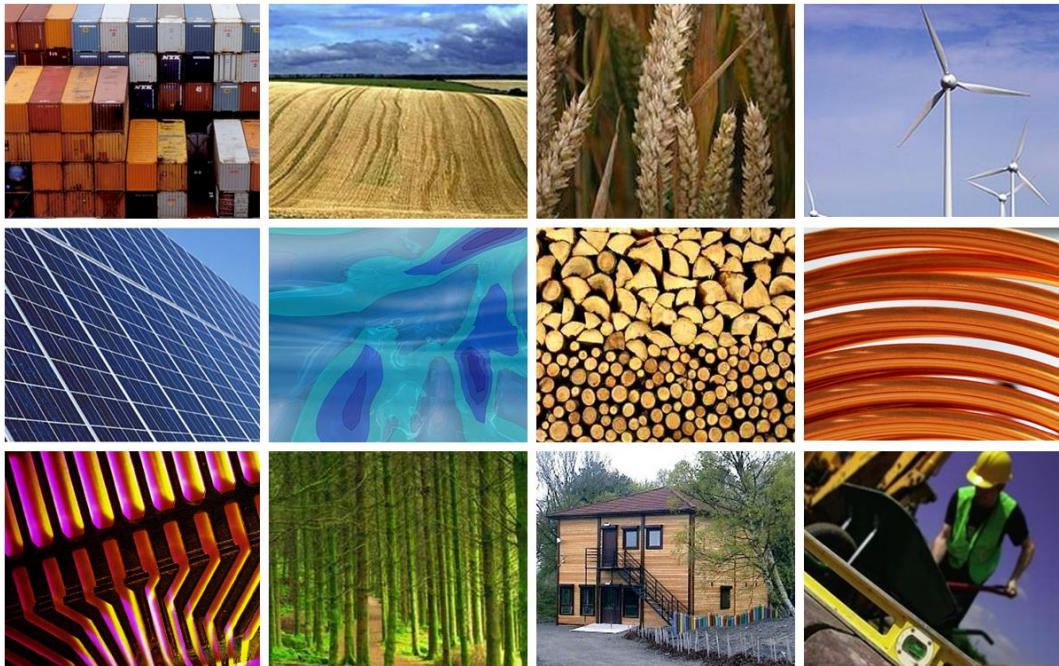


NHS AI lab and UKRI Sciencewise

Public Dialogue on Data Stewardship for AI Research

Evaluation report

October 2023



Quality Management

URSUS Consulting Ltd has quality systems which have been assessed and approved to BS EN ISO9001:2000 (certificate number GB2002687).

Creation / Revision History

Issue / revision:	V1
Date:	12.10.2023
Prepared by:	Anna MacGillivray
Authorised by:	Hilary Livesey
Project number:	u.191
File reference:	NHS AI Lab/final evaluation report

URSUS CONSULTING LTD
www.ursusconsulting.co.uk

57 Balfour Road
 London N5 2HD
 United Kingdom
 Tel. 07989 554504



Contents

Glossary and Acronyms

- 1 Introduction.....4
- 1.1 Background 4
- 1.2 Policy Context..... 4
- 1.3 Dialogue objectives..... 5
- 1.4 Framing of the public dialogue 5
- 1.5 Dialogue and evaluation approaches 6
- 1.5.1 Evaluation approach 7
- 2 Impacts.....8
- 2.1 Dissemination of findings..... 8
- 2.2 Impacts on the NHS AI Lab research competition so far 9
- 2.3 Potential for future impact on NHS Data Stewardship approaches 11
- 2.4 Making existing delegated approaches more participative..... 11
- 2.5 Economic benefits likely to exceed the dialogue costs 12
- 2.5.1 A relatively low cost public dialogue 12
- 2.5.2 Potential for significant longer term economic benefits..... 13
- 3 The factors which contributed to the dialogue meeting its objectives..... 14
- 3.1 Meeting the project objectives..... 14
- 3.2 How the design and delivery contributed to meeting objectives 15
- 4 Conclusions and emerging lessons for future dialogues..... 19
- 4.1 Conclusions 19
- 4.2 Lessons learnt and recommendations..... 19

- Annex A** Dialogue methodology
- Annex B** Oversight Group and stakeholder contributors
- Annex C** Assessment of how far the dialogue has met its objectives
- Annex D** Assessment of delivery against Sciencewise best practice principles
- Annex E** Evaluation feedback from participants

Glossary and Acronyms

ALB	Arm’s Length Body
CAG	Confidentiality Advisory Group
COPD	Chronic obstructive pulmonary disease
DAC	Data Access Committee
Data stewardship	The process of collecting, maintaining and providing access to data
De-identified data	Information at a person-level, where identifiers have been obscured (e.g. replacing names with pseudonyms) so that it is not easily identifiable
5 Safes	Five Safes framework is a set of principles developed by ONS which enables data services to provide safe research access to data by control of who, where and how the data can be used.
GPDRP	General Practice Data for Planning and Research
HDRUK	Health and Data Research UK
HRA	Health Research Authority
ICHP	Imperial College Health Partners
ICO	Information Commissioner’s Office
IGARD	Independent Group Advising on the Release of Data
MHRA	Medicines and Healthcare products Regulatory Agency
NDG	National Data Guardian
NDOO	National data opt out service allows patients to opt-out of their confidential patient information being used for research and planning. NHS Digital and all health and care organisations must comply.
NID	National chest imaging database
ODI	Open Data Institute
POD	Personal Online Datastore
SDE	Secure Data Environment
ToR	Terms of Reference
TRE	Trusted Research Environment
UDP	Understanding Patient Data

Executive summary

This report summarises the findings of an independent evaluation of a public dialogue on data stewardship on behalf of NHS Transformation Directorate's Artificial Intelligence (AI) Lab with support from the UKRI Sciencewise programme.

Context

There has been an increasing interest between database owners, academics and the charity sector in encouraging more participatory approaches to data stewardship for de-identified data in the health and other sectors. Review of a wide range of approaches – including delegated, collective and individual - suggest that greater transparency and public agency in deciding how data is used, by whom and for what purposes increases trust in the process and makes it more likely that individuals will be willing to opt-in to data sharing for the public good.

Dialogue aims

The overall aims of this public dialogue were to get a deeper understanding of the risks and benefits of different data stewardship models from the perspective of the public, and to understand their expectations for how patients and the public should be engaged in decisions about data access for AI research in the health sector. The outcomes from the process were expected to help inform the next phases of NHS AI Lab's research competition to scope out and prove the concept of more participative approaches to data stewardship of large medical imaging databases.

Dialogue approach

This small wholly online dialogue process was designed and delivered by Ipsos in partnership with the Open Data Institute (ODI) and Imperial College Health Partners (ICHP) and ran between Spring and Autumn 2022. The work was overseen by a 13 person Oversight Group (OG) and managed by a small core team including NHS AI Lab staff, Sciencewise, the delivery contractors and independent evaluator.

The dialogue brought together 47 individuals broadly reflective of the UK population, including some heavy users of the NHS. Participants met online (Zoom) for 12 hours over two weeks in a mix of three weekday evenings and one Saturday morning session. During this time they heard from two specialists on health data, AI and how it might be used in research (case studies) and then spent time in small groups (7-8 participants re-mixed each session) to discuss their hopes, concerns, and thoughts on nine theoretical data stewardship scenarios (spanning delegated, collective and individual models). In the final session participants co-created a set of nine principles which could be used as criteria for assessing the pros and cons and trade-offs involved in different stewardship models.

Dialogue impacts

The outcomes (cross-cutting issues, reflections on scenarios and the principles) were presented in a final dialogue report. The NHS AI Lab team has drawn heavily on these findings

in drafting an invitation to tender for Phase 1 research (scoping and discovery) and received a good response with a long list of 25 bids partly on the basis of how well they had integrated the dialogue findings in their approach. The winning research team is now developing and testing models bringing together elements of those considered in the dialogue. The dialogue principles will be incorporated in a multi-dimensional analysis for refining the models. The ethos of involving the public has been carried forward to refining and testing the models and useability of tools. Background information created for the dialogue will help to get participants up to speed on the topic.

In the medium term the dialogue has potential to help shape how the NHS develops data stewardship for AI purposes of large medical imaging databases. Positive impacts will depend on a number of decision points: at least one workable participative model emerges from Phase 1; that Phase 2 proves the concept workable; and that NHS teams are willing and able to implement the approach in the prevailing funding and political context from late 2024 onwards.

In the longer term, a successful participative model being applied by the NHS could have significant economic benefits: patients and the public may feel more trust and opt-in in higher numbers; the resulting AI research may help improve the efficiency and effectiveness of diagnosis and treatments, reducing costs and improving health outcomes. More broadly a supportive atmosphere for AI research could help strengthen AI development and the sectors' contribution to the UK economy. Such benefits would greatly outweigh the initial costs of the dialogue (£149.9K financial and an estimated £30K in-kind contribution in commissioner, OG and stakeholder staff time).

Lessons on good practice

The key elements which contributed to this dialogue's success and the credibility of the findings included:

- **An experienced delivery team which incorporated both dialogue and topic (health, data and AI) expertise.** The team were able to pull in key stakeholders and to rapidly design and deliver a 'no frills', but effective, dialogue within a challenging time frame.
- **A small and efficient core project management team** were able to make timely decisions, respond to changing circumstances and keep their policy colleagues on board.
- **A well chaired Oversight Group** was able to contribute a full range of health, data and ethical perspectives and feed in valuable inputs at the right moments. While some members were less involved in the later stages, several went on to be involved in the follow-on research stages (selection and steering) ensuring continuity with the dialogue.
- **A cost and time efficient online meeting process.** Four large workshops on Zoom (47 participants) at convenient times over an intense two week period kept up momentum and helped minimise drop-out rates.
- **Continually remixing the small groups exposed participants to the full diversity of backgrounds and lived experience within the wider group.** A combination of staged thankyou payments, the warm and encouraging atmosphere created, and plenty of opportunity for discussion with each other kept the group focused on a potentially dry topic and a difficult task (reviewing many subtly different scenarios).

- **Pitching the level and amount of information provided by specialists just right.** Since the topic was not particularly contentious just two specialists were recruited to share background information. This helped control the quality of the information and ensured that all participants could hear the same answers.
- **Elicitation exercises based on theoretical, but realistic, scenarios for data stewardship took participants on a journey from the status quo (delegated approaches) through more participative approaches.** Although repetitive, the exercises worked well to draw out overarching principles used to assess the pros and cons of different models. This has proved one of the most valuable outputs from the dialogue.
- **Allocation of time and resources for simultaneous transcription during workshop sessions and rapid analysis after each workshop allowed the contractors to share a timely and very detailed interim report.** The OG and commissioners comments were then able to help the team build this quickly into a well-written final report with a strong narrative that put the participant's voices front and centre.

1 Introduction

1.1 Background

This final evaluation report, submitted by URSUS Consulting Ltd to NHS AI Lab and UKRI Sciencewise, summarises the evaluation findings of a public dialogue on the benefits and risks of different approaches to data stewardship for AI purposes. The report builds on an interim report and focuses on the impacts of the public dialogue six months after the findings were shared via an NHS AI research competition (*chapter 2*). It also summarises the factors which have contributed to the dialogue meeting its immediate objectives (*chapter 3*) and the conclusions and recommendations for future public dialogues (*chapter 4*).

1.2 Policy Context

The NHS Transformation Directorate’s Artificial Intelligence Laboratory (AI Lab) was set up to enable the safe, ethical and effective use of AI-driven technologies in health and care. In February 2021, the AI Lab launched the [AI Ethics Initiative](#) to support research and practical interventions to help strengthen the ethical adoption of AI by bringing together government, health and care providers, academics and technology companies across the sector.

In the last few years a wealth of information has been collated on public attitudes towards use and sharing of health data (summarised in the [Understanding Patient Data library](#)), but this has stopped short of investigating practical models for data stewardship. Recent research by academic and policy experts has explored different models for more participative models for data stewardship¹.

Within the health and care system, several initiatives have started to involve patients and the public in ethical review processes for proposed research and development or as part of Data Access Committees (DACs). Members of the public have, for instance, been part of the [Confidentiality Advisory Group](#)² (CAG) and the [Independent Group Advising on the Release of Data](#)³ (IGARD). However, the recent backlash against the [General Practice Data for Planning and Research \(GPDPR\)](#) proposals – which many patients chose to opt-out of – suggests that much more needs to be done to develop models that address the public’s key concerns.

In order to understand the public’s concerns better and help develop some practical models to address them, the AI Lab will run a research competition between 2023 and mid-2024

¹ These include those introduced in the US ([Open Humans](#)), Europe (My Data, [Flanders](#)), [India](#) and globally by the [World Economic Forum](#). NHS AI Lab’s recent work with the Ada Lovelace Institute on [algorithmic impact assessments](#) (2022) highlighted the importance of public participation in data stewardship for AI based research. Likewise, the [Goldacre Review](#) identified the need for greater public participation in the design and running of health data infrastructure.

² An independent body, including members of the public, which provides advice on the use of confidential patient information to the Health Research Authority and the Secretary of State for Health.

³ an independent group which makes recommendations to NHS Digital about data access requests whose membership is a mixture of both specialist and ‘lay’ members

when the initiative wraps up. This public dialogue was the first step in that process, with dialogue findings designed to inform the research competition terms of reference (TOR) and the approach taken by successful bidders in carrying out their research. The dialogue insights are also expected to be of wider interest in the health research and AI fields.

1.3 Dialogue objectives

This public dialogue was the first step in trialling a bottom-up approach to deciding priorities for investment that can further the AI Lab’s aim to strengthen the ethical assurance of AI-driven technologies in health and care. The specific objectives were the following:

1. To engage a diverse and broadly reflective section of the public in exploring different models of data stewardship.
2. To understand attitudes towards these different models, and which models are trustworthy.
3. To understand the values and principles that underlie dialogue participants’ views on data stewardship models.
4. To guide the AI Ethics Initiative in prioritising its next tranche of research funding based on the findings.
5. To inform the data stewardship approach to national [medical imaging] databases based on the findings.

The aim initially was to agree a procurement pathway and seek research bids by the end of January 2023, with a view to appointing two to five research grantees to carry out scoping and then pilot their preferred approaches. However, in order to make best use of available procurement pathways the research competition has been divided into two stages: a discovery and scoping stage in mid-2023; with piloted of recommended models in from late 2023 to mid-2024. Largely due to restructuring within the NHS, the Phase 1 research was delayed until June 2023.

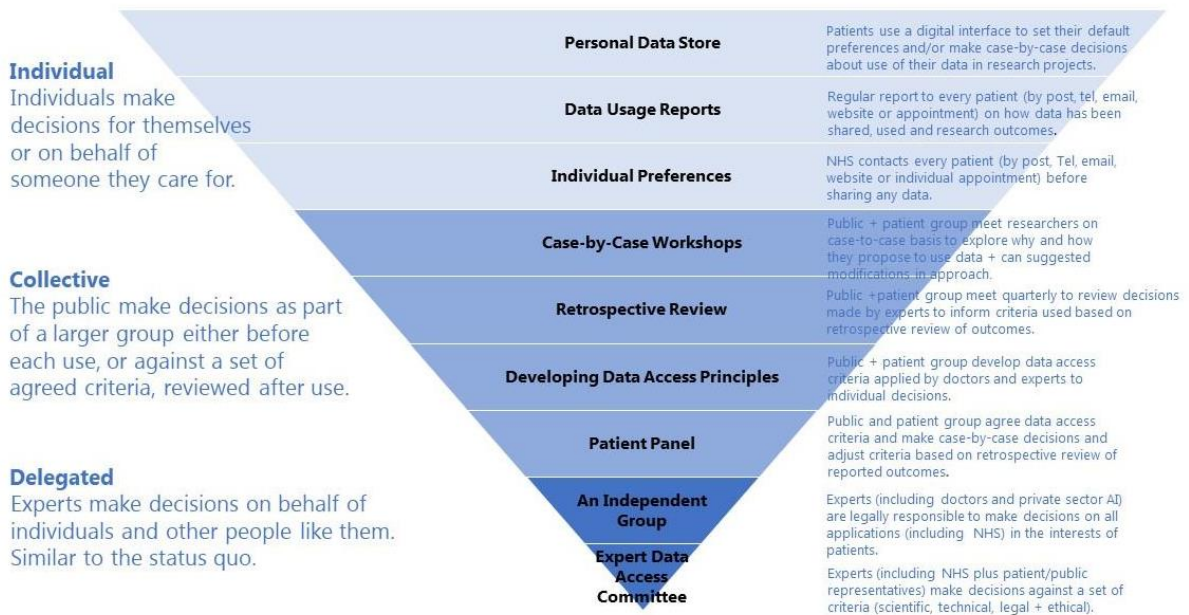
1.4 Framing of the public dialogue

- The framing agreed by the commissioners and Oversight Group (OG) was intended to build on previous dialogues (e.g. on patient data and ethics of AI) and to generate insights which would be applicable for national imaging databases – which are different in complexity and size to other types of health database.⁴ The commissioners and OG members were keen that discussions should be rooted in reality and therefore initially planned to use the National COVID-19 Chest Imaging Database (NCCID) as a case study, which would then also be the focus of the research competition. However, as NHS priorities moved away from the NCCID, the dialogue shifted focus to a hypothetical imaging database.

⁴ Imaging databases include massive datasets including scans, x-rays (involving very large file sizes) alongside medical and demographic data. In order to be really useful for research they need to be fully representative, including of Black and ethnic minority communities. The regional Breast Imaging database (OPTIMAM) is one of very few large scale examples in the UK. The National Chest Imaging Database is one of very few national datasets.

- In order to bring the concepts alive, specialists presented several case studies showing different partnerships between the NHS, academics and businesses to use AI for research and analysis of imaging data. Examples included an NHS hospital working in partnership with a tech company to review eye scans, and an NHS hospital and university working together to screen for the causes of dementia and strokes.
- The design also explored models for data stewardship (see *Figure 1.1*) spanning the delegated approaches (the status quo), collective and individualised. The nine scenarios were drawn from both the health and non-health settings identified through a literature review, stakeholder consultation and the experience of oversight group members. Each scenario was broadly realistic but with limited detail so as to allow space for participants to think creatively.

Figure 1.1: Three models and nine hypothetical scenarios for data stewardship



1.5 Dialogue and evaluation approaches

Dialogue approach

- This relatively small public dialogue (£149K) was designed and delivered by Ipsos in partnership with the Open Data Institute (ODI) and Imperial College Health Partners (ICHP). Together they brought experience in running public dialogues and expertise in managing health data and data stewardship approaches.
- A diverse group of 50 public participants were recruited to be broadly reflective of the UK and attended four online (Zoom) workshops for a total of 12 hours spread over three weekday evenings and a Saturday morning. A detailed description of the methodology is shown at *Annex A*.
- The project was managed by a core team led by NHS AI Lab, and including UKRI Sciencewise, the contractors and the independent evaluator and overseen by an Oversight Group (OG) chaired by Natalie Banner of Genomics England. Together they brought together a diverse group of health, data and AI experts from NHS, academic and

non-governmental (NGO) backgrounds. Further expertise in data and AI from other sectors was contributed by a Stakeholder group. OG members and stakeholders are listed at *Annex B*.

- Links to the final dialogue report are available at [NHS AI Lab](#), [Sciencewise](#) and [Ipsos](#) websites. In addition a set of anonymised transcripts from the workshops was shared with AI Lab.

1.5.1 Evaluation approach

The evaluators were not commissioned until part way through the dialogue, after the initial scoping phase, and the evaluation has therefore been mainly summative. The findings draw on observations from attending project management, OG, stakeholder and public workshops (*see Annex C*), and detailed participant feedback (*Annex E*). During the interim and final evaluation stages we have carried out interviews with the core team, OG members and the successful research consortium leader. This final evaluation report also draws on desk research and a short online survey completed by a handful of OG members.

2 Impacts

2.1 Dissemination of findings

- In order to meet a challenging deadline to feed into the next rounds of research, the delivery team made great efforts to share findings as they emerged after each workshop, to develop a very detailed interim report and to produce a high quality final report by November 2022.
- A high quality report put the participant’s voices front and centre; it provided a nuanced understanding of what matters to the public, trade-offs that will need to be considered and the implications for three broad data stewardship approaches (delegated, collective and individual) and nine specific models (or blended approaches). NHS AI Lab and OG members interviewed agreed that “*We got the insights we needed to go to the next stage and test different models.*”

Table 2.1: Dissemination of dialogue report and finding

Target audience	Audience	Dissemination approach
Research groups interested in applying for the research competition (25 bids received)	<ul style="list-style-type: none"> • Research organisations (universities, trusts, commercial) 	<ul style="list-style-type: none"> • NHSE website blog: Aims of research and implications of dialogue findings • AI Lab ‘show and tell’ webinar for potential researchers
	<ul style="list-style-type: none"> • Personal online datastores (POD) ‘solid’ community including ODI, Inrupt and international universities in US (MIT) and Belgium 	<ul style="list-style-type: none"> • Report shared to encourage them to bid for the research competition to integrate dialogue lessons into proposed POD approaches
Those interested in the findings and views of the public	<ul style="list-style-type: none"> • National, regional or trusts setting up or running safe research environments and data access agreements • Policy makers developing policy for research environments and data access • Learned societies such as Royal Society⁵. Findings resonated with RS findings on public trust and willingness to share data 	<ul style="list-style-type: none"> • NHSE website blog: Findings and implications for those working on research environments • NHS AI Lab ‘show and tell’ event to potential audience of 2000 stakeholders and members of the AI Virtual Hub • NHS Transformation Bulletin (~6000 readers in NHS, arm’s length bodies (ALBs), industry & academia) • NHS AI lab speaker pack and slide deck for sharing at internal & external events • NHS AI Lab webpage update • Ipsos, ODI and ICHP shared with findings with Health Data Research UK (HDR) on Trusted Research Environments workstream
Those interested in using novel methodologies for informing research investment	<ul style="list-style-type: none"> • Funding organisations and other public sector bodies or depts within NHS • Comms, engagement and public participation specialists in NHS, Civil Service & wider 	<ul style="list-style-type: none"> • The NHS AI Lab team continue to share the findings informally with these audiences

⁵ [Creating resilient and trusted data systems: The public perspective and recommendations for action, 2023, Royal Society and HVM](#) findings that public trust in data systems comes from clarity of purpose for that data, transparency in data flows, and knowledge of who owns, controls and governs.

- Since the findings were mainly intended for an internal audience there was no external launch event but the report and findings have been shared by the AI Lab project manager and OG members with both internal NHS and external stakeholder audiences (see *Table 2.1*).
- The report has been well received. One interviewee described the dialogue as: "*the start of a conversation with the public to create a social licence for safe use of health data and what is an acceptable level of risk- with research and policy developing around this.*"

2.2 Impacts on the NHS AI Lab research competition so far

Table 2.2: How far the dialogue has impacted on NHS AI Lab research competition

Dialogue impact objective	How far the findings have contributed to the objective
To guide the AI Ethics Initiative in prioritising its next tranche of research funding based on the findings	<p>Influence on the bidding process (November to June)</p> <ul style="list-style-type: none"> • AI Lab took the findings into account in designing the competition. Cross-cutting issues, principles and scenarios/models were fully cited in research TOR for phase 1 (conceptual feasibility). • OG guidance that individualised models should reflect the existing context were reflected by including descriptions of existing participatory approaches (e.g. CAG and IGARD). • 25 bids received and shortlisted by a small selection committee (including 3 OG members) using how well the dialogue findings had been integrated as one assessment criteria. All shortlisted bids were considered to have integrated the dialogue findings well. <p>On the methodology of Phase 1 (underway)</p> <ul style="list-style-type: none"> • Phase 1 has been strongly influenced in terms of the models being considered, the criteria being used for multidimensional analysis⁶, and in involving the public and patients in the co-creation and testing process. <p>Potential for impact on the models to be piloted in Phase 2 (from October 2023)</p>

- An initial draft of the Phase 1 Terms of Reference (TORs) was shared for review with all OG members in November and their suggested amendments were reflected in the final brief (December) which went live in February 2023.⁷
- [The TOR](#) explicitly referred to and included a link to the dialogue. Cross-cutting issues, the nine principles and nine data stewardship scenarios (including pros and cons identified by the participants) were included in the text and annexes. Recognising that no single scenario had emerged as a clear favourite, the brief also encouraged 'combined' options or novel approaches (described in the dialogue report as "mix and match" approaches).
- Bidders were made aware that they would be expected to take the public dialogue into account, would be assessed on how well their approach integrated the findings and expected to explain the impact of the dialogue in discussions with the evaluator.⁸ Effective integration of the dialogue findings was a criterion for whittling down a long list of 25 bids to a short list. The selection panel judged that all shortlisted bids had done

⁶Some principles may have been more linked to national databases (e.g. costs of contacting large populations, getting timely responses, and risks of high opt-out which might compromise the quality of data for research).

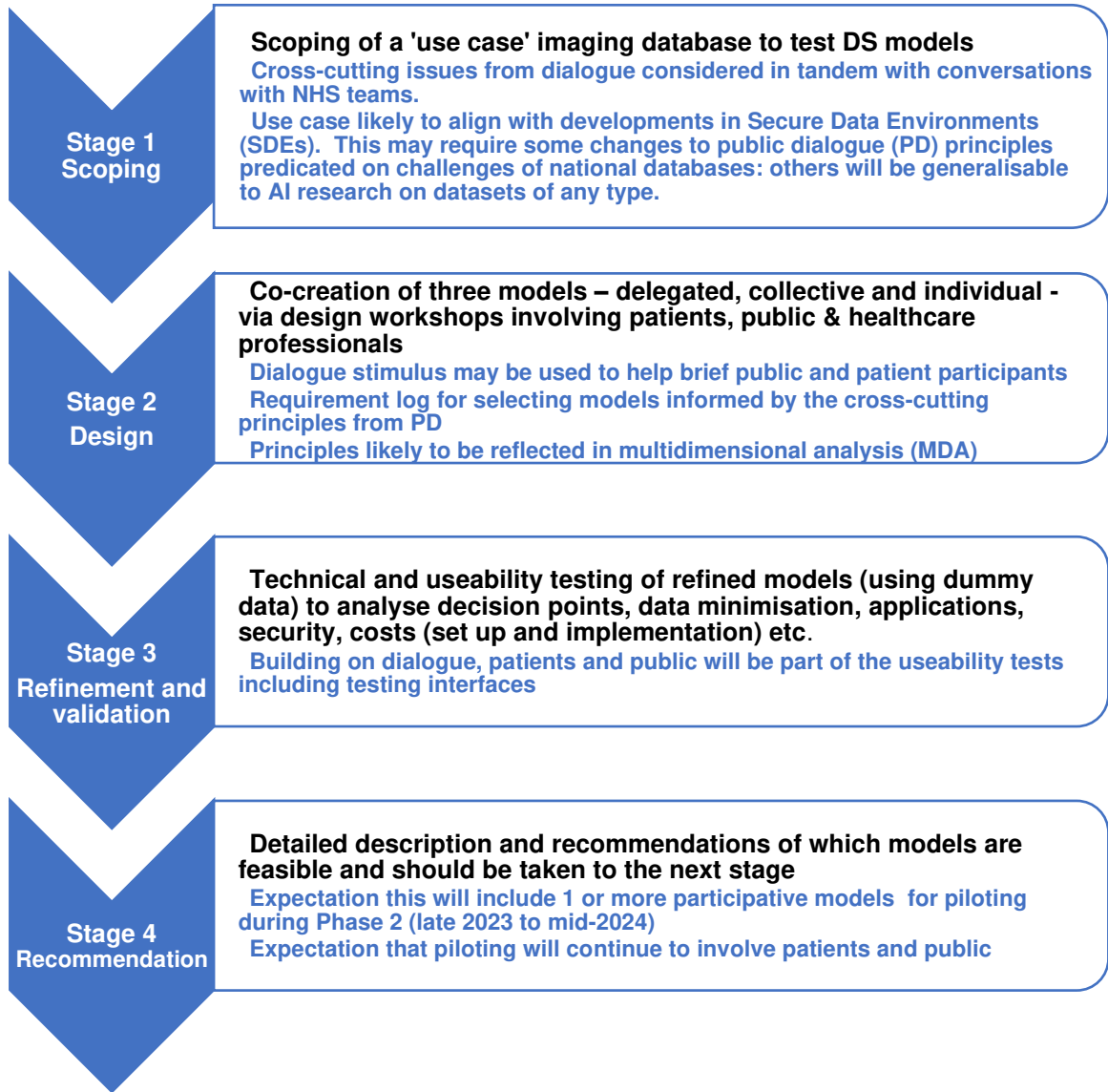
⁷ <https://www.digitalmarketplace.service.gov.uk/digital-outcomes-and-specialists/opportunities/19837>

⁸ "*The successful bidder will be required to participate in an interview with an independent contractor as part of an evaluation of the public dialogue we conducted.*" ([service description, Feb 2023](#)).

this well. As a result the commissioners felt the dialogue has already had a positive impact on the overall quality of the process: *“The dialogue findings have been very helpful in advancing our thinking and improving the quality of the bids.”*

- Although the successful bidder (Atkins) has only been in place since June, the value of the dialogue findings is already evident. *Figure 2.1* summarises how the findings and proactive public engagement ethos have been threaded through the research approach.
- In addition, the research steering group will include a number of organisations⁹ represented on the OG so guaranteeing some continuity with the dialogue and increasing the likelihood that this will feed through to Phase 2.

Figure 2.1: How public dialogue (PD) findings are informing phase 1 research



⁹ Membership includes internal stakeholders (AI Lab, Secure Data Environment, Data for R&D teams) and key external stakeholders including: Health Research Authority (HRA); Medicines and Healthcare products Regulatory Agency (MHRA); Health and Data Research UK (HDRUK); Understanding Patient Data; Information Commissioner’s Office (ICO); and the National Data Guardian (NDG).

2.3 Potential for future impact on NHS Data Stewardship approaches

The final dialogue objective - to impact the approach that the NHS takes to data stewardship of national medical imaging databases – has not yet been met. The potential for impact is considerable but will not be evident for several years. In order for positive impacts to be seen a number of decisions need to align: points:

- Positive recommendations emerging from Phase 1 for participative data stewardship model(s).
- Funding to pilot these approaches is still being considered a priority in late 2023.
- Successful commissioning of one or more research groups to pilot participative models, and willingness of the researchers, public and patients to be involved in detailed piloting to prove the concept
- The ability of internal NHS teams and other steering group members to keep policy makers on board and implement one of the successful participative models to relevant databases in the funding and political climate prevailing from late 2024 onwards. The switch from a national chest imaging database (strongly related to Covid) to a hypothetical non-national database may help keep policy makers on board (e.g. they may feel more able to buy into the dialogue approaches without feeling bound by participants' preferences for existing databases).

2.4 Making existing delegated approaches more participative

In the health sector two organisations who took a leading role in the OG (Genomics England and the Royal Surrey-led [OPTIMAM](#) (regional breast imaging database) valued the opportunity to take part in the process. They appreciated the fact that participants identified ways of making existing delegated approaches more participative and transparent by widening public and patient membership of committees. Both organisations currently operate data access/review committees which they hope might benefit from the dialogue insights. One interviewee valued how the dialogue: "*highlighted the benefits of in-depth dialogue as opposed to tick box approaches to patient and public involvement and engagement (PPIE)*". Both organisations are considering how more meaningful public engagement could be built into both their current stewardship and any future database sharing models.

Other non-health OG members were encouraged that the participant's journey took most to a place where they recognised the benefits of more participatory approaches. One interviewee was pleased to note that "*participatory approaches were as important to members of the public as they are to experts working in this space*." Others appreciated the efforts participants made to grapple with the complexities and the nuanced suggestions they made about how such models could be developed: "*the public consulted had a very clear understanding and appreciation of the techniques and challenges associated with granular individual approvals*."

Table 2.3: Potential for wider impact of the dialogue findings

Stakeholder	Intended audience	How findings have been used
Genomics England	Internal teams involved in Whole Genome Sequencing	Diverse Data Initiative e.g. on whether models which allow greater individual agency and visibility of personal data might encourage higher opt-in rates. Access Review Committee (which includes patient and public representatives) could explore a more nuanced approach (e.g. opt-in/out for specific projects).
OPTIMAM (regional breast imaging database) with 80 vendors and 240 research applicants	Findings shared with internal colleagues to inform current stewardship approaches and ensure more participative approaches are considered rather than tick box PPIE	Integrated relevant PD findings to refine existing Data Access Committee (a delegated approach) e.g. by reviewing committee membership and 'assumed consent' model for data access requests. Will use findings to help shape data stewardship approaches for new shared databases.
British Medical Association (BMA)	Findings shared with internal colleagues	A reminder of the disparities in understanding about the topic amongst the public and for public engagement to reflect this Greater likelihood of using public dialogue as an engagement tool in the future
UKRI	House of Commons Parliamentary committee on Governance of AI to assess	This dialogue is cited in UKRI's written evidence on governance of AI (Dec 2022) as part of the case for public dialogue as a key input to development of future AI applications and the regulatory regime. UKRI evidence reflected in the Government White paper: A pro-innovation approach to AI regulation, Updated 3 August 2023

2.5 Economic benefits likely to exceed the dialogue costs

2.5.1 A relatively low cost public dialogue

- This was a relatively small public dialogue with financial costs of £149,950 excluding VAT plus an estimated additional £30K of in-kind contributions in time from the NHS Transformation Directorate, OG members, specialist contributors to the workshops and stakeholder participants in workshops. OG members interviewed felt that the time they had invested was commensurate with the value they took from being part of the process: indeed, three individuals (or their colleagues) were happy to also sit on the research competition selection committee. Organisations represented on the OG have gone on to help steer the next stages of the research.
- This initial investment has already significantly influenced the next £150,000 worth of follow-on research. The ground already covered in the dialogue has allowed the contractors to make a rapid start, reduced time needed to produce background explanatory information and demonstrated the benefits of involving the public in the research. If Phase 2 goes ahead, this research will provide a robust starting place for the successful research team(s) applied research. In the unlikely event that no new participative approaches are recommended, having designed in this decision gate will enable the NHS team to reallocate the earmarked funds to other priorities.

2.5.2 Potential for significant longer term economic benefits

If Phase 2 goes ahead and proves the concept of more participative model(s) the knock-on economic benefits for the NHS and UK economy could be considerable.

- Recent research (Royal Society, and large scale polling reported in the [BMJ](#)), [confirm the](#) general support for sharing de-identified data for research to the NHS, academia, and the charitable sector found in the dialogue. And confirm that trust in transparent and fair approaches to data sharing are key to encouraging high rates of opt-in. As noted in *Section 3*, increased transparency and an increased sense of agency have already encouraged a few public dialogue participants to rethink their previous decisions to opt-out of national health data sharing.
- The Government's AI White Paper (2023)¹⁰ highlights how successful AI research on medical data ([Mia mammography intelligent assessment](#), NHS England, 2021) could help improve diagnosis, find new treatments, develop drug and diagnostic tests and suggest further research and trials. Such outcomes could reduce costs for the NHS and improve health outcomes for patients.
- Proving the benefits of more open and transparent health data stewardship approaches in the health sector may also encourage wider support for AI applications (already estimated to contribute [£3.7 Bn in Gross Value Added](#)¹¹ (GVA) to the UK economy) and cement the UK's position as a global leader in this area.

¹⁰ [Government White paper: A pro-innovation approach to AI regulation, Updated 3 August 2023](#)

¹¹ [AI Sector Study 2022](#), Department for Science, Innovation and Technology and Office for Artificial Intelligence

3 The factors which contributed to the dialogue meeting its objectives

3.1 Meeting the project objectives

Despite this being a challenging and potentially emotive topic, this rapid ‘no frills’ design was able to recruit and engage a diverse and inclusive group of public participants in a dialogue that has really added value in deepening understanding of attitudes to different data stewardship models and values that underlie them. *Table 3.1* summarises how the immediate dialogue objectives were met and *Section 3.2* summarises the key factors that have contributed to success. Fuller analysis is included in *Annexes C to E*.

Table 3.1: How far the dialogue has met its immediate objectives

Dialogue objective		Explanation
1. Engage a diverse and broadly reflective section of the public in exploring different models of data stewardship	√√√	<ul style="list-style-type: none"> • 50 participants were recruited to reflect national demographics and include lived experience of heavy users of the NHS. • Participants attended 12 hours over 4 sessions with minimum drop out (47 attended all sessions). • A combination of scheduling, a clear design, lively and warm atmosphere, and staged thank you payments kept almost all participants fully engaged despite the challenges of an abstract topic and very detailed tasks.
2. To understand attitudes towards different models, and which models are trustworthy	√√√	<ul style="list-style-type: none"> • Expertise and perspectives drawn from the delivery team, OG and stakeholders ensured information shared with participants gave them enough knowledge to consider 9 different data stewardship scenarios. • Scenarios shifted from framing around a national chest imaging database to a hypothetical model: this took up some valuable design time, but also freed participants to think more broadly. • Repeated elicitation exercises to explore the pros and cons of each model generated useful insights. Most participants increasingly saw the benefits of more participative approaches but did not favour any one model. The suggestions they made for improving delegated, collective and individual models were sufficiently detailed to inform the research and to make OG members think about existing delegated models.
3. To understand the values and principles that underlie dialogue participants’ views on data stewardship models	√√√	<ul style="list-style-type: none"> • Analysis of values that emerged in a first round of reviewing the 9 scenarios was played back to participants as principles¹² in the final workshop. Single words such as ‘transparency’ were then used as criteria to reassess the 9 scenarios. • By unpicking and refining these terms the group co-created a set of principles which all were able to agree on. These proved useful in structuring the dialogue report, framing the Phase 1 research Terms of Reference (TOR) and will be used as part of the multidimensional analysis of three models.

¹² Expertise, transparency, feasibility, choice, accountability, consistency in decision-making, inclusivity, completeness of the data and speed of research happening.

3.2 How the design and delivery contributed to meeting objectives

Experienced delivery, project management teams and a well chaired Oversight Group were quickly able to design a simple but effective process to a challenging time frame

- The inclusion of topic expertise on health data and stewardship alongside experience of running Sciencewise-funded dialogues allowed the team to hit the ground running.** The wholly online process was based around large Zoom workshops (up to 50 participants) and printed background information shared in advance. Although the design did not include other digital tools (such as a dedicated micro-site or polling), evaluation interviewees agreed that the team had *"delivered something very complex in a very tight timeline to a very high quality."*
- A two person NHS AI Lab project management team was able to make timely decisions and respond to changing circumstances and priorities within the NHS.** The team made considerable efforts to keep NHS colleagues on board during a time of internal structuring, which is likely to have positive impacts in generating support for Phase 2 and implementation of any preferred model. This may have also been made easier by moving away from the national chest imaging framing towards a hypothetical database.
- A diverse Oversight Group (13 members including clinical, data science and ethics expertise) met three times online and was able to help shape the design.** The OG's inputs to framing, recruitment sampling, suggesting plausible data stewardship scenarios and as specialists and observers helped ensure the outputs were robust and credible. Some OG members moved posts or became less engaged in the final stages but three have continued to support AI Lab in selecting and steering Phase 1 research. OG members interviewed valued the opportunity to take part: *"It was great to hear fresh perspectives on a complex topic."* And *"It was interesting to see that members of the public were able to bring up points that experts are discussing on a day-to-day basis."*

An online recruitment process, with just one set of workshops scheduled at convenient times over an intense two-week period, proved a cost and time efficient approach to bringing a national group together and for participants to hear a diversity of views

- The mix of 47 participants was both diverse (closely reflecting national demographics) and inclusive (including those with lived experience as heavy users of the NHS with chest related conditions).** Most participants found the weekday evenings and weekend morning sessions convenient, and of about the right length (although a few found the 3-hour evening sessions rather long and we noted their attention flagging). High retention rates meant that all participants were able to make the journey from thinking as patients to thinking as citizens. Over the journey most participants developed a growing sense that more participative approaches would have real benefits over the status quo.
- Participants enjoyed exposure to different life experiences and the diverse views as a result of remixing groups for each session.** For many this was a highlight of taking part. Mixing up the groups also helped facilitators to manage a few more dominant individuals and created space for quieter individuals (including young people and those from Black and minority ethnic community) to share their views. This added to the overall

richness of the findings but did not allow any disaggregation of the findings (as AI Lab had hoped): this would have needed to be consciously built into the design (e.g. in how groups were organised and polled) from the outset.

- The overwhelmingly positive feelings that participants expressed about taking part are summarised in the Word Cloud below** (Figure 3.1). In evaluation feedback (See Annex E) almost all participants said that they understood the scope of their influence (informing the next phase of research rather than how the NHS makes health care decisions). Nonetheless, participants felt it was important for the public to be involved. They enjoyed the workshops and the opportunity to learn more about health and data access issues. A few participants reported that having a clearer understanding of how delegated decisions are currently made had already led them to rethink their previous opt-out from national health data sharing: one told us that *"I feel I can [now] opt-in if I know people with opposing views have scrutinised the decisions going forward."*

Figure 3.1: How participants felt about having taken part (47 participants)



Participants received enough, accessible information to enable them to engage in discussions without feeling overwhelmed

- Sparing use of specialists to share introductory information ensured that all participants heard the same background information (health data and AI) and the same answers to their initial questions.** The choice of calling on only a small number of specialists (2) to present the materials made best use of the available design time to polish and print the stimulus materials so they could be sent in advance by post. Participants appreciated having hard copies of the materials to hand and often referred back to them in small group discussions.
- Broader perspectives were introduced via case studies and data stewardship scenarios grounded in real life examples suggested by OG members and**

stakeholders. Printed and digital versions of the scenarios followed the same format but participants and facilitators sometimes appeared to struggle to differentiate between them, resorting to reading out the wordy text in small group discussions. In some cases participants would have liked more information (such as how many requests and records might be involved, what the cost of different approaches might be to the NHS) but this also left space for them to suggest how models could work in practice: this will be helpful in the next phases of the research. In retrospect, PowerPoints could have been less wordy and made greater use of pictures, icons, colour coding etc. to help participants navigate between them and home in on the essential aspects of each.

- **Data specialists within the core team attended all workshops as observers.** Their role helped with the analysis but could have also been enhanced if facilitators had been able to call on their expertise to answer questions which arose in the two final workshops.

A clear structure moving through discovery, deliberation and co-creation exercises addressed allowed plenty of time for information sharing and small group discussions

- **Case studies (on uses of AI in relation to eye conditions and for predicting dementia/ stroke risks) helped bring the topic alive** and demonstrated the potential long-term benefits of AI research using large imaging data sets. However, the examples may also have contributed to some participants (mainly older individuals with health conditions) continuing to think more in terms of AI for individual diagnosis/care decisions, rather than for large scale research.
- **Elicitation exercises - built around the 9 scenarios – proved helpful in surfacing and then testing underlying values.** Going through elicitation exercises nine times was rather repetitive, but as participants became more familiar with the material and the process the exercises proved a useful way of refining underlying values into a set of principles with clear trade-offs between them. Ultimately participants found the co-creation process in the final workshop satisfying.
- **The roving role of ‘Chair’ played by the lead dialogue designer gave her a helicopter view to inform an adaptive design approach.** Freed from the role of small group facilitation, the Chair was able to move between small groups and observe what was working well, where the challenges were, and where reframing or amendments were needed. The insights gained fed into the design of later sessions based around the emerging principles. Seeing their insights used in this way helped engender confidence that these would be useful to the commissioners.

Sympathetic and professional facilitation created space for all to contribute

- **A pool of experienced and professional facilitators created a warm and welcoming atmosphere in the small groups and provided continuity between most sessions.** A clear design guide and support provided by assigned notetakers in each small group, left facilitators free to focus on involving all participants, keeping discussions on track, and probing the underlying reasoning and values behind participants’ views.
- **The encouraging role played by the Chair in closing each workshop helped to keep participants motivated and engaged.** In opening and closing plenary sessions the Chair reflected on what people had said, mirroring their language and acknowledging

the complexity and importance of the tasks. This encouraging tone helped keep participants motivated and we observed them becoming increasingly engaged with a topic that initially may have seemed dry and abstract.

- **The option to take time out in a neutral space if they found the topics emotional was only taken up by a few: but the thought was appreciated by all participants.**

Participants' views were captured, analysed and fully reported in a timely report

- **Allocation of resources for simultaneous transcription and rapid analysis immediately after each workshop allowed the Ipsos Chair to feedback the findings to participants and the commissioners in a timely fashion.** The emerging findings were able to inform a very detailed interim report which allowed OG members and the core project management team to make helpful comments. These fed into a high quality final report with a strong narrative: it required very few iterations. Quotes from participants expressing both majority and outlying views, and illustrating the values that underlay them, were used to good effect to put the participants' voices front and centre. As one OG interviewee remarked: "*The report does a really good job in drawing out what really matters to people and what implications this has for the different models.*"

4 Conclusions and emerging lessons for future dialogues

4.1 Conclusions

This small public dialogue was designed and delivered to a high quality within a very tight timescale. The design was stripped back to the essentials in the interests of meeting the commissioner’s deadline. The dialogue has fully met its immediate objectives and is already having the desired impact on Phase 1 of the NHS AI Lab’s follow-on research competition.

Participants were satisfied to have taken part in the process and OG members found it valuable to hear that the public was willing and able to grapple with the key issues and develop nuanced views on the different options. Both stakeholders and participants appreciated the importance of involving the public in this area and making the topic of data stewardship of de-identified health data more transparent. The ethos of public involvement is being carried forwards in the scoping and piloting phases of the research. Several OG member organisations involved with stewarding large health databases seem keen to apply the learnings from the dialogue, including through making their own (delegated) data/research access committees more participative.

In the medium term (late 2024 onwards) – assuming piloting of workable models results in proof of concept and the NHS is willing and in a position to apply the results to medical imaging data - the benefits for the NHS could be considerable. Greater transparency and trust in how data is used in AI research could result in higher opt-in rates, and improvements to diagnosis, treatment and research which could deliver cost savings and improved health outcomes for patients. Successful application in the NHS could have wider knock on effects on the development of AI research for the UK economy in general.

4.2 Lessons learnt and recommendations

For commissioners on project management and governance

- Recruit all contractors at the same time so that the independent evaluator can be in place and make formative inputs during the initial design stages of the project.
- Convene the OG soon after the inception meeting. Set aside time before the meeting to brief the chair and agree key inputs/decision points so that they will be most useful to the design and delivery of the project.
- Agree from the outset what is and is not possible in terms of segmented analysis of the findings: if disaggregated data is required for specific groups this will need to be consciously built into the design and analysis tasks (e.g. through grouping of participants, in-workshop surveys and coding frameworks).
- Build in at least one review point to assess overall progress with the Sciencewise specialist or evaluator.

Design and delivery for contractors

- **Scheduling and timing of workshops.** Consider how timings for online workshops can best work for all participants: in this case the mix of weeknight and one weekend session

worked well, but some shorter weekday evening sessions (2-2.5 hours instead of 3) and a longer weekend session might have helped hold the attention of all participants.

Ensuring all participants can share their views and feel supported

- Consider mixing small groups up between each session unless there are compelling reasons to keep some sub-groups together (such as a need for more segmented analysis). In this case mixing groups allowed all participants to benefit from exposure to the diversity of opinions and backgrounds across the wider group: this was also important in addressing the risk that a small number of strong personalities would steer the conversation in a particular direction.
- If participants are likely to find the topic emotional, make provisions for individuals that need to take time out from small groups. Consider ensuring that an empathetic listener is available (e.g. in a private chat room) and provide links to other sources of support in the participant pack.

Facilitator, chair and notetaker roles

- Where the budget allows, consider budgeting for a roving Chair/lead designer to overview all groups. In this case the Chair was able to use the oversight they gained to encourage participants, synthesise emerging views and adapt designs (reframing issues and pulling out themes to structure final workshops).
- Where the timetable for analysis and reporting is tight, consider budgeting for simultaneous notetakers in some or all small groups: in this case instant transcripts made it easier to identify emerging themes and build them into the design of a co-creation exercise in the final session.

Information shared with participants

- Where possible share printed materials (or digital versions) with participants in advance so they have them to hand during workshops.
- Ensure sufficient time during initial discovery workshops for participants to consider all they have heard and ask any outstanding questions so that everyone has the same understanding of the framing and key concepts.
- For a small, rapid dialogue where the topic is not too contentious, consider picking just a few specialists to deliver information and answer participants' questions. This can help ensure that participants receive just the right amount of well-pitched background, so reducing the risk of them feeling overwhelmed by too much detail or jargon.
- Where participants are expected to consider a large number of subtly different scenarios (e.g. in an elicitation exercise) make sure that how they are presented makes best use of images, colour and icons so that participants can easily distinguish between them.
- Where there is subject expertise within the core team, consider whether this could be called on as a resource by facilitators to help answer participants' questions in small groups. Separate briefing or training sessions may be required to develop the role and share tips on how best to talk to the public on an equal footing.

Annex A: Methodologies

Project management

- The core project management team of about seven included two individuals each from the NHS AI Lab, UKRI Sciencewise, the contractors plus the independent evaluator. The team met weekly during the scoping and field work stages and was joined by others (ODI and ICHP) at key points.
- A two-person NHS team convened and coordinated the OG, framed the dialogue, and provided technical input to information shared with participants. The team had delegated responsibilities to sign-off designs and the final report. Several NHS staff also contributed as specialists or attended workshops as observers.
- Ipsos, ODI and ICHP led on dialogue and materials design, recruiting participants, briefing specialists, running workshops, data analysis and reporting. ODI also coordinated a stakeholder group who met twice to feed into the design of materials and to share the interim findings.

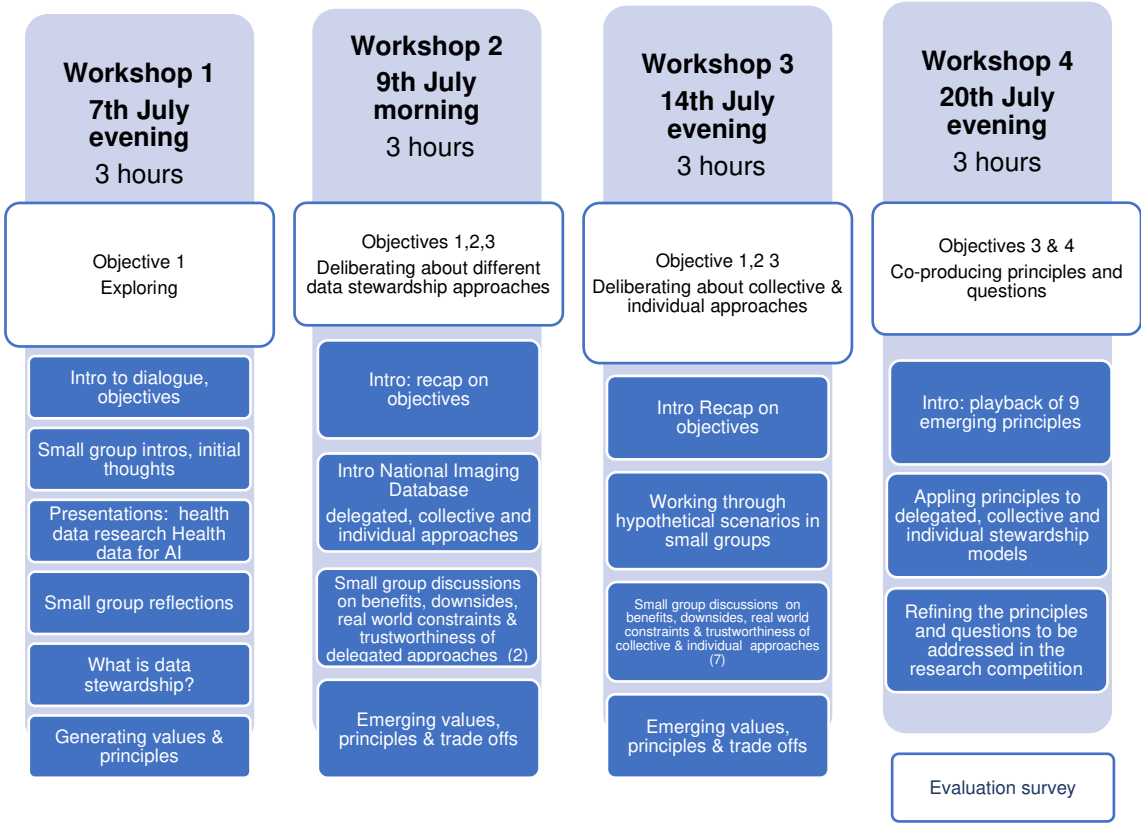
Project Oversight

- A group of about 14 individuals from health, data and ethics backgrounds oversaw the process (see *Annex B*). Natalie Banner (Genomics England) chaired the group which was broadly representative of public health practitioners, academia, government regulators, and third sector. Commercial AI or research interests which might be involved in the research competition were purposefully not included.
- The group met three times online. At the first meeting they helped inform the overall framing to meet the project objectives, the participant recruitment brief, and in suggesting scenarios for different data stewardship models. At the second meeting they reviewed interim findings and advised on key themes and structure for the final report. A final meeting focused on the key findings, their implications for the research competition and how findings might be disseminated. Individuals provided comments on a close-to-ready version of the final report. Half a dozen OG members contributed to the evaluation via interviews and/or an online survey. Three OG members volunteered to be part of the research competition assessment panel.

Dialogue design

- The dialogue was designed by Ipsos, with ODI and ICHP leading the development of stimulus materials based on their specialist knowledge of the data stewardship and health data fields, respectively.
- ODI drew on a stakeholder group (see *Annex B*) representing points of view – including data businesses – beyond the OG. The group of specialists and civil society representatives helped shape stimulus materials. The group was re-convened in mid-September to share the interim findings and explore how they might be disseminated or used in research.
- *Figure A.1* illustrates how four workshops (12 hours in total over the course of two weeks in July) built towards delivering the objectives.

Figure A.1: Key elements and coverage of the workshops



- Each 3-hour session was run online (Zoom), with two 10- minute comfort breaks on a mix of weekday evenings and a Saturday morning.
 - **Workshop 1**, introduced health data and AI research, including the National Data Opt-Out (NDOO) and the Five Safes framework for ensuring the protection of health data and used two case studies (eye scans and dementia/stroke risks) to prompt discussion on the potential benefits and risks of healthcare data for AI research.
 - **Workshop 2** introduced the NDDIC, high level information about delegated, collective and individual approaches and explored two delegated scenarios in detail.
 - **Workshop 3** – explored collective (4) and individual (3) scenarios in detail.
 - **Workshop 4** – presented criteria/principles which had been identified in previous workshops, tested them against the 9 scenarios, and refined them to apply to delegated, collective and individual models, respectively.

Hard copies of stimulus materials were shared in a participant pack, sent in advance of the workshops. The pack included the presentations by specialists and the theoretical data stewardship models explored in workshops 2, 3 and 4. Participants were not expected to complete formal homework tasks between workshops and this project did not involve a dedicated microsite for reviewing materials online.

Participant recruitment

- A specialist agency recruited 55 Participants (for a target of 50 with 5 reserves in case of dropouts for workshop 1) against quotas agreed by the core project team and OG to be broadly reflective of UK demographics (gender, ethnicity, age, socio-economic grade and locations) and attitudes towards how their data is used.
- The group also included a small number of individuals with lived experience of long-term chest-related health conditions such as asthma, bronchitis or Chronic Obstructive Pulmonary Disease (COPD).
- Anticipating that discussions about the NHS and chest conditions could be triggering so soon after COVID-19, the Ipsos team made provision for participants who needed to take time out from their small groups and talk to an empathetic listener (the Ipsos Chair).
- Each participant received a staggered incentive payment for attending all four workshops to remove any financial barriers to their participation.

Role of specialists

- Two specialists from Genomics England and the NHS Transformation Directorate gave introductory presentations on health data and how it related to AI and medical images (Workshop 1) and data stewardship approaches (workshop 2). They and the commissioners were also on hand to answer participant questions in workshop 2 in plenary and rotating around small breakout groups.

Analysis and reporting

- All plenary and small group discussions were transcribed by notetakers and raw data was organised into a thematic framework for analysis and reporting.
- The research team engaged with two patient representatives who reviewed the workshop materials and provided feedback on the draft principles before they were finalised.

Project outputs

- Between late August and mid-November the core team worked closely on several iterations of the interim report (shared with the OG for its final meeting in September) and the final report and executive summary.

Annex B: Oversight and stakeholder participants


Oversight group membership



Name	Organisation
Natalie Banner	Director of Ethics, Genomics England
Vicky Chico	Senior Privacy Officer, Office of the National Data Guardian
Mark Halling Brown	Head of Scientific Computing, Royal Surrey County Hospital
Jasmine Leonard	Technology Consultant & Advisor
Susheel Varma	Head of AI & Data Science, Information Commissioner’s Office (ICO) (now left)
Sinduja Manohar	Senior Manager, Public and Patient Involvement and Engagement, Health Data Research UK
Phil Booth	Coordinator, medConfidential
Joseph Savirimuthu	Senior Lecturer in Law, University of Liverpool
Sophie Brannan	Senior Policy Advisor (Medical ethics), British Medical Association
Kira Allmann	Public Engagement Researcher, Ada Lovelace Institute
Dr Margaret Charleroy	Head of Strategy for CIDC, NHS Transformation Directorate
Joseph Watts	Head of Cross Government Data Strategy, NHS Transformation Directorate
Ruth Keeling Laurence Thorne	Former Head of Data Strategy, NHS Transformation Directorate

Stakeholder group members

Name	Organisation
David Stone	Managing Director, Kaleidoscope Consultants
Jessica Morley	Director of Policy, Bennett Institute for Applied Data Science, University of Oxford
Nikita Japra	Senior Manager, Data & Society Accelerator Program, The Patrick J. McGovern Foundation
Ryan Dunn	Head of Data Science, DWP Digital Newcastle
Soujanya Sridharan	Research Analyst, Aapti Institute
Carly Kind	Director, Ada Lovelace Institute
Wen Wha Lee	Chief Executive Officer at Action Against AMD & Chair, Data Trust Advisory Board at HDRUK INSIGHT Hub
Miranda Marcus	Global Mental Health Databank Product Lead, Wellcome Trust
Tim Davies (reviewed material offline)	Research Director, Connected by Data
Kasia Odrozek (reviewed material offline)	Director, Mozilla Insights

Annex C: How the dialogue objectives have been met and what has contributed

Objective	Status	Comments on how and why objective have been met
<p>Engage diverse and broadly reflective section of the public in exploring different models of data stewardship</p>		<p>This objective was fully met with a diverse mix of about 47 participants – including those that brought lived experience of relevant health conditions – attending all sessions.</p> <ul style="list-style-type: none"> • The participant mix reflected the agreed recruitment brief and was therefore reflective of the UK population with strong representation of individuals from Black and minority ethnic backgrounds and age groups including young people. Participants demonstrated a broad range of starting knowledge about AI and data stewardship and underlying attitudes towards data security, privacy, commercial uses and trust/concerns for the NHS which shaped their deliberations. • The pace (intensive two weeks before the summer holidays), welcoming environment and exposure to people from different backgrounds (by constantly re-mixing small groups of 6-7) created a lively atmosphere. Staggered incentive payments probably helped keep participants involved (47 of the target 50 - 94% retention rate). • As the journey took them from thinking as patients (workshop 1) to citizens discussing wider social and ethical issues (workshop 2 onwards) most participants became very engaged. • Participants unanimously agreed that they understood the objectives (reiterated at the beginning of each workshop) and how the findings would feed into the research competition. By the end, participants shared their satisfaction that they had helped to co-produce useful principles for developing future data stewardship approaches. • Almost all felt it is important to involve the public in this type of research. A number told us that they really valued the opportunity to learn about the NHS AI Lab team’s consultative approach and to help influence NHS policy and practice in the longer term. • Many also reported that they had found the health data topic of personal interest. They talked in their evaluation feedback, of how they valued the learning opportunity and how they felt that individualised approaches would give more people opportunities to understand their health data: <i>“The information given about the different types of data usage was highly valuable, especially the individual decision-making.”</i> • In a few cases participants also reported that by learning more and deliberating in their small groups they had changed their own opinions about sharing their health data. A number of individuals told us, for instance, that they had rethought their position on the national data opt-out scheme: <i>“I got lots of info and I feel I can [now] opt in if I know people with opposing views have scrutinised the decisions going forward.”</i> • The final dialogue report was enriched by describing the diversity of participants’ views: this gave the commissioners and OG members confidence that the findings are robust. The topic was new to many but many brought their experience of sharing data/setting preferences in other online settings.

		<ul style="list-style-type: none"> Although the sample included individuals bringing a perspective as patients with relevant chest conditions, they were mixed in with the general sample and no separate analysis was undertaken of whether their views were different from those of the wider population.
<p>To understand attitudes towards different models, and which models are trustworthy</p>		<p>This objective was fully met. Subject expertise embedded in the team (ODI and ICHP) and inputs from the OG and a stakeholder workshop made it possible, within a challenging timetable, to design sessions and materials which gave participants a broad enough grounding that they were able to identify what would be important in making different models for data stewardship (the status quo, collective and individual approaches) trustworthy.</p> <ul style="list-style-type: none"> A uniform approach to describing the 9 hypothetical scenarios gave participants enough information to identify the pros and cons without constraining their ability to suggest how they might work. All small groups explored all 9 models: this was repetitive but generated enough evidence to describe the pros and cons of each in the final report, while stopping short of identifying a preferred option (which was never the intention). Most participants were able to move from thinking as patients to citizens after the first workshop, but it was clear that some continued to think of AI being used to make treatment decisions rather than for research. Despite the chair and facilitator's best efforts to reiterate the research context, evaluation feedback and observations shows that a minority continued to think this way and it may have coloured their perceptions of collective and individual approaches and particularly the emphasis they placed on having medical experts involved vs individual choice. By the end of the dialogue many participants favoured a 'mix and match' approach bringing together what they considered the best elements of different models (e.g. a mix of experts, public and patients, and opportunities for retrospective review).
<p>To understand the values and principles that underlie dialogue participants' views on data stewardship models</p>		<p>This objective was very well met. Transcripts of small group discussions in workshops 2 and 3 were analysed before workshop 4 to identify 9 underlying values and principles¹³. Workshop 4 successfully turned these into a set of co-produced principles by replaying them to participants and getting each small group to test them as criteria for assessing a sub-set of scenarios. This process wasn't intended to lead to a consensus on a preferred scenario: instead, it teased out meanings to different groups and, with minor amends for clarity, resulted in co-produced principles that NHS AI Lab and OG members considered useful for future research.</p> <ul style="list-style-type: none"> The exercises also made clear that individuals' preferences between models depended on the weight they placed on each value (e.g. those valuing individual choice and transparency preferring individual approaches while those valuing experts and feasibility/cost for the NHS preferring collective approaches). The values in

¹³ expertise, transparency, feasibility, choice, accountability, consistency in decision-making, inclusivity, completeness of the data and speed of research happening.

		turn reflected their views on how trustworthy they considered the NHS or researchers. No clear trends linking these values to factors such as age, education levels or ethnicity emerged.
To guide the AI Ethics Initiative in prioritising its next tranche of research funding based on the findings	Mostly	<p>This objective has already been partly met. The interim report identified emerging principles (objective 3) and started to identify some cross-cutting issues: these have been further drawn out in the final report and are quoted verbatim in the draft research competition brief. The brief also describes the 9 theoretical models but stresses that there is scope for bidders to propose ‘mix and match’ approaches. The OG and stakeholder group praised the richness of data underlying this analysis, the principles and cross-cutting themes and some of the research questions to ensure that the research competition findings (due 2024) will be practical and policy relevant.</p> <ul style="list-style-type: none"> • However, some OG members would like to see the research brief describe existing data stewardship models (such as IGARD), reference existing research standards (e.g. ICO suggested its Guidance on AI and data protection and Explaining decisions made with AI) and to draw out more specific questions relating to the cross-cutting issues to be explored through the research.
To inform the data stewardship approach to national [medical imaging] databases based on the findings.	Not yet	<ul style="list-style-type: none"> • It is too early to assess this impact. The extent to which the NHS is able to take on board findings from the research competition will not become clear until mid-2024 onwards. Interviewees so far feel that some of the findings are generalisable to other settings and a few suggested areas are summarised in <i>Section 2</i>.

Annex D: Assessment of how the design and delivery has met good practice

	Evidence
Retention of an engaged mix of participants	<ul style="list-style-type: none"> The 47-50 participants in 8 small groups (6-7) reflected the recruitment brief with a good mix of locations, a wide age range and good participation from Black and minority ethnic background individuals. Mixing small groups allowed for cross-fertilization of views and to manage dynamics between dominant, quieter and a few less focused individuals. A few (mainly quieter) individuals said they would have preferred to stay in one group throughout. By the final workshop, most participants were fully engaged (with the facilitation team having followed up offline with those who appeared to be experiencing tech problems, were distracted, or not actively participating).
All participants treated with respect and their support needs catered for	<ul style="list-style-type: none"> A quiet room and the opportunity to have a chat with an empathetic listener was provided for those who needed time out. A few individuals did become upset during workshop 2 but did not need to seek external support. By workshop 3, participants were starting to build on each other's points and as the exercises were repeated conversations started to feel more natural, with everyone chipping into the assessment process rather than individually making their points via the facilitators. Participants reported that they mostly felt comfortable expressing their views, while recognising that some others had a tendency to be overbearing. <i>"Some participants had an "I know what's best, this is how it's always been, you're stupid if you disagree" attitude, however, otherwise I felt mine and others' views were respected and appreciated overall."</i>
Participants provided enough unbiased information to feel informed and able to contribute	<ul style="list-style-type: none"> Participants mainly agreed that they received enough information (28, 87% about right) while a small minority (4, 13%) felt it was too much. Most also felt that the level of detail was about right. PowerPoints included some legal and medical jargon, but mainly well-explained in everyday language. Several noted that having the information in advance was helpful, but many found it <i>"a lot to absorb"</i> before talking materials through in small groups. At points several said they felt a bit overwhelmed. Participants often referred back to their hard copy information packs and we saw them drawing on key concepts (such as the Five Safes framework, de-identified and pseudonymized data and the National opt-out) in their small group discussions. By the end of the four workshops the majority of participants felt well enough informed to contribute useful opinions: <i>"After several discussions with different sets of people together with the facilitators I felt I was able to make contributions to the topic."</i> Just one participant still felt ill-equipped, noting that <i>"I feel if people like me were involved in making any kind of decision it could be disastrous for many."</i> As noted above, some participants were confused about the different uses of AI and the case study on eyes (which talked about using AI for diagnosing conditions) may have contributed to this (by suggesting that machine learning would replace doctors). Despite the best efforts of the Ipsos chair and facilitators to bring discussions back to AI for research, some continued to talk about personal health care decisions until the final sessions: this framing may have had a knock-on effect on the emerging values/principles with perhaps more weight placed on making sure that medical expertise would be involved in every data access decision and some participants feeling that individuals would not be equipped to make decisions themselves. <i>I suspect there would have been value in having more of a worked example presentation from an AI practitioner about how they get, use and report on their use of data to build AI models. (stakeholder observer blog).</i> Scenarios Presentation of the nine scenarios followed the same format and covered the same information. One-page summaries (used in both hard copies and on screen) were quite dense and lacked visual clues (numbering, colour coding, icons, images, analogies from other contexts) that might have made it easier to distinguish between them.

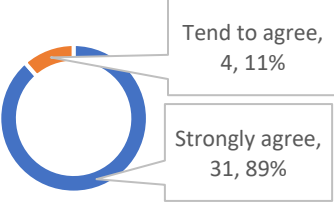
	<ul style="list-style-type: none"> • Almost all participants felt that the scenarios of potential data stewardship models were helpful in thinking through the wider implications. However, the presentation was dense and wordy: slides would have benefitted from more images, icons, colour coding and analogies from non-AI settings to help distinguish between key features (especially during Workshop 3 when they reviewed 7 case studies over two breakout sessions). A tendency for facilitators to read out the slides verbatim (even though participants had them in their packs and on screen) in some cases led to time pressures and the need to move on mid-discussion so that all options got covered.
<p>Workshops structured to deliver dialogue objectives</p>	<ul style="list-style-type: none"> • The balance of time spent sharing information in plenary vs in small group discussions and focusing this on elicitation exercises worked well. • Repeating the exercise of reviewing different data stewardship scenarios nine times in each small group was rather repetitive but it proved important that all individuals reviewed every scenario in the same order, as this mirrored the participant journey. As they worked their way through delegated approaches (the status quo, workshop 2) to collective and individual approaches (workshop 3), almost all participants' views evolved from finding the status quo broadly acceptable to feeling there would be benefits in terms of agency, transparency, accountability etc. from more participative approaches. • In workshops 2 and 3, while exploring potential benefits and drawbacks of each scenario, participants quickly started to deliberate about how approaches would work in practice; and how far they addressed societal concerns about choice, transparency, accountability, and inclusivity. • Generic values that emerged from workshops 2 and 3 were first presented as single words (such as transparency) and then used as criteria to assess a smaller sub-set of scenarios in each group. The vagueness of the criteria was initially difficult for participants to assess and most groups struggled to eliminate or add new concepts. But the approach gave them space to tease out what each meant, make the language more precise or identify where more information was needed, and allowed them to explore trade-offs between principles. • Revisiting the scenarios for a second time in workshop 4 (where they were rated red, amber or green) also generated more nuanced views on the relative strengths and weaknesses of each: these insights have proved useful in shaping the follow on research.
<p>Participants had enough time and opportunities for meaningful deliberation</p>	<ul style="list-style-type: none"> • Three quarters of participants found the timings (3-hours with two 10-minute comfort breaks) about right overall, but a significant minority found the length and early start time and late finish (6-9 pm) too long on a weekday evening. The lack of time to eat may have contributed to concentration problems we noted amongst a few participants, particularly in the final half hour of evening sessions. • A mix of times (say some 2.5 hours for evening sessions and 4 hours for weekend sessions) could have worked well in this case. One OG member interviewed felt that workshop 2 (on a Saturday morning) where participants were still being asked to take on a lot of new information would have benefitted from more time for them to digest what they had heard and ask questions from specialists. If they had more time to get their questions answered and clarify any misunderstandings (such as whether the discussion was about using AI for making health care decisions) it might have avoided confusion in the later sessions. We noted a handful of participants never quite grasped the difference and tended to stray back to talking about personal health care for the remainder of the dialogue. • Mixing up groups each session helped to manage a few individuals who tended to dominate and allowed quieter individuals (including a few younger people from Black and minority ethnic backgrounds) more space to express their views, adding to the overall richness of findings.
<p>Professional and independent facilitation</p>	<ul style="list-style-type: none"> • The role of a lead designer/workshop chair was really effective: within each session the chair introduced the objectives, ways of working and played back findings from previous workshops, and then rotated around the small groups as an observer, allowing them to summarise and set up the next session at the end of each workshop. This approach: <ul style="list-style-type: none"> • demonstrated that participants were being heard (using their own language). • acknowledged where topics or tasks were difficult and kept people motivated. • examples of how insights informed the structure of the final workshop gave confidence that their views were important.

	<ul style="list-style-type: none"> • allowed small groups to hear what others in the room were thinking. <i>"...wonderful chair for explaining everything so clearly and concisely. Thank you!"</i> • We saw – and all participants agreed – that facilitation was mainly professional and effective. We saw facilitators bringing in all participants (including a few that seemed a bit distracted), bringing them back on topic, probing their underlying reasoning, and helping them express thoughts as principles. <i>"They were very good at answering questions and interpreting the meanings of phrases or ideas. Also analysing our answers and teasing out our ideas."</i> A few participants would have liked facilitators to nudge quieter participants to involve them more. • Several participants noted that facilitators had been empathetic and good at teasing out meanings of what participants said. One commented that facilitators <i>"did an excellent job of bringing a diverse cross-section of the public together, educating them and most importantly keeping everyone on topic."</i> • However, in a few cases we observed facilitators with a tendency to talk too much or over participants and on one occasion pressing the group to a view by constantly repeating a question (in this case on remuneration for taking part in collective stewardship models): this annoyed the group and left them less time to discuss other scenarios. • In another case a participant would have welcomed a more encouraging approach: <i>"Facilitators could clarify points in the gaps within discussions, rather than during someone's point- to not discourage a member of a group"</i>
<p>Participants were able to interrogate specialists and get their questions answered</p>	<ul style="list-style-type: none"> • Given the compressed timescale it proved efficient to only ask two specialists (Genomics England and NHS in workshops 1 and 2 respectively) to present information to participants and answer their questions. This enabled the specialists to be well briefed, the overall design and discussions to flow well and prevented participants being overwhelmed with unnecessary clinical/technical detail. • PowerPoints presented in plenary (health data, data sharing and security, and AI) were accessible and prompted many questions during workshop 2 on general (not specifically medical imaging) data security questions which AI Lab answered during workshop 3. Wider perspectives were incorporated via case studies and scenarios which drew on wider stakeholder inputs. • Participants appreciated the role specialists played describing them as <i>"full of knowledge and professional"</i>, while one appreciated that they <i>"allowed me to hear thoughts outside of my own and provided clarity."</i> None told us that they had not heard broad enough perspectives. • Core team specialists (ODI, ICHP and NHS AI Lab) also rotated around small groups, but mainly in silent listening mode and we did not see any answering questions in workshops 3 or 4. With suitable pre-briefing they could have played a more active role in answering participant questions in their areas of expertise. • Nevertheless, almost all participants agreed that specialist contributors to the workshops had been helpful in answering their questions and helping them think through the issues, with one participant describing them as "full of knowledge and professional" and another appreciating that they <i>"Allowed me to hear thoughts outside of my own and provided clarity."</i>
<p>Analysis and reporting</p>	<ul style="list-style-type: none"> • During the reporting stage it became clear that the commissioners would have appreciated more segmentation in the analysis of findings (such as patients with lived experience as heavy users of the NHS, or of groups traditionally under-represented in medical research such as individuals from Black and minority ethnic communities). However, this would have needed to be identified and built into the design from the outset (e.g. in the sample framing and by keeping those with similar characteristics together throughout). The small size of the sample does not really support such analysis. Furthermore, some of the richness of public dialogue in describing participant's journeys as they negotiate different views would have been lost. As a compromise, full transcripts of discussions in all small groups were shared with the commissioners.

Annex E: Evaluation feedback from participants

(Feedback via SurveyMonkey completed by 35 out of 47 participants)

<p>I understand why NHS AI Lab wants to understand public views on access to research data for Artificial Intelligence (AI).</p> <table border="1"> <tr><th>Response</th><th>Count</th><th>Percentage</th></tr> <tr><td>Strongly agree</td><td>29</td><td>83%</td></tr> <tr><td>Tend to agree</td><td>6</td><td>17%</td></tr> </table>	Response	Count	Percentage	Strongly agree	29	83%	Tend to agree	6	17%	<ul style="list-style-type: none"> • <i>Lab wants to gain trust from public.</i> • <i>Great to have been involved- I might opt in now. As long as the AI is tested equally among different skin types/colours otherwise I feel my data would benefit others more than myself despite providing similar levels of information. Inclusivity = Trust.</i> • <i>The medical data belongs to us as individuals and therefore strongly agree that the public should be involved.</i> 						
Response	Count	Percentage														
Strongly agree	29	83%														
Tend to agree	6	17%														
<p>The quantity of information I was sent before the workshops felt:</p> <table border="1"> <tr><th>Response</th><th>Count</th><th>Percentage</th></tr> <tr><td>About Right</td><td>28</td><td>87%</td></tr> <tr><td>Too much</td><td>4</td><td>13%</td></tr> </table>	Response	Count	Percentage	About Right	28	87%	Too much	4	13%	<ul style="list-style-type: none"> • <i>Enough to prepare me for discussion.</i> • <i>I read through it before I started the research and it gave me an indication of what the workshops were about.</i> • <i>At first too much until it was went through on zoom and I could get my head around it.</i> • <i>I didn't get sent anything in the post, but had it emailed instead.</i> • <i>I did not have any information in the post.</i> 						
Response	Count	Percentage														
About Right	28	87%														
Too much	4	13%														
<p>The level of detail in the information I was sent before the workshops felt:</p> <table border="1"> <tr><th>Response</th><th>Count</th><th>Percentage</th></tr> <tr><td>About Right</td><td>28</td><td>85%</td></tr> <tr><td>Too much</td><td>4</td><td>12%</td></tr> <tr><td>Too little</td><td>1</td><td>3%</td></tr> </table>	Response	Count	Percentage	About Right	28	85%	Too much	4	12%	Too little	1	3%	<ul style="list-style-type: none"> • <i>It did give headings to cover the content of discussion.</i> 			
Response	Count	Percentage														
About Right	28	85%														
Too much	4	12%														
Too little	1	3%														
<p>The balance of time spent learning about the topic and discussing the issues felt about right.</p> <table border="1"> <tr><th>Response</th><th>Count</th><th>Percentage</th></tr> <tr><td>Strongly agree</td><td>16</td><td>46%</td></tr> <tr><td>Tend to agree</td><td>11</td><td>31%</td></tr> <tr><td>tend to disagree</td><td>7</td><td>20%</td></tr> <tr><td>neither agree or disagree</td><td>1</td><td>3%</td></tr> </table>	Response	Count	Percentage	Strongly agree	16	46%	Tend to agree	11	31%	tend to disagree	7	20%	neither agree or disagree	1	3%	<ul style="list-style-type: none"> • <i>I learnt more from Zoom than on my own due to it being read out and guided by people who could answer questions.</i> • <i>I strongly agree as I learnt a lot about AI</i> • <i>3 hours too long.</i> • <i>2 hours a session instead of 3.</i> • <i>It could have spent longer to go deeper into discussion.</i>
Response	Count	Percentage														
Strongly agree	16	46%														
Tend to agree	11	31%														
tend to disagree	7	20%														
neither agree or disagree	1	3%														

<p>The facilitation has been professional, independent, and effective.</p>  <p>Tend to agree, 5, Strongly agree, 30,</p>	<ul style="list-style-type: none"> • <i>Great facilitators and wonderful chair for explaining everything so clearly and concisely. Thank you!</i> • <i>They were very good at answering questions and interpreting the meanings of phrases or ideas. Also analysing our answers and teasing out our ideas.</i> • <i>Facilitators did their best in each group and tried with different situations to encourage thoughts - some more effectively - but all made an effort with those who spoke less in the group settings.</i> • <i>Very professional but spoken in lay terms.</i> • <i>Very professional [well] trained people.</i> • <i>Facilitators could clarify points in the gaps within discussions, rather than during someone's point-to not discourage a member of a group?</i>
<p>Facilitators made sure I felt comfortable expressing my views.</p>  <p>Tend to agree, 4, 11% Strongly agree, 31, 89%</p>	<ul style="list-style-type: none"> • <i>Totally agree.</i> • <i>They made you comfortable and welcomed.</i> • <i>Definitely felt comfortable most of the time.</i> • <i>They were empathetic and understanding to different viewpoints.</i> • <i>Some participants had an "I know what's best, this is how it's always been, you're stupid if you disagree" attitude, however, otherwise I felt mine and others' views were respected and appreciated overall.</i>
<p>Hearing from specialists has been helpful in answering my questions and helping me think through the issues.</p>  <p>neither agree or disagree, 2, 6% Tend to agree, 7, 20% Strongly agree, 26, 74%</p>	<ul style="list-style-type: none"> • <i>Allowed me to hear thoughts outside of my own and provided clarity.</i> • <i>Specialists were full of knowledge and professional</i>
<p>The case studies on potential data stewardship models were helpful in thinking through the wider implications.</p>  <p>Tend to agree, 13, 37% neither agree or disagree, 1, 3% Strongly agree, 21, 60%</p>	<ul style="list-style-type: none"> • <i>Could see how the information was relevant.</i> • <i>The case studies helped me understand the potential data stewardship.</i>

<p>I think it is important that the public gets involved in helping to inform this type of research and practice.</p> <table border="1"> <tr> <td>Strongly agree, 26, 76%</td> <td>Tend to agree, 6, 18%</td> <td>neither agree or disagree, 1, 3%</td> <td>Tend to disagree, 1, 3%</td> </tr> </table>	Strongly agree, 26, 76%	Tend to agree, 6, 18%	neither agree or disagree, 1, 3%	Tend to disagree, 1, 3%	<ul style="list-style-type: none"> • <i>Of course, the public play an important part in all this.</i> • <i>Yes, for transparency reasons involvement of public is necessary.</i>
Strongly agree, 26, 76%	Tend to agree, 6, 18%	neither agree or disagree, 1, 3%	Tend to disagree, 1, 3%		
<p>By the final workshop I felt sufficiently informed to make a useful contribution to the topic.</p> <table border="1"> <tr> <td>Strongly agree, 27, 79%</td> <td>Tend to agree, 6, 18%</td> <td>neither agree or disagree, 0, 0%</td> <td>tend to disagree, 1, 3%</td> </tr> </table>	Strongly agree, 27, 79%	Tend to agree, 6, 18%	neither agree or disagree, 0, 0%	tend to disagree, 1, 3%	<ul style="list-style-type: none"> • <i>Definitely and hopefully I could answer others' questions with a balanced viewpoint.</i> • <i>After several discussions with different sets of people together with the facilitators I felt I was able to make contributions to the topic.</i> • <i>I feel if people like me were involved in making any kind of decision it could be disastrous for many.</i> • <i>It was a learning project for all participants.</i>
Strongly agree, 27, 79%	Tend to agree, 6, 18%	neither agree or disagree, 0, 0%	tend to disagree, 1, 3%		
<p>I feel confident that the NHS will take our opinions into account in informing their research. (% confidence - mean=72)</p> <p>The scatter plot shows individual data points for confidence levels, ranging from approximately 10% to 100% on the x-axis and 0 to 150 on the y-axis. The mean confidence is 72%.</p>	<ul style="list-style-type: none"> • <i>[without public involvement] I would not want more of my data shared, which I currently allow.</i> 				

What, if anything, was most valuable about the public dialogue workshops?

Important for shaping NHS approaches and gaining trust

- *It is the way forward to get public involved to gain trust of public – a little bit more time spent would have been beneficial.*
- *I think it has been very important in shaping the future of the NHS and the involvement of AI.*
- *Engaging general public on values in relation to public health data to be used for research.*
- *The transparency to the public.*
- *Getting a public view.*
- *The fact that there were no wrong or right answers.*

Learning more about a topic with both societal and personal implications

- *Knowing more about the subject and the objective of the whole workshop.*
- *The fact that we were involved in a generally new but still personal topic.*
- *The information given about the different types of data usage was highly valuable, especially the individual decision making.*
- *Got lots of info and I feel I can opt in if I know people with opposing views have scrutinised the decisions going forward.*
- *Learning about research.*
- *It was very informative overall.*

- *Information that I didn't know, I learnt a lot.*
- *Learning about different options.*
- *Learning about the challenges of AI health data*

Listening to the views of a wide cross section of the public

- *Different opinions*
- *The interaction in chat rooms*
- *Hearing everyone else's views on the subject.*
- *Being able to hear different viewpoints.*
- *I really enjoyed hearing everyone's views and various opinions.*
- *Input from non tech ppl.*
- *Understanding the differing viewpoints.*
- *Talking to everyone.*
- *It was good to see everyone's individual understanding.*

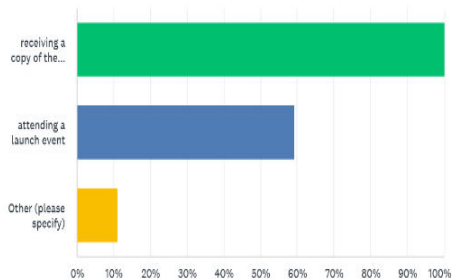
Suggestions for what might have been done differently

18 comments from 35 participants of which 7 were to say that they would not have changed anything and three suggested shorter sessions.

- *I feel it might have been better being with the same people throughout*
- *It was structured correctly and timely*
- *Differ the variety of people in each chatroom*
- *It would have been nice to ask certain candidates about their health statements e.g. someone who said that diabetes affected his foot suddenly and another man saying that he is now old has not had many challenges. I would have liked to ask him how he managed that.*
- *I think you did an excellent job of bringing a diverse cross section of the public together, educating them and most importantly keeping everyone on topic.*
- *Facilitator nudging silent participants - a bit lacking in session 4.*
- *Maybe meeting up physically*

Please let us know if you'd like to be involved in any of the following ways:

Answered: 27 Skipped: 8



- *Attending the launch of an NHS AI system.*
- *Further access panels to give opinions, I really feel a part of the structure overall now.*
- *Be a panellist.*

Additional comments on the content of the dialogue

- *I felt it was going to cost far too much money, far too much work & at the end of the day it's up to the Consultant and patient to decide/discuss what will happen. For the cost this would cost, I'd much rather the money be put straight to the NHS.*
- *Whilst I can see it may be helpful, I honestly just can't see it happening. Everyone should have a say in their care, but I guess I'm old fashioned and I'd much rather discuss everything about me with the consultant. No two cases are the same but the Consultant knows best (that's his job). I guess what I mean is a bit like too many cooks spoil the broth.*
- *The public's involvement might add to more specialist information - research data being gathered in midwifery why there is miscarriage in pregnancy and other genetic disorders as there are still some gaps in data.*