge to other agents within their sphere of influence i.e., the companies they go on to work for.

9 Note: face-to-face training was not undertaken during lockdown; the count reflects Covid restrictions, and includes online classroom training

All these activities, as well as objectives to increase measurement skills factoring into wider strategic programmes (e.g., National Timing Centre and Sharing in Growth), address shortages by increasing the pool of people with measurement skills in the UK workforce. The availability of these new skills enables firms to either employ these talents in-house or outsource the work to firms with relevant measurement skills. Thus, the increase in the pool of measurement skills enables firms to reduce technological uncertainty and increase productivity. This knowledge underpins industry's ability to demonstrate conformance against standards or empower them to work towards generating industry standards. As an example, Rolls-Royce have mandated that certain roles in their organisation undertake the Dimensional Measurement training courses. The requirement ensures consistency of skills, knowledge level and measurement comparability across their organisation, not just within the UK but globally (including America and Asia).

NPL has provided training courses to those in the public and private sector involved with defence. NPL has accredited Lockheed Martin to deliver NPL's Dimensional Measurement Training programme. Lockheed Martin now have three members of staff (Calibration technicians) to conduct training courses to their employees at Ampthill in Bedfordshire. This is because it was far more efficient to deliver these courses on-site. NPL's framework gave Lockheed Martin the confidence to make the right measurements which has reduced general waste and repeat measurements for Lockheed Martin - saving them both time and money.

Vestas, the world's largest manufacturer of wind turbines, had a common problem in manufacturing: a disconnect between the employees responsible for introducing new technology and processes, and those using them. A 'Prototype Team' was established for validating new products and processes i.e., provide a link between the Production Business Unit and the shop floor. To ensure the team had the right skills, Vestas put the team through a range of training courses, including NPL's Dimensional Measurement training programme. This knowledge helped the prototype team evaluate new technologies and identify the best ones for the job. The training has given them the confidence to make full reports and recommendations back to the engineers, thus, increasing the overall productivity of the firm.

The NMS' skills and training assets provide measurement foundations for a consistent knowledge base in industry, and this is extended by leading and supporting studentships and PHD programmes. Bespoke training serves to transfer knowledge to address specific knowledge gaps or enable workforces to meet requirements for standards and accreditation. By addressing knowledge gaps in the market, the NMS provides the skills to reduce the cost of uncertainty and improve productivity of industry through efficient and robust measurement techniques in the workforce.

## ANNEX B: THE DEVELOPMENT OF INDICATORS FOR OUTCOMES AND IMPACTS

### Output Metrics

The UK Measurement Strategy (2017-2020) was monitored using output metrics related to the four mechanisms: "*Research*", "*Trade and Standards*", "*Innovation*" and "*Knowledge Transfer*". The existing metrics are counts of observable outputs from the programme that calculated every year. In the calculation of the existing output metrics (Part-A Metrics), administrative data produced by the programme's own activities is readily used.

### Impact Metrics

The new UK Measurement Strategy (2022) puts greater emphasis on the contribution of the NMS to addressing four national challenges: '*Prosperity*', '*Health*', '*Environment*', and '*Security*'. As a response to the challenge-focused strategy, new impact metrics (Part-B metrics) will need to be created. These impact metrics will be focused on tracking the programme's contribution to major socio-economic objectives, whilst also maintaining the existing mechanism-based categorisation of the current outcome metrics. They will be grounded by the empirical evidence of causal links between outputs and outcomes and will be used to gauge the strategic fit between the program and government objectives. The evidence used to justify the new metrics will draw from econometric analysis, as well as quantitative and qualitative research.

The mechanisms to be considered against each of the challenge areas will include research (inhouse R&D), collaborations (collaborative R&D), knowledge transfer (through training and consultancy), and trade and standards.

### Approach: the establishment of working groups

With a view at discussing the revisions to the current metrics, three different working groups (WGs) for the challenges will be established in 2022. Above these working groups, an overarching group will be set-up, whose purpose will be guiding the general development of the indicators.

These working groups would be open to representatives from all NMS labs, and it's expected that they will be mainly comprised of scientists and economists from across the programme expert groups (PEGs). Discussions on the development of the new indicators will take place during 2022 and the adoption of the new metrics will follow in coming years.

The list of possible metrics for Part-A and Part-B will be suggested to Working Groups consisting of engineers and analysts from the labs, once the suggested metrics are agreed. The metrics will be grounded by the empirical evidence of causal links between outputs and outcomes and will be used to gauge the strategic fit between the program and government objectives.