Unity Insights

NHS Health Education England

AI Roadmap
Methodology and findings report

January 2022

This report has been produced in partnership between Health Education England (HEE) and Unity Insights.

Health Education England

Health Education England (HEE) exists for one reason only: to support the delivery of excellent healthcare and health improvement to the patients and public of England by ensuring that the workforce of today and tomorrow has the right numbers, skills, values and behaviours, at the right time and in the right place.

At any one time HEE supports more than 160,000 students and trainees whilst working closely with partners across the NHS locally, regionally and nationally on shared priorities.

In 2019, HEE were commissioned by the then Secretary of State to deliver the Topol Review recommendations looking at the impact of leading edge digital technologies on the workforce. The Digital, Artificial Intelligence and Robotics Technologies in Education (DART-Ed) programme picks up from this in 2021 to explore the linkage between mature evidenced AI and its workforce impact and required training and education.

Unity Insights

Unity Insights offer bespoke analytics and evaluation services to the NHS, the AHSN Network, academia, innovators, and industry. It was formed in 2021 following several years of sustained growth as the analytics and evaluation function within Kent Surrey Sussex Academic Health Science Network (KSS AHSN).

Unity Insights has extensive experience of working with the NHS, from supporting measurement capture of small-scale local programmes to providing analytics and evaluation for innovations rolled out nationwide. The team has a clear understanding of the healthcare system governance, commissioning and challenges. This knowledge enables them to tailor their service options to each individual project, ensuring that meaningful insights can inform better decision making for the benefit of the system.

KSS AHSN, in collaboration with the AHSN Network, originally conducted a survey and developed a report in 2018 entitled "Accelerating Artificial Intelligence in health and care: results from a state of the nation survey" at the request of the Department of Health and Social Care (DHSC) and NHS England (NHSE) to help map out and define the state of the AI landscape in the UK, its definition and various forms. This new roadmap in 2021 picks up on the earlier version and work done since by NHSX to update and capture the latest progress in this area and present it in an easy, readable and user-friendly format.

Table of contents

Foreword

Almost three years have passed since the publication of the Topol Review (Topol), setting out a vision for preparing the healthcare workforce to deliver the digital future. During the review there was a sense of a digital revolution, with anticipation of the impact of technologies such as Artificial Intelligence (AI), Digital Medicine and Genomics on the functions and roles of the current and future healthcare workforce. Predictions were made on the impact of these technologies, and the skills and capabilities we would need to build for healthcare staff and learners to work safely and effectively with these technologies in a digitally transformed health and care system. Little did we know that in early 2020, a confluence of pressures driven by the Covid-19 pandemic would prove to be the impetus for the development and implementation of these technologies at a faster pace and scale than imagined.

The emergence of the NHS AI in Health and Care Award (AI Award), and the vision set out in the Government's National AI Strategy, represents the start of step-change for AI, recognising the power and potential for AI to increase resilience, productivity, growth, and innovation which can be applied to the health and care system. However, for AI and data driven technologies to have the impact we hope for, we must ensure workforce skills and readiness are a core focus of ongoing education reform and healthcare policy.

This report and its associated dashboard allow us to understand existing AI and data-driven technologies which currently exist within our healthcare system, the taxonomies that they sit within, their spread and adoption and ultimately the potential workforce impact of these technologies. The case studies within this report provide a more detailed understanding of the impact of two AI technologies which are more mature in practice and use. Overall, this provides a direction for ongoing discussion on, and review of training and workforce needs for all our healthcare professionals.

This report expands on the legacy of Topol, extending our knowledge into how far along we are in recognising the impact of AI to augment the delivery of healthcare and to release 'time to care' for our healthcare workforce. This report is a culmination of significant work and collaboration across our organisations, and provides valuable insight for leaders in AI policy, education, regulation, innovation and digital transformation, and workforce strategy. We would like to thank everyone who's input has shaped this work and look forward to the next steps in the further development of the AI Roadmap, case studies and dashboard and their wider implications.

Patrick Mitchell
Director of Innovation,
Digital and Transformation
NHS Health Education England

Marie-Anne Demestihas
Chief Services Officer
Unity Insights

Executive summary

Context of the commission

Alongside key delivery partners, Health Education England (HEE) has established a Digital, Artificial Intelligence and Robotics Technologies in Education (DART-Ed) programme which brings together a number of projects with the aim to further explore and build on the findings from the Topol Review (DART-Ed, 2021). The AI Roadmap work is one of two contributions exploring the use and impact of AI in the NHS, Yorkshire and Humber AHSN and University College London Partners are the AHSNs leading on the other.

HEE commissioned Unity Insights (previously part of Kent Surrey Sussex Academic Health Science Network) to develop a roadmap of AI-driven technologies nearing or ready for market as well as understand where and when they are due to have an impact. The perspective is focused on the workforce groups affected but also the effects on the clinical pathways and point of care and their respective transitioning requirements. The project ran between March and October 2021. The Figure 1 below presents the main activities and deliverables of this commission.

***Figure 1:** Timeline, activities and deliverables of the project*

March – May 2021

Dataset analysis and profiling

- Analysis of **existing datasets** (NHSX State of the Nation AI survey, NIHR Horizon Scan for instance) to **explore associations** between parameters
- Documenting **gaps and limitations**

May – July 2021

Development of the database

- Determining the technology taxonomy
- Establishing the template for the database
- Defining the **impact on the workforce framework**
- **Horizon Scanning exercise** to populate the database

July – October 2021

Interactive dashboard and case studies

- Building of the **interactive dashboard**
- Case study selection (Oxehealth and Optellum) and interview with innovators and/or NHS staff
- Writing of the **methodology and findings report**

The audiences for the work commissioned are:

- Internally, within HEE, to inform and influence education and training activity to ensure the workforce is able to embrace and effectively use AI and digital healthcare technologies as part of new ways of working.
- Stakeholders involved in setting education and training standards for the wider workforce, including Royal Colleges, regulators and faculties responsible for setting curricula for the registered workforce. The AI Roadmap will be used to engage and have strategic conversations with stakeholders to make relevant changes to reflect the training needs created by the use of AI and to enable the workforce to actively drive its use in the NHS.
- Other Arms-Length Bodies are both stakeholders as well as an audience for the AI Roadmap. To inform and support the National AI Strategy for Health and Social Care being produced by the NHS AI Lab, the Evidence Standards Framework-AI for ongoing guidance and regulation at the National Institute for Health and Care Excellence (NICE) and the potential scaling of AI technologies across the NHS through the NHS Accelerated Access Collaborative (AAC). It has been important and valuable to have their input in the creation of this roadmap.

Purpose of the report

The report aims to present the activities and the methodology used to deliver the AI Roadmap. In chronological order, the report covers the work conducted as well as how stakeholder engagement was used to shape the approach and amend the deliverables. Finally, the report also explores the limitations of the work and how these could be addressed in the future.

Key findings

The key findings are based on publicly available information only. Namely, they are based on the value proposition claims of the companies (self-reported not validated) and any complementary publications available. Moreover, the interactive dashboard presents the breadth of information collected as part of the Horizon Scanning exercise and enables the user to explore the data by filtering by technology type, point of care, clinical area, workforce group and geographical spread.

Summary of key findings

The findings from the activities related to i) the initial analysis and profiling, ii) the creation of the database and iii) the building of the dashboard, are summarised as follows:

- **Estimated time to deployment:** with 60 technologies (25%) estimated to be ready for large scale deployment within a year, the AI Roadmap highlights how relevant and timely training the NHS workforce for the use of AI technologies is. Indeed, technologies in that one-year category are for the majority already implemented in NHS sites and are impacting workforce groups ranging from medic general practice, adult nurse, healthcare scientist or diagnostic radiographer. With 78% of these technologies used in secondary care, 22% in primary care, and 7% community care, the roadmap shows the need to have a holistic and cross-organisational strategy to adapt the education provided on usage of AI in the NHS. It also underlines the need to monitor the spread of AI to ensure a fair access to innovative products across regions, points of care and types of sites.
- **Distribution of AI technologies:** out of the 240 technologies included in the database, 'Diagnostic' was the most represented type with 34%, closely followed by 'Automation/Service efficiency' at 29%. 'P4 Medicine' and 'Remote monitoring' technologies accounted for 17% and 14% of technologies respectively. Within the 'Diagnostic' type, 'Imaging' solutions and 'Cardiorespiratory and neurology' solutions were at respectively 49% and 27%.
- **Key clinical areas using AI:** the database included 67 clinical areas. After 'Multiple clinical areas' (selected for 23% of technologies), the most selected options were 'Clinical Radiology' (11%), 'Cardiology' (9%) and 'General Practice' (8%). The percentage of Automation/service efficiency technologies can explain why 'Multiple clinical areas' was so often selected, as they can be used in a multitude of settings.
- **Most affected workforce groups:** 155 workforce groups (developed based on the occupational codes) were used to describe the NHS workforce. The top 5 workforce groups who have been identified as direct users of AI technologies are 'Medic, Clinical Radiology' (with 15% of technologies), 'Medic, General Practice' (13%), 'Non Clinical, Admin' (10%), 'Diagnostic Radiographer' (8%) and 'Medic, Cardiology' (8%).

Limitations and recommendations

Whilst the roadmap presents a comprehensive overview of the AI landscape in the UK to-date, the limitations of the exercise have been documented for transparency. The 'Limitations and recommendations' section describes in detail the caveats of the work and makes suggestions on how to correct them, a summary of the key limitations is listed below.

- **Publicly available information was used to populate the database.** It was not within scope to engage with the companies listed in the database to obtain supplementary information therefore only information accessible to the general public, self-reported by the companies, was used for the Horizon Scan exercise.
- **No validation of the value proposition claims of the AI solutions** was done as part of the exercise. The purpose of the roadmap was not to evaluate AI technologies but to map the AI technologies currently on the market in England, to understand the distribution between the type of technologies, the pathway and the workforce impacted. Therefore, the assessment of the impact on the workforce was based on the impact claimed by the company.
- **Further analysis needed to keep the taxonomy up-to-date and relevant.** A periodic review of the taxonomy to assess its relevance and to update it in-line with academic publications as well as common usage is recommended.
- **The dashboard does not contain workforce data.** Options on what data could be integrated to the database were explored during the design phase and HEE supports revisiting them in the future to understand how available workforce data could utilised.
- **The impact on the workforce framework only captures the direct users and the documented impacts**. To avoid formulating too many assumptions, the researchers reported the impact on the direct users as presented by the AI technologies but did not try to infer what these impacts would mean for the indirect users.
- **The dashboard only presents a snapshot of the AI landscape at a given time**. The intelligence collected yields most value at the time of the dashboard release and will become obsolete if not updated.

"This dashboard is a useful step towards measuring the impact of AI technologies to the NHS. The blend of sources of information, including clinician experience, provides meaningful insights. We hope to build on this work in the future strategy for AI in health and care including strengthening data from social care and ensuring that our governance and commercial frameworks from across the digital health ecosystem, such as NICE's Evidence Standards Framework, and adoption data from AAC backed technologies, are brought to life."

Leanne Summers - Head of AI strategy, NHS AI Lab

As keeping the roadmap up to date is critical to keeping the dashboard relevant and insightful, there are many avenues to address the limitations listed, to hone in on the positive elements and to action some of the recommendations suggested by the authors. Keeping the collaborative spirit observed during this commission, with regular engagement with system partners in the AI space such as the NHS AI Lab, NICE, the AAC and the AHSN Network, will be key to ensure the roadmap is in alignment with, and adds value, to the healthcare system.

Acknowledgments

The authors would like to thank a number of people who have provided helpful comments and insights during the production of this report. These include the core HEE team Alan Davies, Richard Turnbull, Hatim Abdulhussein, Lucy Dodkin; the extended HEE stakeholder group Henrietta Mbeah-Bankas, Paul Sadler, Nicola Calder, Adrian Brooke, Don Liu. It also includes our AHSN colleagues Kasia Zawadzka, Kelly Bradley, Elizabeth Graham, Lucy Brock, Suzanne Ali-Hassan, Soraya Jenney, Helen Hoyland and the people who engaged with us as part of the case studies Hayley Bolton, Charlotte Wood and Rhiannon Lassiter. We would also like to thank Leanne Summers at NHS AI Lab, Emma Hughes at AAC and Jeanette Kusel from NICE for their support and input into this work.

Notes to the reader

Throughout the report, the authors aim to be clear and refrain from the use of jargon to accommodate a wide range of stakeholders. To this effect, please note the following considerations with regards to vocabulary:

- **AI technology:** throughout this work a technology was classified as an AI technology if communications by the company or other publicly available information resources used the expressions 'deep learning', 'machine learning', 'deep neural networks', 'artificial intelligence', or 'AI' to describe it (Hinton, 2018).

Taxonomy for technology types

- **Automation/Service efficiency:** Technologies within this type refer to the use of automation in the form of control systems and advanced technology to eliminate or decrease the need for manual tasks. It is usually applied to repetitive tasks, such as data entry, maintenance of records, and patient health monitoring. The solutions range from automated data to feedback collection to patient triage systems.
- **Diagnostic**: Technologies within this type refer to the use of AI tools to supplement and enhance the process of using medical images to deliver high-quality patient care across a wide variety of diseases and organ groups.
- **P4 Medicine:** P4 Medicine is an approach to make medicine more Predictive, Preventive, Personalised and Participatory. Its two major objectives are to quantify wellness and predict and prevent disease. It incorporates a range of technologies from predicting the likelihood of a patient developing a long-term condition by analysing patient records to predicting patient response to medication, allowing to create a personalised plan.
- **Remote monitoring:** Technologies within this type include monitoring devices that collect data which can be shared with healthcare staff to monitor patients inside or outside hospitals and allow for earlier interventions if a patient's condition is worsening. They may be used to monitor patients after surgery or hospitalisation or for patients to manage a long-term condition.
- **Therapeutic:** Technologies within this type includes technologies which deliver evidence-based therapeutic interventions to patients that are driven by high quality software programs to prevent, manage, or treat a medical disorder or disease. It includes technologies ranging from mental health apps to technologies used in radiotherapy.
- **Other:** Technologies within this type are technologies which do not fit into any clear category such as AI solutions used for medical education purposes or a health information platform for children.

1

Initial analysis and profiling

Description of datasets

The following datasets were utilised for initial analysis to understand the technology characteristics and how they corresponded with the scope of the AI Roadmap:

- The 2021 AI State of the Nation survey (conducted by KSS AHSN and commissioned by NHSX)
- The AI Horizon Scan (conducted by the National Institute for Health Research, NIHR, and commissioned by the AAC)
- The Innovation Pipeline National Master Data (managed by Health Innovation Manchester)
- The list of applicants for the AAC AI Health and Social Care Award Round 2

The 2021 AI State of the Nation survey

The 2021 AI survey builds on the two initial baseline surveys conducted in 2018 and 2019 to achieve the following objectives:

- To engage with developers and procurers of AI-driven technologies in the UK to gain their perspective on the maturity of the landscape and progress to-date, alongside new opportunities and risks identified
- To reflect on the awareness and engagement the NHS AI Lab has had with developers and procurers in its first year.

The survey targets both companies that are developing AI solutions and commissioners that are procuring these solutions. For the AI Roadmap, only developer respondents were of interest. Given the length of the survey, many responses were incomplete (75 out of 183); 108 developers fully completed the survey. The survey questions are presented in Appendix A.

The NIHR AI Horizon Scan

In November 2020, the AAC recommissioned NIHR to re-run a landscape review for AI to inform the AI Award team planning and the AI Lab annual review. The Horizon Scan sought to ascertain the global pipeline of AI technological interventions that are in development or commercialised in health and social care. Appendix B describes the fields included in the Horizon Scan. The dataset included 153 UK-based technologies (out of 801), and these were included in the analysis.

The Innovation Pipeline National Master Data

The Innovation Pipeline has been developed to help select and adopt solutions for impact at pace and scale, it does not have a particular AI focus. At the core of its process are structured decisions gateways to enable prioritisation. Against these decision gateways, close-ended questions and a typology have been proposed. At the time of the analysis, the Innovation Pipeline was being tested across the AHSN Network and was building its minimum dataset. This dataset was mostly used to compare the list of care settings and clinical areas used.

The list of applicants for the AAC AI Award Round 2

The AI in Health and Social Care Awards aim to support technologies across the spectrum of development: from initial feasibility to evaluation within the NHS. Funding is made available for winners of the awards to accelerate the testing and evaluation of the most promising AI technologies which meet the strategic aims set out in the NHS Long Term Plan. The list of shortlisted Round 2 applicants was used to ensure these AI technologies had not been missed.

Initial metrics of interest

A number of measures relevant for the roadmap were identified through the analysis of the datasets, they are described in Table 2. These criteria were amended throughout the analysis, and the subsequent changes are explained in the report.

Table 2: *Measures of interest identified from the dataset*

Criteria	Sub-criteria
Organisation profile	▸ Type of organisation ▸ Size of organisation
Solution profile	▸ Solution type ▸ Clinical area ▸ Technology type
User type	▸ User affected
Maturity	▸ Innovation stage ▸ Time to deployment ▸ Regulatory approval / CE marking
Setting	▸ Point of care
Workforce	▸ Workforce group affected
Effect on tasks / decision-making	▸ Outcomes for users ▸ Specific tasks / decision-making affected

Analysis and findings

Criteria

Initial analysis was conducted across the datasets, where possible, to understand variances in the measures described in Table 2. The discrepancies in available measures across the four datasets are described in Table 3.

Table 3: *Variation of the sub-criteria within the datasets*

Sub-criteria	AI Survey 2021	Horizon Scan 2021	Innovation Pipeline HinM	AI Awards List
Type of organisation	✓			
Solution type	✓	✓	✓	
Clinical area		✓	✓	
Technology type	✓			
User affected	✓			
Innovation stage		✓	✓	✓
Time to deployment	✓			
Regulatory approval / CE marking	✓	✓		
Point of care	✓		✓	
Professional group affected				
Outcomes for users	✓		✓	
Size of the organisation	✓			

Due to the number of respondents and the granularity of the questions, the 2021 AI State of the Nation survey was chosen to be the primary dataset for the analysis, with the other datasets used for validation purposes.

Preliminary findings: Criteria

One should note that in the AI survey the number of respondents varied between questions, which explain why the total number of responses is different from one question to the next.

Solution type

Findings from the AI survey and the NIHR Horizon Scan revealed that 'diagnostic' was the most represented solution type within both datasets; almost 50% out of 183 respondents to the question selected diagnostic to describe their solution type. Whilst the options were different for the NIHR Horizon Scan, the equivalent category 'Diagnostic and Treatment' represented 45% of the technologies (69 out of 153).

Clinical area

The NIHR Horizon Scan included 24 clinical areas which can be found in Appendix B. The most frequently reported clinical area in the dataset was 'Medical condition unspecified' which accounted for 43% of technologies (66 out of 153). Following was 'Oncology' and 'Cardiology', both encompassing 10% of technologies (15 out of 153).

User affected

The 2021 AI State of the Nation survey allowed respondents to specify the users of their technologies. Out of 181 responses for this question, the most selected user was 'Clinician', with nearly 80% of respondents selecting, followed by 'Person with long term condition' accounting for almost 40% of respondents.

Innovation stage and Time to deployment

The NIHR Horizon Scan included a dimension named 'Innovation stage' with four phases (N=153):

- **Phase 1** - Proof-of-concept stage: 2%
- **Phase 2** - Prototype: 8%
- **Phase 3** - Tech validated/demonstrated in relevant environment: 75%
- **Phase 4** - Commercialised (i.e. regulatory approved/ready to market): 16%

Within the State of the Nation survey, respondents were asked to report on the likelihood of being ready for deployment at scale across three-time horizons (N=108). Part of the results are presented in the Table 4.

Table 4: *Key results from the State of Nation survey*

Readiness for deployment	5 years	3 years	1 year
Percentage of technologies declared that they were 'very likely' to be ready for deployment at that time horizon	More than 70%	More than 50%	Less than 35%
Percentage of technologies declared that they were 'likely' to be ready for deployment at that time horizon	More than 10%	Less than 25%	Less than 20%

Regulatory approval / CE marking

Both the AI survey and Horizon Scan included measures relating to regulatory approval. Out of 127 answers to this question, 19% reported that they were already classified, with 44% reported to be in the process of being classified or intend to seek classification, the 37% remaining declared they were not classified. Out of 153 technologies within the Horizon Scan, 14% were reported as having CE marking or MHRA marking.

Point of care

Out of the 160 respondents to the question asking about the point of care for deployment of their technology, the most frequently selected option was 'Secondary care' with over 70% of respondents. The second most selected was 'Primary care', accounting for over 55% of respondents. Multiple answers were allowed for this question.

Outcomes for users/system

Out of 160 responses to this question, the most selected outcome was 'system efficiency' with nearly 75% of respondents selecting this outcome. 'Faster diagnosis' was the second most frequently selected outcome, accounting for over 60% of respondents.

Preliminary findings: Early profiling

Grouping approach

In order to understand relationships between technology characteristics and to begin to curate technology profiles, a grouping approach was utilised for three key dimensions of interest found in the 2021 AI State of the Nation survey:

- Solution type
- User affected
- Outcomes for users

Four groups, and an 'Other' were created for 'Solution type' as displayed in Table 5.

Table 5: *The four solution type groups*

Automation	Diagnosis and treatment	Prevention and health promotion	Remote solutions
• Triage • Decision support (specified in other)	• Diagnosis • Therapeutic • Care-based • Self-care	• Health promotion • Population health • Screening (specified in other)	• Remote monitoring • Remote consultation

Within the 'User affected' dimension, three groups, and 'Other', were created with specified rules as presented in Table 6.

Table 6: *The three user affected dimensions*

Clinician	Commissioner	Patient facing (only if respondent did not select Clinician or Commissioner)
Any respondent who selected Clinician	Any respondent who selected Commissioner	Person with long-term condition Parent/Carer Person with a physical disability Person with a cognitive or learning impairment Person with broad care needs Person interested in monitoring their health (e.g. fitbit) Person wishing to access ad-hoc services (e.g. video consultation) Person seeking mental health support

The "Outcomes for users" were divided in four groups "Improved patient outcomes", "Service Improvement", "Diagnostic Improvement" and 'Other'. The first three groups are described in Table 7.

Table 7: *The three outcomes for users groups*

Improved patient outcomes	Service improvement	Diagnostic improvement
Improved Quality of Life Improved independence/ autonomy Prevention of ill-health/ improvement of health	System efficiency Better experience of health services Better experience of care services Better access to health services Better access to care services	Faster diagnosis, Faster identification of care need More accurate diagnosis

Findings

To validate the grouping approach used and explore relationships, an initial statistical analysis was conducted using the grouping for the 'User affected' measure and for two other key dimensions which were not grouped:

- Solution type
- Point of care

The analysis was conducted using the "MRCV" package in the R programming language, specifically designed for evaluating multiple response categorical variables like those seen in this dataset. For each question, a number of assumptions were made:

- For the entire dataset, "Other" responses were removed, as this type of analysis would not provide any insight for these response types.
- For the question "Which group of health and care system users is your AI-driven technology for?" it was assumed that any technology that does not have the carer and/or commissioner as the main user fit the category of "Patient". Any results that have patient and clinician/commissioner were assumed to have the latter as their main user. Response type "commissioner" was removed, as the number of these was too small to provide meaningful results in the analysis.
- For the question "At which point of care do you expect your AI-driven technology to be deployed?" it was assumed that community care and secondary care were the same response and could be joined together, as it was observed on a sample of the dataset that they were often selected together. Response "For the purpose of population screening" was removed.

Once these assumptions were made, the data was imported to R for analysis, producing the tables below. The responses analysed were those for which the main user was listed as "clinician" (n=125). Tables were then produced to test for significant relationships between the variables in the other two questions.

***Table 8:** The statistical significance between solution type and point of care*

How do you classify your AI-driven technology?	At which point of care do you expect your AI-driven technology to be deployed?				
		Point of care			
	Solution type	**Primary Care**	**Secondary Care**	**Tertiary Care**	**Individual**
	Diagnostic	1	1	1	1
	Therapeutic	1	1	1	1
	Population Health	1	1	1	1
	Care-Based	1	1	1	1
	Triage	0.6616	1	1	1
	Self-Care	0.8341	1	1	**0.0003***
	Health Promotion	0.1203	1	1	**0.028***
	Remote Monitoring	1	1	1	**0.0168***
	Remote Consultation	0.9848	1	1	0.4411
	Social Care	1	1	1	1

**The significant relationships have been bolded in the table.*

In Table 8, the numbers smaller than 0.05 indicate a statistically significant correlation between the two parameters compared.

Following the significance test, a correlation diagram was plotted to establish whether the relationships above were positive or negative:

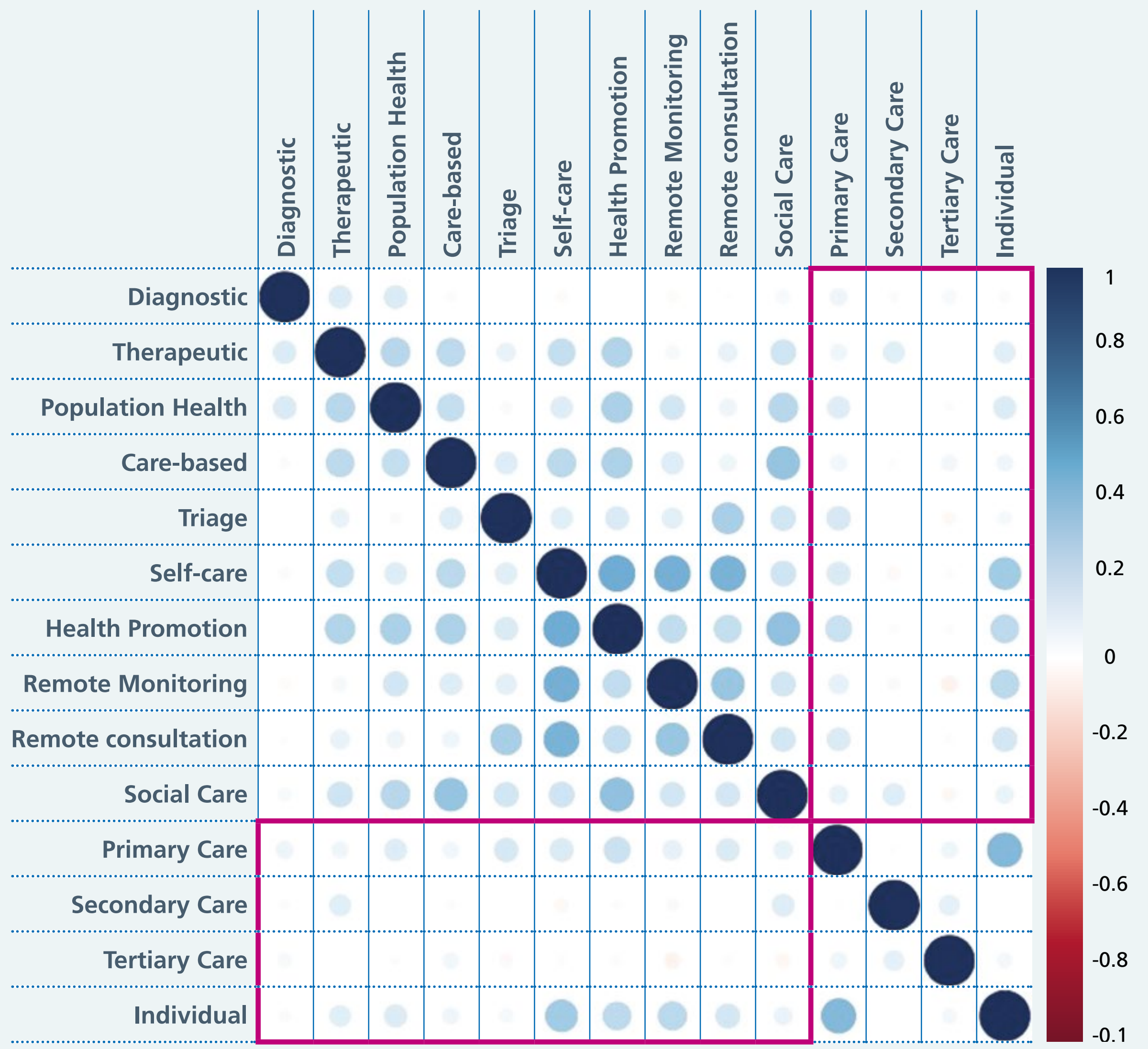


***Figure 2:** The correlation between solution type and point of care*

Figure 2 displays pairwise correlations between each of the response variables, with the correlations related to the above Table 8 highlighted in green. This supports the table, but it is also noted that the strongest correlations are between the responses from the same question, particularly between "Self-Care", "Health Promotion", "Remote Monitoring" and "Remote Consultation". Nonetheless, we can infer that the responses paint such a varied picture that no clear trends emerge – it could be due to data quality, sample selection or some other limitation or it could be that the landscape is broad and variable. At this point, there is not enough evidence to suggest that any of the care settings (except individual care) is linked to any particular type of solution.

Additional analysis was conducted using the grouping approach on all measures. Given the multiple-choice format of the survey, and even with the grouping approach, a significant combination of answers was found in each measure. Table 9 displays the top 10, out of 28 found, combinations for solution type.

Table 9: *The top ten combination of answer groups*

Solution type combinations	Count	%
Diagnosis and treatment	33	21%
Diagnosis and treatment, Remote solutions, Automation	11	7%
Diagnosis and treatment, Automation	11	7%
Diagnosis and treatment, Remote solutions	11	7%
Automation	10	6%
Diagnosis and treatment, Prevention and health promotion, Remote solutions	10	6%
Other	9	6%
Remote solutions	7	4%
Prevention and health promotion	7	4%
Diagnosis and treatment, Prevention and health promotion, Remote solutions, Automation, Other	7	4%

Further exploration was conducted to evaluate the relationships between the grouped measures. Table 10 presents the top five, out of 68 found, most frequently selected combinations of answers across the two measures.

Table 10: *The top five most frequently selected combinations of outcomes*

Solution type combinations	Outcomes for users combinations	Count	%
Diagnosis and treatment	Improved patient outcome, Service improvement, Diagnostic improvement	14	9%
Diagnosis and treatment, Remote solutions, Automation	Improved patient outcome, Service improvement, Diagnostic improvement	9	6%
Diagnosis and treatment	Service improvement, Diagnostic improvement	8	5%
Diagnosis and treatment, Prevention and health promotion, Remote solutions	Improved patient outcome, Service improvement, Diagnostic improvement	7	4%
Diagnosis and treatment, Prevention and health promotion, Remote solutions, Automation, Other	Improved patient outcome, Service improvement, Diagnostic improvement	6	4%

Similarly to the previous statistical method used, there was such a variability in responses there was no evidence of any real clear signals. One should note that this could indicate that it is a very diverse, variable landscape – which is an important insight in itself – however, more evidence is needed to confirm this.

Tree approach

It was agreed that a tree approach would be a pragmatic method to represent and explore relationships between criteria of interest:

- Type of users
- Type of technology
- Point of care
- Time to deployment

The levels of the tree approach are displayed in Table 11.

***Table 11:** Tree analysis levels and dimensions*

Level	Dimension
Level 0	Type of technology
Level 1	Type of users
Level 2a	Point of care
Level 2b	Time to deployment

The grouping of the branches of the tree was used to form the preliminary use case profiles and constitute the taxonomy. For the purposes of the point of care dimension, level 2a, grouping was harder to interpret as many respondents picked multiple options. Additionally, as the answers to the time to deployment were self-reported by the innovators, level 2b in the tree would potentially be biased by the innovators' optimism. Therefore, it was agreed that a set of more objective questions would be used to determine solution maturity and its time to deployment.

Out of the 157 responses used in the analysis, only 22 did not include a combination of diagnostic, triage, remote monitoring or population health as their solution type. Consequently, the four solution types were created as the four groups for technology type and allowed for combinations between them. Figure 3 displays the resulting tree, with technology type as the root of the tree.

Level 0	Diagnostic + triage + remote monitoring	Diagnostic + triage	Diagnostic + remote monitoring	Diagnostic	Remote monitoring	Population health	Triage
Level 1	Health care professionals	Health care professionals	Health care professionals	Patients + health care professionals	Patients + health care professionals	Patients + health care professionals	Patients + health care professionals
Level 2a	Primary care + secondary care	Primary care + secondary care	Primary care + secondary care	Primary care + secondary care	Primary care + secondary + self care	Population screening	TBD with examples
			Community + secondary care	Secondary care	Community care + self care		
Level 2b	1 year	1 year	3 years	1-3 years	1 year	1 year	1 year

***Figure 3:** The tree analysis results*

Limitations in the datasets

A number of limitations were identified in the initial analysis of the datasets.

The 2021 AI State of the Nation survey

The AI survey was a self-reported survey with multiple choice options for a number of questions. As such, respondents often selected multiple options when describing their benefits, the point of care or their user. There was a potential for biases across the survey with developers being over-optimistic when describing the functionalities of their solution. These issues raised questions regarding the reliability of the data.

An additional key limitation of the multiple-choice format was the significant number of respondents selecting multiple options, for example selecting all points of care for implementation. This created difficulties when trying to identify relationships between measures given the volume of different options for each question and raised questions regarding the reliability of the data.

An additional limitation to the survey was the option for developers to remain anonymous. Consequently, there was uncertainty around duplicated responses to the survey and made it challenging to cross-reference solutions between datasets. After the initial analysis stages of the work, the decision was made to remove all anonymous responses to the survey, to allow for the validation of answers.

The AI Horizon Scan

The AI Horizon Scan was an international dataset and therefore the purpose was not to understand the landscape of AI technologies in England, which was the key focus of the roadmap. Additionally, the scan was more clinically focussed and hence may have underrepresented a number of technologies. Since one aim of the roadmap was to understand in full the different types of technologies ready for or nearing deployment and their distribution, it was not possible to solely rely on the list of technologies within the Horizon Scan.

The Innovation Pipeline National Master Data

The key limitation of the Innovation Pipeline within the work of the AI Roadmap was that the dataset was not AI specific.

Validation process

To address some of the limitations identified during the initial analysis, 15% of each technology type was used to validate the taxonomy for technology type. This was conducted for technologies from the survey who had declared the company name within the survey. Example technologies included in the validation were:

- Diagnostic + Triage + Remote monitoring: Biomind, Kemuri, DocMe, Feebris
- Diagnostic + Triage: qXR, Limbic.ai, Skin Analytics
- Population health: MySense, Scaled insights
- Triage: Visiba Care, PinPoint Test

The validation was conducted through desktop research on the technologies themselves. Findings of the validation demonstrated that the technology type declared often did not match the result, in particular for combinations of technology types. It was agreed that desktop research would be conducted across all technologies included in the database to ensure accuracy and consistency across the datasets. Additionally, it was agreed that further analysis and research was to be conducted to improve the list of technology types.

"This database is a really useful compilation of the existing AI technologies in development or use within health care in England as there is no central reference for this currently. If kept up to date and evolved it has the potential to inform national, regional and system strategies around the technologies that should be rolled out more widely across the health and social care."

Emma Hughes, Senior Manager – Innovation, Research and Life Sciences, NHS Accelerated Access Collaborative

Use case profiles and taxonomy

The use case profiles were developed around the technology type to present the diversity across the different measures and technology characteristics, for example the spread across the different time to deployment horizons. The six use case profiles of the different technology types were:

- Automation/Service efficiency
- Diagnostic
- P4 Medicine
- Remote monitoring
- Therapeutic
- Other

Following analysis of the datasets and desktop research, a taxonomy for technology type and subtype was created. The taxonomy is described in full in Appendix C.

2

Creation and population of the database

Aims of the database

The database was created through an extensive Horizon Scanning exercise in order to provide a reliable information source for 240 AI technologies researched. The database was used to build the dashboard which illustrates the diversity of the AI technologies and how the characteristics differed across the technology types. An additional key objective of the database was to understand the variations in the distribution across the technology types and workforce groups affected as this would support the selection of the technologies for the case studies.

Database template

The full database template can be found in Appendix D.

Technology type and subtype

The technology type and subtype included in the database were taken from the taxonomy which was an output of the initial analysis and tree approach used to understand the characteristics of the technologies.

Clinical area

The list of clinical areas that were included in the first draft of the database template were taken from the list utilised in the NIHR Horizon Scan which included 24 clinical areas that can be found in Appendix B. Following the presentation of the first draft of the database template to the HEE team, the list of clinical areas was replaced with a list provided by HEE, which included 66 clinical areas, all mapping back to NHS occupation codes and can be found in Appendix D.

Primary user

Four options were utilised in the database template for primary users: Healthcare Professional, Patient, Carer, Commissioner. These four options were used in the initial analysis, taken from the State of the Nation AI survey 2021.

Point of care

Two columns were included for the point of care and the options were taken from the State of the Nation AI survey 2021. One should note that:

- 'Tertiary care' was excluded from the list of options and captured within 'Secondary care'.
- 'For the purposes of population screening' was excluded due to the ambiguity around what areas of care this could be; in most cases they were deployed in Primary Care in General Practice.
- 'Research' and 'Education' were added to the list of options, as several technologies were to be deployed in these areas and these points of care were not covered within the original list.

Additionally, as the exercise progressed, the researchers encountered technologies which impacted a point of care outside of the one they were used at. Therefore, a third column within the point of care dimension named 'Secondary point of care' was added with the same list of options.

Sites

Four columns were included in the database to indicate known implementation NHS sites of the technology. The number of columns, e.g. the maximum number of sites captured by the database, was determined using a small sample of technologies and the number of sites that they have been deployed in. Additional sites were noted down for any technologies that had been implemented in more than four known sites. The list of sites included all Clinical Commissioning Groups (CCGs), Acute Trusts, Mental Health Trusts and Community Trusts.

Workforce groups affected

In the first iteration of the database template, the list of workforce groups was curated using the preliminary analysis on the technologies and the NHS Health Careers website. Following the first presentation of the database template to the HEE team, it was agreed that HEE would formulate their own list of workforce groups to be used, so that they could be traced back to occupational codes. The list of workforce groups that was provided by HEE contained 155 workforce groups. A column titled 'Other' with a free text box was also included in the database for technologies that specified a workforce group or healthcare professional as being directly impacted, that did not fall into the list provided by HEE.

Impact on the workforce

In order to define the impact on the workforce framework, two engagement sessions with a wider HEE team were held. The aim of these sessions was to present findings from desktop research on a selection of technologies from different technology types and brainstorm the different impacts on the workforce. Additionally, the sessions' aim was to consider how these impacts can be grouped into a framework that can be applied to all the technologies in the database.

The output of these sessions was a list of impacts which were grouped into three dimensions: the impact on the workforce groups affected, impact on the pathway and impact on the system. The impacts found in the engagement sessions were complemented by some desktop research focusing on literature relating to the impact of AI on the workforce and required changes in order to fully benefit from the technologies.

Based on the list of impacts identified in the engagement sessions, impact groups were formulated to create nine impacts overall, three in each impact dimension, to complete the impact framework. The table in Appendix E presents how the impacts were further grouped. Given the breadth of technologies within the database and presented in the engagement sessions, some of the listed impacts did not fit into any of the nine groups and related more to impact so they were excluded from the framework. Examples of these impacts include:

- Change in user's wellbeing, work life
- Change in the ergonomics

In the first version of the framework, the scale used for each impact group was: 'Reduction', 'No change', 'Increase'. However, upon further discussions it was agreed that some impact groups would require a different scale. Two other options were added to the scale; 'Early-stage solution – undetermined' was included for technologies that did not present their value proposition or benefits; 'No direct impact' was included for technologies that had little or no interaction with healthcare services and did not pose any direct impact on the workforce groups, the pathway or the system. Technologies could not have one of these two impact options on only one of the impact groups; it had to be selected as the impact option across the whole framework. The impact framework with the three dimensions and scales for each impact group are displayed in Appendix E.

Time to deployment

The time to deployment is defined as the time until the technology is ready to be deployed at scale (regional or national roll-out), e.g. a technology with a time to deployment of one year would be already implemented in NHS sites. The time to deployment in the State of the Nation survey 2021 used a 3-point scale; within 1 year, within 3 years, within 5 years and it was a self-reported estimate by the technology themselves. The time to deployment criteria was created to provide a more objective assessment for the time until the technology is deployable at scale across multiple sites and provide a set of criteria based on information that could be collected through the Horizon Scanning exercise for all the technologies in the database. The criteria use a Yes/No scale and includes a weighted score with a minimum score required in each of the 3 timescales. Some technologies did not meet the minimum score for five years so in the final database the time to deployment was a 4-point scale (with the addition of over 5 years to deployment). The Appendix F presented the criteria used and the scoring logic used.

Rules and assumptions

A set of rules and assumptions was considered for the Horizon Scanning exercise in order to maintain consistency in the database with the different researchers. The rules and assumptions were presented and approved by the HEE team.

Primary User

Within the primary user column:

- 'Clinician' was selected for technologies where a healthcare professional was a directly using the technology themselves
- 'Patient facing' was selected for technologies where patients were the direct user of the technology; clinical team may review data collected by the solution
- 'Carers' was selected for technologies where the primary user was a nonpaid carer
- 'Commissioner' was selected for technologies that integrate into the system but don't directly interact with staff or patients

Point of care

The point of care captured within the database is the care setting where technology is deployed currently or where it will be deployed in the first instance. Two columns have been provided for point of care however this was only utilised for a small number of technologies where there was complete certainty that the technology would be deployed in two different care settings, for example both primary and secondary care.

Sites

When the database template was first formulated and following the research relating to technology implementation on the sample of technologies, it was proposed that 'Widely used' would be selected in the 'Site' column for technologies that did not specify the sites that they had been implemented in but reported that they had partnered with multiple NHS organisations. Following discussions with HEE around the potential for misinterpretation and the ambiguity around using 'Widely used', it was agreed that 'Unspecified site' or 'Unspecified multiple sites' would be options for such technologies.

Workforce groups affected

It was agreed that only workforce groups directly impacted by the technology would be captured in this column and workforce groups indirectly affected would not be captured in the database but would be investigated via the case studies. In addition to the list provided by HEE, it was agreed that three additional options were to be added to the list;

- 'Multiple roles affected' to be included as an option for technologies that did not directly impact a specific workforce group, but multiple roles.
- 'Undetermined' to be included as an option for technologies that did not specify or indicate which roles would be impacted or used terms such as 'clinician'.
- 'No workforce group directly impacted' to be included as an option for technologies that did not pose any direct impact on the workforce.

Impact on the workforce

The impact on the workforce framework was reliant on self-reported claims made by the technology themselves and these claims were not validated as part of the exercise. There is an ambition to validate the claims made by the technologies, as described in Section 6.

Time to deployment

It was agreed that if the information required for each of the criteria in the time to deployment framework was not found the first two pages of a Google search, then, as researchers had a time cap on how long they could spend on each technology, 'No' would be selected for that criterion.

3

Case studies

Selection of the technologies and engagement

Shortlisting of the technologies

The Table 12 below presents the criteria and the weighing used to shortlist the technologies.

Table 12: *List of criteria for the shortlisting of the technologies*

Level	Dimension
Time to deployment: within a year	5
Solution impacting a clinical area prioritised by NHSE/I or other national programmes: Yes	2
Existing networks/initiatives supporting the solution type or the clinical area impacted by the solution: Yes	1
Documented workforce opportunity in the healthcare professional group using the solution: Yes	1
Solution spread across regions: Yes	1

All technologies with a score of 7 or above were selected to be part of a shortlist. A total of 28 solutions were shortlisted, the list was presented to the HEE team so they could select the two technologies which would constitute the case studies. The priorities when choosing the case studies were for HEE to ensure that the solutions had been used or tested in the NHS and that different clinical pathways, workforce groups and point of care were represented. Based on these considerations they picked:

- **Oxevision, from Oxehealth**, is a non-invasive vision-based monitoring device which is deployed in inpatient mental health wards to help staff plan care and intervene proactively. It is used in Mental Health Trusts by adult and mental health nurses and has been implemented across 3 AHSNs footprint.
- **Optellum's Lung Cancer Prediction AI** helps physicians make optimal clinical decisions by providing a diagnosis of incidental pulmonary nodules. It is currently being implemented on the U.S. market, including leading medical centres and nationally recognized leaders in lung cancer. It has been selected as one of the Round 2 AI Award solutions.

Engagement activities

Once the technologies were selected, Unity Insights contacted AHSN colleagues who had worked with the companies to present the project and ask if introduction could be made. The Appendix G presents the questionnaire developed to conduct the case studies. The information collection should be done by using the evidence and documentations available from the company as well as by interviewing a member of staff involved in the deployment of the technology.

The Oxford AHSN team provided an introduction to Charlotte Wood, Director of Oxehealth. After an introductory meeting with Charlotte to present the project and the aim of the case studies, we were introduced to Hayley Bolton, ward manager at Cumbria, Northumberland, Tyne and Wear NHS Foundation Trust, one of the sites using Oxevision. For Optellum, the HEE team was able to make an introduction and Rhiannon Lassiter, Head of Marketing and Communications, attended an introductory meeting. Engagement to develop the case studies with front line staff proved more tricky than anticipated due to clinical pressures and it was decided that the case studies would focus on presenting the learnings of the company to date as well as the dimensions they wish to explore through the AI Award.

One should also note that the companies are at different stages of implementation. Indeed, Oxehealth has gathered real-world evidence of outcomes and economic impact, whilst Optellum is looking to develop a body of real-world evidence with their AI Award implementations. This is reflected in the case studies.

Oxehealth

What is Oxehealth?

Oxevision, from Oxehealth, is a non-contact vision-based patient monitoring system which is deployed in inpatient mental health wards to help staff plan care and intervene proactively. Oxevision measures pulse rate (through skin micro blushes), breathing rate (through chest/diaphragm movements) and activity of patients in their rooms. The data collected by Oxevision provides staff in the wards with alerts to early warning signs, reports on risk factors and the ability to take pulse and breathing rate spot-check measurements. Staff can access this information on the viewing screens that are installed at nursing stations and on dedicated tablets for when moving around the ward as displayed in the diagram in Figure 4.

Other features of Oxevision includes the ability to replay the incident to understand what happened.

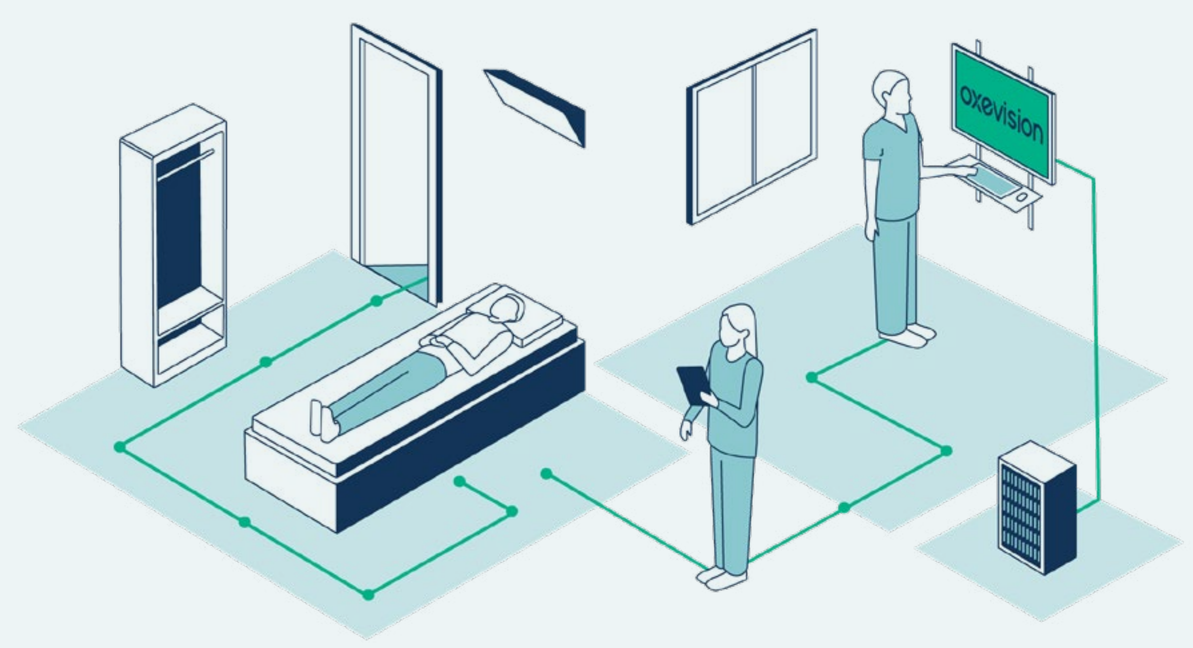


Figure 4: *Diagram showcasing how nurses can use Oxevision to access patient information within the ward*

Problem addressed

In acute inpatient mental health units, it is compulsory for patients to have nursing observations at night, sometimes as often as every 15 minutes (Barrera et al., 2020). These checks are to ensure patients are safe and breathing, however they can also disturb patients while they are sleeping which can negatively impact their recovery and prolong their stay in the ward. In their study, Malik et al. found that in comparison to those without sleep disturbances, patients with psychiatric diagnoses and co-morbid sleep disturbances were 99% more likely to report suicidal behaviours (Malik et al., 2014).

Moreover, in older adult wards, patients can also easily come to harm in between the regular checks. Assistance may be required for dementia patients in case of falls. These incidents often require multiple members of staff and are unpredictable and often unwitnessed. Furthermore, data collected by the Care Quality Commission (CQC) revealed that 224 people died of self-inflicted injuries between 2010 and 2016 in mental health hospitals in England (Care Quality Commission, 2018).

Evidence to date

- **Clinical and economic evidence:** In older adult wards, clinicians making use of Oxevision have managed to reduce falls at night and harmful falls by respectively 48% and 82% (Wright et al., 2021). This translated into a positive economic impact, with a £6.94 return for every £1 invested in the deployment of Oxevision in older adult wards. In acute wards, clinicians making use of Oxevision has resulted in a 44% reduction in self-harm in bedrooms and a 66% reduction in ligatures in en-suite bathrooms, with a £2.40 return on investment for every £1 invested. Additionally, in Psychiatric Intensive Care Units (PICU) a 26% reduction in assaults in bedrooms and a 40% reduction in rapid tranquilisations related to assaults was observed, with £3.70 return for every £1 invested (Royal College of Psychiatrists, 2020; York Health Economics Consortium, 2021). An additional study revealed that in a Quality Improvement project the rate of obtaining clinically accurate vital signs (pulse & breathing rate) of patients in seclusion increased by 12.3 times compared to real-world baseline and 5.5 times compared to policy guidelines (Clark et al., 2021).
- **Qualitative findings:** across five NHS Mental Health Trusts, 8 out of 10 patients agreed that Oxevision helps provide a better sense of safety and 7 out of 10 patients agreed that Oxevision helps improve sleep (Oxehealth, 2021). Staff working in Trusts where Oxevision has been deployed also shared positive views; in data collated from seven NHS Mental Health Trusts, 94% of staff agreed that Oxevision helps improve patient safety. Furthermore, in data compiled from six Trusts, 73% of staff agreed that Oxevision helps better manage their own safety (Oxehealth, 2021).

Case study methodology

In addition to the publications, reports and case studies provided by the Oxehealth team, insights of the impact on the workforce were also gathered through an interview and a focused group discussion.

- Hayley Bolton, the ward manager at Longview Ward within Cumbria, Northumberland, Tyne and Wear NHS Foundation Trust, was interviewed. Longview ward is an acute admission ward for females over the age of 18 years, with a mental illness who require assessment and treatment in hospital. It has a total capacity of 18 and the average length of stay is 33 days. Oxevision was deployed in the Trust in February 2020.

- The Digital Health London Accelerator Programme organised a Virtual Showcase to look back at the 2020/2021 cohort. One of the sessions was dedicated to Central and North West London NHS Trust (CNWL) which has recently rolled out Oxevision at scale. It presented three perspectives from across the organisation including the Associate Director of Nursing, the Business Strategy & Transformation Manager and Modern Matron about their experience.

The transcript of the recorded interview and the notes taken during the webinar were compiled and thematic analysis was used to present the learnings from past implementations and the impact on the workforce.

Learnings from past implementations

Oxevision is currently deployed in 40% of NHS England Mental Health Trusts and over 100 wards, supporting over eight million hours of patient care (Oxehealth, 2021).The lessons learnt listed below are based on the published results of the Vaughan Thomas Ward implementation (Barrera et al., 2020), learnings from implementation in Essex Partnership University NHS Foundation Trust (Essex Partnership University NHS Foundation Trust, 2021), as well as the data collected during the interview and webinar.

- **Early engagement with staff members and patients** to gain support amongst staff and help manage change. At the Longview Ward, it was reported that the ward held focus groups with patients to share information regarding the system and ensure communication around it prior to the 'go-live date'. At Vaughan Thomas Ward, Standard Operational Procedures (SOPs) and patient leaflets were developed with clinical champions and patient advocates to support a clear communication. Similarly, when reflecting on who should be involved earlier in the roll-out, CNWL reported that service users and carers should be involved in the process "from day one". Learnings from Essex Partnership University NHS Foundation Trust also aligned, with an emphasis on frequent and persistent patient engagement by ward staff.
- **Ward champions supporting the launch of the solution** were key in all sites to enable a successful deployment. To support the progressive transition, Longview ward introduced five 'super users', who were trained first and shared the learning with their teams (train-the-training model). This process enabled 'peer-on-peer feedback' and helped the system to be adopted organically faster than if the implementation were managed only by the company or the senior leadership team. Similarly, CNWL also noted that they have frontline champions from each ward, they attend a monthly group meeting to discuss Oxevision and share any challenges and learnings.
- **Estates and Information Technology (IT) teams are key stakeholders**. The CNWL team highlighted the importance of involving these teams as early on as possible. Indeed, the tablets which allow staff to view insights produced by Oxevision while moving around the ward rely on good Wi-Fi coverage across the ward. This is all the more important as it was reported that if staff can only view insights on the screens placed the ward office, Oxevision is less likely to be used frequently.

Oxehealth
Impact on workforce

- **Impact on workforce groups.** Whilst mental health nurses are the workforce group most impacted by the use of Oxevision, other teams on the wards have access to the platform and therefore are indirectly impacted. These are occupational therapists, medics, the psychological team and the peer support team.
- **Change in workload**. When it is clinically appropriate, staff can review the patients' observations on the platform rather than physically visiting their rooms to collect the patients' vitals, thus saving staff time which can be reinvested in patient care. The economic savings reported by York Health Economics Consortium are a £3.64 return for every £1 invested, with 71% non-cash releasing and 29% cash releasing. Overall, the cash releasing benefits were predominantly delivered through reducing 1:1 observations, a proportion of which were cash releasing from reducing agency or bank spend. Non-cash releasing time efficiencies are delivered through a reduction in 1:1 observations within safe staffing levels, reduction in incidents and faster night-time observations (York Health Economics Consortium, 2021). Similarly, the ward manager at Longview ward stated that 'it does create more capacity for staff to be available to do other things', with the main time savings being at night for staff completing observation rounds. She also declared that the solution is quick and effective to use, with little training time required. However, she noted that it was difficult to ascertain the true impact on staffing levels due to COVID-19.
- **Change in role**. As Oxevision is embedded into everyday practice at Longview Ward, it is now part of the daily review of patient care meeting, where staff decide when and why to use Oxevision with patients. A recurrent theme in the interview was the increase of positive risk-taking thanks to clinicians making use of Oxevision. Indeed, the staff can review the patients' vitals on-the-go on the tablets and are able to intervene more quickly. Moreover, they have access to additional patient's information such as activity report, this means that they have more confidence in their clinical decisions. Thus, the use of Oxevision has shifted roles towards a risk-based and evidence-based approach leaning on well-reported observations. It has enhanced decision-making for nurses. With regards to a change in system inputting, staff are currently required to manually transcribe observations and insights from Oxevision onto the electronic care record. A future improvement of Oxevision will intend to increase interoperability between electronic health care records and the Oxevision platform.
- **Change in team dynamic**. It was reported in the Longview Ward interview that there is a change in team dynamic with the team described as 'more collaborative at times'. Since many teams have access to the Oxevision platform, it leads to more collective discussions about clinical decisions and a more holistic care plan. The change in interactions between staff was also highlighted in findings from across eight Mental Health Trusts, as a staff nurse reported that Oxevision eases handovers between shifts, and further facilitates communication regarding behaviours and decisions (Oxehealth, 2021).

Oxehealth
Impact on the pathway

- **Change in patient outcomes.** The use of Oxevision was reported to positively impact patient safety and patients' outcomes, with a reduction in falls, self-harm, and assaults for Trusts deploying Oxevision in their wards (Oxehealth, 2021). A clinical study by Coventry and Warwickshire Partnership NHS Trust showed a 48% reduction in falls at night and a 68% reduction in demand for A&E services (Wright et al., 2021). The improvement in patient outcomes was also evidenced through enhanced patient experience. In addition to an improved sense of safety, 7 out of 10 patients agreed that Oxevision helps improve wellbeing. Besides, one patient reported that knowing Oxevision was monitoring them gave them 'peace of mind knowing that if anything happens to [them] the staff are alerted, and they can come in and do what's necessary' (Oxehealth, 2021). In Acute wards, there was a 66% reduction in ligatures in en-suite bathrooms and 44% reduction in self harm in bedrooms. In PICU, a 26% reduction in assaults in bedrooms and a 40% reduction in rapid tranquilisation related to assaults were reported (Royal College of Psychiatrists, 2020). An improvement in patient outcomes aligns with the interview testimony as the ward manager stated that Oxevision is enhancing recovery and enabling staff to use engagement time more proactively and have more meaningful interventions. Subsequently, the ward manager felt that patients' recovery journeys were shortened or that their journey was more positive (higher the quality of care).
- **Change in support, communication or education provided to patients.** A key topic that arose, in both the Longview Ward interview and the CNWL webinar, was the importance of communication to ensure positive experiences of Oxevision. When patients are admitted to the ward, staff will present Oxevision and its benefits. At the Longview Ward patients are not able to opt-out and Oxevision is an integral part of care. For a few patients who are presenting with paranoia, as noted in the interview, regular communication about the system and its functionalities is essential to reassure them. At CNWL, patients can opt-out when admitted and Oxevision can be turned off in their room.

  A key element noted by CNWL was communicating the benefits of Oxevision to patients and explaining the practicalities around improved sleep and safety, as well as it being a tool to supplement nursing practice rather than replacing it. Moreover, they highlighted the importance of repeated communications to build the patients' trust in the system.

  A key challenge, as noted from CNWL is staff confidence to communicate with patients about Oxevision. They reported that Oxehealth provided training sessions with staff on how to communicate about Oxevision and conducted practice conversations to ensure staff felt confident. This coincides with Longview ward's experience, as they stated that staff communication with patients was an added pressure in the early stages of implementation. The key element in overcoming this challenge was understanding the technology and being able to communicate with other staff and patients what it does and why it is there. It is apparent that Oxehealth understand the importance of engaging with staff to ensure confidence as noted in Essex Partnership University NHS Foundation Trust insights of their Oxevision trial, where the open communication and support from Oxehealth meant 'staff gained a sense of ownership' (Essex Partnership University NHS Foundation Trust, 2021; UCLPartners, 2021).

- **Change in waiting time, time to diagnosis, treatment or referral.** In some cases, AI technologies may have an impact on waiting time, time to diagnosis, treatment or referral. Although Oxevision delivers insights to assist clinical staff with care and planning, there is no direct change in waiting time, time to diagnosis, treatment or referral. In the interview, the ward manager did not report a change due to clinical staff using Oxevision on the wards.

Oxehealth
Impact on the system

- **Change in delivery of integrated care.** A change in the delivery of integrated care means the use of a solution may improve collaboration between care settings by enabling data sharing across primary and secondary care for example. Findings from the Longview Ward interview did not report that the deployment of Oxevision had an impact on the delivery of integrated care.
- **Change in access to care.** A change in access to care means that the use of a solution is either improving the access to care by engaging with hard-to-reach groups or reducing it by making it harder for certain patient populations to access NHS services (using a digital service for patients with low digital literacy for instance). The Longview Ward interview did not report a change in access to care. It was noted that Oxevision may indirectly improve access to care for patients who were previously refusing physical checks and who are now being better monitored thanks to the system.
- **Impact on the system performance, efficiency, or resilience.** The impact on the system performance and efficiency have been explored in the previous dimensions of the framework. Indeed, for every £1 invested the weighted average return is £3.64. Additionally, insights produced by Oxevision improve patient care and staff confidence on the wards, with 90% of patients reporting that they felt staff provided better care with the Oxevision system installed in rooms (Oxehealth, 2021).

Optellum

What is Optellum?

Virtual Nodule Clinic, from Optellum, is an AI powered clinical decision support software to assist clinicians in diagnosing early-stage lung cancer. This is achieved by identifying and tracking at-risk patients who present suspicious lung nodules. The objective is to treat patients before the disease has metastasized and subsequentially improve survival rates. The features of the Virtual Nodule Clinic can be utilised for two key functions in the early lung cancer diagnosis pathway. The first function is the coordination of care in hospitals by automatically reviewing radiology reports and identifying patients with lung nodules flagged in any CT scan. Subsequentially, a dashboard is populated and updated to present the information gathered, allowing clinicians to track patients and ensure timely action by clinical teams. The second function of the Virtual Nodule Clinic is diagnostic support for radiologists and chest physicians. Support is provided through automatic analysis of a user-selected nodule on a CT scan and assigning a clinically validated nodule-specific Lung Cancer Prediction (LCP) score which indicates the risk of a nodule being cancerous. The steps to compute the score for nodules of interest are displayed in Figure 5.

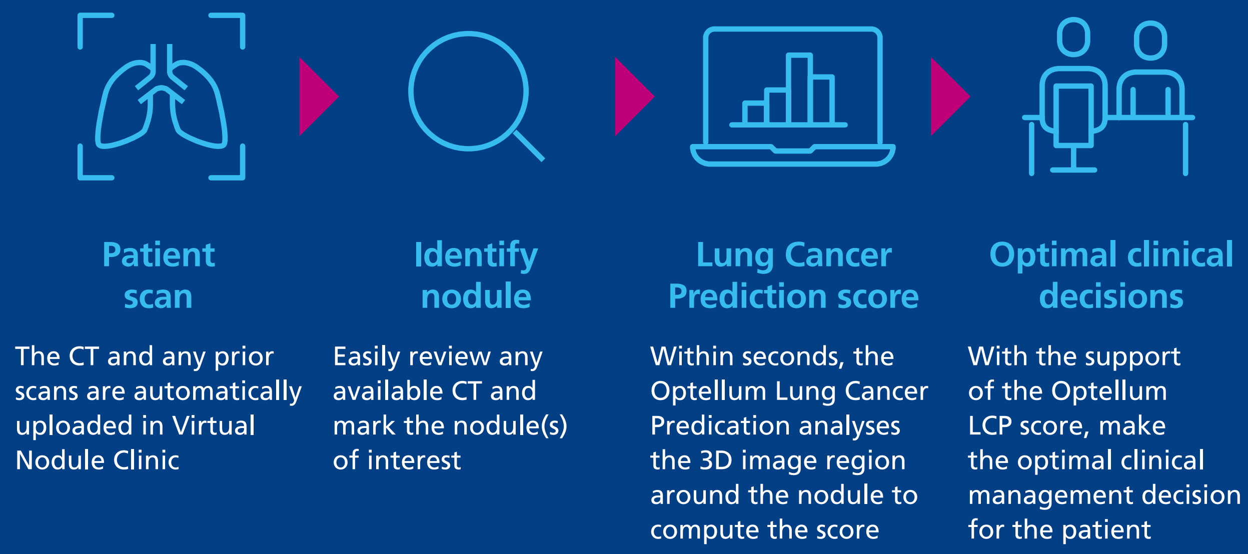


Figure 5: *Optellum user journey*

Problem addressed

In the UK, lung cancer has the highest mortality rate of all cancers and its current yearly cost to the NHS is £307m (Digital Health, 2020). Each year, 47,000 people are newly diagnosed and prevalence rates have risen 23% since 2004 (British Lung Foundation, 2012). Furthermore, over 35,000 people die from lung cancer each year and survival rates have not shown much improvement in the last 40 years, with the current 10-year survival rate being 9.5% (Cancer Research, 2017). Survival rates of lung cancer are heavily dependent on stage of diagnosis, with the 5-year survival rate of stage 1 and stage 4 lung cancer being 57% and 3% respectively (Cancer Research, 2017). It was revealed by experts that late-stage diagnosis is common due to GPs missing signs despite repeated visits. Up to 56% of people in some parts of England are only diagnosed with lung cancer when visiting A&E. Such people are five times more likely to die within a year than those whose cancer was diagnosed by a GP or through a cancer screening programme (UK Lung Cancer Coalition, 2020).

The NHS highlighted the need for a faster pathway from referral to diagnosis for lung cancer in the 2018 handbook for local health and care systems 'Implementing a timed lung cancer diagnosis pathway' (NHS England, 2018). The handbook underlines how to achieve diagnosis from referral within 14 and 28 days and highlights the key role of cancer alliances in delivering large scale transformation across whole systems. A key issue identified in the lung cancer diagnosis pathway is the stage of diagnosis when compared to other cancers. Figure 6 displays the stage of diagnosis for lung cancer compared to all other cancers, which may correlate to the 1-year survival rate. The most common stage of diagnosis for lung cancer in 2016 is stage 4, in comparison to stage 1 for all other cancers. It is also common in the diagnostic stages of lung cancer for biopsies to be conducted unnecessarily, with between 20% and 40% of lung biopsies estimated to be performed on patients with benign lung nodules (Grogan et al., 2011).

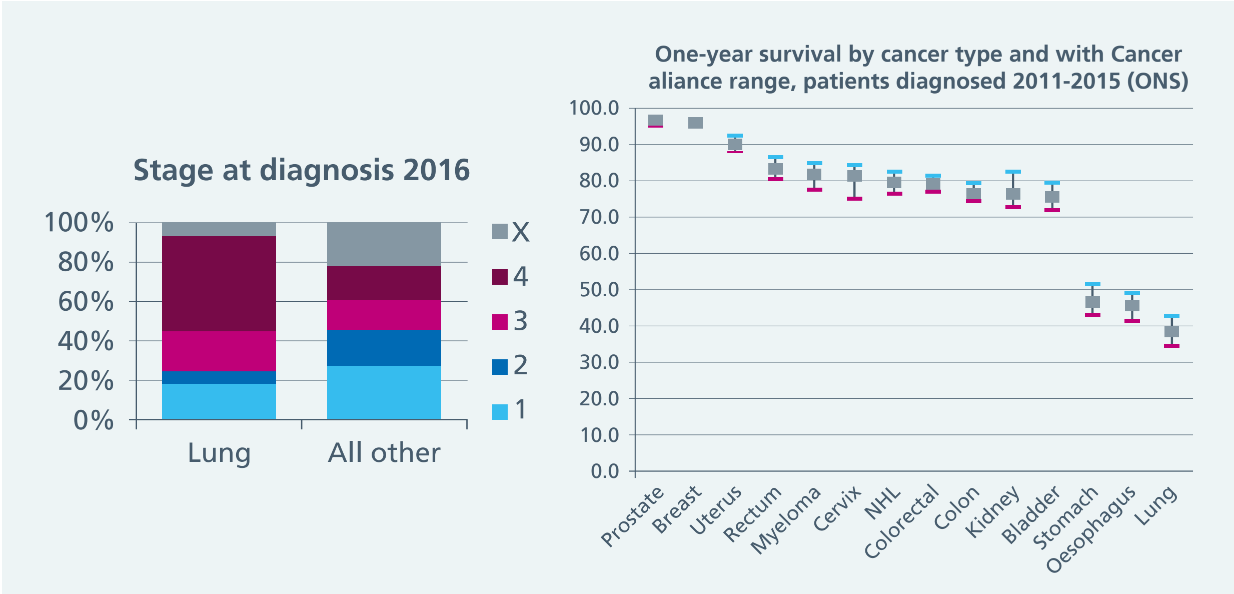


Figure 6: *Cancer diagnosis stage and 1-year survival rate comparison (NHS England, 2018)*

Evidence to date

Optellum received FDA 510(k) clearance in March 2021, being the first such application of AI decision support for early lung cancer diagnosis to be cleared (Optellum, 2021). They are currently seeking CE marking. The Virtual Nodule Clinic has been clinically validated in both the US and UK in multi-centre studies to assess the accuracy of the AI driven technology in the early-stage lung cancer pathway:

- External validation of a convolutional neural network artificial intelligence tool to predict malignancy in pulmonary nodules. The validation datasets used in these studies were independent of the training datasets with a total of 1187 patients with lung nodules from multiple NHS Trusts across the UK. The LCP was used for overall classification accuracy and as a test to rule-out patients from high and intermediate risk groups. This was done as a retrospective validation. Results included a difference in overall classification accuracy of the LCP compared to the standard-of-care Brock model, with 89.6% compared to 86.8% respectively, a statistically significant difference (Baldwin et al., 2020).
- Assessing the Accuracy of a Deep Learning Method to Risk Stratify Indeterminate Pulmonary Nodules. The study used independent data from the UK (Oxford University Hospitals) and the US (Vanderbilt University Medical Center) with a total of 543 patients with nodules used as part of this non-training dataset. Improved accuracy highlights included 61% of benign nodules being additionally and correctly reclassified as benign compared with the standard-of-care Mayo risk model (Oxford data) using pre-defined thresholds. 31% of cancers were additionally and correctly reclassified as cancers compared with the standard-of-care Mayo risk model (Oxford data) using pre-defined thresholds (Massion et al., 2020).

Case study methodology

Publications, reports and information was provided by the Optellum team to understand the potential impacts on the workforce. Due to time constraints and current pressures on the NHS workforce, it was not possible to conduct an interview with clinicians using the technology to gain further insights. A future update of the case study would benefit from including interviews with users at implementation sites.

Learnings from deployment of AI in the NHS Lung Cancer Screening Programme

Optellum is currently being tested as part of research programmes and trials in several NHS Trusts, with funding from the NHSX AI Awards 2021. The lessons learnt listed below are based on early-stage piloting of the Virtual Nodule Clinic.

- Integration of AI with the deployment process to better suit the NHS system and procedures. Since Optellum was first piloted in the US, it has been noted as an adoption challenge due to the difference in healthcare systems.
- Improvement of user training and sharing technical information relating to how the algorithm is trained and validated with the clinicians.
- Improvement in the display of insights produced by the technology. Changes were made to enhance how the LCP score was presented on the user interface and how it could be exported.

A potential change in lung cancer diagnostic services

Earlier diagnosis of cancer has become a focus for the NHS, with the Long Term Plan 2019 ambition setting out of diagnosing three in four cancers at stage 1 or 2 (NHS, 2019). In the UK there is currently no national lung cancer screening programme. However, the NHS launched Targeted Lung Health Checks in some areas of the UK since autumn 2019. The NHS England Targeted Lung Health Checks programme is the first large-scale lung cancer screening initiative substantially supported by AI. Nineteen pilot lung checks have been rolled out as reported in August 2021, with 15 using an AI lung nodule management solution. The aim of the screening pilot is to identify approximately 3,400 cancer cases and save hundreds of lives over four years, which will require analysis of 200,000 additional CT scans by the radiology workforce (Aidence, 2021).

It was announced in 2020 that Optellum would be part of a research programme with other AI health technology companies, NHS Trusts, and a team of academics to accelerate pathways for earlier diagnosis of lung cancer and other thoracic conditions (Digital Health, 2020). The research programme has been conducted alongside the NHS England Lung Health Check programme and is combining clinical, imaging, and molecular data. The objective of the research programme is to improve diagnostic accuracy and time to diagnosis, in addition to reducing the number of invasive clinical procedures. In order to identify at-risk patients to be selected for screening, the programme will link to data from primary care and assess risk in the general population. Selected patients will be invited for a screening which will include CT scans and blood tests. These scans and tests will subsequentially be analysed with the assistance of AI technologies including Optellum.

A unified approach between care settings is essential for successful implementation of lung cancer screening programmes. Patients for the NHS England Lung Health Check programme are identified through a complete national electronic register, usually based on broad demographic data. The screening infrastructure and capacity to deliver the programme includes community facilities which use mobile CT scanners and primary care facilities to support assessments for eligibility and health checks. Patients will also be referred to secondary care services with additional tests and treatments. (NHS England, 2019). A successful lung cancer screening pathway is likely to require changes in the workforce's role, support and training. It was noted within the 'Implementing a timed lung cancer diagnostic pathway' handbook that to ensure a faster pathway, workforce utilisation features included (NHS England, 2018):

- Workforce development for teams to support new ways of working across the whole pathway
- Co-location of medical, nursing, navigator, and support staff to improve communication, aid business intelligence, reinforce team integration, and enable effective day-to-day working
- Patient navigators for administrative support and value in tracking patients for improved flow

By implementing Optellum's software in the lung cancer diagnosis pathway through screening programmes such as the research programme, significant changes are made upstream in the pathway and a more integrated care approach is likely to be required. It is hoped that the research programme will define a new set of standards for lung cancer screening, which will allow for earlier diagnosis and at a stage where treatment is more likely to be successful. Differentiating between cancers and non-cancers at an earlier stage in the pathway using initial CT scans has the potential to remove the delay and reduce the need for further scans and invasive tests (Digital Health, 2020). Such a programme would create additional pressure in the first instance, due to large volumes of patients being screened and the subsequent interpretation of CT scans required. However, it is hoped that technologies like the Virtual Nodule Clinic will help relieve pressure on the workforce. Indeed, improving the efficiency of the pathway upstream has the potential to reduce pressure on the lung cancer pathway downstream and improve survival rates.

Optellum
Impact on the workforce

Due to the implementation stage of the Virtual Nodule Clinic, results from a real-world evaluation were not available to understand the impact on the workforce. Hence, an analysis of the intended impact is described below instead. It should be noted that the analysis may not be fully representative of the Virtual Nodule Clinic, as other impacts may be uncovered in the current pilots which will ascertain the impact on the workforce.

Impact on the workforce groups

Radiologists and chest physicians have been noted as the workforce groups that are directly impacted by the deployment of Optellum. However, lung cancer screening programmes that use the Virtual Nodule Clinic will also indirectly involve GPs since referrals to the programme come from primary care.

- **Change in workload.** A change in workload for the workforce groups directly impacted by the Virtual Nodule Clinic may be a reduction in daily tasks or reduction in time spent on daily tasks. Stated benefits from papers or articles of Optellum with regards to users are not focused around reducing workload. They are related to enhancing care and decision making. As such, it is not expected the use of the Virtual Nodule Clinic will directly change in workload, pending demonstration and validation of evidence across sites.
- **Change in role.** The Virtual Nodule Clinic provides a nodule-specific Lung Cancer Prediction (LCP) score which indicates the risk of a nodule being cancerous. Radiologists and chest physicians are expected to use the score as support in their decision and care planning with improved accuracy. As such, there is a change in role where healthcare professionals use the technology to validate their decision. The change in role is anticipated to be a significant impact of the Virtual Nodule Clinic, pending demonstration and validation of evidence across sites.
- **Change in team dynamic.** One of the Virtual Nodule Clinic's key functionalities is a dashboard that is automatically populated and updated to present the information gathered from radiology reports and patients with lung nodules flagged in any CT scan. This allows clinicians to track patients and ensure timely action by clinical teams. It is possible that there may be a change in team dynamic, pending demonstration and validation of evidence across sites, however it is not expected to be the most significant impact. Increased integration between care settings may also impact the team dynamic through patient navigators.

Optellum
Impact on the pathway

- **Change in patient outcomes.** Identifying patients with lung cancer at an earlier stage can improve the success rate of treatment and therefore improve survival rates. A key objective of Optellum is to provide a tool to help healthcare professionals identify these patients and ensure a better chance of survival. The 5-year survival rate of stage 1 and stage 4 lung cancer being 57% and 3% respectively (Cancer Research, 2017). An improvement in patient outcomes is likely to be a key impact of implementing the Virtual Nodule Clinic in the lung cancer diagnosis pathway. Current pilots and trials will confirm the real-world impact on patient outcomes.
- **Change in the support, communication or education provided to patients.** By implementing the Virtual Nodule Clinic in the lung cancer diagnosis pathway, patients may be informed of the technology and the functionalities. Due to the implementation stage, findings from real-world settings are not available and thus it is not possible to ascertain the expectation of clinicians in communication about the technology to patients.
- **Change in waiting time, time to diagnosis, referral or discharge.** Improving clinical decisions by accurately identifying malignant nodules will reduce the need for additional CT scans invasive procedures such as biopsies. Consequently, there will likely be a reduction in time to diagnosis for patients with lung cancer and earlier confirmation for patients with benign nodules that no additional tests will need to be conducted. A reduction in waiting time to diagnosis is a key objective for Optellum's Virtual Nodule Clinic and will likely be a significant impact of deployment in the lung cancer diagnosis pathway, pending demonstration and validation of evidence across sites.

Optellum
Impact on the system

- **Change in delivery of integrated care.** The move to lung cancer screening programmes with the use of technologies like the Virtual Nodule Clinic will require a move to an integrated approach to care. This will include primary, secondary and community Care. Referrals to lung cancer screening programmes will likely come from primary and secondary care. Patient data will be shared to secondary and community care where screening programmes will be conducted, which will require an increase in the delivery of integrated care, pending demonstration and validation of evidence across sites.
- **Change in access to care.** A change in access to care means that the use of a solution is either improving the access to care by engaging with hard-to-reach groups or reducing it by making it harder for certain patient populations to access NHS services. It is possible that patients will be identified through their GP to be included in screening programmes, who may not have been flagged by as 'at-risk' patient previously, subsequently increasing access to care. However, this would be an indirect impact of the Virtual Nodule Clinic in the lung cancer screening pathway.
- **Change in the system performance, efficiency, or resilience.** Identifying cancers and non-cancers at an earlier stage will reduce the need for additional tests, and consequently relieves pressure on the system. It will ensure that resources in the lung cancer pathway are dedicated to those with confirmed lung cancer. Additionally, with the Virtual Nodule Clinic supporting radiologists and chest physicians in diagnosing lung cancer earlier, treatment costs are likely to be reduced. It was estimated that diagnosing all cancers as early as the best in England could save the NHS £210m every year (Cancer Research, 2014). It is likely that the Virtual Nodule Clinic will increase system performance, efficiency, or resilience pending demonstration and validation of evidence across sites.

4

Dashboard

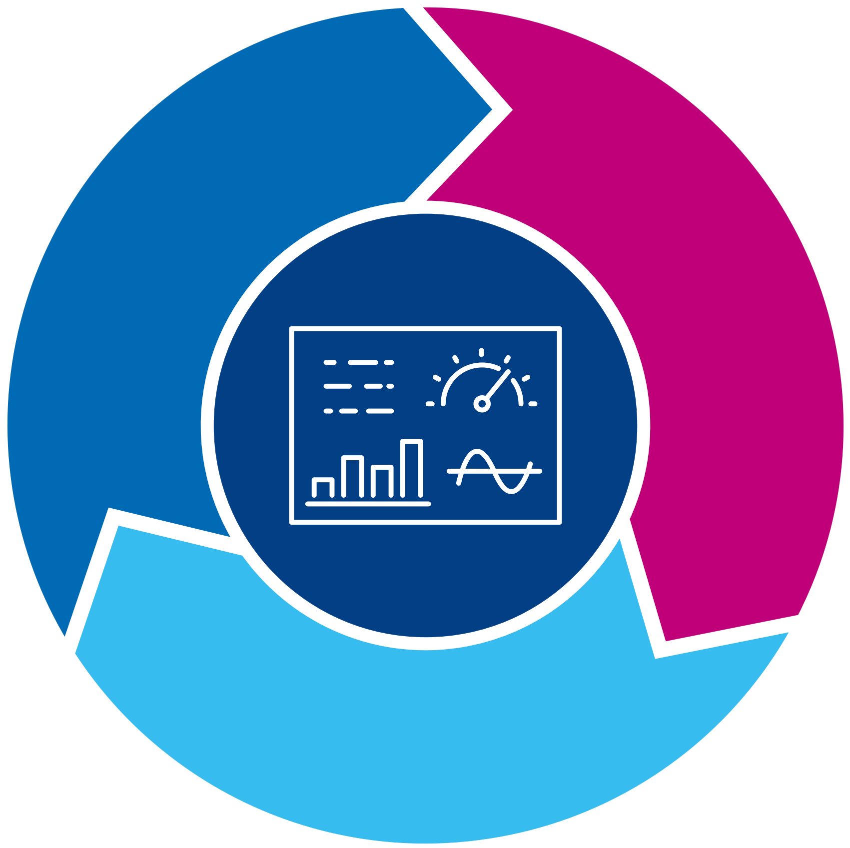
Methodology

Datasets

The database that was generated in the Horizon Scanning task of the 240 technologies was the single data source of the dashboard.

Mock-up

A mock-up of the AI Roadmap dashboard was designed on PowerPoint initially. This was to provide an idea of the visualisations and help to understand the best structure and flow of the information presented. The mock-up was presented to the HEE group for feedback and was approved.



Feedback and iterations

The AI Roadmap dashboard went through an extensive internal and external feedback cycle with three main versions of the dashboard being produced for review.

Version 1

Version 1 of the dashboard was created on Tableau and published to the server for feedback from the HEE team. All feedback was collated and included points around:

- Improving the structure of the dashboard and ordering of information to be more digestible for the user
- Not enough clarity on the limitations of the dashboard and database
- More information and explanation required about what the different charts represent

Version 2

Following the feedback, version 2 of the dashboard was created with the changes made including but not limited to:

- Restructuring of the pages: moving the ‘Use Case Profiles’ and ‘Spread of AI technologies’ to the front of the dashboard and the ‘Overview of AI technologies’ and ‘Impact on the workforce’ to the end
- Clearer caveats put on every page and extra caveats on the Overview of AI technologies and Impact on the workforce pages
- Icons with tooltips with clear explanations of what each chart represents and includes added where necessary

Version 3

Version 3 of the dashboard was created mainly to improve the aesthetics of the dashboard with the changes made including but not limited to:

- Improvement of the homepage with navigation buttons next to the description of each page
- Methodology and ‘How to use’ instructions included in tooltips of icons to make more space for the charts and graphs
- Supplementary resources developed
  - Definitions
  - Time to deployment criteria
  - Impact on the workforce framework
  - General limitations and caveats

Version 3 was presented to HEE and considered the final version of the AI Roadmap Dashboard.

Structure of the dashboard

Home page

The home page of the dashboard is split into two parts. The first part provides context around the project in three headings: Purpose, Methodology, Considerations. The second part of the home page provides a short description of what is included on each section of the dashboard with a navigation button that takes users directly to that page.

Distribution

The Distribution page of the dashboard displays the taxonomy for technology and subtype using the tree approach. A short description of what each technology type and sub-type encompasses is included within the chart for users to view. Additionally, the distribution of the technology type and subtype within the database and the technology type respectively are presented on the chart.

Use Case Profiles

The ‘Use Case Profiles’ page presents six different profiles for each of the six technology types within the taxonomy and aims to display the diversity of the products encompassed in each technology type. It combines multiple measures in the database including:

- The top five workforce groups affected with the percentage of technologies within the database that have identified that workforce group as being impacted.
- The top five clinical areas, with the percentage of technologies within the database that sit within that clinical area, are presented.
- The spread across the different points of care and time to deployment.
- Information relating to the implementation of the technologies within the technology group have been included: the number of sites the technology type has been implemented in and the NHS region they are most commonly implemented in.

Spread of AI technologies

The ‘Spread’ page displays all information collected relating to the implementation of the technologies and includes:

- The geographical spread of the technologies is displayed on a map using information on implementation in known sites; one site being a CCG or Trust (Acute, Mental or Community).
- The number of known sites where technologies are implemented so users can explore how many technologies are deployed in each number or sites up to five.
- The workforce groups that have been identified as being impacted by the technologies.

Overview of AI technologies

The ‘Overview of AI technologies’ page allows users to explore four key areas of the Horizon Scanning exercise:

- The clinical area that the technologies lie within, ordered from most to least represented.
- The workforce groups that have been identified as being directly impact by the technology.
- The spread of points of care that the technology has been deployed in or is expected to be deployed in, sized by the percentage of technologies within each point of care.
- The spread of technologies across the different time horizons in the estimated time until the technology is deployable at scale.

Supplementary resources

A supplementary resources section has been included for users and contains four parts:

- Definitions by technology type and subtype
- Time to deployment criteria
- Impact on the workforce framework
- Limitations and caveats (including recommendations for innovators to present information regarding their technologies)

5

Key findings

The Horizon Scanning exercise brought insights to the landscape of AI technologies that are nearing and ready for market across a number of defined characteristics. The database that was generated through the exercise is based on factual publicly available data but also value proposition claims made by the technology which were not validated as part of the exercise. In this section, the key findings of this work were compared to those presented in a few reference documents such as the State of the Nation AI Survey 2021, The Topol Review and the state of AI based FDA approved medical devices and algorithms (Benjamens, Dhunnoo, & Meskó, 2020). They are presented below in three sections:

- **Expected findings:** confirming what was known already and helping to understand the scale and distribution of technologies for these known characteristics
- **Unexpected findings:** new insights or emerging areas that have not been at the forefront of previous work
- **Requires further evaluation:** areas or impacts that may be underrepresented within the database and would require more exploration

Expected findings

Percentage of Diagnostic technologies

Out of the six technology types used in the database, the most represented technology type was 'Diagnostic', with 34% of the technologies. This result also coincides with the findings of the State of the Nation AI Survey where diagnostic was also the most frequently selected technology type. Within the 'Diagnostic' category in the database, the sub-type 'Imaging' represented 49% of the technologies.

Percentage of technologies within 'Clinical Radiology'

A list of 66 clinical areas was provided by HEE to be used for the Horizon Scanning exercise, of which 30 were used when populating the database. 'Clinical Radiology' was the second most selected clinical area, after 'Multiple clinical areas'. Technologies within 'Clinical Radiology' in the database are primarily diagnostic technologies, indeed, 31% of technologies in this technology type have the clinical area 'Clinical Radiology'.

Most frequently identified workforce groups

The list of workforce groups provided by HEE contains 155 workforce groups, of which 48 were used when populating the database. Echoing the previous remark, medical roles in 'Clinical Radiology' were the most frequently identified in the database as being directly impacted by the technologies. 15% of technologies within the database and 37% of technologies in the 'Diagnostic' technology type impacted directly medical roles in 'Clinical Radiology'. Non-clinical admin roles were the third most frequently impacted workforce group and were identified by 10% of technologies. 37% of technologies within the 'Automation/Service efficiency' technology type identified these roles as being directly impacted by their deployment. The top 10 workforce groups that were identified as being impacted by technologies in the database are presented in Figure 7.

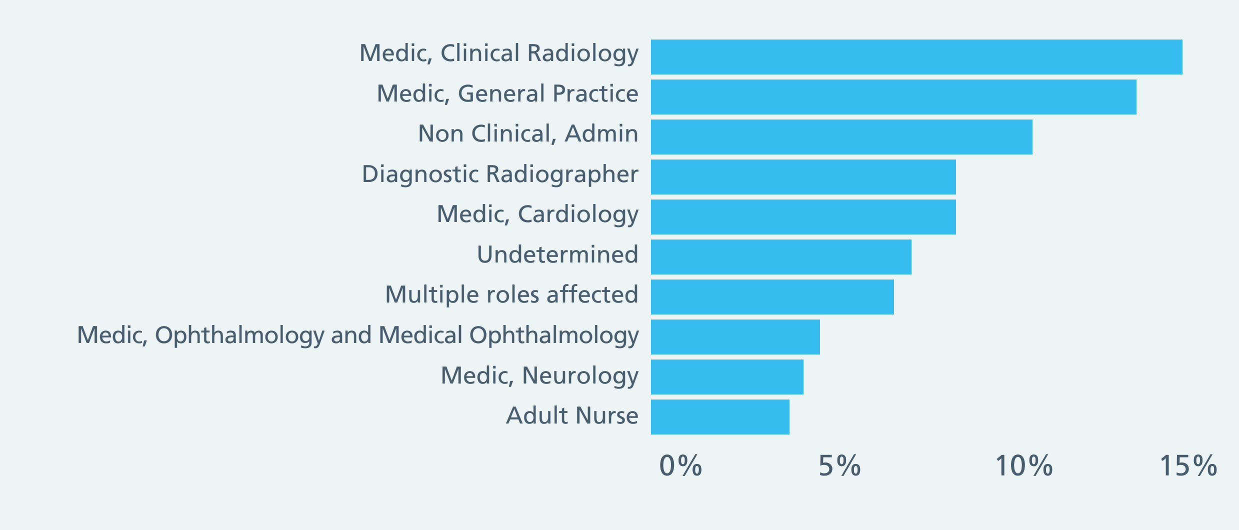


Figure 7: *Bar graph showing the workforce groups most affected by the use of AI technologies*

Percentage of technologies with a reduction in workload

The impact on the workforce framework has been divided into three dimensions: The impact on the workforce groups, the impact on the pathway and the impact on the system. Within each dimension there are three impact groups that are measured on different scales. The most identified impact in the 'Impact on the workforce groups' dimension of the framework was a reduction in workload. Using technology claims and assessing against the framework, 53% of technologies led to a reduction in workload for workforce groups directly impacted by the technologies, 17% no change and 1% increase. The proportion with a reduction in workload increased for technologies identifying medical roles in Clinical Radiology as being impacted, with 69% reducing the workload, and the only other impact recorded being 'Undetermined'. This finding was coherent with the Topol Review where it was estimated that eliminating the need for a second radiologist for mammography scan reading, by using AI, would reduce time spent reviewing mammograms by 30% (HEE, 2019).

Percentage of technologies with a change in role

Furthermore, results of the Horizon Scanning exercise showed that 49% of technologies result in a change in role for workforce groups directly impacted and 23% with no change in role. The necessity of a change in role for the workforce was highlighted in the Topol Review as one of the four conditions required for clinicians to fully benefit from AI (HEE, 2019). Technologies within the 'Remote Monitoring' technology type contained the highest proportion of technologies resulting in a change in role, out of the six technology types, with 67% of technologies resulting in this impact, and 12% with no change in role.

Percentage of technologies with a reduction in waiting time, time to diagnosis, treatment, referral or discharge

There are three dimensions in the impact on the pathway category of the framework, namely:

- Change in patient outcomes
- Change in support, education or communication provided to patients
- Change in waiting time, time to diagnosis, treatment, referral or discharge.

50% of the technologies within the database led to a reduction in the waiting time, time to diagnosis, treatment, referrals or discharge, 20% with no change and 2% with an increase. The majority of technologies operating in the clinical area 'Clinical Radiology' indicated a reduction in waiting time, time to diagnosis, treatment, referrals or discharge (69%). This echoes the findings of the Topol review as it was estimated that annually, the potential impact of AI technologies on diagnostic radiology equates to the equivalent of approximately 890,000 hours of radiologist time which allows for pathways to be sped up and reduces the time to diagnosis for cancer, for example, following medical scans (HEE, 2019).

Unexpected findings

The percentage of P4 technologies compared to Remote Monitoring technologies

The distribution of the technology types is displayed in Figure 8, with 'Diagnostic' and 'Automation/Service efficiency' the first and second most represented respectively. An unexpected finding of the Horizon Scanning exercise is the percentage of 'P4 Medicine' technologies, which is the third most represented technology type in the database, before 'Remote monitoring'. This technology type accounts for 17% of technologies, with Remote monitoring encompassing 14% of technologies within the database. The State of the Nation AI Survey results reported that 34% of technologies identified as being Remote monitoring, or having remote monitoring capabilities, which was the second most represented technology type selected in the survey (NHSX, 2021).

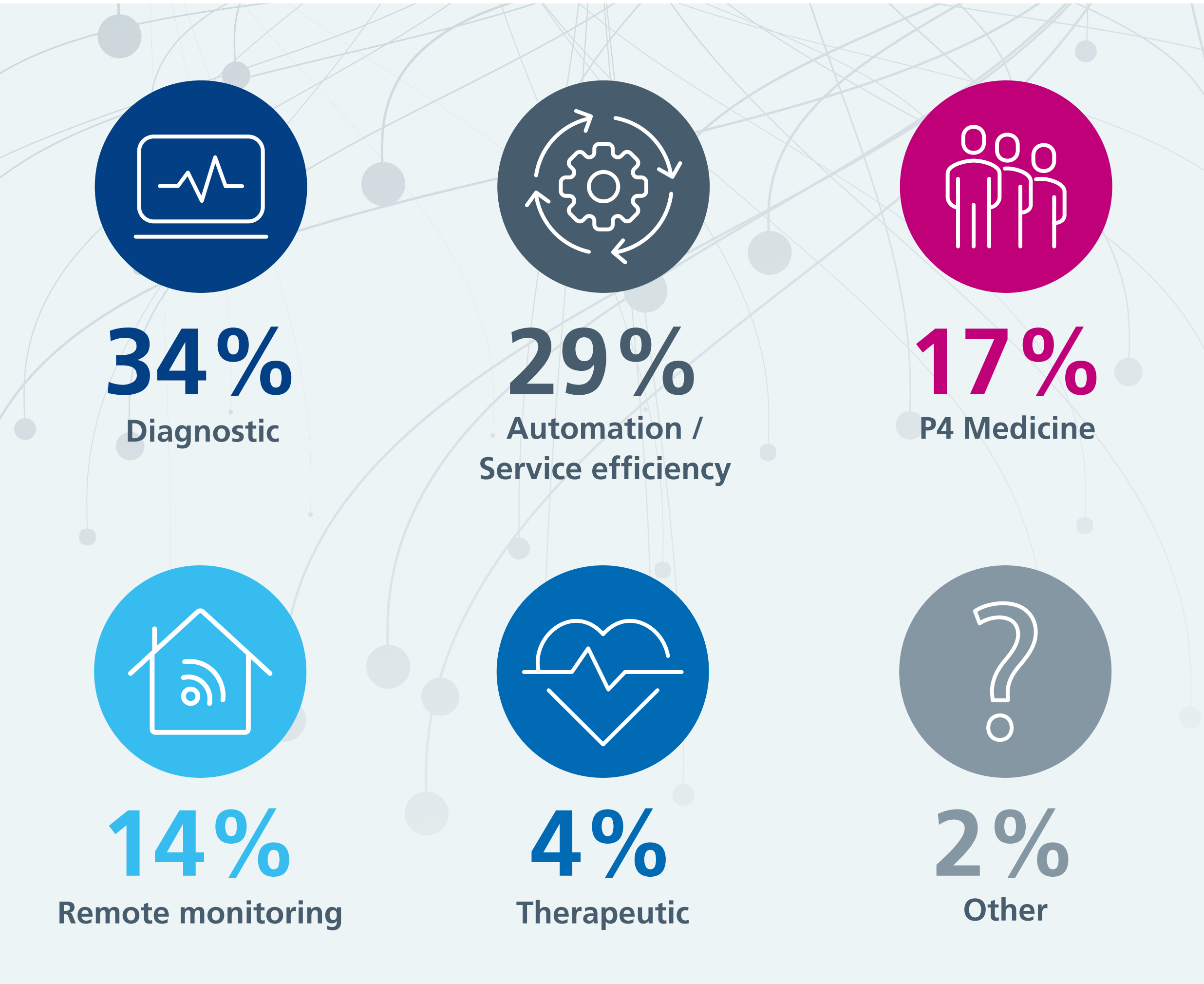


Figure 8: *Distribution of AI technologies*

Distribution of technologies with 'Multiple clinical areas'

An additional unexpected finding is the proportion of the technologies within the database that were unable to fit into a specific clinical area, and could be implemented in multiple settings. 23% of technologies have been categorised as 'Multiple clinical areas' which is the most selected within the 'Clinical Area' characteristic of the database. This was likely due to 74% of technologies in Automation/Service efficiency being categorised as having 'Multiple clinical areas'.

Estimated time until deployable at scale

The estimated time until the technologies are deployable at scale was calculated using 13 criteria displayed in Appendix F and includes four time to deployment scales: within 1 year, within 3 years, within 5 years, more than 5 years. The State of the Nation AI Survey reported that, for the same question, 54% of respondents believe their product will be ready for deployment at scale in one year (NHSX, 2021). The most common time to deployment in the database was within three years (39% of technologies), followed by within 1 year (25% of technologies). The technologies within the 'Automation/Service efficiency' category presented a shorter timescale for deployment than the overall database with 35% of technology ready within 1 year and 34% within 3 years as presented below in Figure 9.

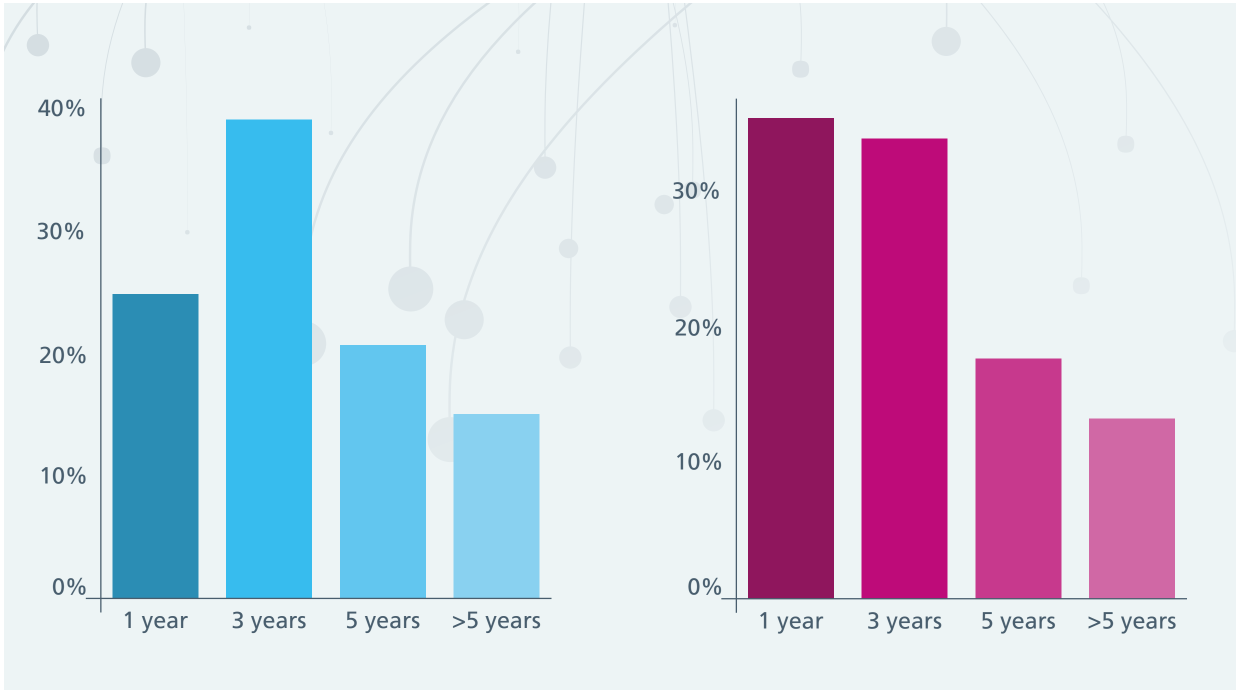


Figure 9: *Bar chart showing the readiness for large scale deployment in the AI technologies represented in the database (left) and only 'Automation/Service efficiency' technologies (right)*

Low proportion of technologies with an improvement in patient outcomes

One of the impact groups that technologies were assessed against within the impact on the pathway dimension of the impact on the workforce framework was the impact on patient outcomes. Using technology claims and assessing against the framework, 14% of technologies were found to improve patient outcomes and 58% with no change in patient outcomes. Only direct outcomes of the technologies were captured in the Horizon Scanning exercise and, in many cases, an improvement in patient outcomes presented as a secondary impact and as a result of direct improvements such as a reduction in time to diagnosis.

Geographical spread of technology implementation

Information relating to the implementation of technologies was collected during the Horizon Scanning task. 68% of technologies reported that they had been implemented in an NHS site and 54% of technologies reported the NHS sites that they had been implemented in which are displayed on the map in Figure 10. London had the highest number of technologies implemented out of the seven NHS Regions with 66 technologies implemented, equating to 40% of technologies with known implementation sites. The South West NHS Region presents the least implementation with only 14 technologies implemented in the region. In addition to the regional variation of technology implementation, the number of known sites the technologies have been implemented differs across the database; 29% of technologies are reported to have been implemented in one known site, and 5% of technologies are reported to have been implemented in five or more sites.

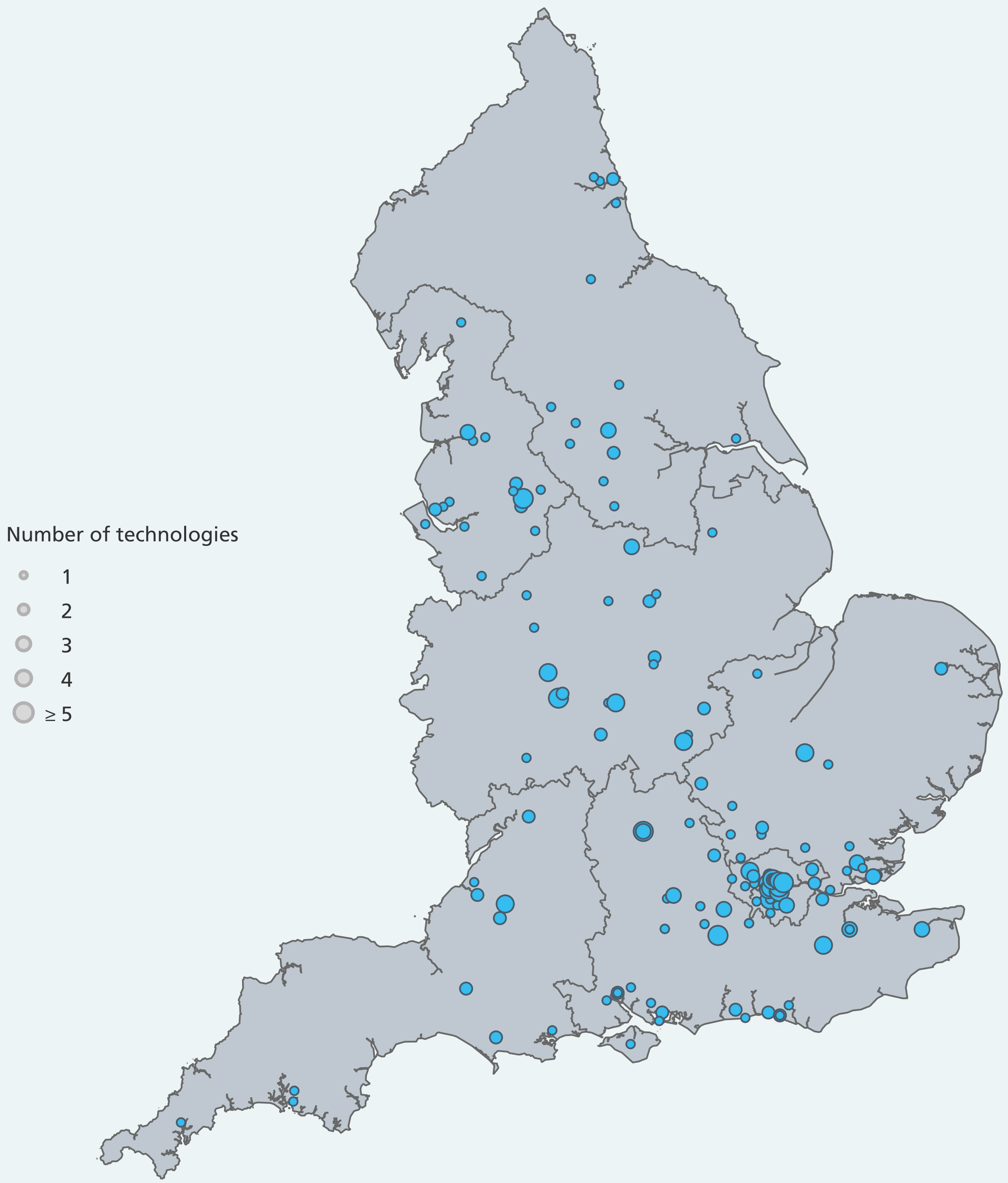


Figure 10: *The number of technologies deployed in known NHS sites*

Requires further evaluation

Due to the nature of the Horizon Scanning exercise and the use of publicly available data to populate the database, there are a number of areas and findings which may require further evaluation to be confirmed or inferred.

Lack of or unclear information regarding the impact on the workforce for P4 Medicine technologies

'P4 Medicine' was found to be the third most represented technology type out of the six types within the database, with 17% of technologies. The expected time to deployment for technologies in P4 medicines was not estimated to be longer than other technology types, with only 5% of technologies to be deployed within 1 year and 54% within 3 years. However, value proposition and impact on the workforce was not clearly stated for the majority of P4 Medicine technologies, the most frequently selected option for the workforce group was 'Undetermined' which accounted for 27% of the solutions.

For technologies where they were early stage with no clear value proposition, 'Early-stage solution – undetermined' was selected for every impact group in every dimension in the impact on the workforce framework. Out of the 240 technologies in the database, 25% fell into this impact. For the 41 technologies in the P4 medicines, 56% were labelled as 'Early-stage solution – undetermined' due to the lack of clear value proposition and defined benefits by the innovators.

Underrepresented workforce groups

An additional key area needing further evaluation the workforce groups which are underrepresented in the database. Nurses and midwives currently represent 29% of all FTE staff in the NHS (Nuffield Trust, 2021), but they accounted for only 3% of the workforce groups impacted by the technologies listed in the database. Similarly, Healthcare Scientists appear underrepresented in the database, with Healthcare Scientists in Genetics and other Healthcare Science roles accounting for 1% and 0.3% of the workforce groups impacted by AI solutions.

Low percentage of technologies with a change in team dynamic

As presented earlier in the report, 49% of technologies reported to trigger a change in role and 23% suggested there was no change in role. One could think that a change in role would have implications on team dynamic. According to the Topol Review, workplace support is one of the four conditions to maximise the potential of the technology (HEE, 2019). Nonetheless, results from the Horizon Scanning exercise showed that only 24% of technologies within the database would lead to a change in team dynamic and 48% would not lead to a change in team dynamic. Since the Horizon Scan relies on the evidence collected to date by innovators, many technologies may lack the perspective needed to document changes in team dynamic, as this requires for the solution to be embedded into practices for a long period of time.

Low percentage of technologies with a change the support, education or communication provided to patients

An additional impact that may require further evaluation is the impact that deployment of AI will have on communication between the patients and the workforce. The importance of citizen involvement in the design and implementation of AI technologies is emphasised in the Topol Review (HEE, 2019). Findings from the Horizon Scan showed that 42% of technologies did not report changing the support, education or communication provided to patients and 29% showed an increase of communication needed. Understanding how communication will change and how it needs to change in order to facilitate the uptake of these technologies and best benefit the patients and the workforce should be explored further.

6

Limitations and recommendations

Throughout the report, the authors have described the methodology, the assumptions chosen and their rationale. The following section summarises the limitations of the work and proposes avenues to address them in future iterations of the roadmap.

Principles and methodology

Publicly available information was used to populate the database. It was not within scope for the researchers to engage with the companies listed in the database to obtain supplementary information. Therefore, they were not able to add any relevant commercially sensitive information. For instance, innovators were often vague when disclosing the names of the NHS sites their solution was used at. If they were not specified by the company, the researchers had to use the "Unspecified NHS site(s)" option.

> ***Recommendation:*** *More communication and engagement with innovators is recommended in the next phases of the roadmap. This could include sharing some of the outputs of the phase 1 and presenting the benefits for innovators to engage with the team in the future. The innovators could then contribute to the update on the database to provide complementary information, to check its accuracy but also to express their opinion on how the database design could be more inclusive and impactful. Similarly, signposting the dashboard's users to a short survey could be an effective way to collect their views on the design, how it can be improved and give innovators a mechanism to report any inaccuracy or questions they may have.*

Some available information may have been missed by the researchers if it was not displayed on the technology's website nor referenced in the first two pages of the search engine. Indeed, due to the number of technologies included in the Horizon Scan exercise, the researchers had to cap the amount of time spent on each technology.

> ***Recommendation:*** *Engagement with innovators could also help to address this limitation. A favourable outcome of the communication strategy would be for AI innovators to display more clearly a number of information which were hard to get in this phase one (list of healthcare workforce group affected, number of sites using the solution, evidence to date, etc.). It would be ideal to have a single platform, co-owned and managed by NHS and ALB organisations, collecting information from innovators and validating the claims and evidence submitted. This will be a powerful tool to better understand the pipeline and requirements from innovators.*

No validation of the value proposition claims of the AI solutions was done as part of the exercise. The purpose of the roadmap was not to evaluate AI technologies but to map the AI technologies currently on the market in the England, to understand the distribution between the type of technologies, the pathway and the workforce impacted. Therefore, the assessment of the impact on the workforce was based on the impact claimed by the company. This should be noted by the user exploring the dashboard, indeed because of the innovator's bias, the impacts presented may be more numerous and far-reaching than what an independent evaluator would have presented.

> ***Recommendations:*** *A mitigation for the user exploring the current dashboard is to consider the impact on the workforce in conjunction with the time to deployment of the solutions. If the latter is greater than 3 years, the company would not have gathered many of the following: regulatory approval, proof of efficacy, usage in an NHS site, listed on a procurement framework. As a consequence, the value proposition claims are likely to be estimated or based on testing in a controlled environment, rather than demonstrated in a real-world setting and the impact on the workforce should be viewed with a critical eye.*

Discussions are currently ongoing around a proposed validation of the claims of the companies as part of the future iterations of the dashboard. One should note that an independent body, such as NICE, would be required to validate evidence claims of the AI technologies. Some of the parameters already collected such as presence of published results of a clinical trial, CE/UKCA marking, presence of an economic evaluation could be a good starting point for this.

Further analysis is needed to keep the taxonomy up-to-date and relevant. The need for further analysis on the technology types and subtypes in the taxonomy became apparent when the researchers were populating the database. Due to the complexity of some of the AI technologies and given there was no engagement with the technologies to clarify the features, it proved difficult at times to categorise all the technologies into six technology types. For example, some solutions allowing to monitor patients in an acute setting could be classified as "Remote Monitoring" although they do not operate in a residential/community setting (which is often as a characteristic of remote monitoring solutions), or they could be classified as "Therapeutic" as they provide crucial information for patient management. Additionally, the technology types "P4 medicine", "Therapeutic" and "Other" do not have subtypes.

> ***Recommendation:*** *A periodic review of the taxonomy to assess its relevance and how it aligns with academic publications as well as common usage is recommended. Determining the place of digital therapeutics (DTx) in the taxonomy as well as revisiting the need to create subtypes for "P4 medicine" and "Therapeutic" are also recommended.*

Close collaboration with the NHS AI Lab, NICE and the AAC teams to keep the language consistent would also be beneficial. In particular, it will be important to incorporate the findings of the Academic Collaboration on Artificial Intelligence, jointly commissioned by NICE and the NHSX AI Lab. The aim of the project is to update the Evidence Standards Framework for Digital Health Technologies, so it fully incorporates data-driven artificial intelligence technologies, including those that use adaptive algorithms (NICE, 2021).

> *"The AI Roadmap is a valuable resource to help understand and navigate the AI technology landscape. The database and dashboard are practical tools we can use when planning resources for NICE guidance production and will provide a source of information for the multi-agency advisory service to monitor new digital health technologies. We see this work as an example of how collaborating with strategic partners can support the regulation and evaluation of AI technologies."*
>
> **Jeanette Kusel,** Director – Scientific Advice, NICE

Database and dashboard

The dashboard does not contain workforce data. Including workforce statistics was considered during the design phase. For example, it would provide the ability to show the headcount or FTE of workforce groups who are currently impacted by AI technologies or likely to be impact in the near future across England at a site level, thus giving a different measure of the impact of AI technologies on the healthcare system. Moreover, the inclusion of workforce data on staff shortages would highlight the opportunity to alleviate workforce pressure by supporting the AI technologies which impact these workforce groups.

> ***Recommendation:*** *These views were shared by the HEE team and whilst the data could not be made available for the phase 1, these options could be revisited for future iterations. In complement to visualising workforce in the dashboard, the opportunity to use system dynamic modelling to understand resource allocation depending on AI availability across sites could be explored.*

The impact on the workforce framework only captures the direct users and the documented impacts. To avoid formulating too many assumptions, the researchers reported the impact on the direct users as presented by the AI technologies but did not try to infer what these impacts would mean for the indirect users or for the workforce groups working with the direct users. Consequently, non-clinical staff, and in particular healthcare scientists, are likely to be underrepresented in the database, as they are rarely presented as the direct users of AI solutions, despite the impact that AI solutions have on their role. Similarly, the impacts of AI technologies presented in the dashboard are likely to be more centred around clinical staff and medic groups.

> ***Recommendation:*** *We recommend that further work is undertaken to strengthen the impact on the workforce framework and to explore further into what the different impacts mean for different workforce groups. The framework as it currently stands is effective for evaluating the potential impact, or intended impact, of the technologies at a high level. However, it should not be assumed that a change in role, for example, will have similar meaning across the breadth of workforce groups that are expected to be impacted by the deployment of AI technologies. The improvement of the framework should be driven by feedback from a representative sample of workforce groups, engagement with sites implementing AI technology as well as discussions with innovators about how and why clinical professions should not be the sole focus of studies of AI implementations. Some current initiatives to rethink the use of data in healthcare and the professionalisation of healthcare scientists and analysts such as the Goldacre Review should be reviewed once they publish their findings, as some of the learnings may be useful to help shape how to report on these workforce groups (Department of Health and Social Care, 2021).*

The lists of sites for the AI technologies does not differentiate between pilot sites and sites of adoption. In the current database, pilot sites and sites where the solution is adopted and part of business as usual. This difference has implications on the maturity of the AI solution but also it potentially means that the database lists sites which are no longer using the solution.

> ***Recommendation:*** *Listing separately pilot sites and sites of adoption is recommended to bring more granularity to the 'usage in the NHS' dimension.*

The database does not capture the IT infrastructure needed to implement the solutions. As implementation considerations are not the main focus of this work, the database does not document the IT requirements for the use of the solution. However, changes in IT and hardware systems have long been presented as a recurring challenge when implementing digital or AI technologies (Singh et al., 2020).

> ***Recommendation:*** *The IT infrastructure requirements could either be explored through the case studies by adding a question specifically to understand the changes in infrastructure and IT systems needed to implement the solution. It could also potentially be captured in the database through the use of the NASSS framework: Non-adoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability (Abimbola et al., 2019). AI technologies could be categorised between off-the-shelf, integrated or complicated solutions. Should these requirements be documented in the database, one must remember that this is likely to be an at-site-*

level consideration, as what could be very easy to implement in one site could prove very difficult in another site. Indeed, it is dependent on the site starting point with regards to connectivity, Wi-Fi coverage, state of the IT systems, etc. Therefore, it could be more appropriate to capture this through the case studies.

The visualisation of the "Use Case Profiles" may be overwhelming for the reader. The aim of the 'Use Case Profiles' page of the dashboard is to provide an illustrative profile of the six different technology types included in the taxonomy. Data and insights from all different areas of the Horizon Scanning exercise are presented. Given the volume of characteristics and insights included on the page, the flow and appearance are not as clear and structured as the rest of the dashboard.

Recommendation: *Feedback may be provided through the survey link at the end of the dashboard to understand what users find the most useful and drive improvement to the page in future iterations to ensure the illustrative aspect with the combination of multiple characteristics remains but with improved visualisations and flow.*

The dashboard only presents a snapshot of the AI landscape at a given time. The intelligence collected in the database and illustrated through the dashboard yields most value at the time of the dashboard release, should the dashboard be consulted in a year's time without being updated, there is a risk for the data to become obsolete and inaccurate.

Recommendation: *Updating the database and corresponding dashboard is recommended to conserve the value of the roadmap developed. Indeed, the power of this tool will really be around the continuous monitoring of the landscape's trends to understand growth and changes occurring over a longer period of time. It is about providing the healthcare system with a forecasting tool that helps with decision-making by setting priorities and understanding their implication at a place-based and workforce level.*

Next steps

This section focuses on recommendations for the future dissemination of the dashboard and overall communication around the work:

- Formalising a process to engage with the different audiences of the roadmap. This includes innovators to ensure accuracy of the information collected and to create an incentive for better transparency around evidence. Royal Colleges, commissioners and providers should also be considered to share the intelligence around the pipeline and spread of AI technologies and drive local and national conversations about how to support the workforce to implement AI in the NHS. Although it not a direct audience for the AI Roadmap, engagement with workforce groups and patients will be crucial to understand potential gaps, bias and take notes of any impact that is not currently captured.
- Liaising with the AHSN Network to align, where possible, the roadmap with the Innovation Pipeline. Should the AI Roadmap be used in an operational capacity not only to understand the landscape but also to identify promising solutions, it would be recommended to work with the Health Innovation Manchester (HInM) team, who have developed and maintain the Innovation Pipeline, to learn from their work, to avoid duplication and utilised synergies. At the start of this commission, Unity Insights engaged with HInM to make them aware of the work and they graciously accepted to share the database for the Innovation Pipeline.
- Utilising the outputs from the AAC AI Award evaluations. As the rounds 1 and 2 solutions are implemented across the country, special attention should be given to the interim reports produced by the Technology Specific Evaluation Teams (TSETs). Indeed, the evaluation requirements include assessing the implementation fit, understand the workforce perspective and the effectiveness implications of the use of the solutions. Therefore, outputs should be reviewed on a regular basis to understand how the elements captured in the database can complement, strengthen or explore further some of the TSETs' findings.

7

Conclusion

Thanks to the breadth of AI technologies captured in the database and the extensive desktop research conducted, the AI Roadmap presents a comprehensive overview of the AI landscape in England, the distribution of the different types of technologies, their spread in NHS sites across the country and the main workforce groups they impact. Building on previous studies such as the NHSX State of the Nation AI survey and the NIHR AI Horizon Scan, and in alignment with the Innovation Pipeline and AAC AI award, the roadmap brings a new dimension to the Horizon Scan exercise by exploring which workforce groups are impacted by the use of AI and by classifying this impact using a bespoke framework.

As stated earlier in the report one should be aware of the limitations of this exercise. The data collection relied on publicly available data and did not seek to validate the claims and evidence publicised by the companies; therefore some benefits may be overstated, especially for technologies at an early stage of development. Furthermore, the database is not free of bias commonly found when it comes to reporting on the workforce, e.g. a perspective centred around clinical groups, less representation of multi-disciplinary teams, allied health professionals and administrative staff.

As maintaining the roadmap up to date is critical to keeping the dashboard relevant and insightful, there are many avenues to address the limitations listed, to hone on the positive elements and to action some of the recommendations suggested by the authors. Keeping the collaborative spirit observed during this commission, with regular engagement with important players in the AI space such as the NHS AI Lab, NICE, AAC and the AHSN Network, will be key to ensure the roadmap is in alignment with, and adds value, to the healthcare system.

> *"The AI Roadmap is an invaluable asset in helping to understand the AI and data driven landscape in healthcare, and the implications this will have on our staff and learners. It is important we achieve transformation through emerging technology, helping scalability to improve patient care throughout the country, and can understand impact on the system, pathways, and users. We need to ensure the workforce is ready to support this aim and the insights from this roadmap will focus our efforts on education and training to achieve this."*
>
> **Hatim Abdulhussein,** Clinical Lead – DART-Ed, HEE

8

References

- Abimbola, S., Patel, B., Peiris, D., Patel, A., Harris, M., Usherwood, T., & Greenhalgh, T. (2019). The NASSS framework for ex post theorisation of technology-supported change in healthcare: worked example of the TORPEDO programme. BMC Medicine, 17(233).
- Aidence. (2021). NHS England lung cancer screening programme: a leading implementer of AI. Retrieved from PR Fire: https://www.prfire.co.uk/press_releases/nhs-england-lung-cancer-screening-programme-a-leading-implementer-of-ai
- Baldwin et al. (2020). External validation of a convolutional neural network artificial intelligence tool to predict malignancy in pulmonary nodules. Retrieved from Thorax: https://thorax.bmj.com/content/thoraxjnl/75/4/306.full.pdf
- Barrera, A., Gee, C., Wood, A., Gibson, O., Bayley, D., & Geddes, J. (2020). Introducing artificial intelligence in acute psychiatric inpatient care: qualitative study of its use to conduct nursing observations. BMJ, p34.
- Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. NPJ Digital Medicine, 3(118).
- British Lung Foundation. (2012). Lung cancer statistics. Retrieved from British Lung Foundation: https://statistics.blf.org.uk/lung-cancer
- Building Better Healthcare. (2019, March 27). Mental health trust cuts falls and saves time with patient monitoring technology. Retrieved from BBH Website: https://www.buildingbetterhealthcare.com/news/article_page/Mental_health_trust_cuts_falls_and_saves_time_with_patient_monitoring_technology/153166
- Cancer Research. (2014). Half of cancers diagnosed at late stage as report shows early diagnosis saves lives and could save the NHS money. Retrieved from Cancer Research: https://news.cancerresearchuk.org/2014/09/22/half-of-cancers-diagnosed-at-late-stage-as-report-shows-early-diagnosis-saves-lives-and-could-save/
- Cancer Research. (2017). Lung cancer survival statistics. Retrieved from Cancer Research: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/survival#heading-Zero
- Care Quality Commission. (2018). More than 200 suicides recorded at mental health units over seven years. Retrieved from The Guardian: https://www.theguardian.com/society/2018/aug/14/more-than-200-suicides-recorded-at-mental-health-units-over-seven-years
- cf.Wiliams. (1985).
- Clark et al. (2021). Non-contact physical health monitoring in mental health seclusion. Retrieved from Journal of Psychiatric Intensive Care: https://www.ingentaconnect.com/content/napicu/jpic/pre-prints/content-napicu_jpi_21009
- DART-Ed. (2021, October 29). Digital, Artificial Intelligence and Robotics Technologies in Education (DART-Ed). Retrieved from NHS HEE website: https://www.hee.nhs.uk/our-work/dart-ed

- Department of Health and Social Care. (2021, February 9). New review into use of health data for research and analysis. Retrieved from GOV.UK: https://www.gov.uk/government/news/new-review-into-use-of-health-data-for-research-and-analysis
- Digital Health. (2020). New £11m programme to explore AI in lung cancer diagnosis. Retrieved from Digital Health: https://www.digitalhealth.net/2020/07/new-11m-programme-to-explore-ai-in-lung-cancer-diagnosis/
- Essex Partnership University NHS Foundation Trust. (2021, October 29). Essex Partnership University NHS Foundation Trust Insights. Retrieved from Oxehealth: https://www.oxehealth.com/resources/essex-partnership-university-nhs-foundation-trust-insights
- Grogan et al. (2011). Thoracic operations for pulmonary nodules are frequently not futile in patients with benign disease. Retrieved from Pubmed: https://pubmed.ncbi.nlm.nih.gov/21760552/
- HEE. (2019). The Topol Review: Preparing the healthcare workforce to deliver the digital future.
- HEE. (2020, December 17). eCollection. Retrieved from HEE: https://www.hee.nhs.uk/our-work/hee-national-data-function/ecollection
- Hinton, G. (2018). Deep learning-a technology with the potential to transform health care. J. Am. Med. Assoc.(320), 1101-1102.
- Malik, S., Kanwar, A., Sim, L., Prokop, L., Wang, Z., Benkhadra, K., & Murad, M. (2014). he association between sleep disturbances and suicidal behaviors in patients with psychiatric diagnoses: a systematic review and meta-analysis. PMID, 3-18.
- Massion et al. (2020). Assessing the Accuracy of a Deep Learning Method to Risk Stratify Indeterminate Pulmonary Nodules. Retrieved from PubMed: https://pubmed.ncbi.nlm.nih.gov/32326730/
- NHS England. (2018). Implementing a timed lung cancer diagnostic pathway. Retrieved from NHS England: https://www.england.nhs.uk/wp-content/uploads/2018/04/implementing-timed-lung-cancer-diagnostic-pathway.pdf
- NHS England. (2019). Targeted Screening for Lung Cancer with Low Radiation Dose Computed Tomography. Retrieved from NHS England: https://www.england.nhs.uk/wp-content/uploads/2019/02/targeted-lung-health-checks-standard-protocol-v1.pdf
- NHSX. (2021). 2020-21: A year in the life of the NHS AI Lab. Retrieved from https://www.nhsx.nhs.uk/ai-lab/about-the-nhs-ai-lab/2020-21-a-year-in-the-life-of-the-nhs-ai-lab/
- NICE. (2021, May 5). Academic Collaboration on Artificial Intelligence. Retrieved from GOV.UK: https://www.contractsfinder.service.gov.uk/Notice/77022e0a-c949-4b25-92f8-7770301b49df
- Nuffield Trust. (2021, September 7). The NHS workforce in numbers. Retrieved from Nuffield Trust: https://www.nuffieldtrust.org.uk/resource/the-nhs-workforce-in-numbers
- Optellum. (2021). LUNG CANCER PREDICTION AI. Retrieved from Optellum: https://optellum.com/lung-cancer-prediction-ai/
- Oxehealth. (2021, 10 29). Benefits of Oxevision. Retrieved from Oxehealth website: https://www.oxehealth.com/benefits/patient-safety#benefits-tabs
- Oxehealth. (2021, October 29). Central and North West London Early Insights Report. Retrieved from Oxehealth website: https://www.oxehealth.com/resources/cnwl-early-insights-report
- Oxehealth. (2021, October 29). Essex Partnership University NHS Foundation Trust Insights. Retrieved from Oxehealth website: https://www.oxehealth.com/resources/essex-partnership-university-nhs-foundation-trust-insights
- Oxehealth. (2021). Improving Inpatient Care. Unpublished, 1.
- Oxehealth. (2021). Patient Experience with Oxevision. Retrieved from Oxehealth: https://www.oxehealth.com/resources/patient-experience-with-oxevision
- Oxehealth. (2021, 10 29). Staff Experience with Oxevision. Retrieved from Oxehealth website: https://www.oxehealth.com/resursers/staff-experience-with-oxevision
- Royal College of Psychiatrists. (2020). QNPICU and NAPICU webinars (COVID-19). Retrieved from Royal College of Psychiatrists: https://www.rcpsych.ac.uk/improving-care/ccqi/quality-networks-accreditation/psychiatric-intensive-care-units-picu/qnpicu-webinars
- Singh, R. P., Hom, G. L., Abramoff, D., M., Campbell, J. P., & Chiang, M. F. (2020). Current Challenges and Barriers to Real-World Artificial Intelligence Adoption for the Healthcare System, Provider, and the Patient. Transl Vis Sci Technol., 9(2)(45).
- The Nuffield Trust. (2021, September 7). The NHS workforce in numbers. Retrieved from The Nuffield Trust: https://www.nuffieldtrust.org.uk/resource/the-nhs-workforce-in-numbers
- UCLPartners. (2021). Enhancing patient safety with innovative technology. Retrieved from ULCPartners: https://uclpartners.com/impact-story/enabling-safety-and-trust-on-mental-health-wards-through-innovation/
- UK Lung Cancer Coalition. (2020). Late diagnosis of lung cancer hitting survival rate, study says. Retrieved from The Guardian: https://www.theguardian.com/society/2020/jan/31/late-diagnosis-of-lung-cancer-hitting-survival-rate-study-says
- Wright et al. (2021). Reducing Falls in Dementia Inpatients Using Vision-Based Technology. Retrieved from Journal of Patient Safety: https://journals.lww.com/journalpatientsafety/Abstract/9000/Reducing_Falls_in_Dementia_Inpatients_Using.98990.aspx
- York Health Economics Consortium. (2021). York Health Economics Consortium (YHEC) Health Economic Evaluation of Oxevision. Retrieved from Oxehealth: https://www.oxehealth.com/yhec-2-pager-oxehealth-download

9

Appendices

Appendix A – Questions of the NHSX AI survey

Survey question	Multiple choice options
If you would like your company to be considered for inclusion in case studies, please enter contact details. • **Name of organisation** • **Name of solution** • **Email address**	
How do you classify your AI-driven technology? Select as many as applicable.	• Diagnostic • Therapeutic • Care-based • Self-care • Population health • Health promotion • Remote monitoring • Remote consultation • Triage • Social Care • Other
Which group of health and care system users is your AI-driven technology for? Select as many as applicable.	• Person with long-term condition • Parent/Carer • Person with a physical disability • Person with a cognitive or learning impairment • Person with broad care needs • Person interested in monitoring their health (e.g., Fitbit) • Person wishing to access ad-hoc services (e.g., video consultation) • Person seeking mental health sup-port • Clinician (e.g., decision-support) • Commissioner/System Manager (e.g., operational efficiency) • Users for population screening purposes • Other

Survey question	Multiple choice options
What category of outcome are you expecting to achieve for your identified 'user'? Select as many as applicable.	‣ Improved Quality of Life ‣ Improved independence/autonomy ‣ System efficiency ‣ Better experience of health services ‣ Better experience of care services ‣ Better access to health services ‣ Better access to care services ‣ Prevention of ill-health/improvement of health ‣ Faster diagnosis ‣ Faster identification of care need ‣ More accurate diagnosis ‣ Other (please specify)
At which point of care do you expect your AI-driven technology to be deployed? Select as many as applicable.	‣ Primary care ‣ Secondary care ‣ Community care ‣ Tertiary care ‣ Individual care of self e.g., user's home/office ‣ For the purposes of population screening ‣ Other (please specify)
Taking into consideration the need to train, validate, evaluate and seek appropriate regulatory approval, how likely is it that your AI-driven technology will be ready for deployment at scale within the next:	‣ 5 years ‣ 3 years ‣ 1 year

Appendix B – Questions in the NIHR Horizon Scan

Horizon Scan questions
‣ **Developer Name**
‣ **Developer Profile**
‣ **Product Name/Other names**
‣ **Source link to product**
‣ **Product Description**
‣ **Type of Scanning/ Medical Imaging (if applicable/specified)**
‣ **Clinical Area (broad)**
‣ **Clinical Condition (specific name)**
‣ **Classification of Technology**
‣ **Country of Development**
‣ **Development Stage**
‣ **Overall Regulatory Approval**
‣ **Trial Start Date**
‣ **Trial End Date**
‣ **Trial Name**
‣ **Trial Identifier**
‣ **Trial URL**
‣ **Article link**
‣ **Additional Comments**

Appendix C – Taxonomy

Type	Definition	Sub-type
Automation/ Service efficiency	Technologies within this type refer to the use of automation in the form of control systems and advanced technology to eliminate or decrease the need for manual tasks. It is usually applied to repetitive tasks, such as data entry, maintenance of records, and patient health monitoring. The solutions range from auto-mated data to feedback collection to patient triage systems.	**Patient impacting:** Technologies within this subtype are technologies directly impacting patients, for ex-ample patient triage technologies and patient chat-bots. **System efficiency:** Technologies within this subtype do not directly impact patients. They include technologies ranging from IT infrastructure to technologies used for feed-back or research.
Diagnostic	Technologies within this type refer to the use of AI tools to supplement and enhance the process of using medical images to deliver high-quality patient care across a wide variety of diseases and organ groups.	**Cardiorespiratory and neurology:** Technologies within this subtype are used to help diagnose conditions such as epilepsy through electreon-cephalograms or heart problems using echocardiograms. **Endoscopy:** Technologies within this subtype may be used in upper GI endoscopies, flexible sigmoidoscopies, and colonoscopies whereby organs inside your body are looked at using an endoscope. **Genomics:** Technologies within this subtype can help identify which genes have been affected by harmful mutations using genomic sequence data. **Imaging:** Technologies within this subtype are used during or after medical imaging to help diagnose an injury or illness, includes a range of imaging technologies. **Pathology:** Technologies within this subtype may be referred to as laboratory medicine and includes the analysis of blood, urine and tissue samples to examine and diagnose disease.

Type	Definition	Sub-type
P4 Medicine	P4 Medicine is an approach to make medicine more Predictive, Preventive, Personalised and Participatory. Its two major objectives are to quantify wellness and predict and prevent dis-ease. It incorporates a range of technologies from predicting the likelihood of a patient developing a long-term condition by analysing patient records to predicting patient response to medication, allowing to create a personalised plan.	
Remote monitoring	Technologies within this type include monitoring de-vices that collect data which can be shared with healthcare staff to monitor patients inside or outside hospitals and allow for earlier interventions if a patient's condition is worsening. They may be used to monitor patients after surgery or hospitalisation or for patients to manage a long-term condition.	**Tertiary only:** Technologies within this subtype are independent of the healthcare system such as devices where loved ones can monitor their family members with dementia. **Within clinician:** Technologies within this subtype allow healthcare professionals to view the data collected and monitor patients.
Therapeutic	Technologies within this type includes technologies which deliver evidence-based therapeutic interventions to patients that are driven by high quality soft-ware programs to prevent, manage, or treat a medical disorder or disease. It includes technologies ranging from mental health apps to technologies used in radio-therapy.	
Other	Technologies within this type are technologies which do not fit into any clear category such as AI solutions used for medical education purposes or a health information platform for children.	

Appendix D – Breakdown of the database

Column Name	Dropdown options	Notes
Technology type	‣ Automation/Service Efficiency ‣ Diagnostic ‣ Other ‣ P4 (population health) ‣ Remote Monitoring ‣ Therapeutic	Some technologies may include features from multiple technology types e.g remote monitoring technologies may have some predictive or diagnostic capabilities; however this column is capturing the primary function of the technology.
Sub-type (if applicable)	**Automation/Service Efficiency:** ‣ System efficiency (no patient) ‣ Patient impacting **Diagnostic:** ‣ Imaging (CT scanning) ‣ Imaging (MRI scanning) ‣ Imaging (PET-CT) ‣ Imaging (non-obstetric ultrasound) ‣ Imaging (DEXA) ‣ Imaging (Plain X-ray) ‣ Imaging (Smartphone as medical device) ‣ Endoscopy ‣ Cardiorespiratory and neurology diagnostics ‣ Pathology ‣ Genomics **Remote monitoring:** ‣ With clinician ‣ Tertiary only	**Automation/service efficiency:** System efficiency (no direct patient impact) e.g. automated feedback collection, Patient impacting (directly impacts patient) e.g. patient triaging system **Remote monitoring:** With clinician (HCPs have access to patient data/monitor the patients), Tertiary only (no involvement with healthcare services e.g. family members can monitor their loved ones with dementia)
Clinical area (if applicable)	‣ e.g Cardiology for example	**Provided by HEE** General Practice has been used for technologies for general health management if there is no other appropriate clinical area

Column Name	Dropdown options	Notes
Primary user	‣ Healthcare Professional ‣ Patient facing ‣ Carer (e.g Family) ‣ Commissioner	This column captures the direct user of the technology: **HCP:** Health care professional (HCP) or non-clinical staff uses the technology directly **Patient facing:** patient uses the technology directly; clinical team may review data collected by the solution e.g. self-management/remote care **Carers:** nonpaid carers, such as a family member monitoring a loved one with dementia without HCP involvement for example **Commissioner:** technologies that integrate into the system but don't interact with staff or patient e.g automated scheduling technology
Point of Care (2 columns)	‣ Primary care ‣ Secondary care ‣ Individual care of self ‣ Research ‣ Community care ‣ N/A	Care setting where technology is deployed currently or where it will be deployed in the first instance There may be some technologies which will be deployed in other care settings in the future – we want to capture the care setting intended for deployment (two care settings will only be captured in a small number of instances where there is evidence for both)
Secondary Point of Care (if applicable)	‣ Primary care ‣ Secondary care ‣ Individual care of self ‣ Research ‣ Community care ‣ N/A	There are some technologies, in particular remote monitoring, that are deployed in a patient's home but will impact another point of care e.g GP practices recommend a technology and advise a patient on setting it up and review the data collected
Sites	‣ List of CCGs, Acute Trusts, Mental Health Trusts, Community Trusts	These columns capture all sites where a technology has been/is being used or piloted

Column Name	Dropdown options	Notes
Workforce Groups affected (if applicable) - 2 columns	‣ Multiple roles affected ‣ No workforce groups directly affected ‣ Undetermined ‣ e.g Medic, Cardiology	**Provided by HEE** The database only captures direct workforce groups affected by technologies. There may be other users affected downstream however we are only capturing direct HCPs affected. For early-stage technologies where there is no information on who is impacted we have selected undetermined. For technologies, most commonly service efficiency technologies, where the technology sits in the background and does not affect a specific workforce group we have selected multiple roles affected. For technologies with no involvement with healthcare services e.g families monitoring their loved ones we have selected no HCP directly affected
Other	‣ Free Text box	Where developers mention a specific role impacted by their technology, which is not included in the list
Impact on the primary workforce groups affected: Change in the workload	‣ Reduction ‣ No change ‣ Increase ‣ Early-stage solution – undetermined ‣ No direct impact	Examples of this impact include: ‣ Change in the volume of daily tasks ‣ Change in the capacity
Impact on the primary workforce groups affected: Change in team dynamic	‣ Yes ‣ No direct impact ‣ No ‣ Early-stage solution - undetermined	Examples of this impact include: ‣ Change in interaction with other staff members ‣ More/less reliance nf the MDT or IT/ informatic colleagues

Column Name	Dropdown options	Notes
Impact on the primary workforce groups affected: Change in role	‣ Yes ‣ No ‣ Early-stage solution – undetermined ‣ No direct impact	Examples of this impact include: ‣ Change in the nature of daily tasks ‣ Change in responsibility for the decision-making ‣ Change in job description ‣ Change in the user's digital literacy (needed) ‣ User to complement and validate the information provided by the solution ‣ Change in the user skill set (widened/more specialised)
Impact on the pathway: Change in patient outcomes	‣ Improvement ‣ No change ‣ Worsen ‣ Early-stage solution – undetermined ‣ No direct impact	Examples of this impact include: ‣ Change in patient outcomes ‣ Change in patient safety
Impact on the pathway: Change in waiting time, time to diagnosis, treatment, referral or discharge	‣ Reduction ‣ No change ‣ Increase ‣ Early-stage solution – undetermined ‣ No direct impact	Examples of this impact include: ‣ Change in the time to diagnosis for patients treated with (or without) the solution ‣ Change in the time to treatment for patients treated with (or without) the solution ‣ Change in the time to diagnosis for patients diagnosed with (or without) the solution ‣ Change in the time to discharge for patients not diagnosed with (or without) the solution ‣ Change in the waiting times/waiting list
Impact on the pathway: Change in the support, communication or education provided to patients	‣ Reduction ‣ No change ‣ Increase ‣ Early-stage solution – undetermined ‣ No direct impact	Examples of this impact include: ‣ Change in the support, communication or education provided to patients

Column Name	Dropdown options	Notes
Impact on the system: Change in access to care	‣ Reduction ‣ No change ‣ Increase ‣ Early-stage solution – undetermined ‣ No direct impact	Examples of this impact include: ‣ Change in access to care (better/worse engagement with hard-to-reach groups)
Impact on the system: Change in the delivery of integrated care	‣ Reduction ‣ No change ‣ Increase ‣ Early-stage solution – undetermined ‣ No direct impact	Examples of this impact include: ‣ Communication and sharing of patient information modified ‣ Enable care in the community or change in point of care ‣ Better utilisation of patient data or change in utilisation of patient data ‣ Change in the delivery of integrated care
Impact on the system: Change in the system performance, efficiency or resilience	‣ Reduction ‣ No change ‣ Increase ‣ Early-stage solution – undetermined ‣ No direct impact	Examples of this impact include: ‣ Performance/efficiency modified ‣ Prioritisation of tasks modified ‣ Triaging of patients modified ‣ Change in patient throughput
Presence of website (Y/N)	Yes N/A No	
News story in the last quarter (Y/N)	Yes N/A No	
Published results of a clinical trial or equivalent proof of efficacy (Y/N)	Yes N/A No	
If automation: Case study/Proof of usage (Y/N)	Yes N/A No	

Column Name	Dropdown options	Notes
Ongoing clinical trial or equivalent study (Y/N)	Yes N/A No	
Early development research (Y/N)	Yes N/A No	
CE / UKCA marked (or equivalent certification: ISO, ICO, etc. (Y/N)	Yes N/A No	
If no CE/UKCA marking: Other international certifications (FDA approved, etc.) (Y/N)	Yes N/A No	
Currently used in NHS trusts, GP practices, etc. (Y/N)	Yes N/A No	
Listed in some procurement frameworks (Y/N)	Yes N/A No	
AAC AI fund awardee (phase 3 or 4) (Y/N)	Yes N/A No	
AAC AI fund awardee (phase 1 or 2) (Y/N)	Yes N/A No	
Published economic evaluation (Y/N)	Yes N/A No	
Time to deployment	‣ 1 year ‣ 3 years ‣ 5 years	Time until deployment at scale (e.g. regional or national). It is automatically calculated using the results from the 13 criteria above and a weighting system.

Appendix E – Impact on the workforce framework

Impact on the user	Impact on the pathway	Impact on the system
Change in the nature of daily tasks	Changed to the physical space used to deliver care to adapt for the ergonomics of the technology	Communication and sharing of patient information modified
Change in the volume of daily tasks	Changed in the time to treatment for patients treated with the solution	Performance/efficiency modified
Change in interaction with patients	Changed in the time to treatment for patients not treated with the solution	Prioritisation of tasks modified
Change in interaction with other staff members	Changed in the time to diagnosis for patients diagnosed with the solution	Triaging of patients modified
Change in responsibility for the decision-making	Changed in the time to discharge for patients not diagnosed with the solution	Solution alleviates/reinforces a shortage of clinical/non-clinical staff
Change in job description	Changed in the time to discharge for patients diagnosed/treated with the solution	Timescale of the change (immediate, within a year, within 3 years)
Change in the user's confidence	Changed in the time to discharge for patients not diagnosed/treated with the solution	Enable care in the community or change in point of care
Change in the user's wellbeing, work-life balance	Change in patient outcomes	Change in hospital admissions
Change in the user's digital literacy (needed)	Change in patient safety	Faster discharge to community
User to complement and validate the information provided by the solution	Change downstream in the pathway	Better utlisation of patient data or change in utlisation of patient data
More patient education needed to enable them to self-declare/fill out health check forms	Change upstream in the pathway	Change in the delivery of integrated care
More/less reliance of the MDT or IT/informatic colleagues	Change in the waiting times/waiting list	Change in patient throughput
Change in the user skill set (widened/more specialised)	Change in the support/communication/ education provided to patients	Change in access to care (better/worse engagement with hard-to-reach groups)
Change on the user capacity		Change in demand planning

Dimensions of the framework
Change on the user workload/capacity
Change in team dynamic
Change in role (skill-mix change, job widening, job deepening, new role)
Change in patient outcomes
Change in waiting time, time to diagnosis, treatment, referral or discharge
Change in the support, communication or education provided to patients
Change in access to care
Change in the delivery of integrated care
Change in the system performance, efficiency or resilience

Appendix F – Time to deployment criteria and scoring logic

The table below presents the criteria included in the calculation of the time to deployment as well as the weighting system.

Weighting	Dimension	Question	Answers	Point if Yes
1	Virtual presence	Presence of a website	Yes / No	1
		News story in the last quarter	Yes / No	1
3	Proof of efficacy	Published results of a clinical trial or equivalent proof of efficacy	Yes / No	1.5
		(If automation) Case study/Proof of usage	Yes / No	1.5
		Ongoing clinical trial or equivalent study	Yes / No	1
		Early development research	Yes / No	0.5
3	Proof of regulatory compliance	CE / UKCA marked (or equivalent certification: ISO, ICO, etc.	Yes / No	1.5
		Other international certifications (FDA approved, etc.)	Yes / No	1
2	Usage in the NHS	Currently used in NHS trusts, GP practices, etc.	Yes / No	2
		Listed in some procurement frameworks	Yes / No	0.5
		AAC AI fund awardee (phase 3 or 4)	Yes / No	1
		AAC AI fund awardee (phase 1 or 2)	Yes / No	0.5
2	Proof of economic impact	Published economic evaluation	Yes / No	1

Appendix G – Questionnaire for the case study

Section 1: Presentation of the solution

- Technology type and subtype and problem addressed
- Main features of the solution
- Clinical areas impacted (clinical area, pathway and point of care)
- Current level of implementation (number of sites, geographical spread)
- What were the main adoption challenges?
- What changes were made as a result of the early-stage piloting or first implementation?
- Is there an ambition to use the solution across different pathways or other clinical area in the future? If so, what is the time horizon (1, 3 or 5 years)?

Section 2: Impact on the workforce

- What training is provided as part of the implementation of your solution? (length of time, training model, workforce group trained, refresher sessions provided)?
- What have the users reported as challenging when using your solution? What gap in knowledge in the workforce could be addressed prior to the implementation of your solution to aid its adoption?
- What different workforce groups are directly and indirectly impacted by the use of your solution? Please be as specific as possible. Please provide an example for each healthcare professional/workforce group.
- How would you qualify the impact on the primary user (workforce group)? [Have as options the subcategories from the framework] Can you provide an example for each impact?
- How would you qualify the impact on the pathway? [Have as options the subcategories from the framework] Can you provide an example for each impact?
- How would you qualify the impact on the system? [Have as options the subcategories from the framework] Can you provide an example for each impact?

Concluding remarks

- Do you have any other comments you would like to add?

www.unityinsights.co.uk
@UnityInsights
enquiries@unityinsights.co.uk

www.hee.nhs.uk
@HEE_DigiReady
dart-ed@hee.nhs.uk

NHS Health Education England | December 2021