Putting Good into Practice

A public dialogue on making public benefit assessments when using health and care data

A Findings Report
Hopkins Van Mil
April 2021

Contents

Foreword

A profound shift is underway in how data is collected, managed and used across the health and care sectors. As we move to more digitised and joined up services, data gathered as part of providing our health and care can be stored in formats that make it more useable and useful for purposes beyond individual care. These range from informing the planning of services to research and innovation in search of new insights and treatments.

At the same time, using data generated from people's interactions with the health and care system carries inherent risks both to individuals and to society, raising questions about privacy, equity, exploitation and fairness. The devastating Covid-19 pandemic has raised the prominence of data use in public discourse – emphasising both the enormous potential but also the risks.

Ensuring health and care data are used in ways that benefit the public is therefore a critical safeguard. By centring the purpose of data use on the benefits it delivers back to patients and service users, the health system or society more broadly, we can hope to tip the balance in favour of the great advantages of using data, and away from those uses that pose the risk of harm. Previous research has emphasised that people expect public benefit to be driving decisions about health data use.

However, the notion of "public benefit" is itself not clear. What counts? What is fair? What different factors need to be weighed up and how, when deciding if a proposed data use counts as being for "public benefit"? These are complex issues that many decision-makers grapple with, from university ethics committees to data oversight groups. We commissioned this project to help inform their thinking, by putting the question to members of the public and bringing their views and values to the fore. We also brought social care data into these discussions as a vital and often overlooked part of the picture.

Despite the complexity of the core question, our public participants did not disappoint. Our sincere thanks to Hopkins Van Mil, the project oversight group, expert contributors and public participants for their enthusiastic contributions to an intensive process that has brought out a set of rich insights.

It is clear from this dialogue report that the public expect more transparency over how health and care data are used and how decisions are made, at a systemic level. There is a strong appetite for this information and a sense of urgency to use data well if, and only if, these transparency needs are met. The public are worried about the potential for data to be misused or manipulated to suit an agenda, and see transparency across the lifecycle of data use as important for mitigating these risks. This points us towards the need for more co-ordinated action across the health and care ecosystem as we look ahead beyond the pandemic.

We are proud that this report seeks to uphold the legacy of Dame Fiona Caldicott, the inaugural National Data Guardian, who sadly passed away shortly before this work could be published. Throughout her career she championed the voices and rights of patients in how data about them is secured, managed and used.

This report is testament to her conviction that good decisions about data are built on trusting people to handle complex information and listening to what they say in return. The Office of the National Data Guardian is now developing statutory guidance informed by the insights from this report, in keeping with this clear commitment from Dame Fiona.

Dr Nicola Byrne
The National Data Guardian for
Health and Social Care

Dr Natalie Banner
Understanding Patient Data Lead,
Wellcome

Executive Summary

In the Putting Good into Practice public dialogue we find that participants are supportive of health and social care data being used for public benefit. The headline findings, to inform policy advice or guidance to be created by the National Data Guardian (NDG), are divided into three categories:

1. Prerequisites for public benefit

Transparency cannot be separated from public benefit. It is not an add-on or nice to have. Health and social care data use requests *only* demonstrate public benefit if they have integrated communications within their application including activity which demonstrates the value of data use to society

To demonstrate public benefit, transparency is required throughout the whole data life cycle (collection, storage, assessment and use), not just at the point of application

Public benefit is undermined if authentic public engagement is not integrated into data assessment. This requires engaging people from a cross-section of society in data assessment processes.

2. Areas that matter most to dialogue participants

Equitable distribution of benefits of data use in health and social care with safeguards to protect against discrimination and geographic disparities

Identifiable and sensitive data should be treated with the utmost care, if it is, it has the potential to bring public benefit. Data was perceived as being particularly sensitive if it is of a personal nature, such as genomics or mental health data, or because greater care is needed in its interpretation, such as qualitative data

Safeguards and provisions in place to protect society from data manipulation, where the outputs from the data use could be interpreted in different ways, for example, to achieve political or financial ends. This includes publication of statements of data users' credentials and sources of funding

Public benefit must outweigh profit with profitable uses of data rigorously scrutinised for demonstrations of public benefit before access is granted. There is a recognition that data use in this context can enable health and social care improvements and innovations

Being ambitious for health and care data use - to realise public benefit from global collaboration; exploratory research driving breakthroughs; and using profit for new developments, such as drugs, treatments and services.

3. Areas that matter least to dialogue participants

Did not feel that the data use needed to remain close to the original purpose of its collection to bring public benefit. They were more concerned about the relevance of the data - should it be used beyond the purpose of its original collection. However, all changes in direction must still be is predicated on the prerequisites for public benefit being in place

The scale of benefits is not a significant factor in determining whether a data use has public benefit as there is inherent value in data use which produces an impact, even if only for a small number of people.

4. Participant expectations

By the end of the process participants had set out two sets of expectations for public benefit assessments when using health and social care data. The first are expectations of public benefit assessment, the second relate to wider expectations of public benefit assessment systems and infrastructure:

1. Public benefit assessment expectations

- Once a clear definition of public benefit is developed **use a case-by-case approach** to public benefit assessment **so that the definition can be tested** against potential unforeseen consequences or harms to individuals and society
- Ensure that **safeguards and protections** are in place so that sensitive data continues to be protected; and the process doesn't exclude certain sections of the population from gaining benefit
- Enable a **balanced level of governance** which is not overly restrictive whilst ensuring data access requests are in line with public benefit, including ethical factors and with safeguards in place
- Data applicants should be required to have **a clear purpose** for their programme of data access and use; however, **this purpose can evolve** as the programme develops to embrace **unexpected findings** and **exploratory purposes**
- Any major alterations in purpose or scope should trigger the data applicant to **go back to the data assessor** as a minimum to acknowledge the change or to seek approval to proceed on the basis of the new focus for the work
- Ensure that the data assessment process has measures in place so that any new data access requests **build on research that has preceded it** and do not re-invent the wheel
- Include in the data assessment process details about **the time factor** – the length of time the data can be accessed; what data users do if the technology improves or changes whilst they are working on the programme (and therefore the project needs to change)
- All data assessment should include an extended **risk assessment** to consider short-term harms against the longer-term benefits (including for future generations).

2. Wider systems and infrastructure expectations

- Don't limit the role of the data assessor to exclusively reviewing data use requests, **the experience and knowledge data assessors have should be used for reviewing the whole research/ innovation/ planning life cycle** to monitor and audit good practice

- Encourage the health and adult social care sectors to **standardise the tools for data collection**, including encouragement and support for the adult social care data collection to digitise and become a more effective data resource
- Ensure that **safeguards and protections** are in place so that data can't be manipulated for political or financial gain
- Publish clear **statements of data users' credentials** and **sources of funding** to protect against data manipulation, potentially non-altruistic motivations and hidden agendas
- Engage in genuine activities to **make the full cycle of data collection, assessment, use and impact transparent** to guard against public mis-trust
- Put steps in place to give assurances on **data quality** determined by factors such as accuracy, representativeness, size of the data set being proportionate to the expected outcomes, method of collection and the impact of data gaps
- Embed **authentic public engagement activities** in the **data assessment and publication process** this could include, for example:
    - A data assessment jury to be drawn on for complex 'edge' assessment cases with, for example an ethical dimension
    - Create a publicly accessible database of approved projects which have gone through the assessment process which can be understood and reviewed by those with no specialist knowledge of the subject
    - Communicate widely, in places and language accessible to the wider population, the impact data use is having on research, innovation and planning for health and social care.
- **Give priority** in data assessment to those who have agreed credentials for demonstrating public benefit
- **Provide a supportive environment** for data use applicants, for example, if their access request is declined data assessors could give constructive feedback including **possible partnerships with other applicants**, suggesting **alternative more appropriate data sets** or giving **guidance on how public benefit might be better achieved** to ensure that great ideas aren't lost because of a mis-judged application.

5. About this report

The Putting Good into Practice public dialogue was commissioned in March 2020 by the National Data Guardian for Health and Social Care (NDG) and Understanding Patient Data (UPD) in partnership with Sciencewise and UK Research and Innovation. It was designed and delivered by the deliberative engagement specialists Hopkins Van Mil and independently evaluated by 3KQ.

Health and social care provision affects everyone in society. Data use is key to research, innovation and planning. But there are a range of views on how data should be used, for what, and by whom. As such Putting Good into Practice aimed to allow the NDG and UPD to understand how people assess public benefit in the use of health and adult social care data for purposes beyond individual care. It set out to explore views on complex questions such as - what if:

- A use of data benefits some groups of people, but not others?
- The research will not benefit the people whose data was used, or their families, but instead people who are distanced by geography and or time?
- It is uncertain what the results of a piece of research might be?

Two practical outputs will result from the public dialogue:

1. The public dialogue findings report and, drawing on these findings,

2. Policy advice or guidance issued by the NDG to support those making public benefit assessments.

This report is the first of these outputs. It is divided in to 2 parts:

Part A: Scene setting	Which sets out the how we designed and developed the process. It will be of particular interest to those wishing to understand the detail of the public dialogue methodology.
Part B: Findings	Which speaks to what we found having discussed these issues through the course of the process.

An accompanying Annex provides the:

1. Desk research document list
2. Process plans for each facilitated workshop
3. Stimulus materials used throughout including the case studies referenced in this report.

In testing understanding of what people consider to be beneficial about the use of health and adult social care data for purposes beyond individual care we found Putting Good into Practice public dialogue participants both endorsed and amplified what is known. They also brought a range of perspectives and new insights on public benefit to inform the NDG's guidance.

Public dialogue is a qualitative methodology, as such the findings in this report do not demonstrate statistically representative analysis. In this way, we present the subtleties and nuances of participants' views, concerns, hopes and aspirations so that they can richly inform the next steps in the development of the NDG's policy guidance or advice.

The public dialogue engaged 112 participants recruited in a 50 mile radius from four locations: Great Yarmouth, Stockport, Plymouth and Reading. Each participant attended five dialogue events, a webinar plus 4 workshops. Dialogue activities were designed purposively for online participation, with groups of 28 in each location supported to work collaboratively, mostly in small group discussions, on Zoom. Reflective tasks were completed in participants' own time in a dedicated online space. Throughout the process participants had the opportunity to interact and discuss the topic with policymakers and specialists in data use in health and adult social care settings.

Part A: Scene Setting

1. Introduction

The Putting Good into Practice public dialogue was commissioned in March 2020 by the National Data Guardian and Understanding Patient Data with support from Sciencewise and UK Research and Innovation. It was designed and delivered by the deliberative engagement specialists Hopkins Van Mil.

1.1 Public dialogue partners

The National Data Guardian (NDG) role was created in November 2014 to be an independent champion for patients and the public when it comes to matters of their confidential health and care information. The NDG advises and challenges the health and care system to help ensure that citizens' confidential information is safeguarded securely and used properly. The NDG has statutory powers to issue advice and guidance about the processing of health and adult social care data in England.

Understanding Patient Data (UPD) aims to make the uses of patient data more visible, understandable and trustworthy. UPD seeks to explain how and why data can be used for care and research, what's allowed and what's not, and how personal information is kept safe. We work with patients, charities and healthcare professionals to champion responsible uses of data.

Sciencewise is an internationally recognised public engagement programme which enables policy makers to develop socially informed policy with a particular emphasis on science and technology. Sciencewise helps to ensure policy is informed by the views and aspirations of the public. The programme is led and funded by UK Research and Innovation (UKRI) with support from BEIS.

Hopkins Van Mil (HVM) facilitates engagement so that voices are heard, learning is shared, and understanding achieved. In practice this means finding the process by which people can explore their hopes, fears, challenges and aspirations for the future. HVM's work enables stakeholders, technical specialists, and a diversity of publics to work together as equals to make actionable, better informed, and powerful decisions.

1.2 Project aim and scope

Putting Good into Practice aimed to allow the NDG and UPD to understand how participants assess public benefit in the use of health and adult social care data for purposes beyond individual care. Its purpose was to produce two practical outputs. This public dialogue findings report and, drawing on these findings, NDG policy advice or guidance to support those making public benefit assessments.

The project objectives were developed in collaboration with the project partners and Oversight Group[1] members and became **the scope** for the dialogue. They were to:

1. Test our existing understanding of what people consider to be beneficial about the use of health and adult social care data for purposes beyond individual care (e.g. research, innovation, planning)

2. Explore how people weigh benefit and disbenefit of the use of data generated from publicly funded health and care, taking into account a number of factors such as:
    a. The identity of those who benefit or are disadvantaged
    b. The scale of any benefit and disadvantage
    c. How close the benefit should be to the original purpose of the data collection
    d. The likely prospect of the benefit or disadvantage
    e. The impacts of the type of data used on public benefit assessments

3. Explore how far the attitudes vary on 1 and 2 when social care data is being used.

The NDG and UPD were keen to explore views on the use of social care data as well as health data. And to look at some more complex questions: what about when a use of data might benefit some groups of people, but not others? What if it's quite uncertain what the results of a piece of research might be? What if it won't benefit the people whose data was used or their families, but instead people who are quite distant, perhaps in another country, or if any benefit is not expected to be available for many years? How do these factors affect people's attitudes towards whether there is enough public benefit to merit data being used? Our work therefore focused on these questions throughout.

This scope was our guide as we designed and developed the public dialogue. It allowed us to focus our process on relevant aspects of health and social care data in relation to public benefit to develop the project partners' understanding of participant views, hopes, perceptions, and challenges. However, as this report makes clear, participants did not limit themselves to the scope in their discussions. There are some aspects of public benefit in this context which they needed to expose before being able to focus in on the detailed aspects of the dialogue. We highlight throughout the report where the findings fit within the scope, and when they are these pre-cursor points.

1.3 The shifting context of 2020

2020 was a year like no other in the social and economic shifts caused by the global pandemic. The COVID-19 crisis shone a light on health and social care in ways which were new to everyone in society. The work on the dialogue project started on 24th March 2020, the day after the first national lockdown was instigated. The dialogue played out against the context of events, media reporting and developments in data and health and social care (figure 1) which were the backdrop to the public dialogue workshops preparation and delivery. These points undoubtedly informed participant thinking as they considered the use of health and adult social care data beyond individual care.

Participants were encouraged throughout the dialogue to imagine a time before the pandemic began affecting daily lives, and to consider a future where COVID-19 is a memory rather than a current reality. In this way we tried to ensure that the findings are not simply a snapshot of

---

[1] The full list of Oversight Group members is listed in Chapter 7 Acknowledgements.

people's views in a time of crisis, but rather a measured and thoughtful response to the dialogue questions around the benefits and harms of accessing health and social care data for purposes beyond individual care in the long-term. We also emphasised the importance of having a practical output from the dialogue in terms of policy guidance or advice to be issued by the National Data Guardian.

COVID-19 announcements

- 2 national lockdowns and the tier system
- Stay at home, protect the NHS, Save lives campaign
- Government & CMO/ CSO briefings
- Increased citizen awareness of data use
- Hands face space campaign

Health and social care

- Escalation of COVID-19 cases in care homes
- Data sharing to support the identification of clinically vulnerable people for shielding & support
- Media reporting on pressures on the NHS
- NHS COVID-19 app and track and trace
- Vaccine developments

The National Data Guardian for Health and Social Care

Annual report 2019-2020

National Data Guardian

- Annual report published including the priority to support public knowledge and understanding
- Launch of the public consultation on the Calidicott Principles and Caldicott Guardians

Figure 1: The public dialogue context

1.4 How to read this report

This report distils the findings from the public dialogue. In the next chapter in Part A we describe the process through which the public dialogue was designed and delivered including the stimulus used, the specialist expertise provided throughout, and detail on the public dialogue participants and the locations from which they were recruited. Part B of this report starts with how we built on what is already known about public attitudes towards the use of health and social care data. We also talk about how attitudes changed throughout the process from early reactions to final reflections on the subject.

From chapter 5 onwards we concentrate on the main dialogue findings by describing attitudes towards the use of data generated through our daily interactions with the health and social care systems. We focus on the core elements of the scope set out in section 1.1. Chapter 5 ends with points made by participants on transparency and trust which, in their view, are a prerequisite to data assessment which fulfils the obligations of public benefit. The report ends with the expectations participants shared when considering public benefits, and a summary of wider system and infrastructure factors which the NDG, the NHS and social care organisations may find valuable in developing and implementing the NDG's policy guidance.

1.3.1 Interpretation of findings

The programme was initiated just as the first lockdown began. Stakeholder interviews, roundtable PPI interviews and the full dialogue process were therefore designed and delivered online. The dialogue fieldwork resulted in over 40 hours of audio recordings from the Zoom workshops which were transcribed for analysis using NVivo software together with:

- Data from the reflective tasks that participants completed in between each workshop
- Results of the online polling questions used live during workshops
- Transcripts from the interviews undertaken for a participant film produced as part of the dialogue outputs.

HVM applies standard principles for social science reporting when reporting on the findings from a public dialogue process, particularly Sciencewise Guidelines for Reporting (July 2019). Throughout the process the experienced HVM coding, analysis and writing team have maintained a rigorous approach to recording and held frequent sense-checking sessions to mitigate against researcher bias. Public dialogue is a qualitative methodology, findings do not demonstrate statistically representative analysis and they are drawn from the data rather than trying to prove pre-determined hypothesis. In this way we present the subtleties and nuances of participants' views, concerns, hopes and aspirations so that they can richly inform the next steps in the development of the NDG's policy guidance or advice.

We use terms such as 'a few', 'many', 'several' or 'some' to reflect areas of agreement and difference. These should be considered indicative rather than exact. Where views only apply to one location we make this clear in the text, we also explain when a vocal few present an argument as this can be indicative of the strength of feeling around a particular issue.

It is important in any dialogue process that the report reflects the voices of participants. As such, we have used quotations taken from transcripts to emphasise main points. Some quotes have been edited to remove repeat or filler words. There have been no other edits, which might distort the meaning intended by participants.

2. Methodology summary

The Project Team included representatives from UPD, the NDG, Sciencewise, UKRI and Hopkins Van Mil. An Oversight Group[2] was established to provide challenge and guidance to the project. It comprised academics; health and social care data and technology specialists; data assessors; and health and social care policy, ethics and lay advisors. The Oversight Group met 4 times over the course of the process. They helped to shape a comprehensive approach to the technical and practical issues associated with the assessment of health and social care data, working with the Project Team to ensure that the project stimulus materials, including specialist presentations, were balanced, presenting a range of views on the subject across both health and social care data.

The dialogue brief had been shaped by the NDG and UPD after a lengthy process of consultation and reflection, long before the onset of COVID-19. The programme follows on from the Citizens' Jury run previously by the National Data Guardian on when it is reasonable for patients to expect patient data to be shared[3]; and by Understanding Patient Data on what constitutes 'fair' when health and care data is used for purposes beyond individual care[4]. The NDG and UPD also held consultation workshops in advance of the public dialogue. It was essential that this deliberative project built on this extensive body of work. We therefore conducted a rapid desk research programme in April 2020, reviewing 43 documents including the reports from public dialogue and Citizens' Jury programmes; academic papers; current data assessment guidance; health and social care data use guidance (Document 1 in the stand alone Annex provides the full list of sources). We followed this up with 15 stakeholder interviews[5] to ensure the design of the programme was informed by existing expertise and to indicate the direction of travel for event speakers.

The programme was independently evaluated by 3KQ, led by Helen Fisher. 3KQ attended all dialogue events, attended Project Team meetings and sought input from the Oversight Group by means of interviews held outside the meetings. These constructive formative inputs and summative findings have helped to ensure the dialogue follows best practice Sciencewise principles[6].

2.1 A deliberative process

Before setting out our approach in detail it is worth reflecting on why the public dialogue approach fulfilled the needs of the project. Public dialogue is not a 'we tell you this and you tell us what you think about it' information exchange. Dialogue works when participants interact on a level playing field with specialists: academics, scientists, and policy makers and shapers. In this dialogue these included data assessors, data use applicants and those using the results of the data use for research, planning and innovation. This specialist evidence is then viewed through the lens of participants' lived experience which leads to rich and powerful insights.

In a public dialogue citizens come together with sufficient time to reflect, to:

- Learn about the issue

---

[2] A full list of Oversight Group members is provided in Chapter 7 Acknowledgements
[3] National Data Guardian, Connected Health Cities, Reasonable Expectations (April 2018)
[4] Hopkins Van Mil, Understanding Patient Data, Ada Lovelace Institute Foundations of Fairness (February 2020)
[5] A full list of interviewees is provided in Chapter 7 Acknowledgements
[6] Sciencewise Guiding Principles (July 2019) and Sciencewise Quality Framework (August 2018)

- Talk with, not past, each other
- Consider diverse points of view
- Discover key tensions and values
- Spark new ideas

This leads to an understanding of what people value, what they see as benefits and harms, their trade-offs and redlines and, in this case, highlights areas of importance when assessing health and social care use in line with public benefit.

Dialogue uses expert facilitators. In this case each online workshop had 4 facilitators, including a lead facilitator. We used a consistent group of facilitators and this number of team members, plus technical support, allowed us to have small groups of no more than 7 participants. Such ratios allow trust to build and a greater depth of exploration of the issues. Facilitators followed workshop process plans designed in discussion with the Project Team and in consultation with the Oversight Group and other stakeholders we interviewed. The process plans and stimulus were tested and refined in pilot workshops in September 2020 with 12 recruited participants.

2.2 The recruitment radius

The dialogue was originally conceived (at the proposal writing stage) as a face-to-face dialogue. A clear rationale was established for selecting four dialogue locations which was retained when the dialogue became an online process. We also wanted to bring the dialogue to new locations, where dialogues on health, care and/ or data use had not taken place before. The recruitment radius spanned out from a specific location, for example, Great Yarmouth, for a radius of 50 miles to take in participants from across the area.

Stockport:
An example of health and social care integration. Proximity to Greater Manchester Connected Health Cities Programme.

Great Yarmouth:
A significant older population and challenges faced by health and social care integration.

Plymouth:
A good example of health and social care integration.
Significant areas of deprivation.

Reading:
A mix of affluent and lower income recruits as well as a range of ethnicities.

Figure 2: Public dialogue locations

2.3 Who took part?

To shape the dialogue, we held 4 roundtable discussions, each of which comprised a 1-hour workshop separated by a few days before the second workshop. The 40 participants for these sessions were recruited by means of existing Patient Public Involvement (PPI) groups reached via announcements in People in Research, UseMyData patient advocates, James Lind Alliance and National Voices networks. These sessions were held in June 2020 so that they could inform the design of the dialogues.

Crucial to the inclusive nature of public dialogue is the way participants are recruited and retained. 28 people were recruited to each of the four dialogue locations and were retained throughout. A specification and screener were used to ensure that as far as possible, participants reflected the demographics of the population of England, sampling for age, ethnicity, gender, life stage, disabilities and socio-economic group. We excluded those who had taken part in deliberative type activities in the previous 12 months. Participants in the dialogue were given a cash honorarium to recognise the time committed. This is standard in Sciencewise public dialogues and means people are not excluded because of their financial circumstances.

The intention of the recruitment specification was to ensure that we over-sampled for people with direct experience of social care services, particularly as there has been little research on social care data use attitudes previously (see Chapter 3). This was a delicate balance in recruitment as we wanted to ensure that people were able to discuss social care as well as health data use meaningfully, but we were aware that those with a lot of exposure to the health and social care sectors might tend to see the benefits of data access for purposes beyond individual care more readily. Facilitators were briefed to challenge an overly-positive attitude to data use as it arose in order to mitigate this risk.

Potential participants also answered attitudinal questions on their current perceptions of data use in very broad terms. We asked:

Here are some of the ways in which data about you is collected every day.

- Store cards/ loyalty cards
- Social media platforms such as Facebook or Instagram

On a scale of 1-5 (where 1=extremely concerned and 5=not at all concerned) please state how concerned you are about your data being collected in these ways?

These questions were used to understand the spread of views on data collection and use rather than as a reason to exclude or include participants in the dialogue.

2.4 What did participants do?

Health and social care provision affects everyone in society. Data use is key to research, innovation and planning. But there are a range of views on how data should be used, for what, and by whom. As such it was essential to the process that participants were given evidence, information and resources to discuss the public benefit aspects of health and social care data use in ways which would help them to explore possible benefits, harms, impacts and scale of data use effectively. We explored what this evidence and information should be in the planning stages of the programme. The desk research, stakeholder interviews and roundtable discussions all indicated that knowledge of health and social care data for purposes beyond individual care would be low. It was felt that the dialogue could get side-tracked into too many issues that were out of scope such as an expectation that personal health records are integrated, or that profit is the only motivation for commercial data access requests. This foundation research suggested that dialogue participants were likely to have an even lower awareness of the existence of social care data use. For these reasons we were advised to design a process which would:

- Start from individual care as the route to collecting data for purposes beyond that
- Ensure that the different landscapes of care and health provision are understood
- Ensure that participants understood that systematic social data collection and use has been less prevalent than in health, and that there was now a drive to improve the situation through investment and learning from the growing pains of health data
- Make it clear that data assessment is not a simple, nor a unified, process.

The COVID-19 context for the dialogue delivery was reflected on by everyone involved in our planning research. There was a widespread view that awareness of health and social care data will have increased as a result of the crisis. It was felt there would be a greater awareness of data use beyond individual care, and possibly a greater scepticism of data validity if the policies put in place to protect the population are not seen to be working.

2.4.1 The workshop content

5 sessions (1 webinar and 4 workshops) were held in each of the 4 dialogue locations. Table 1 sets out the discussion points for each session.

Table 1: The content of the dialogue events

Webinar: Context 1 hour	**Workshop 1: Explanations 2 hours**	**Workshop 2: Explanations**
• Introductions to the team and settling into the subject • Vox pop introductions introducing the project partners and dialogue purpose and jargon busting • Presentation on the scope and practical output • Using the chat for questions, comments and initial reactions to the subject.	• Initial discussions on use of data for purposes beyond individual care • Presentations, films and handouts describing the multi-organisational health and social care landscapes • Presentations from data assessors: IGARD, CPRD, NHS Digital • Initial exploration of benefits and dilemmas raised by the contextual information.	• A discussion to explore views on public benefit • Presentations around three case studies: 1. Academic study on COVID-19 in care homes 2. Using AI to examine large retinal scan data sets 3. The National Cancer Registry • Discussion on the impacts on public benefit decisions of 'What if's' in these case studies • Presentation from the Health Foundation on data use applications • Discussion on opportunities for and dilemmas posed by data use

Workshop 3: Exploration 3 hours	Workshop 4: Culmination 2 hours	Reflective tasks in between workshops
• Case study presentations from: - NHSX Digital imaging Lab; - London Borough of Islington combining health and social care data; - Registered Nursing Home Association - Future Care Capital on the social care landscape • Discussion on the factors that might impact on public benefit such as identity, scale, prospect of achieving outcomes • Preparing for the final discussions.	• A reminder presentation on all the examples, scenarios and presentations given • Discussion on the issues raised throughout the dialogue process and asking participants to consider these issues as if it was November 2019 (before COVID-19) and in five years' time. • Devising specific considerations for data assessors • Participant presentations on the key considerations for the policy guidance or advice • Reflections back from NDG and UPD on what was presented.	Throughout the process participants were asked to go to a dedicated online space after each dialogue event. There they: • Were reminded of elements of the dialogue including the timetable, participant pack (sent in hard copy to everyone in advance of the workshops) • Reviewed the materials they had explored in the workshop, including re-watching filmed presentations • Read and commented on the points made by the small groups they weren't in • Read through answers to questions raised throughout the dialogue and added to the question list.

Examples of the detailed process plans and stimulus materials used throughout the dialogue are available as Documents 2 and 3 in the Annex to this report.

It is an essential part of all Sciencewise public dialogues that participants interact with specialists in the field. This exposes participants to a range of examples and perspectives on the issues. This dialogue had a broad remit to consider public benefit in the context of health and social care data use. It was critical to the programme that social care was integrated throughout as a previously under-researched area. Equally we needed to provide examples which could surface opportunities and dilemmas for participants around the scope points. We therefore worked with a range of specialists[7] to provide expertise in the form of presentations, answering questions, responding to comments with additional evidence, information, and a range of opinions.

2.5 The culmination of the process

The dialogue process does not end with the publication of this report. The dialogue findings will be used by the NDG and UPD to create draft policy guidance or advice. This document will be reviewed by a sample of the public dialogue participants from each of the locations working with stakeholders in a reconvened workshop to ensure that the guidance or advice reflects participant views before it goes out to wider consultation and publication in the summer of 2021.

---

[7] The full list of specialists is listed in Chapter 7 Acknowledgements

2.6 Data capture

Essential to this process is the capture of views to make sure the findings fully reflect the participant voice. All data collection remained robust throughout, despite being different from a face-to-face workshop series. Each facilitator recorded their own small group discussions, and the plenary discussions were audio recorded with the Chat being saved; the latter providing a useful source of real-time commentary in the participants' own words.

At each workshop facilitators took visible notes by sharing their screens whilst typing. This served the function of flip charts in a physical space where participants could amend what was written, review what they had discussed and prioritise key points made as required. As such, these were not part of the data capture process but were useful in understanding the points on which participants had placed particular emphasis and have been drawn on to some extent in our analysis. In addition, comments made in the dedicated online space were captured and analysed. Data capture points were transcribed and integrated in to the NVivo analysis.

Part B: Findings

3. Testing our existing understanding

Summary

We begin this report chapter with a focus on the first point in the project scope – to test our existing understanding of what people consider to be beneficial about the use of health and adult social care data for purposes beyond individual care. It provides a summary of the:

Findings of reviews of and consultation on public engagement activity by the NDG and UPD which sought to understand the factors that public engagement participants have considered in relation to data use to date

The desk research and stakeholder interviews used to inform the planning and design of this public dialogue

The chapter highlights:

The importance of including the previously under-explored area of public benefit from data drawn from adult social care settings

What is known about perceptions of public benefit to date

The dimensions of public benefit and the value propositions applied to it as shown through the desk research

The expectation before embarking on the dialogue that knowledge of health and social care data for purposes beyond individual care would be low

The first element of the dialogue scope was to test our existing understanding of what people consider to be beneficial about the use of health and adult social care data. It is important to highlight three points in relation to this. The dialogue:

1. Was not intended to explore with participants whether data collection and use should be carried out in health and social care - this has already been examined thoroughly in a range of studies (see Document 1 in the Annex)
2. Explored participant views on use of adult social care data use alongside health data use. Much of the existing public engagement had been concentrated on healthcare and little is understood of public attitudes towards what constitutes benefits and harms in the context of the use of adult social care data
3. Needed to give those who make data use assessments a clear view on the factors that the public may use to weigh public benefit, by building up a better picture of how people assess and weigh the public benefits and potential harms of proposed data uses.

We know from the work of the commissioning bodies in the planning for this public dialogue, including a review of previous public engagement work, that some factors have influenced whether participants consider that a use of data has a benefit for the public including:

Table 2: Testing our existing understanding – research to laying the foundations for the dialogue

Public benefit is a necessary but not sufficient criteria

- The widespread willingness to share health records for the public benefit is not unconditional
- Factors such as the trustworthiness of research organisations and the kind of data to be used are also seen as important
- Dialogues and surveys have tended to examine support for benefits to the NHS, rather than a more generalised conception of public benefit.

Commercial access (and gain) may be acceptable if there is public benefit

- But commercial gain should be secondary to public benefit. One study found that commercial uses that did not produce actual health benefits were unacceptable, regardless of any safeguards for the data, for instance even if the data were anonymous and aggregated
- Studies have identified that small but significant numbers of people would be against commercial companies accessing health data under any circumstances.

An equitable spread of benefits matters

- Public dialogues have identified support for both the widest possible public benefit and also explored some dimensions of public attitudes toward data usage which could provide a benefit and/or disadvantage for particular groups
- For instance:
    - where this may lead to improvements in care for small numbers of patients with rare conditions
    - Or to a reduction or change in service provision which might save NHS resources, but impact on a particular population.

The desk research conducted for the dialogue endorsed and amplified the points made above and are summarised in the following sub-sections.

3.1 Definitions of public benefit applied to data assessment

The desk research showed that public benefit had previously been explored almost exclusively in terms of the health rather than social care sectors. We, therefore, addressed this gap in this public dialogue. There were also seen to be gaps in terms of understanding public perceptions of the ‘end goals’ of data use. Improved health outcomes, reduced health inequality and independent living were seen by the public as crucial to realising public benefit. However, less is understood about the benefits that might come from data use beyond these goals such as the development of new treatments and technologies, improved efficiency of services and systems, an empowered and knowledgeable workforce. Desk research at this point revealed a factor which was important for our public dialogue participants – that public benefit is not a static concept. Studies such as the New Zealand Data Commons Project[8] describes it as a ‘feedback loop’ enabled by transparency and personal control over data which has to be sustained over time to avoid an erosion of public trust.

The dimensions of public benefit highlighted in the desk research include:

- Personal: e.g. improved health outcomes

---

[8] Mansell J et al, Data Commons Blueprint , 2017

- Social: e.g. a reduction in health inequalities
- Professional: e.g. greater individual and shared knowledge and expertise
- Economic: e.g. a growth in the life sciences sector; a more efficient NHS.

The value proposition of data initiatives was seen through the desk research as being:
- Efficiencies in service delivery
- Improvement in medicines and technologies
- Generating financial and social value through innovation.

Current data assessment criteria were reviewed as part of the desk research. We found there was an in-built assumption that ‘public benefit’ was understood by those assessing and making data access requests. We build on this in the report because, as we see in Chapter 4 (How attitudes to data changed) and Chapter 6 (A practical output) participants found this term to be highly subjective and difficult to pin down precisely. As one participant said in workshop 1.

> There is a grey area here between public interest and not public interest because you could argue, per researcher, it's quite subjective. | Participant, Plymouth

3.2 Perceptions of what the public dialogue would find

The research conducted in the design and planning phase of the public dialogue anticipated what we might expect to hear from public dialogue participants. In Figure 3 we see what roundtable PPI participants and stakeholder interviewees felt would be reactions to data use, and in some cases articulated how they felt themselves.

Figure 3: PPI Roundtable participant groups’ reflections on data use

Key points made by roundtable PPI participants in these discussions included:
- A lot of data is being collected:
    - Is it being used at all?
    - Is it being used well?
    - Is it accurate?
    - Will it provide a bigger picture that helps wider society?
- Concern that a lack of understanding of data use and data use benefits creates a culture of fear and alarm which prevents public benefit being realised
- Concern that not being transparent about data use does the same.

As we progress through the findings from the public dialogue we find that these points are echoed by participants. They are not all within the scope of the dialogue, but as we see in Chapter 4 (From overarching perceptions to shifts in attitudes), they need to be considered before public benefit can be established.

4. From overarching perceptions to shifts in attitudes

Summary

In this chapter we give a bird's eye view of how the thoughts and views of participants developed and changed during the dialogue process. We begin with their overarching perceptions of benefits and disadvantages where participants highlighted the value of data use for:

Research and development, including understanding population and regional health and social care needs

Planning and implementation, including matching provision to need

There was less focus initially on disadvantages or harms, but where they were raised, the focus was on:

- Data security, and
- The consequences of planning decisions made as a result of the data which might impact adversely on individuals.

We then move to reporting on the **early reactions** to the concept (which was, as the planning research indicated, new to many) of health and social care data being used for research, innovation and planning. We found that:

Most participants are very supportive of data being used for purposes being individual care

But at the same time they had many questions about the **safety and security of the data** and the **credentials of those who applied to use it** and **those who assessed the applications**

In the second section of this chapter, we follow the responses to a question we asked throughout the dialogue process: 'What comes to mind when you think about health and social care data?'. We show how participants' thinking evolved throughout the process via the following responses:

1. Definitions, questions and comments about privacy and other concerns
2. Reflections on the kinds of public benefit participants thought data use could yield and the process for assessing applications

By the end, when asked for a final message to the National Data Guardian when developing the guidance or policy advice for data assessors, participant thoughts had turned to:

3. Transparency to build more awareness and appreciation for data use and its potential for public benefit was uppermost in participants' minds.

4.1 Perceptions of overarching benefits and harms

Throughout the dialogue we gathered broad brush reflections on the perceived benefits and harms of data use, into which we would do a series of in-depth explorations to understand the subtleties and nuances of participants' thinking. The broad reflections are summarised here, before being explored in more depth in the rest of the chapter in the context of how attitudes to data changed.

4.1.1 Benefits

We found participants were more than willing to embrace the benefits of health and social care data use and to insist that the assessment process should acknowledge these. They spoke of data use being vital and those who expressed this view felt there could be no research without data use in both sectors.

> I just think the benefits are pretty self-evident, aren't they? It's all about weighing up the risks versus the benefits. It's about planning for the future, research, development of medicines, of social care, just from a research and technical point of view there must be massive benefits in that. **|** Participant, Reading

Research and development; planning and implementation were data uses for which participants could see intrinsic and specific benefit. Their overarching reflections on these benefits are summarised in Table 3.

Table 3: Overarching reflections on intrinsic benefits

Research and development
Understanding population level health and care needs
Finding cures, new treatments and therapies to improve health outcomes
Understanding regional and national trends in disease and care needs
Understanding pockets of illness or disease related to regions or communities to take preventative measures
Understanding the picture of carers supporting family members and what their needs are.

Planning and implementation
Planning for the future of health and social care
Evaluating health and care provision and making long-term improvements
As evidence for current funding allocations and to plan for future budget rounds
Scoping a consistent quality of care provision across the country
Taking preventive action for population wide challenges such as growing obesity levels.

> I would say it's all about progress. How do we progress unless we follow avenues? And how do we learn other things unless we stumble upon them? So, I think it's a win-win, needs to be done. And anybody at this time, we can't cure cancer overnight, but 50 years down the line, maybe, because we've had it, God forbid anybody. But that's progress, that's life and we have to follow it. **I** Participant, Reading.

4.1.2 Innovation

Innovation was the area of data use that participants said they felt they knew least about. But when they heard about the innovative use of artificial intelligence (AI) and machine learning, for example; or indeed the development of the COVID-19 vaccines for which data was being collected at the time of the dialogue, they felt it was of broad public benefit to use data in these ways. Even when participants felt nervous about a specific technology being used they nevertheless felt it would bring public benefit.

> I'm slightly suspicious of artificial intelligence, the idea behind it generally, but used for something specific like this, it just makes so much sense. It's such a good use of the technology. **I** Participant, Stockport.

4.1.3 Harms

As we have seen, many participants, particularly in the early stages of the dialogue, were far less ready to accept there might be any downsides to data collection and use. This might be because of boosted recruitment samples for people who had experience of the health and care systems, and the COVID-19 context.

> I'm still a bit confused as to what the risks of sharing data are as there seem to be so many positives to it. **I** Online space, Great Yarmouth.

However, participants did explore this throughout the dialogue.

In general terms participants' focus, in relation to harms, was, as in previous dialogues and research, on commercial gain when perceived as counter to public benefit objectives. Some participants also raised the potential harm of data being used which could mis-represent the 'real picture' either through mis-interpretation or because it is not in political or commercial interests to give the 'true' picture shown in the data. Other points raised in some locations and by some participants include:

- If data breaches occur, or if pseudonymised data can be linked back to the original data source and causes harm – particularly in social care related to child protection issues
- If the consequences of the data use are that a service is closed and in so doing harms are caused to (particularly already disadvantaged) individuals or groups

> Sometimes it feels like they provide a benefit to the public but on the proviso that something else is going to be taken away, which I see as an issue. So, if we identify this, this is going to be good for this, however it does mean we're going to restrict this service. So, it sometimes feels like there's always a consequence for it. **I** Participant, Stockport.

- For many the harms in not using data for research, planning and innovation purposes is far greater than any risk in doing so.

Participants were introduced to UPD's Spectrum of Identifiability[9] early in the process to help develop understanding that the focus for the dialogue was mainly (with exceptions) on de-personalised and anonymous data. However, it is worth noting at this point that for many participants it was very difficult to disentangle 'my' data from data use more broadly. It was hard to conceptualise big, aggregated data sets and move away from discussions which focused on individual care. Facilitators challenged participants throughout to remain within the scope of the dialogue discussions. By the final workshops we were clear that the focus was mainly on the assessment of requests to use de-personalised and pseudonymised big data sets, although participants also understood and accepted that where necessary access to personally identifiable data could be requested and granted. In section 5.6 we summarise views on the impact of the types of data used on the potential for public benefit, with some data being perceived as having a more fragile link to public benefit. There is also understanding that some data collection falls under the current legislative framework[10], but the use of anonymous data falling outside the legal framework has ethical implications which are explored.

4.2 Early reactions

Widespread agreement that the use of health and social care data for purposes beyond individual care was a good thing - and of public benefit - was expressed from the start of the process. Anonymity of data was a sufficient factor for many participants to express early support for data being used for research, planning and innovation that primarily benefited the public.

> If any data can help anything, especially if it's anonymised, for me, personally, there is no reason why it shouldn't go into purposes beyond my care. I Participant, Plymouth

When we asked participants the extent to which they had thought about data use beyond their own care, a quarter said they never had and over half said occasionally. In this context, it is not surprising then that many participants needed to explore basic needs, such as robust information and data governance, and the safety and security of data (in the style of Maslow's Hierarchy of Needs), before going on to explore the higher/ psychological needs of public benefit.

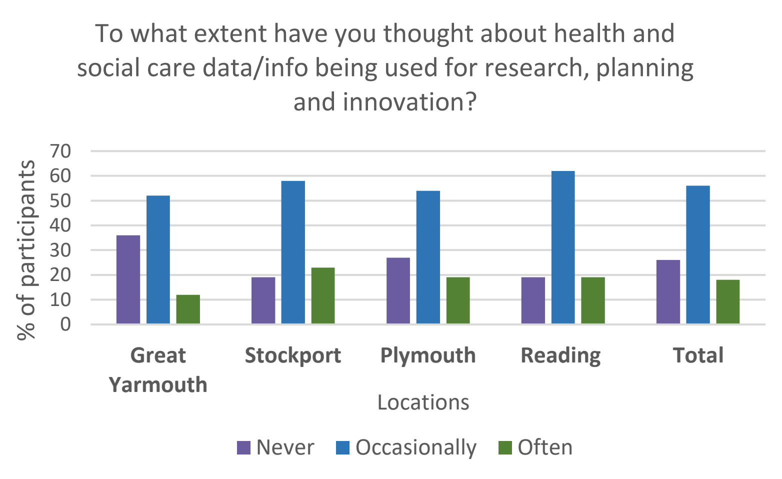


Figure 4: Extent of thinking about health/ social care data.

In line with the points made in the PPI roundtable discussions, frequent questions raised early on in the workshops and online space included:

- Who is data shared with?
- How is data accuracy checked?
- Is the data comprehensive (e.g. private care data; lifestyle data)?
- How is social care data collected?
- What are the qualifications/capacity of the data assessors?

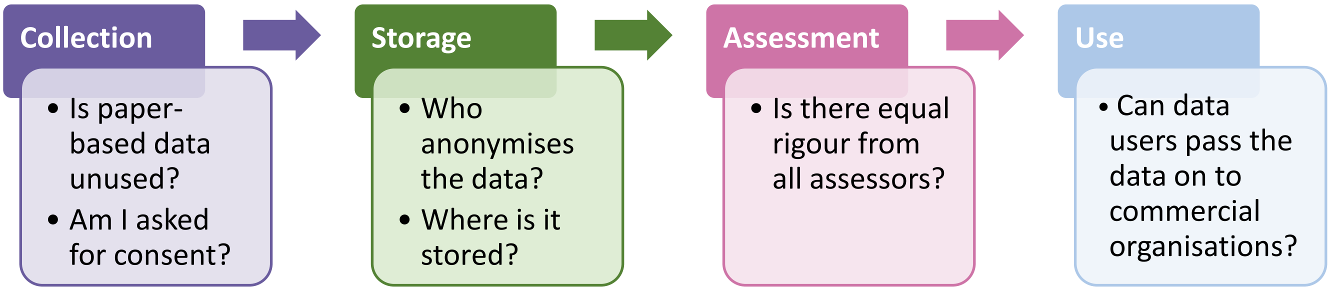
---

[9] https://understandingpatientdata.org.uk/sites/default/files/2017-07/Identifiability%20briefing%205%20April.pdf

[10] For example section 251 of the National Health Service Act 2006 or the Data Protection Act 2018

- Why aren’t the public more informed about data use?

But as well as seeking clarity on the nature and security of data, participants were exploring the implications for public benefit throughout the life and use of health and social care data:



Collection
- Is paper-based data unused?
- Am I asked for consent?

Storage
- Who anonymises the data?
- Where is it stored?

Assessment
- Is there equal rigour from all assessors?

Use
- Can data users pass the data on to commercial organisations?

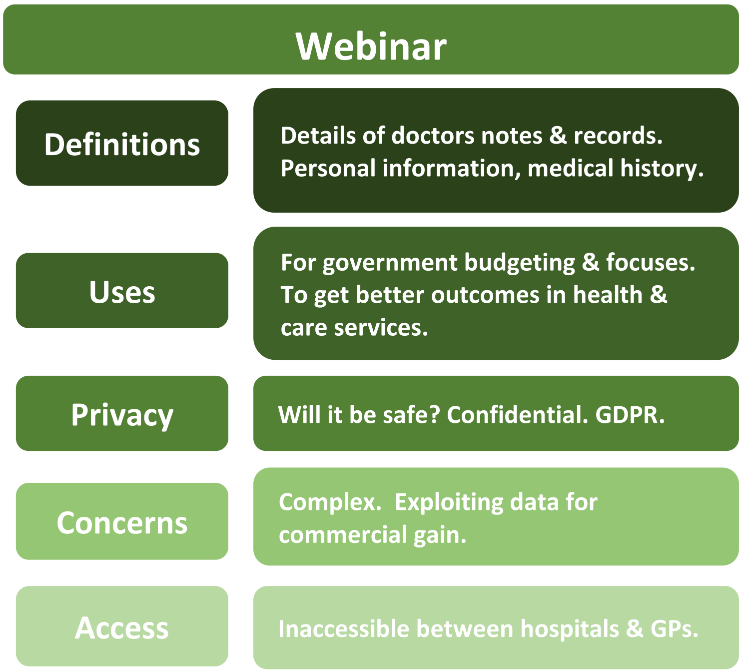
Figure 5: Implications for public benefit

Initial wariness about data use beyond their own care was felt by some participants to stem from their experience of:

- **Frustration** with poor data sharing within the health sector (e.g. GPs – Hospital) and between the health and social care sector for their own care.
- **Concern** that, in their experience, their own/family members’ health/social care data wasn’t accurate so how could benefit be drawn from flawed data?
- **Lack of knowledge/sight of their own data** to judge how beneficial it could be to society.

4.3 How attitudes to data use changed during the process

At each stage of the dialogue process, starting with the webinar, we asked participants, ‘What comes to mind when you think about health and social care data?’. The graphics below illustrate how what came to mind evolved over the course of the process. Each topic (such as Definitions, Uses and Privacy) directly quotes participant responses. The topic at the top of the graphic shows the most dominant participant comments: for example, in the Webinar, definitions and uses of data dominated what came to mind, with fewest comments around access.



Webinar

Definitions Details of doctors notes & records. Personal information, medical history.

Uses For government budgeting & focuses. To get better outcomes in health & care services.

Privacy Will it be safe? Confidential. GDPR.

Concerns Complex. Exploiting data for commercial gain.

Access Inaccessible between hospitals & GPs.

At the start of the process, factual definitions of what health and social care data is, and its uses, were at the forefront of participants’ minds.

Privacy and confidentiality were words frequently used. Some expressed concerns about complexity, commercial exploitation and government ‘snooping’.

There were also a handful of mentions about health and social care data not being shared between service providers for their individual care.

Figure 6: Webinar reactions to data use

By the start of workshop 2, thoughts about health and social care data had shifted predominantly to the benefits of its use and the process of granting data access, seen as robust by many.

Concerns put forward by a few participants included the misuse of data, e.g. by data users cherry-picking information to suit their intended outcome or agenda. As participants learnt more about its uses and heard the experiences of data applicants and assessors, some described the duality, the pros and cons, of data use:

> When used carefully - a huge potential for public benefit - but with a risk of unforeseen consequences. |
> Participant, Great Yarmouth.

Workshop 2	
Public benefit	**Opportunity for multiple areas of research, with the potential to make public services better.**
Process	**The lengths you have to go to, to get it.**
Concerns	**Manipulation of data.**
Pros & cons	**Dangerous and the potential for great public good.**
Transparency	**Public should be more involved.**

Figure 7: Workshop 2 reactions to data use

As workshop 3 began, most participants' comments were again about public benefit, but compared to previous workshops, we heard more emphasis on the importance of its use, words such as: vital, necessary, powerful, crucial. Definitions also became more expansive, to include research, planning, worldwide collaboration through data access and profitability in many areas.

Workshop 3	
Public benefit	**Positive for progressive research and global data access agreements.**
Pros & cons	**There is enormous benefit, it's really essential for progress, but the ethics are sometimes difficult to work out.**
Transparency	**There could be public good, but it needs to be clearer what those benefits are.**
Concerns	**Lots of considerations involved.**
Process	**Essential, appears well managed.**

Figure 8: Workshop 3 reactions to data use

We also heard more about the pros and cons of data use – with emphasis on the proviso that data could only be beneficial to the public if it is accurate and used with care. The few concerns raised tended to mention an appreciation of the complexity of deciding public benefit. As knowledge of and confidence in the process of data assessment grew, fewer comments on this were shared, but there was one request to:

> Put as few obstacles as possible. |
> Participant, Reading.

At the start of the final workshop, participants' responses to the question: 'What comes to mind when you think about health and social care data?' largely mirrored the themes from workshop 3. Descriptions of public benefit tended to be more future focused, perhaps as participants looked beyond their involvement in the dialogue to the legacy of data uses:

> Imperative for future generations. **|** Participant, Stockport.

> Progress for greater good. **|** Participant, Plymouth.

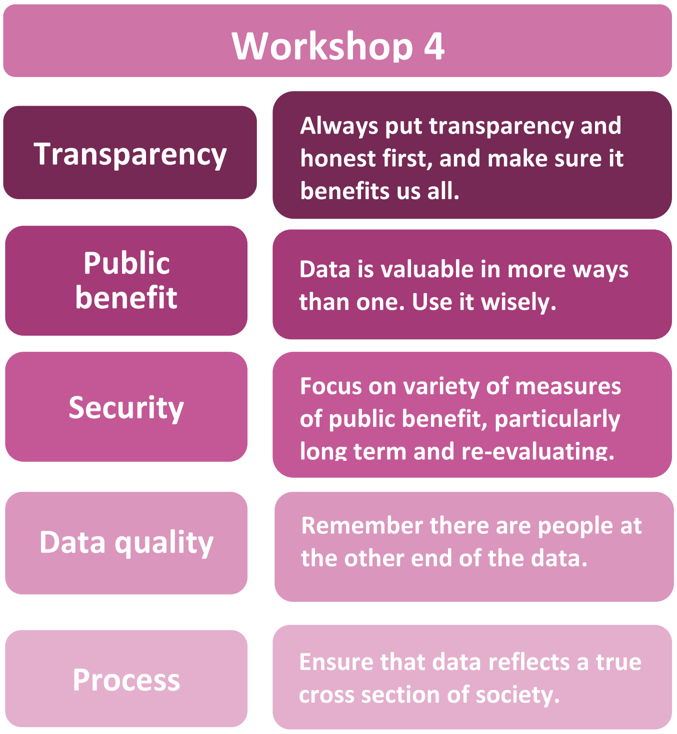


Figure 9: Workshop 4 reactions to data use

In the closing moments of the final workshop, we asked participants to give, 'One final piece of advice to the National Data Guardian in devising the policy advice or guidance based on these discussions'. Here we saw a significant shift in focus, as the advice most frequently put forward in all locations was for **transparency**,

> Data is precious and should always be treated with respect, ethics, positive outcomes, in the best interest of public benefit. Trust and transparency are essential. **|** Participant, Great Yarmouth

Advice about **public benefit** emphasised this as the priority for data assessors, with specific caveats around not just favouring data by the number of people it benefits and not being influenced by media, political or commercial agendas.

Comments on the **process** of data assessment were diverse and included: using standardised procedures, streamlining for regular data applicants, reviewing the process regularly – revisiting decisions as necessary, monitoring for public benefit throughout the life of data usage, not just the application stage:

> Don't just focus on the approval process. Follow the data and audit it as it is used, analysed and turned into outcomes. Use your own or independent auditors to assess whether public good is really at its heart at every stage. **|** Participant, Great Yarmouth

For a small number of participants **data security** and the **quality of data**, particularly in terms of its completeness, were still concerns that they wanted the NDG to hear.

5. Attitudes towards the use of data

Summary

In this chapter we highlight participant views on the key areas for consideration for public benefit data assessment. We demonstrate that participants:

Did not feel that the data use needed to remain close to the original purpose of its collection to bring public benefit. They were more concerned about the relevance of the data - should it be used beyond the purpose of its original collection

Were positive about **exploratory research** which they felt was worth taking the risk of finding nothing. Participants repeatedly referenced **major scientific breakthroughs arising from accidents** in support of this argument.

Felt that **re-collection of data in different contexts** would prove a **barrier to potentially beneficial research**, **innovation and planning** taking place

Felt there is **inherent value in data use** which produces an impact, even if only for a small number of people; as such they were less concerned about the **scale of benefits**, although cost-effectiveness was raised as an issue

Were keen to express their support for protecting the **identity of those who are disadvantaged**, with safeguards in place to protect against discrimination and marginalisation of vulnerable groups through use of data

Were concerned about the **scale and prospect of harms** and discussed the need for a **balance** between benefits and harms to derive public benefit from data use

The chapter unpacks how participants perceived the **impacts of type of data used** on public benefit, it explores perceptions of the:

- Use of personal data, with some data seen as **more 'fragile' in terms of public benefit**
- **Quality** and **accuracy of data**
- Lack of parity between the systems for the collection of **social care data** and **health data**, which might make the analysis of social care data less insightful currently

The chapter looks at how participants conceptualised **how data findings are used**, including across the research lifecycle and for profit. The chapter ends with participant views on **transparency** and **trust** as integrated with their definition of public benefit including:

- **Authentic public engagement** in the data assessment process
- The use of **transparent communications** in spaces well used by people throughout society
- Trust in data applicant's credentials and how studies are funded; and
- Concern about data manipulation and its impact on public benefit.

5.1 Proximity to original purpose

Across the board, participants were not concerned about how close the reason for using data was to the original purpose of its collection e.g. chest scans taken during COVID-19 treatment being used to explore treatments and innovations in understanding other lung conditions (see case studies used in Document 3 in the Annex). This was in the context of anonymised data discussed in the dialogue and highlighted in the stimulus materials.

Other factors were consistently felt to be more important in understanding and assessing public benefit. Some participants felt that the data routinely collected in the health and care system was a collective resource that could be used for public benefit, including for uses which may become apparent subsequent to data collection.

> Data that's been collected will be a benefit for future generations, and I believe that historical data that is perhaps accessible by researchers decades later, decades after it's been collected, can inform research and development into new medical solutions, so it's not all about here and now, is it? It's about keeping a safe data bank for future generations. **|** Participant, Reading

> I don't think it matters how close it is as long as the data that is being used is applicable to whatever study is being carried out. As long as it can be used in an effective manner, I don't see why it needs to be close. **|** Participant, Plymouth

There was also a sense among many participants that health and social care should be seen more holistically, with data collected in one context being an important resource to improving care elsewhere. Participants perceived that different areas of care (such as different departments or illnesses) were currently unhelpfully separated and felt there were opportunities to improve care by looking at the interfaces between different conditions or lifestyle factors such as diet. Linked to this, participants felt that re-collection of data in different contexts would be onerous and prove a barrier to potentially beneficial research, innovation and planning taking place.

The only aspect where some participants indicated a concern surrounding proximity to original purpose of data collection, was where data collected in non-health or care contexts might be used for medical purposes. For example, one participant gave an example of an idea to use mobile phone data to detect if people had developed a new and continuous cough, to understand the spread of COVID-19, which they were concerned could lead to technology being trained in one context being used in the future to ‘understand people's everyday emotions.’. However, this was outside the scope of this dialogue which focused on data routinely collected in the health and care system.

5.2 Identity

For the most part, participants showed a greater level of interest in the idea that data use would genuinely come to benefit people across society, rather than concerns about the specific identity of who would benefit. They did though, want to know which organisations and groups would have access to health and social care data and whether they gained benefit from it. When prompted to think about the identity of who could benefit or be harmed by health and social care data use, participants considered: national and local government, charities, pharmaceutical companies, the media, insurance companies, the NHS, people with health needs and those in social care settings.

Participants thought data use should benefit everyone equitably and there should be open access or equal access to the benefits. There was optimism in many discussions that a benefit for one would result in a benefit for all, as ultimately any benefit would trickle down through society by means of, for example, new treatments, improved service provision, financial savings and improved resource allocation. This was strengthened by views on future uncertainty regarding personal illness and disease: although data use may not be of direct benefit to individuals now, it may be of benefit in the future. Some participants felt impartial towards who held the data, and who would benefit from its use, with the belief that any developments that would go on to improve treatment or care were worth pursuing. However, participants found it challenging to explore potential trade-offs in terms of benefits to one group of people causing harms to others and found it difficult to pinpoint the identity of certain people benefiting as being of a greater public benefit than others.

> If health and social care services could become more efficient or financially more effective, then that is less of a drain in other areas, there is more money that will benefit other things. | Participant, Reading

Across all locations, participants agreed that data use agreements should have built-in safeguards to protect those vulnerable to discrimination, at risk of geographic disparities and protect against security breaches. Many participants put an emphasis on the importance of identity when it came to potential harms rather than benefits: they felt it was essential that marginalised and disadvantaged groups are safeguarded against harms of data use with one participant describing it as a moral issue. Particular concern arose over discrimination based on class and ethnicity and the safeguarding of children, those with mental health conditions and elderly people. An example raised by one participant was the consequences of the closure of vital services such as bus services. This might seem at first glance to be a simple financial decision in relation to transport systems, but participants then reflected on the impact this action might have on mental and physical ill health such as an increase in social isolation for the elderly in their community. To counter this, participants in Stockport thought there needed to be an independent body as a second pair of eyes overseeing any closure of services as an outcome of health and social care data use and look for unintended consequences.

> The overriding caveat must be the protection of the vulnerable in society | Participant, Stockport

5.2.1 Safeguarding children

Participants felt particularly strongly about the need to safeguard children identified using data as being at risk, advocating for health and social care data to be linked but also reflecting on the importance of safeguarding needing to be on a case-by-case basis. Discussions flagged up concern of being overly judgemental of a person’s situation and felt that data cannot provide the whole truth of individuals’ circumstances.

A number of participants were more uncomfortable when exploring the ‘What if’s’ of a case study example of combining housing and health data to make a case to fund housing improvements to prevent health issues such as asthma. The ‘What if’ of the data identifying a child in the same household as an adult with substance misuse issues was seen as a use of data which came too close to identifying an individual. Participants felt there was potential to draw inappropriate conclusions from large data sets.

> The most likely scenario would be pulling the child out of that situation, and it could be making it worse for the child because, even though it wasn't great, it was a little bit of stability for them… it could cause more problems later on. | Participant, Plymouth

Many participants felt that data assessors should put more protections in place against individual harm the closer the project outcomes are likely to come to conclusions about an individual's personal situation (see Figure 10 and Document 3 in the Annex).

5.2.2 Marginalised groups

During workshop 2, participants were given the example of an application for data use from the Health Foundation on migrant use of health services. This sparked conversations on the role of the media in increasing harms for certain groups by amplifying discrimination and disparities and reinforcing negative stereotypes (trust in the media is explored further in section 5.8). Participants weighed up the potential benefits and harms, with some concluding that if certain groups were to benefit, data use should go ahead regardless of how the media report on it, whereas others were more cautious. They were concerned that media reports could add fuel to the fire of judgemental people who already discriminate against minority groups, with a small number of participants fearing the possible motivation of hate groups. This led to further suspicion from some participants over possible manipulation of data (section 5.8.2) by the media to further their own agenda.

Some participants saw the influence of the media as a separate issue irrelevant to whether the data use is of public benefit. These participants thought there should be another layer of authorisation to ensure the potential harms and benefits have been strategically weighed in the assessment process, with extra safeguards around who has access to the data. They felt this could address fear of data use for political gain and misrepresentation of ethnic minorities in the media.

> What the press does to demonise people is another completely separate arena that, to me, is outside of this particular discussion. The fact is if we use the data correctly and let the right people use the data and it helps people's lives, how can that be a bad thing? **|** Participant, Plymouth

Participants did not want to see people from minority backgrounds either misrepresented or under-represented in data. Concerns were raised that any harms caused by data use may disincentivise people from opting to share their data in the future and the implications this would have if whole minority populations decided to opt out.

> I think the identity of the groups most likely to be disadvantaged are the ones that probably have less ability to represent themselves in the first place. **|** Participant, Great Yarmouth

Rather than cause harm, participants wanted data use to particularly benefit certain groups who may otherwise be marginalised in the current health and care system. Participants spoke of a potential opportunity for data to be used to equalise disparities and reduce socio-economic vulnerabilities. Participants used COVID-19 as an example of where data can be used to help specific populations who are known to be at higher risk. Participants also thought there was a potential use for health and social care data to identify where discrimination already exists in the health and social care system, further amplifying participants eagerness to minimise discrimination of marginalised groups to maximise public benefit.

> You could see if there was privilege, and prejudice and discrimination. If there was ever something underhand going on where you're like, 'Oh, it's really weird because everybody in this area's golden. They're not getting affected by the same stuff… Oh, it's because this person wasn't implementing strategies in the right way or some sort of discrimination going ahead.'. **|** Participant, Plymouth

5.2.3 Geography

Participants reflected on whether the geographic location of those who benefit or are harmed by the use of data was an important factor in public benefit. Some participants were content with data being shared with different geographic locations, particularly if it would lead to improved service. Others were uncomfortable and termed it a postcode lottery, with some wary of the North/South UK divide, describing it as problematic that certain areas might benefit from data use when others would not. There was concern that some areas may intentionally be left on the back foot and that some regions would see greater benefits whilst others suffered more harms. Conversations centred on the need for fairness in the spread of benefits across the UK with benefits to be shared equitably. Participants felt it was important that data is understood within the context of where it was collected with an awareness of the difference in demographics, giving an example of data collected in council housing being irrelevant to more well-off parts of the country.

> It's almost become a bit of a postcode lottery in some cases, where say some data is shared in some regions but not in others and then obviously the infrastructure and facilities are made for that region and not others. | Participant, Stockport

5.2.4 Accessing data internationally

Where data is accessed through global partnerships, participants felt that it could bring benefits world-wide, and to the UK health and social care systems. They specifically reflected on the points in Table 4:

Table 4: Overarching reflections on global collaboration on data use

Points reflecting attitudes global collaboration on data use

Health and social care environment, regulations and policies may differ in other countries and therefore impact on the relevance and accuracy of other countries accessing data derived from UK sources

The ability to share information across country borders on what has or hasn’t worked in health and social care may prevent unnecessary repetition of research on areas which are already well understood

The potential to help people around the world, particularly those in developing nations, is of public benefit to the UK as well as those countries which are supported

There is a need for a mutual agreement between countries and reciprocal data access with an obligation to share the outcome of data use.

> If you're sharing it with different countries, then it brings in different factors. Say like environment. How Coronavirus can change in different environments in the way that it's treated. | Participant, Plymouth

Some global data access agreements were seen as important as there was a belief that advancements and innovation in technology and treatments through research would eventually benefit all. When, for example, data was used for the purposes of research and innovation on COVID-19, the identity of who would benefit based on geography wasn’t an important consideration as the benefit was seen to be for the broader international public benefit. Questions of trust were also raised in relation to global data access which are picked up in section 5.8.

> This would work for the NHS and also larger organisations like WHO, observing health trends, whether that be mortality, morbidity rates or the instances of certain diseases. It's really important to maintain data like that so you can keep up to date with current affairs when it comes to health. **|** Participant, Stockport

5.4 Scale of benefits or harms

There was some debate amongst participants about whether the scale of benefit or harm was an important consideration. For many, small-scale studies were still seen to have the potential for significant public benefit. One caveat to this was how much money and time was spent on the use of the data: participants felt this needed to be proportional to the number of people who would benefit and said it must be cost-effective. A small number of participants thought that benefits that impact a larger number were preferable over only a small number of people benefiting. Where the outcome of data use was lifechanging, even if only for a few, most participants agreed that this was as important as large-scale benefits. One group said that a more important factor to consider than scale was the relative impact of data use.

> If a single life can be saved by using that data, for me it's more than zero **|** Participant, Reading

For those, few, participants who felt scale was an important consideration, the impact on overall population health was a factor in how much public benefit a use of data was deemed to be, with some declaring: the bigger the better. Some framed this as reaching a more diverse group of people whereby the larger the group, the more likely there will be greater diversity in those who benefit. Participants often found it difficult to distinguish between scale and identity, as with the above point on reaching a diverse group of people, thus throughout this section, participants views' often touch on both scale and identity interchangeably.

One group of participants in Plymouth discussed the difficulty in ascertaining the scale of benefit considering that scale may change over time. They drew on their experience of the pandemic and how medical and healthcare advances may have a greater benefit over time given the prevalence of new and existing diseases.

Other conversations described scale as relative and benefits as subjective to those receiving them, making it difficult for participants to deem whether the scale of benefit to a smaller group was of greater overall value to the public than benefit to a larger group. Scale was of greater importance to participants when in relation to harms, although most found it more challenging to conceptualise scale of harms in their discussions than scale of benefits. For most it was easier to envision a situation where a large number of people benefit from the use of data than are harmed.

> I feel strongly that all considerations should be given if a data release can help even a small number of people. **|** Participant, Plymouth

Some participants spoke of data use having a neutral benefit whereby no one should be harmed, but for it to be termed a fair use of data, a decent number of people must benefit, recognising this as a difficult and complex balancing act. Participants contemplated how a big impact for a small number of people may have larger benefits for society than first recognised but if a lot of people are going to be disadvantaged that's an important consideration.

> The numbers of people benefited is not necessarily the same as the amount of benefit to be gained or lost by society, as well as those individuals. So, for instance, if you disadvantaged a small number of people, that might actually cost society a huge amount in resources to look after them, whereas a small thing, which apparently only benefited them, would, in fact, benefit society a great deal. | Participant, Great Yarmouth

Participants termed it uncomfortable that by attempting to benefit the largest number of people, some groups - such as those with rare conditions – who aren't strongly represented in large-scale datasets may be disadvantaged or overlooked despite a significant need. They worried that within a commercial data use context, without the significant profit margin that would likely be present for more prevalent diseases, commercial entities like pharmaceutical companies may not consider it worth their time. Participants also thought about this within the context of service closures, contemplating the balance between scale of harms and benefits if data use led to the closure of some services to benefit others. It was felt that benefits to a large group of people shouldn't come at the expense of smaller, more vulnerable groups.

> Even if data would only be for the aid of a rare disease or small population in the case of social care, it is important not to prevent the use of data, as these people may be overlooked when they have a greater need. | Participant, Stockport

Participants also thought about scale within the context of a cost-benefit analysis, contemplating the impacts of AI and investment of money into new technologies, and the potential benefit for a large number of people in the future. Conversations also turned to the financial implications of benefits that only reach a small number of people. Participants often linked scale to funding, believing that smaller scale benefits were less likely to be funded than data uses with large scale benefits. Some thought that this was often a better use of money, with one Reading participant saying: we don't have a bottomless money pit.

> If you're doing this for 1% of people that it does correlate… [its] not worth it for 1%. If it's for 10-20% of people then I think it's worth it but if you're pumping millions into AI, it's a cost benefit analysis. | Participant, Plymouth

As discussed in section 5.2.3, participants also explored the importance of geography in the context of data use and public benefit. Participants thought not only about geography in terms of identity but also scale. They largely felt positive about the scale of benefit being to many across the world, particularly when it came to research, and used the example of COVID-19 vaccines to articulate their points. Participants also weighed up whether it was worth risking potential data breaches of large datasets if only a small number of people would benefit, one participant said this was a bit hazy and described it as preferable where large sets of data are used for overall population health, however for others this was irrelevant.

> I would hope that we would share not just with our European partners, there are other organisations from other countries that are doing a lot of research. | Participant, Stockport

> It's about whether it's appropriate to give access to confidential data if perhaps the scale and benefit of a piece of work is minimal, but then it may be appropriate if it's only going to benefit 100 people but actually the advantage to them is huge. So, it's about weighing the balance between them and the use of the data. | Participant, Great Yarmouth

5.5 Prospect of benefits, disadvantages and harms

Participants were in favour of research, planning and innovation taking place, including more exploratory research, even in cases where there was a low prospect of benefit. This was because they conceptualised increased knowledge as benefit in itself and they took a long-term view of how benefits might be generated. Participants often explored prospect of benefit by citing the purpose of research and the scientific method, to prove or disprove a hypothesis, with both outcomes felt to be equally valid and beneficial, leading to greater knowledge. The belief that the contribution to knowledge as public benefit meant for many participants any use of health and social care data was considered likely to lead to public benefit.

> It happens a lot where studies don't find what they were hypothesising to find. It's still of public benefit because if you find no link that's one area less to explore. You can rule that out. **|** Participant, Plymouth

Participants were positive about exploratory research taking place, even if it was a shot in the dark, as it was felt that this sort of research could lead to significant innovations in health and social care, and so was worth taking the risk of finding nothing. Participants repeatedly referenced major scientific breakthroughs arising from accidents in support of this argument. Participants also weighed up the likelihood of benefit against the scale of benefit. This supported the view of some that small-scale research could have a substantial benefit, for example in finding a treatment for a rare disease. However, the majority of participants stressed the importance of having an intended purpose, however broad, to being able to assess public benefit in the first place. Without some initial evidence or hypothesis to base some exploratory research on, participants characterised using data in this way as a fishing expedition and were not as supportive as with exploratory research with a remit.

> Research, they have to take risks because it has to be innovative, otherwise it's all been done before, so I don't think you can start from a negative point and say, 'Yes, but what if this happens? What if we don't get the results, like you say, that we wanted?' There's still a chance that it could be successful, and you're not going to know unless you decide broadly what you want to achieve and try to provide that data. **|** Participant, Great Yarmouth

Where research became exploratory later down the line, for example if an unexpected finding indicated another direction of inquiry, there were mixed views amongst participants about obligations on researchers. In general, participants felt that researchers should at the very least report back any change in direction to the data assessor, if not seek full re-approval in order to explore another line of inquiry. Participants were also keen to ensure that changes of direction did not mean that the original study was not completed.

> If other areas come up I think it's important to get the most out of the data and yes benefit would have to be assessed. If a new area was identified as a by-product of the first look at it, then you would have to assess again the public benefit in that new area. **|** Participant, Plymouth

Despite this general sense that public benefit assessment shouldn't curtail more exploratory uses of data, there was an argument amongst a few participants who framed their conversations about likelihood of benefit in terms of value for money. These participants felt that uses of data for which it was less certain whether there would be a benefit should be minimised. This was especially the case if they were to be funded through public money, which participants felt was limited and therefore there was a ‘duty’ to spend it on uses that were more likely to lead to a benefit. (These participants were not thinking of *knowledge* as a benefit in itself here.)

> I think because everyone's aware there's not loads of money being shared about at the moment they want to make sure. Especially taxpayers making sure it's going to the right places and being spent in the correct way rather than being spent on a research project that doesn't impact anyone. **|** Participant, Plymouth

An alternative view was put forward by a few participants who felt it was more important for there to be a high likelihood of benefit for profit-making companies to be able to use the data. This was not a cost-effectiveness argument but was instead linked to a sense amongst many participants that the bar for access to data should be higher for profit-making companies (see more on trust in section 5.8). There was a sense that industry needed to balance promoting the public benefit aspects of their data-driven work alongside their commercial aims.

Linked to this, some participants identified time for benefit to occur as another factor by which public benefit could be assessed, additional to those which they were presented within the workshops. Participants were open-minded to the possibility of uses of data which might take a long time to see a benefit. This was explained in several ways, including:

- The possibility of future research or technologies (e.g. AI) making use of existing research, thereby bringing a benefit later down the line, through incremental progress
- That some ‘harms’ might be short term but in the longer term there would be benefits (especially in relation to a case study in which one scenario of data use was that some care homes might need to shut down if they weren’t able to meet the costs of providing a certain standard of care)
- Taking into consideration benefits over the lifecourse, or a prevention mentality, to understand how they (or future generations) may experience benefits in the future.

> Long-term benefits, not on everything but when we look at most things, I think research is always about long-term, I don't mean doing the research but the actual benefits from the research. **|** Participant, Stockport

There was some interaction between time and profit, for example if a company was making a profit, some participants felt there was a need to see the benefit in a shorter time for this to be considered ‘public benefit’, because otherwise they feared any company could claim that what they wanted to do would lead to public benefit in the future.

> Otherwise, they can all say that there will be public benefit, but if it's 60 years down the line, it's more or less moot. **|** Participant, Reading

Whilst the majority of discussion was about the prospect of benefit rather than harms, there was some discussion of a need to ‘balance’ benefits and harms. This was tied to participants’ own appetite for risk. However, participants found it difficult to think of examples of ‘harms’ that could result from data use and so these conversations were more abstract and caused more disagreement than about benefits. It was accepted that uses of data which had a high likelihood of causing harm should not be allowed. For some, if there was no or low likelihood of a harm occurring, it was important that research and other uses which had the potential for benefit to occur were able to go ahead. The most risk-averse felt that any likelihood of harm occurring was unacceptable and even negated public benefit.

5.6 Type of data used

In exploring a range of case studies on health and social care data use, participants discussed the impact of the type of data on the potential for public benefit. Some data was seen as **more 'fragile' in its useability for public benefit**, either because of the personal nature of the data (e.g. genomics, mental health) or because it needed greater care in its interpretation (e.g. qualitative, spotlighting specific populations). Figure 10 sets out the key points that participants made when thinking about the use of specific types of data for public benefit:

Genomic	Mental Health	Housing	Mobile Phone
•Different to 'normal health data' •How do you anonymise? •Fear of eugenics •Use by insurers •Lots we don't know •It goes to your very soul. Participant, Reading •Use beyond health, e.g. intellect •Risks of sharing overseas •Huge potential for good: cancers, conditions linked to ethnicity e.g. sickle cell •Future of medicine	•More sensitive than physical health data •More open to interpretation •Less useable at scale as individual circumstances need to be understood in context •Diagnosis and description of mental health can be more subjective: A GP's attitude to someone who has a problem could be very different to another. Participant, Reading	•Risk of not getting the full picture if private rented/owned properties not included in the data •Anonymised health data linked to unique property identifiers leading to identifying an individual issue feels invasive •Risks disincentivising discussing drug/alcohol issues with GP •Use to identify flaws in data collection	•Phone knows you better than anyone •Mistake curiosity for mental health risk (e.g. googling 'suicide') •A little bit too invasive and just for me, that's a bit like big brother state. Let's see what they're eating, how much exercise you have, how many people have you spoken to this week. A bit too much for me. Participant, Stockport
Highly sensitive, but public benefit if used with great care			**Highly sensitive: no use without explicit consent**

Figure 10: Sensitive types of data

Participants questioned how qualitative data could be used for public benefit. It was often characterised as subjective and open to interpretation and they wondered how conclusions could be drawn from it when gathered together.

> I think if it's qualitative where you don't have many numerical values, then how is that going to be interpreted if it is used for a wider purpose. Like we've said it can be quite subjective. Whereas if we have quantitative data based off numerical values then I see it being a lot more practical. I think there does need to be a bit of a distinction here. | Participant, Reading.

Data that spotlights specific populations (prompted by the migrant case study presented by Health Foundation expert witnesses in Workshop 3) was raised in relation to political/racial sensitivities around the findings which could prevent its use. Some participants were concerned that these sensitivities could be detrimental to public benefit, as it could lead to neglect in some health areas by not exploring them.

> Things like TB is on the rise again which we believe is coming in through migrants and that sort of thing. Should we ignore that because we don't want to be perceived to be prejudiced in any way? | Participant, Stockport

5.6.1 Data quality & accuracy

The quality and accuracy of data mattered to participants as they discussed their impact on the type of public benefit that could be achieved. Data quality was determined by factors such as accuracy, representativeness, size of the data set, method of collection and the impact of data gaps.

Some participants in every location gave examples of how, when they had seen their health or social care data, had found inaccuracies such as out-of-date weight data and conditions they had never had.

The size of the data was seen as being important for representativeness, but also raised concerns about whether it could be at risk from double counting data (COVID-19 data was used as an example). The point was raised that a large data set with inaccuracies would be of less public benefit than a small data set with higher quality data.

Data collected on paper was seen by some as a less reliable source of public benefit because its translation to digital and shareable format could be subject to the Chinese Whispers effect.

> Collecting data on pieces of paper, pass it to one person, and another person inputs it into a machine. It seems to me like Chinese Whispers. By the time it gets into the machine, is it the same as what was originally thought by the first person? | Participant, Plymouth

5.6.2 Data gaps

Data gaps concerned participants. Many were surprised that it is not mandatory for GPs to share data into a national data set for use in research and planning. As the main way that people access the NHS, the lack of data from some practices worried many participants. They were concerned that areas not included would not have their needs accurately represented when planning services or lead to mistakes in evaluating the prevalence of a health condition.

> If there are 1,000 GP practices in an area, and only 10 of them actually give that data out. It's not really a good way of doing it. | Participant, Stockport

> When there is a great public benefit in having more data should they not be compelled to release that data? | Participant, Plymouth

Some participants worried that paper data was sitting in care settings, unexamined and therefore unmined for its use – meaning that service improvement opportunities could be being missed, particularly in areas such as social care.

> I would really like something more to be done in some kind of digital format, for them not to just have these piles and piles of papers across the country. Something needs to be set up and organised more efficiently. | Participant, Great Yarmouth

Other potential data gaps in data raised by participants included the move towards more reliance on pharmacies for health advice and paper notes. They assumed that these interactions with pharmacists went unrecorded and therefore created gaps in service use and health conditions/treatment data.

5.6.3 Assessment of data quality

During the workshops participants shared their uncertainty and looked for greater reassurance about how **data quality** was assessed to ensure any public benefit delivered by data use was based on sound evidence.

> I'm still quite intrigued about how they monitor the accuracy of it... It has been brought up several times, but how is the accuracy of the data monitored that they're getting, and the data that they're giving out? | Participant, Stockport

Some participants looked to technology to help with data accuracy and thought that Artificial Intelligence (AI) would play a greater role in the future.

> In terms of five years in the future, sample size and accuracy may not be as much of a problem if AI has developed further. It might be easier to collect more accurate data by then. I don't know, they might have equipment that can collect chest scans, retina scans, much quicker and much more accurate than we can now. | Participant, Reading

5.6.4 Social care data

Given the UK's aging population, **social care data** was seen by participants as a vital resource for planning and delivering care effectively. Participants were keen to see social care data transformed from its current characterisation as a fractured mosaic with fragmented information and many missing pieces, into digitised data, collected in a standardised way, but that still has care at its core (not allowing tick boxes to override nurture). The hurdles to achieving this were particularly clear to those who had family or employment experience of social care.

> One of the problems is because social care doesn't come under one umbrella. It can come under the council. It can come out of social care companies. They all gather their own information. My feeling was, when my mother was in social care, it was never joined up. There was never any joined up information. I think most governments have got the same problem because they just don't know what to do with social care. | Participant, Plymouth

Some participants raised concerns about how social care data could be collected reliably and comprehensively when people working in social care may have less training (compared to NHS staff) or that commercially minded private sector providers may pick and choose the data they share.

> They may not even know it's available for them to use, or wouldn't know how to use it because they don't have the staff with the background of using data in the past to then be able to use that data to help them for the future. That may be something that needs looking at a different level. | Participant, Plymouth

The nature of much social care data was seen as harder to categorise than health data, less clear cut and more experiential and context related. Some participants worried that proper funding wasn't being given to social care because it is harder to measure and harder to prove where time is spent and with what benefit.

> We live in a society now that's ripe on ticking boxes and evidence to justify what time they've spent doing this, that, and the other. Social care are predominantly people caring,

> nurturing, supporting, in a verbal way. It's harder for them to prove where their time and energy has been spent. That's why the funding doesn't get given. **|** Participant, Stockport

Examples of public benefit of using social care data were discussed by participants, they included the examples given in Figure 11.

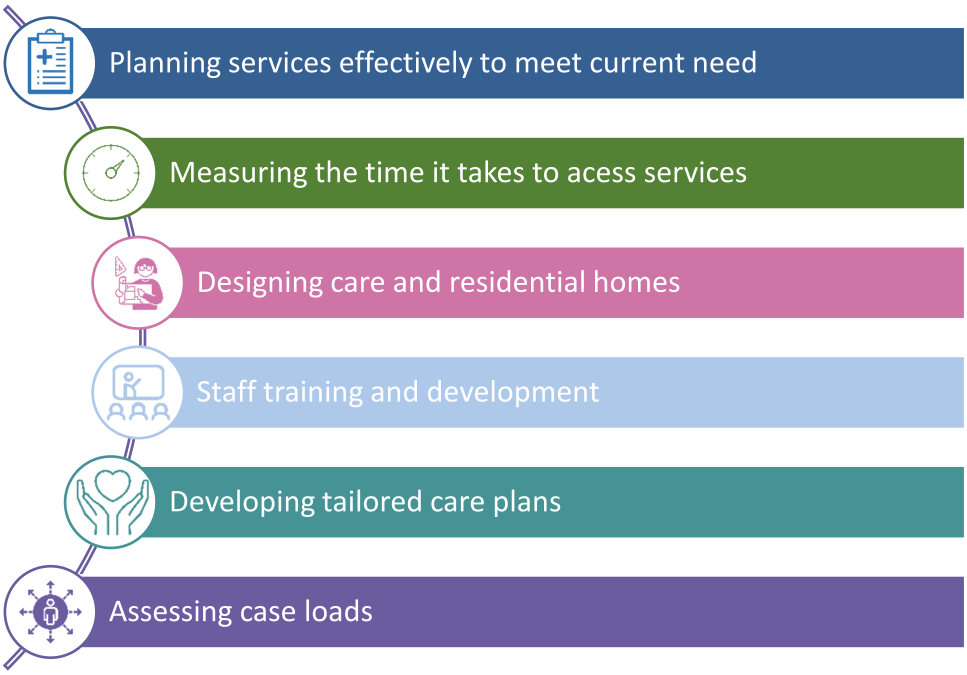


Figure 11: Examples of public benefit derived from social care data use

Some participants could point to examples of where care services, such as respite care and personalised care in sheltered housing, had improved over time and attributed this to good use of data.

> My father had a really rough two years in his last two years, and I thought the sheltered housing he was in was exceptional but even... over that period we could see how they'd stepped up to improve the situations that he was dealt with. So, I'm sure that would then get across and gather that data and improve other people's situations. **|** Participant, Reading

5.7 How findings are used

How findings are used is the ultimate proof of public benefit according to participants. In this context, participants discussed the process for checking that the health and social care data is being used appropriately and effectively and how public benefit is as valid if it saves money and time as if it saves lives (one can lead to the other). The time it takes for public benefit to be achieved was also discussed - that we should be patient, and not feel the need to favour short term projects over longer term ones (see 5.4).

Before thinking about how findings are used, many participants wanted to know more about how the data is kept secure after use has been granted. What protocols, standards and restrictions are in place to ensure data is used only by those who are authorised? Are they stronger and more persuasive than the possible profits/benefits the organisation could gain from selling on the data to unauthorised organisations?

> Once you've accrued this amount of data it's worth a lot of money, is it going to be passed onto organisations that are not so reliable? It seems most people are willing to give data if

> it means it's going to help somebody, as long as it's consensual and it's not passed on. | Participant, Stockport

There was strong interest in monitoring the use of data through to outcomes. Some participants imagined a case worker approach – someone who is responsible for monitoring research projects, their outputs and outcomes. They also expected to see a link between the data users, data assessors and the bodies that approve new drugs/treatments, so that data assessors are informed of how the data has contributed to new treatments. Participants felt that staying close to how findings are used would lead to better informed decisions on data use for public benefit.

> Checking back…IGARD do but others aren't quite as rigorous, but there needs to be a rigorous system everywhere whereby we've given something out, you're supposed to be doing this, where have you got to? What's happening? | Participant, Plymouth

When assessing data for public benefit, some participants expected to see predictions of both positive and negative outcomes and consideration given on how to mitigate the negative. They realised that the nature of research meant these predictions could not be definitive, but an awareness of a range of potential outcomes was important.

> I think you've just got to make it very clear, if you can, what you're trying to achieve, how you think that's going to benefit the public, and then take into account what factors may cause problems. They'll never truly know but if they go in and an assessor can see something will clearly have a negative effect, whether that be financially, for the NHS or for elderly care homes, all that stuff, that should be clear enough for them. At least everything is transparent and there's a clear understanding of what can happen. | Participant, Plymouth

5.7.1 Mental health data

Some participants saw the use of findings from mental health data as different to those of physical health data. They felt it would be harder to generalise findings, when each person's situation (their physical and mental health, the treatment they are receiving, their living situation and background) is unique.

> I think you would have to go in deeper into each individual service. You couldn't just look at a piece of paper or graph and go, 'Right, we're going to close that.' I think you would have to investigate a lot more. | Participant, Plymouth

Just as some participants thought mental health data needed to be used alongside contextual information, a similar sentiment was attached to the use of data to design service provision, particularly the closure of services. Fears were expressed of numbers being used and even manipulated to make decisions, without taking the wider context being taken into account. This was seen as a betrayal of public benefit.

> I know that when austerity happens, they were just cutting things without a great deal of evidential thought. There was some there but where they're using data for other purposes, to decrease investment. i.e. make cuts, that has to be added to something else and the process needs to be transparent to service-users and the electorate. If data isn't used for purposes such as research into conditions and that kind of thing or safeguarding, if it's used as a blunt instrument, I think that's going to dampen public enthusiasm. | Participant, Plymouth

The affordability of interventions that are developed using health and social care data, be they drugs or service design, was discussed by participants. They raised concerns of how unfair it would be if, for example, some local authorities could afford a social care service developed through data use, but others couldn't. Likewise, with drug therapies that could end up being unaffordable. However, participants who discussed this prospect of affordability, didn't think this should be a block for data use. Instead, it was the health and social care system that should adapt to achieve fairer service delivery across the country.

> I think across the wider spectrum, we've discovered that if we do this we can cure etc, etc. But it's going to cost an extremely large amount of money, where are we going to get it from? Resources aren't infinite. **|** Participant, Plymouth

> > So, regardless of the outcome of the study, it might be that it's difficult to take up the findings. **|** Facilitator, Plymouth

> Yes, which then is, so what is the public benefit? On the other hand you wouldn't want that as a constraint when you are deciding what the public benefit is, if it's going to be too expensive, we're not going to let you do it. **|** Participant, Plymouth

5.7.2 Saving time/money as public benefit

Data used to improve people's health and wellbeing was seen as core to public benefit. But some participants also wanted to emphasise that outcomes such as time and money saving should also be woven into the tapestry of public benefit as they could free up much needed and scarce resources to achieve these outcomes.

> I don't think you can say, unless it's a positive health outcome, it's not good. Because a cost-saving outcome or a time-saving one is arguably as good because you've got time to spend on other patients, etc. **|** Participant, Great Yarmouth

5.7.3 Commercial use

When thinking about who uses health and social care data and who benefits from it, the topic of **profit made by commercial organisations** came up and was discussed at every workshop. At the start of the discussions, participants tended to fall into three points of view:

1. Profit and public benefit are at odds, the former is always a threat to the latter.
2. Profit is a risk to public benefit and needs to be scrutinised when it is part of a data use application but **shouldn't automatically prevent the request to use data from being granted**.
3. Profit is a standard and necessary part of any industry, including health and social care (pharmaceutical companies and private care home operators were most mentioned) and their profitable status enables them to exist, employ and innovate.

As the workshops progressed some, but not all, participants who held the first point of view shifted towards the second and third. This shift took place in the context of discussions around the market forces that drive research, innovation and service provision and around the thinking that the data providers (NHS etc) could receive something in return for the data (profit share, cheaper medicines, payments for data).

Public benefit must outweigh profit, was a consistent call from most participants. Profit was seen as potentially causing perverse incentives that could lead to manipulation of data in favour of the commercial organisation which in turn, could cause harm to the public.

> It's not so much the profit is the issue, it's the profit maximising. It's fine that private companies want to make a profit. The issue is there might be various solutions in the data and their incentive is to choose the solution that is the most expensive and the most profit-maximising one. And that's where you need to be careful with the terms in which you provide the data. **|** Participant, Great Yarmouth

Other profit related fears included:

- The **power of knowledge** that data would give to a commercial organisation who could then hold the NHS or local authorities 'over a barrel' and name their own price
- The **abandonment** of research into an important health condition by a pharmaceutical company **if the costs of research would make the treatment unprofitable**
- Social care operators identifying affluent areas to set up care homes, leading to a **lack of provision in lower income areas**.

Some participants pointed out that these practices happen already, but that vigilance in who data is shared with and how its use is monitored have an important part to play in dealing with these outcomes.

Commercial organisations applying for data declaring up front the levels of profit they expected to make from the use of the data was suggested by some participants. But others said that because you can't see into the future, estimating profits would be difficult for most commercial applicants to do at that point in the process.

Participants who were accepting of profit from data use, often caveated their points by saying that excessive profits were not in line with public benefit. Examples of pharmaceutical companies charging high prices for drugs were used by participants to illustrate this point:

> At the end of that study, there still needs to be a public benefit. They can't say, 'Oh, well, this will benefit the public but we're going to raise our prices,' and so that the public might not actually benefit from it because they won't be able to afford it. **|** Participant, Stockport

Throughout the workshops, some participants raised the concept of profits generated by use of health and social care data was and could be used directly and indirectly to benefit the public. Examples of this included:

- A percentage of the profits made being put into further research that leads to other new treatments/services
- Pharmaceutical companies funding hospital developments in the UK and countries in the developing world
- The tax from profits made by commercial companies using health/social care data supports the economy
- Sharing research outcomes with other public health organisations

> I was thinking about, 'The company makes a profit,' and how I feel about it. I think your instant reaction is resentful, but then the companies wouldn't be here if they didn't make a profit in

> some way. It can be seen as a positive if we keep a company going by making a profit, so they can look into more research. **|** Participant, Great Yarmouth

During workshops 3 and 4 in particular, participants made the point that they didn’t want profit to be seen as incompatible with public benefit. Many emphasised that drugs, treatments and services that were more effective and efficient was a significant benefit to the public and that it was unreasonable to expect companies to do this for free/not make a profit.

> If you look at all the drugs for any disease, it has only been found due to research and the data that they've found. You can't get new drugs and new treatment without evidence based practice. So, if the data is not shared or the data isn't released to be looked at, you will not get new services, you won't get any drugs, yes, the pharmaceutical companies are making goodness knows how much profit, but without that we're not going to get new drugs, we're not going to get new treatments. So, I don't see how we move forward if the data isn't shared. **|** Participant, Plymouth

Again, in the closing stages of the dialogue process, some participants talked about the possibility of it being in the public benefit for the NHS and other public service providers to be compensated for the use of data, beyond just cost recovery. Some participants said they thought the UK was not good enough at mixing public and private for public benefit,

> I accept that there's an issue in some areas like drug pricing, other countries and so on where commercial organisations don't come out of it looking good ethically speaking, but I'm absolutely of the view that we're lost if there's no cooperation or more cooperation between public and private. I think there's a real question to be asked about whether it's desirable or even legally permissible to build something into agreements about using data, just in terms of the extent of the profit that could be made, for example. I'm not against something being inserted into contracts, as long as it's legally binding. **|** Participant, Reading

Other participants talked about how it was only fair that health and social care data that contributed to commercial profitability should be exchanged for payments, profit share or cheaper/free access to treatments and services

> If the data is used by care companies to model their business plans, then they should pay a fee for access to the data. We have to remember, even though they have care in their name, it is still a business, and the bottom line is profit. Should we let these people access data free for their own gain?, I don't think so. **|** Participant, Plymouth

5.8 Transparency and trust

5.8.1 Transparency

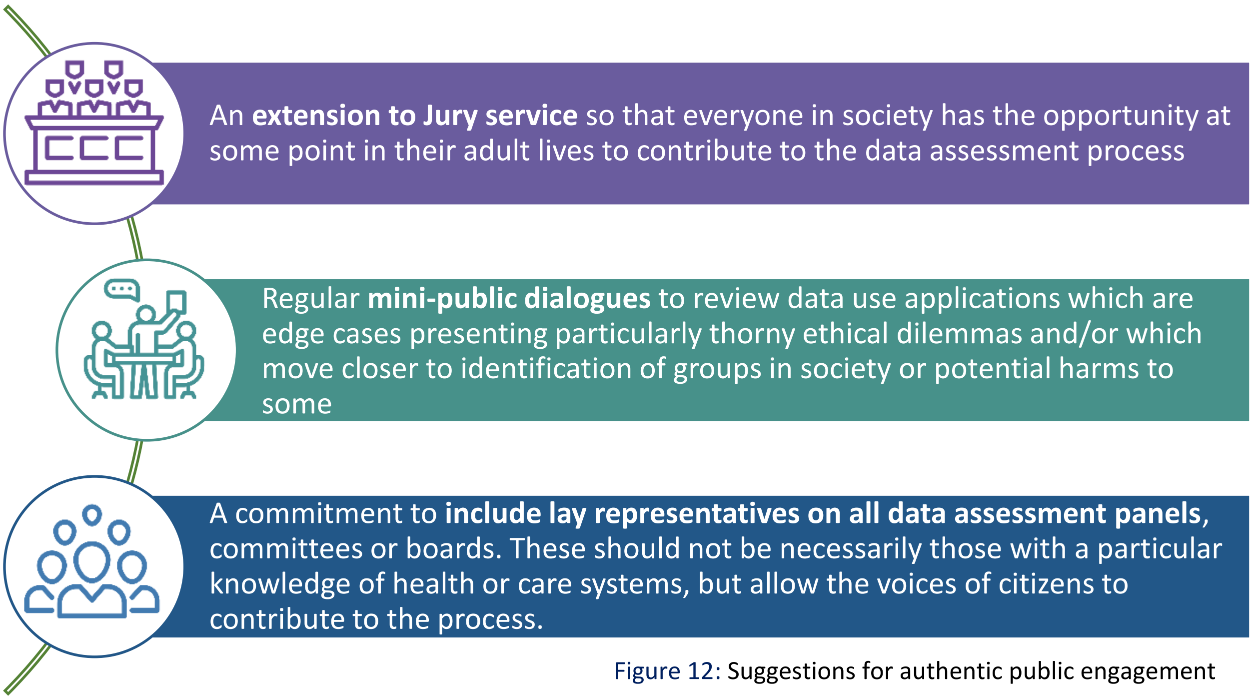
As we have indicated throughout this report - for most participants transparency is central to any discussion about the public benefit realisable from the use of health and social care data. Whilst recognising testing trust in the process was not part of the explicit project scope, it was raised spontaneously and consistently by participants in all locations as a pre-cursor to an effective system of data use assessment, without which public benefit could not be demonstrated or robustly assessed.

There was a shared view that knowing how such data is used, by which organisations, with what benefits being gained is essential for a publicly funded system.

> Transparency is something I think we all agree is needed. It might not really be practical to get individual consent about our data being shared, but it’s important to keep people in the loop as to who is accessing it and what it’s being used for. | Participant, Plymouth

> There has to be some form of transparency [in data use] so that people know that the data being collected is being used for good. | Participant, Great Yarmouth

Many participants across the dialogue felt that having a form of *authentic public engagement* in the data assessment process would really demonstrate a robust assessment of public benefit. Participants welcomed the opportunity of hearing from the Independent Group Advising on Release of Data (IGARD). They were impressed that lay members were involved in the data assessment process and were keen for this involvement to develop and flourish as a concept across health and social care data use assessment. They suggested making sure any database summarising projects are made publicly accessible in terms of language and content. They also proposed ideas such as set out in Figure 12 including:



An **extension to Jury service** so that everyone in society has the opportunity at some point in their adult lives to contribute to the data assessment process

Regular **mini-public dialogues** to review data use applications which are edge cases presenting particularly thorny ethical dilemmas and/or which move closer to identification of groups in society or potential harms to some

A commitment to **include lay representatives on all data assessment panels**, committees or boards. These should not be necessarily those with a particular knowledge of health or care systems, but allow the voices of citizens to contribute to the process.

Figure 12: Suggestions for authentic public engagement

It should be stressed that **participants did not want *only* citizen voices making these decisions**. But they felt that to build trust in society for data use; to broaden understanding and knowledge; and crucially make the process transparent, would benefit from some of these suggested additions to the process – working alongside those with specialist knowledge of health and social care data.

A small number of participants (predominantly in Great Yarmouth) suggested that once the National Data Guardian’s guidance is published, experts, with specialist knowledge should be able to assess data access requests without the need for detailed public engagement in the process. There is a sense among these few that society has only been recently aware of data use, and that society would not be as concerned about data use in these contexts prior to these events. Interestingly they do not site this moment in an increased knowledge of data use in relation to COVID-19, but rather in 2016 with a set of data scandals related to social media and national referendum and elections.

> For me, it's not 2020 and COVID changing people's opinion on data, it was 2016 and social media and Facebook, that's the thing that changed people's opinion on data. That coincided with the start of GDPR and all the rest of it. People are scared now of how data can be misused and I think it was certainly with 2016, with Brexit, with the Trump election, etc. where it was shown how your data could be manipulated. **I** Participant, Great Yarmouth

Whether or not public engagement was front of mind for participants, there was consensus on a desire for communicating widely about how data is used for public benefit. Suggestions included campaigns which shared knowledge about what data about specific diseases had done to improve treatments; or how improvements in social care provision were achieved using data about admissions to adult social care services.

Without these transparent communications there is sense amongst participants that wider society will think there is something to be hidden in this process.

> I think transparency was something we all felt was needed so nothing's hidden, otherwise we feel like they're only telling you half of what's going on. **I** Participant, Stockport

They felt the communications should be widely distributed and displayed on and off-line in public spaces used every day by everyone such as GP surgeries and websites, libraries, local authority websites and newsletters and community venues. It was understood that data assessors currently published information on what data uses had been accepted, but it was felt that these sources of information were not known about or written in accessible language.

> When we were talking about how guidance is published in the end, I think it's really important, in a very traditional way, to ensure that there are posters and leaflets in public-facing places where the public access health and social care. GP practices, councils, even more traditional places, like post offices and libraries. Some authority says, 'Oh, well, it's in paragraph 5b of section 38 at this particular website,' and it's news to the consumer. This data's about them and it's about solutions for them as the public, so it should be accessible at their fingertips everywhere. I Participant, Plymouth
>
> I appreciate all this information is on websites like CPRD and IGARD, but if you don't know it exists you don't how to look for it. **I** Participant, Stockport

5.8.2 Trust

Related to transparency, building public trust in the system of assessment and in the organisations making use of data was seen as essential. As seen in previous deliberative programmes such as Foundations of Fairness[11] (HVM: 2020) a high level of trust was placed on data release applications made by universities and publicly funded centres of research and innovation. Behind this lies the idea that researchers in these institutions are credible, that their credentials have been checked and that they have behind them an ethos of public benefit in their work as a matter of course. This led to participants welcoming what they heard from data assessment speakers: that checking that data applicants' track records are valid and appropriate for delivering the objectives of the data access request is part of the current assessment process.

---

[11] Hopkins Van Mil, Foundations of Fairness: views on uses of NHS patients' data and NHS operational data, February 2020

> Reputable organisations such as universities wouldn't be doing something if it wasn't for public benefit. I think it is important to look into where each of the organisations are coming from, what they've done before, and in that case, if they are doing something that is a little bit out of the box or could be deemed a risk, then just to bear in mind that they wouldn't be doing it for a bad reason. **I** Participant, Plymouth

Again, linked to transparency, there was a view expressed by many that this credibility test should also look at how the research or innovation study is being funded. They suggested that research could only be trusted if there was clearly no hidden agenda,

> A question that surely we should ask is who's funding any piece of research? Because sometimes funding can be hidden, so there could be an agenda behind that. It could be someone in industry funding it, it could be a political faction, it could be all sorts of things. It could be a newspaper who's going to use it for non-scientific reasons. **I** Participant, Reading

This leads us to a term which frequently appeared in discussions in each of the locations: manipulation. It was seen by a number of participants as likely that there were those (participants cited government, the media and industry) who would exploit or even falsify points that emerge from data use programmes for their own ends. For government this was seen to be *pushing* a policy for political gain; for the media to sell stories to boost readership figures; and for industry to boost profits through sales of a medicine or technological solution.

This point should be seen in the context of the COVID-19 backdrop for the dialogue fieldwork. Participants frequently mentioned the government COVID-19 statistics on mortality which state that figures are collated from those who have died within 28 days of a positive COVID-19 test. They felt that this mis-represented the data. They equally mentioned press coverage of the pandemic being used to boost media sales

> The press companies would hugely benefit from being able to release healthcare statistics, especially in the climate that we are now, where everyone wants to know how safe we are, how vulnerable we are, things like that. In terms of private entities, the press and the media companies would hugely benefit in terms of profit, because stories sell more. **I** Participant, Great Yarmouth

As we have seen in section 5.7.3, profit motive was seen throughout the dialogue as a challenge to public trust in data use. Data use *purely* for financial gain was not seen as altruistic enough to be included in the concept of public benefit. Part of this relates to a sense that data is less secure once used by commercial entities.

> For me I believe now that data is so valuable to everyone in this world at the moment, more valuable than money to a lot of people. With data being so valuable it, takes away from the use being beneficial to everyone. Companies will then use that data for whatever needs they have. They're using the data for what they find best for themselves not best for all of us. **I** Participant, Great Yarmouth

As a result, some participants suggested that to gain public trust more scrutiny would need to be applied to data access applications from industry than from others who could more visibly demonstrate their public benefit interests and credentials.

A smaller number of participants in each location felt that industry organisations are in a position to offer public benefit, because they invest in the creation of new medicines and equipment which are seen as life-saving and for the public good. A note of caution was sounded by these participants to make sure that data assessment didn't impose blanket bans on requests from industry who had the funds to make contributions to research, planning and innovation in both health and social care.

> We need to make sure that we don't see it as though the NHS are the goodies and the pharmaceutical companies are the baddies, because they do invest in a lot of research and they do put money into the medications that we take every single day. **I** Participant, Stockport

It was interesting to see that participants went into quite some detail on questions of trustworthiness and integrity. Many, in raising points about trusted organisations, said that they felt the scrutiny of data applications should not stop with an assessment of the organisation making the data use application but should also include questions of those who will be analysing the data and making sure they have the right skills, experience and knowledge. They also wanted data assessment to test whether the methods proposed for the data access project had appropriate research protocols in place to ensure the work is conducted ethically and, as far as possible, without bias.

> There may be something around the competence of those who are analysing it because you could have a company or an organisation that's reputable but those who are tasked with doing the analysis may not be. So, I think the public would need to be confident that people with the right skills, not just the right motivations, are actually analysing their data and responsible for their data while they've got it. **I** Participant, Plymouth

A final point on trust in this context. Some of the case studies given to stimulate discussion (see Document 3 in the Annex) included 'What ifs' around data access across country borders. Participant reflections on this are described in section 5.2.4. In relation to trust, a few participants drew a red line at data access globally unless such agreements could demonstrate that the outcomes would have some public benefit in the UK. This was tied to a belief that data access agreements between countries could increase risks of data breaches, loss or mis-analysis.

> I wouldn't be comfortable sharing the data with other countries, just because I wouldn't-, there's an element of trust, really. Because the more countries that access the data, the more it's open to, I would say, abusing the reports or changing the reports. **I** Participant, Reading

However, many more participants felt that data access agreements across country borders could lead to more innovations such as the COVID-19 vaccines and learning from best practice. This was not without a due diligence data assessment in place, but public benefit could be seen to be demonstrated through such cross-border collaboration using data collected in the UK.

6. A practical output

Summary

This chapter begins with the consideration participants gave throughout the dialogue when thinking about the data assessment litmus test of public benefit. We find that:

Participants **define public benefit expansively**, including direct and indirect benefits and, frequently, in opposition to data use which is seen to be **too close to a profit only** motive

We develop the motif on which the previous chapter ends – that **transparency** and **trust** in the process of releasing data has emerged as an important precursor to public benefit.

The considerations set out in the chapter include:

Participant work towards a **definition of public benefit**, with a focus on **health and social care outcomes**, **knowledge development** and **data sharing**

Reflections on the potential for **payment for data access**

A call for **knowledge** and **awareness** of the data access request process to be increased in wider society so that efforts to bring **public benefit are not undermined**

The fact that participants did not **limit their reflections to the specific moment of data assessment**, they consider public benefit assessment as having a role across the full research lifecycle.

The chapter continues with specific expectations participants have of the data assessment process. These are seen in two sub-sections:

1. Public benefit assessment expectations
2. Wider systems and infrastructure expectations

It was understood that some of these points may not be part of the NDG's policy guidance or advice document but might form broader advice in the wrap around aspects of data use. The chapter ends with an overview of the findings to inform next steps in the process and concluding remarks such as – dialogue participants:

- **Endorsed** and **amplified what is known** as well as bringing new perspectives on **social care data** and the **prerequisites for public benefit**
- **Placed importance on systems and infrastructure** around data use as well as reflecting on specific aspects of the data assessment process
- Demonstrated a belief in the public benefit derived from **both** health and social care data use, seeing the data as a **valuable resources which should be treated with respect**.

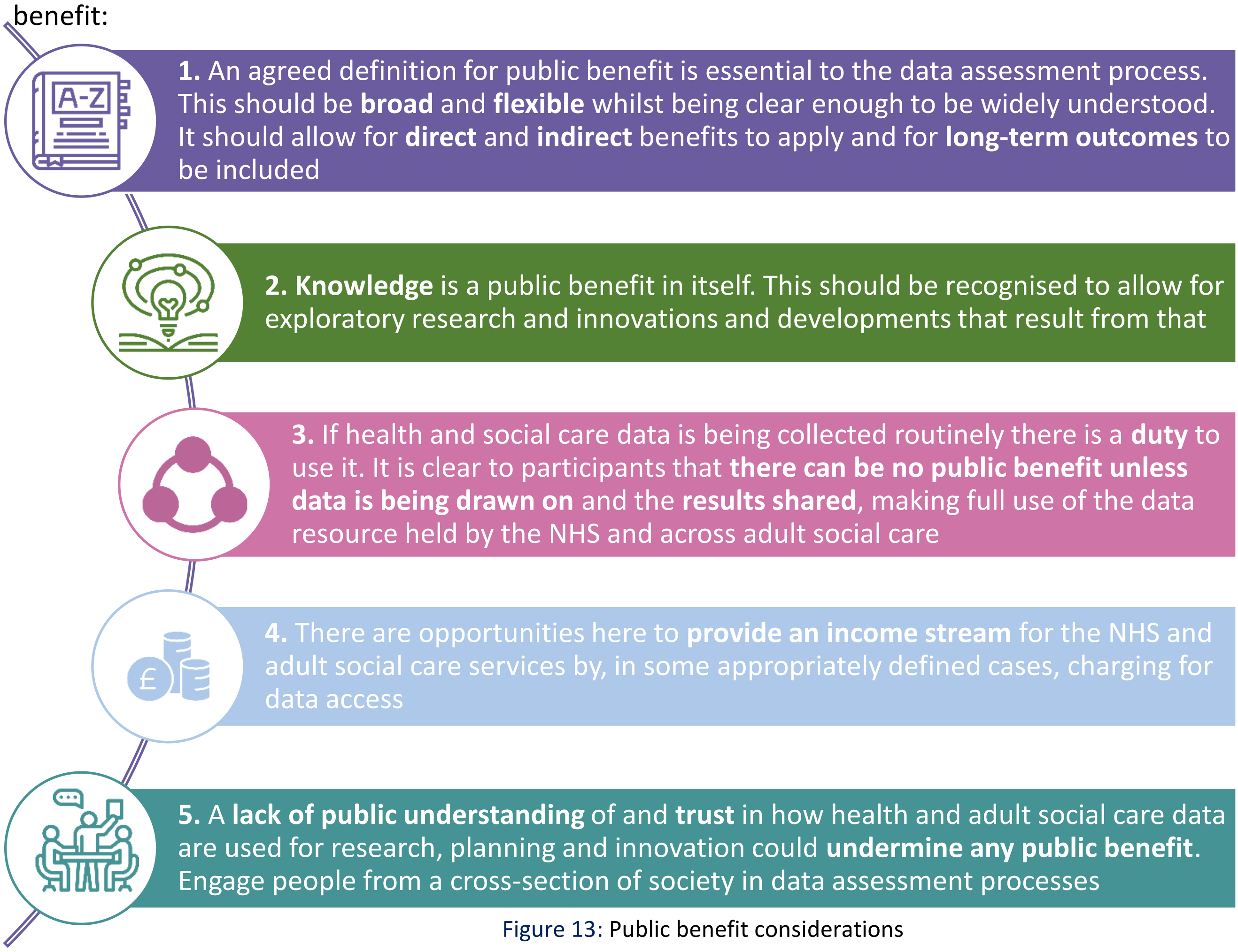
6.1 Public benefit considerations

The previous chapter summarises findings on attitudes towards data use. For the NDG to create policy guidance or advice to support the assessment of public benefit in health and social care use, it is also important to report the considerations participants had of public benefit.

6.1.2 How participants approached talking about public benefit

Most participants had not spent much time thinking about public benefit and what it meant to them. At times throughout the dialogue, participants struggled to express something which they felt to be self-evident or taken for granted. Some participants therefore tried to grapple with the overarching basis on which decisions should be made about public benefit. For example, they wondered whether this was a moral judgement about what was right or wrong, based on morality and philosophy; or a technical and legal judgement.

Others characterised public benefit decisions in terms of science versus politics, with reference to an example given of a study about migrant use of health services having been turned down. For some participants, this was felt to be a case of the collection of beneficial knowledge being curtailed due to political sensitivities about immigration. Others questioned whether public benefit was linked to the challenges of the day (and could therefore change over time), for example, arguing that uses of data connected to cancer or obesity might be considered of particular public benefit due to the current prevalence of these conditions. In Figure 13 we capture the considerations participants used to conceptualise public benefit in this context, before moving into more specific participant recommendations about how it should be assessed in the next section. We begin with the question of what should be in scope in a definition of public benefit:



1. An agreed definition for public benefit is essential to the data assessment process. This should be **broad** and **flexible** whilst being clear enough to be widely understood. It should allow for **direct** and **indirect** benefits to apply and for **long-term outcomes** to be included

2. Knowledge is a public benefit in itself. This should be recognised to allow for exploratory research and innovations and developments that result from that

3. If health and social care data is being collected routinely there is a **duty** to use it. It is clear to participants that **there can be no public benefit unless data is being drawn on** and the **results shared**, making full use of the data resource held by the NHS and across adult social care

4. There are opportunities here to **provide an income stream** for the NHS and adult social care services by, in some appropriately defined cases, charging for data access

5. A **lack of public understanding** of and **trust** in how health and adult social care data are used for research, planning and innovation could **undermine any public benefit**. Engage people from a cross-section of society in data assessment processes

Figure 13: Public benefit considerations

6.1.3 A clear, broad and flexible definition

Participants thought long and hard about definitions of public benefit which they felt was a subjective concept, largely based on qualitative judgements based on experience of what feels right for society. For some this meant that attempting to provide guidance to data assessors was a near impossible task, because each case was so different. However, most participants felt that the subjectivity of the term made clear guidance around it even more necessary to achieve standardisation between different data access regulators.

> I think one of the big problems in judging public benefit is it's intangible. It's not like a mathematical equation. It's an intangible thing where you make a judgement call depending on your set of values. **|** Participant, Stockport

Through the dialogue process, participants defined public benefit expansively in the context of uses of health and social care data beyond individual care, including direct and indirect benefits. The most direct benefits included improving or saving lives, developing better care or better outcomes for people. Most participants conceptualised public benefit according to helping ‘others’, rather than themselves (especially in the context of anonymised data, which some felt was no longer ‘my data’). Although participants felt that anyone (including themselves) could potentially benefit in the future as a result of these uses of data, for example if they were to develop an illness, or need a service or care later in life.

> I think it's really important that data does get collected and analysed for the benefit of others. That, I think, is the only way we learn, isn't it? **|** Participant, Plymouth

Some participants found it easier to talk about what was *not* public benefit and in these cases it was most often defined in contrast to private profit. As has been discussed in Sections 5.7.3, participants disagreed to some extent about whether making a profit affected whether a particular use of data was public benefit. In these cases, participants frequently spoke about an organisation’s motivation for using the data in order to decide whether they felt it was genuine public benefit.

> One of the things that affects public benefit decisions for me, is having public benefit at the heart of the reason for doing it. So, I mean there can be side benefits like earlier it mentioned about companies making a profit from what they do with the data. I think that’s fine if that’s a side shoot, in my opinion. But it's having that public benefit as central, not having that as the side benefit of doing it. **|** Participant, Great Yarmouth

6.1.4 Knowledge is a public benefit

More indirect benefits included contributing to scientific knowledge or efficiency in the health and social care system. Knowledge was itself seen as a benefit, so that a specific change or improvement in the short term was not necessarily required for something to qualify as public benefit; equally just confirming that something is already working could represent public benefit. Some participants framed efficiency in the health and social care system as public benefit, including innovation and planning that freed up clinician time.

> Everybody benefits because the more information there is, even if it's not a direct benefit to yourself, from a health point of view, if somebody could learn from that and feed it back into the community, we can all gain from it. **|** Participant, Stockport

6.1.5 A duty to share

Some participants also thought about the counterfactual argument that there couldn't be any possibility of public benefit if data wasn't shared and so felt there was a duty on government to make use of the data resource. This was particularly apparent in the case of social care, where participants felt there had been less use made of data for purposes beyond individual care, despite big challenges such as planning care for a growing number of over 85s in the population.

> So, is the data being shared elsewhere? I'm not suggesting it's been overlooked, but why is that data, when they are the key holders of it, not being put into practice to prepare for [the ageing population]? **|** Participant, Great Yarmouth

6.1.6 Provide an income stream for the NHS

Participants also identified examples of *secondary* public benefit within the data access process itself, such as the potential for payment for access, in order to help finance the NHS. Some participants defined this as public benefit, as this funding could help improve or extend care so that more people benefitted. Some participants also saw this as a way to balance the potential of private companies making high profits from the use of the data e.g. by charging companies a higher rate to access, or some sort of profit share scheme. Some participants felt it was important to value the data resource more comprehensively in terms of the cost of its collection in the first place, and therefore charges for access should reflect this.

> I think [a profit share arrangement] would underline for such organisations that this is ultimately about public benefit. Whether it's in this country or any other country. Secondly, it might put organisations off, I don't know, spurious requests for accessing data or duplication of data. Thirdly, it might actually help the funding of the provision of the data in the first place. **|** Participant, Reading

6.1.7 Founded on public understanding and trust

Finally, many participants felt that transparency and trust in the process were important precursors to public benefit – and therefore that a lack of public understanding and trust in how health and social care data was being used could undermine any public benefit. This is further discussed in Section 5.8.

> If there's transparency, then there's trust. I think, for public benefit, there has to be both of those things. **|** Participant, Plymouth

6.2 Participant expectations

Participants feel that the policy guidance or advice produced by the NDG should not be limited exclusively to the data assessment process. The majority felt that this snapshot in time could not be divorced from the full process of data collection through to the impacts the data use could have on research, planning and innovation. It was felt that only taking this holistic approach to data assessment could public benefit be truly determined,

> Don't just focus on the approval process. Follow the data and audit it as it is used, analysed and turned into outcomes. Use your own independent auditors to assess whether public good is really at its heart at every stage. **I** Participant, Great Yarmouth

There was a sense that to deliver public benefit, the outcomes and outputs of the data use project should be reviewed by data assessors and the results communicated widely, including in formats appropriate for those without specialist subject knowledge. Participants want to know that at the end of the programme of work – even if years later – those involved in the data assessment process can answer the question, Has the project delivered on its public benefit ambitions?

As a result of this wrap around approach to data assessment, in addition to the considerations set out above, participants created two sets of expectations for data assessment which are summarised on the following page.

1. Public benefit assessment expectations

- Once a clear definition of public benefit is developed, **use a case-by-case approach** to public benefit assessment **so that the definition can be tested** against potential unforeseen consequences or harms to individuals and society
- Ensure that **safeguards and protections** are in place so that sensitive data continues to be protected; and the process doesn't exclude certain sections of the population from gaining benefit
- Enable a **balanced level of governance** which is not overly restrictive whilst ensuring data access requests are in line with public benefit, including ethical factors and with safeguards in place
- Data applicants should be required to have **a clear purpose** for their programme of data access and use; however, **this purpose can evolve** as the programme develops to embrace **unexpected findings** and **exploratory purposes**
- Any major alterations in purpose or scope should trigger the data applicant to **go back to the data assessor** as a minimum to acknowledge the change or to seek approval to proceed on the basis of the new focus for the work
- Ensure that the data assessment process has measures in place so that any new data access requests **build on research that has preceded it** and do not re-invent the wheel
- Include in the data assessment process details about **the time factor** – the length of time the data can be accessed; what data users do if the technology improves or changes whilst they are working on the programme (and therefore the project needs to change)
- All data assessment should include an extended **risk assessment** to consider short-term harms against the longer-term benefits (including for future generations).

2. Wider systems and infrastructure expectations

- Don't limit the role of the data assessor to exclusively reviewing data use requests, **the experience and knowledge data assessors have should be used for reviewing the whole research/ innovation/ planning life cycle** to monitor and audit good practice
- Encourage the health and adult social care sectors to **standardise the tools for data collection**, including encouragement and support for the adult social care data collection to digitise and become a more effective data resource
- Ensure that **safeguards and protections** are in place so that data can't be manipulated for political or financial gain
- Publish clear **statements of data users' credentials** and **sources of funding** to protect against data manipulation, potentially non-altruistic motivations and hidden agendas
- Engage in genuine activities to **make the full cycle of data collection, assessment, use and impact transparent** to guard against public mis-trust

- Put steps in place to give assurances on **data quality** determined by factors such as accuracy, representativeness, size of the data set being proportionate to the expected outcomes, method of collection and the impact of data gaps
- Embed **authentic public engagement activities** in the **data assessment and publication process** this could include, for example:
    - A data assessment jury to be drawn on for complex 'edge' assessment cases with, for example an ethical dimension
    - Create a publicly accessible database of approved projects which have gone through the assessment process which can be understood and reviewed by those with no specialist knowledge of the subject
    - Communicate widely, in places and language accessible to the wider population, the impact data use is having on research, innovation and planning for health and social care.
- **Give priority** in data assessment to those who have agreed credentials for demonstrating public benefit
- **Provide a supportive environment** for data use applicants, for example, if their access request is declined data assessors could give constructive feedback including **possible partnerships with other applicants**, suggesting **alternative more appropriate data sets** or giving **guidance on how public benefit might be better achieved** to ensure that great ideas aren't lost because of a mis-judged application.

We anticipate that the NDG will include some of the recommendations in the policy guidance or advice, others will be of more relevance for action by other elements of health and social care systems.

6.3 Concluding remarks

This public dialogue set out to inform the practical guidance to be produced by the National Data Guardian to support those making public benefit assessments. The avenues for exploration with public dialogue participants were clearly established in the project scope (Section 1.1). What emerged from the dialogue is an understanding that participants can see far-reaching and powerful benefits for health and social care data use and wish to see a consistent and rigorous approach to data use assessment. The Putting Good into Practice public dialogue headline findings, to inform policy advice or guidance to be created by the NDG, are divided into three categories:

1. Prerequisites for public benefit

Transparency cannot be separated from public benefit. It is not an add-on or nice to have. Health and social care data use requests *only* demonstrate public benefit if they have integrated communications within their application including activity which demonstrates the value of data use to society

To demonstrate public benefit, transparency is required throughout the whole data life cycle (collection, storage, assessment and use), not just at the point of application

Public benefit is undermined if authentic public engagement is not integrated into data assessment. This requires engaging people from a cross-section of society in data assessment processes.

2. Areas that matter most to dialogue participants

Equitable distribution of benefits of data use in health and social care with safeguards to protect against discrimination and geographic disparities

Identifiable and sensitive data should be treated with the utmost care, if it is, it has the potential to bring public benefit. Data was perceived as being particularly sensitive if it is of a personal nature, such as genomics or mental health data, or because greater care is needed in its interpretation, such as qualitative data

Safeguards and provisions in place to protect society from data manipulation, where the outputs from the data use could be interpreted in different ways, for example, to achieve political or financial ends. This includes publication of statements of data users' credentials and sources of funding

Public benefit must outweigh profit with profitable uses of data rigorously scrutinised for demonstrations of public benefit before access is granted. There is a recognition that data use in this context can enable health and social care improvements and innovations

Being ambitious for health and care data use - to realise public benefit from global collaboration; exploratory research driving breakthroughs; and using profit for new developments, such as drugs, treatments and services.

3. Areas that matter least to dialogue participants

De-identified data being used for purposes beyond the original data collection is broadly acceptable because knowledge gain is seen as integral to public benefit. Recognition of this allows for innovations and developments to result from exploration. However all changes in direction must still be is predicated on the prerequisites for public benefit being in place

The scale of benefits is not a significant factor in determining whether a data use has public benefit as there is inherent value in data use which produces an impact, even if only for a small number of people.

As we conclude this findings report, it is worth stressing four key points:

1. In testing understanding of what people consider to be beneficial about the use of health and adult social care data for purposes beyond individual care we found Putting Good into Practice public dialogue participants both endorsed and amplified what is known. They also brought a range of perspectives on public benefit to inform the NDG's guidance.

2. We began the process with an approach to deliberation focused on key questions related to the National Data Guardian's remit, of identity; scale of benefits or harms; proximity to the original purpose of the data collection; the likely prospect of benefits or harms; and the impacts of the data used on public benefit assessments. All these questions were indeed explored in the dialogue, but we also found how important the systems and infrastructure

around data use are to participants and that exploring those was equally important when reflecting on issues of public benefit.

3. Transparency and trust in the full process of the collection, storage, assessment and use of health and social care data cannot be separated from public benefit.

4. Participants found that data, whether derived from health or adult social care interactions, is a valuable resource which can realise great public benefit now and for future generations if treated with respect and assessed with the voice of dialogue participants in mind.

   Data is precious and should always be treated with respect, ethics, positive outcomes, in the best interest of public benefit. Trust and transparency are essential. Please keep our voices ringing in your ears as the guidance is written. **I** Participant, Plymouth

7. Acknowledgements

Hopkins Van Mil is enormously grateful to all those people from Great Yarmouth, Stockport, Plymouth, Reading and surrounding areas who took part in the public dialogue. Their commitment to the process, their interest in public engagement in the data assessment process, their lively and passionate contributions have been much appreciated. We also thank the roundtable participants who helped us ensure that our use of language and design tools were appropriate for discussions with lay people with little or no specialist knowledge of health and social care data before embarking on the dialogue journey. The fact that all these participants gave their time in such challenging circumstances is particularly noteworthy.

Many thanks too to the specialist stakeholders who agreed to be interviewed in the design phase of the project. Dame Fiona Caldicott, the National Data Guardian, was one of our first interviewees. She was incredibly generous with her time and showed great concern for the wellbeing of all those involved in the dialogue, particularly as it was to be conducted during the COVID-19 crisis. She specifically asked HVM Directors to take care of the participants and the facilitators in these complex and challenging circumstances. We are immensely proud that this project will be part of her legacy in seeking the views of wider society on trustworthy uses of data.

The full list of interviewees who contributed to the shaping of this public dialogue are:

Name	Role	Organisation
Adam Steventon	Director of Data Analytics	Health Foundation
Ian Hulme	Director of Regulatory Assurance	ICO
Sarah Markham	Patient Rep	Use My Data
Dr Tony Calland	Chair	Confidentiality Advisory Group (CAG)
Stephanie Coughlan	Clinical Lead for Innovation	RCGP
Dr Chris Bunch	Chair	UK Caldicott Guardian Council Oxford Universities Hospitals Trust
Phil Booth	Coordinator	MedConfidential
Dam Fiona Calidicott	The National Data Guardian	
Prof. Tim Hubbard	Professor/ Associate Director	HDR UK (&King's College London/Genomics England)
Claire Sutton	President	National Care Forum
Dr Cait Taylor	GP Rep	Caldicott Guardian Council
Kirsty Irvine	Chair	Independent Group Advising on Release of Data (IGARD)
Ian Turner	Chair	Registered Nursing Home Association
Dr Peter Short	Clinical Lead GP Data for Planning and Research	NHS Digital
Rhian Hortin	Communications and Marketing Manger	Clinical Practice Research Datalink (CPRD)

The Oversight Group, chaired by Vicky Chico, were exceptionally generous with their time and in applying their expertise, giving guidance, challenge and support at each step in the process.

Name	Role	Organisation
Richard Ballerand	Digital Health Engagement Lay Adviser	Health Data Research UK IDG Board

Sophie Brannan	Senior Ethics Adviser	British Medical Association
Dr Vicky Chico (Chair)	Data Policy Adviser	Health Research Authority
Gary Cook	Deputy Director	Office for Life Sciences
Amy Darlington	Executive Director	Imperial College Health Partners
Kirsty Irvine	Chair	Independent Group Advising on Release of Data (IGARD)
Dr Indra Joshi	Director of Artificial Intelligence	NHSX
Jacob Lant	Head of Policy, Public Affairs, Research and Insight	Healthwatch England
Philippa Lynch	Senior Data Analyst, Care and Health Improvement Programme	Local Government Association
Jenny Firth	Acting Deputy Director, Data Policy	NHSX, Department of Health and Social Care
John Marsh	Advisory Group Member	Use my data
Ben Moody	Associate Director of Health and Public Services	TechUK
Annemarie Naylor	Director of Policy & Strategy	Future Care Capital
Dr Joseph Savirimuthu	Senior Lecturer in Law	University of Liverpool
Ian Turner	Chair	Registered Nursing Home Association
Prof. James Wilson	Panel Member and Senior Lecturer, Department of Philosophy	National Data Guardian University College London

Throughout the dialogue evidence and information was given to participants in case studies and presentations. This range of perspectives, and the opportunity for participants to discuss them, is an essential part of a Sciencewise dialogue. We are very grateful to the specialists listed who gave their time and expertise to the process.

Name	Role	Organisation
Jenny Westaway	Head of the Office of the National Data Guardian	National Data Guardian
Dr Natalie Banner	Lead for UPD	Understanding Patient Data
Kirsty Irvine	Chair	IGARD (Independent Group Advising on Release of Data)
Chris Carrigan	Expert Data Advisor	Use MY Data
Fionnuala Ratcliffe	Dialogue and Engagement Specialist	Sciencewise
Dr Janet Valentine	Director	CPRD
Dr Peter Short	Clinical Lead for GP data	NHS Digital
Richard Welpton	Senior Data Manager	The Health Foundation
Ellen Coughlan	Programme Manager	The Health Foundation
Richard Brine	Senior Data Manager	The Health Foundation
Dr Sarah Dougan	Chief Analytical Officer	London Borough of Islington
Anne Marie Naylor	Director of Policy & Strategy	Future Care Capital
Dominic Cushnan	Head of AI Imaging	NHSX
Ian Turner	Executive Chair	Registered Nursing Home Association
Brhmie Balaram	Head of AI Research & Ethics	AI Lab, NHSX

HVM actively seeks to work on projects where close co-design with the commissioning partners and project funders is possible. This project was a delight to work on with the Project Team's expertise, guidance and calm approach in such an extraordinary year. As such we are grateful to have had the opportunity to work in close collaboration with Jenny Westaway and Rajoo Veeren, Office of the National Data Guardian; Natalie Banner, Tom Harrison and Grace Annan-Callcott, Understanding Patient Data; Fionnuala Ratcliffe, Sciencewise; Philippa Lang, UK Research and Innovation; and Helen Fisher and Rhuari Bennet, 3KQ who have provided independent evaluation to the process.

Hopkins Van Mil report authors

Henrietta Hopkins, Director
Suzannah Kinsella, Senior Associate
Grace Evans, Researcher
Sophie Reid, Associate

Hopkins Van Mil
6a Dean’s Yard
London SW1P 3NP
info@hopkinsvanmil.co.uk
www.hopkinsvanmil.co.uk