DARE UK

# Building a trustworthy national data research infrastructure: A UK-wide public dialogue

May 2022

ADR UK

HDR UK Health Data Research UK

UKRI UK Research and Innovation

Kohlrabi Consulting

# DARE UK


## Authors

**Dr Fran Harkness** Kohlrabi Consulting

**Dr Jo Blodgett** Kohlrabi Consulting

**Cornelis Rijneveld** Kohlrabi Consulting

**Elizabeth Waind** DARE UK

**Michelle Amugi** DARE UK

**Fergus McDonald** DARE UK



## Acknowledgements

The authors would like to thank the members of the public who participated in this dialogue for sharing their views with the research team. We would also like to thank the scientists, facilitators and note takers who gave their time to the workshops, the Community Researchers for their support with recruitment, and the members of the Oversight Group for their guidance on the dialogue design and this report. Thank you also to all others who provided their input and feedback during the process.

DARE UK (Data and Analytics Research Environments UK) is funded by UK Research and Innovation (UKRI), the UK's largest public funder of research and innovation. DARE UK is a multi-phase programme, with Phase 1, 'Design and Dialogue', led by Health Data Research UK (HDR UK) and ADR UK (Administrative Data Research UK).


ADR UK

HDR UK Health Data Research UK

UKRI UK Research and Innovation

Kohlrabi Consulting


DARE UK 2022. DOI: 10.5281/zenodo.6451935

# Foreword

**Trustworthiness is integral to the way in which the UK's data research infrastructure must be managed. It is the guiding principle that should run through all elements of its operation, from the security procedures in place to protect data, to the governance processes enabling access to data for research.**

This is essential if the public is to feel confident that their data is kept safe and secure and used appropriately. And in order to do this, we must understand what trustworthiness means to the public.

Furthermore, if the true value of data for public benefit is to be realised, the public must be taken along the journey and invited to contribute in a meaningful way. They should be part of the conversation, and their needs and interests should be properly understood and addressed in data research processes.

Phase 1 of the DARE UK programme, which is led by Health Data Research UK (HDR UK) and ADR UK (Administrative Data Research UK) and runs from July 2021 to August 2022, is all about speaking to stakeholders – researchers, technologists, the public and others – about what's working well and where there are unmet needs relating to the use of sensitive data for research. This Public Dialogue forms a key element of this, and I am delighted to see a clear set of recommendations emerge from our discussions with the public.

We will now be taking these recommendations forward in our ongoing work to make sure research using sensitive data is done in a way in which the public can trust, and in a way which is designed to meet the needs of different groups and communities up and down the UK. I hope the findings of this Dialogue will also provide valuable insights to the wider data research community regarding how trustworthiness can be demonstrated in the handling and use of sensitive data.

**May 2022**

**Hans-Erik G. Aronson**
Director of DARE UK Phase 1

# Executive Summary

**In January and February 2022, DARE UK (Data and Analytics Research Environments UK) and Kohlrabi Consulting carried out a public dialogue to explore views towards how the UK's data research infrastructure could work in a more joined-up, efficient and trustworthy way.**

44 members of the public from across the UK took part in a series of online deliberative workshops, and their discussions revealed six areas of public interest. The findings have been worked into six recommendations for actions that could be taken forward to ensure the views captured are appropriately addressed in the implementation and management of a more coordinated national data research infrastructure.

## Recommendations:

**1** **Proactive transparency should be practiced by those handling and using sensitive data for research**

Data custodians and researchers should go further than making information about sensitive data use in research publicly available to those who seek it out. Clear, understandable and engaging information should be proactively brought into people's lives through a range of accessible channels to raise general awareness of data research practices.

**2** **Public involvement and engagement should be inclusive and meaningful**

At all levels, public involvement and engagement with research using sensitive data should be inclusive, from its design through to its recruitment and reporting. A diverse and inclusive public should be involved in initiatives, and this requires proactive and targeted outreach, particularly for groups currently left out of activities. Furthermore, involvement and engagement should be meaningful – the public's input should be sufficiently informed, documented and addressed, and the those handling and using sensitive data should be upskilled to have these conversations in an effective way.

**3** **Efforts should be made to raise awareness of security processes to protect data, and make sure those processes remain fit for purpose**

Considerable resource should be dedicated to increasing public understanding of the safety and security processes in place to protect sensitive data for use in research. However, public trust in this security should not be assumed: sufficient safety nets such as an independent monitoring body and considerable monitoring of individual researchers using data should be maintained. Security processes should also be regularly reviewed to ensure they continue to be fit for purpose as technology advances.

**4** **The processes and systems supporting data research across the UK should be unified in their approaches where possible**

The four nations of the UK should be unified in their use of sensitive data for research where possible within the legal frameworks, and while maintaining flexibility in response to country-specific needs. For example, they should take standardised approaches to access, accreditation and data security and governance, and these approaches should be agreed in collaboration with multiple members of the public from each nation.

**5** **Where feasible, processes enabling access to sensitive data for research should be standardised and centralised**

To maximise public benefit from sensitive data research, data access processes should be centralised and standardised across the UK where possible. This would help to ensure adherence to security and ethics best practice, and improve the efficiency of data access so that research is not unduly delayed by differing complex processes.
The type or perceived sensitivity of data should not affect the procedures for data access – as long as the data is held safely and securely.

**6** **Sensitive data should be made available for research when it is in the public benefit**

If public benefit is established as the main motivation – i.e. when assessed via an independent panel which includes members of the public – and the end-to-end processes surrounding data use are well monitored and transparent, sensitive data should be made available to approved researchers.
This should not be limited to particular types of researcher and should include those affiliated with academia, industry or government.

# Contents

# 1 / Introduction

**DARE UK (Data and Analytics Research Environments UK) is a programme initiated by UK Research and Innovation (UKRI) – the UK's largest public funder of research and innovation – to design and deliver a national data research infrastructure that is joined-up, demonstrates trustworthiness and supports research at scale for public good. The scope of DARE UK covers linkage and analysis of sensitive data from across different sectors – including data relating to education, health, the environment and much more.**

During Phase 1 of the multi-phase DARE UK programme – 'Design and Dialogue', which runs from July 2021 to August 2022 – a key element is exploring public views towards how the UK's data research infrastructure could work in a more joined-up and efficient way, and particularly how trustworthiness can be demonstrated in the way it is run and managed.

This public dialogue builds on at least a decade of public involvement and engagement activities exploring how the public feel about the storage, linking, and sharing of sensitive data for research (Aitken et al. 2019, Atkinson et al. 2017, Elias 2021, Kispeter 2019, Scott 2018, Waind 2020).

Previous literature has established that, broadly, the public is supportive of sensitive data research practices as long as the data is kept secure and privacy is protected, the research is carried out for public benefit, and there is transparency around all data processes between data custodians, researchers and the public.

However, public support should not be taken for granted; research has found that the specifics of any given project have an impact on public expectations of the processes governing data use in research (Waind 2020). The importance of continuing to understand what the public want – particularly since the beginning of the COVID-19 pandemic and associated developments in how data is accessed and used for research – and meeting their expectations, cannot be underestimated when making use of sensitive data.

# Definitions

## For the purpose of this report, the following key definitions are used:

### Data custodian
A data custodian is the person or organisation who is responsible for holding sensitive data and keeping it safe and secure.

### Data research
Sensitive data is often used for research to better understand how society works. For example, researchers analyse sensitive data to develop new insights into issues such as education, health, the environment, the economy and more. These insights can help to inform public policy and services.

### Data research infrastructure
Data research infrastructure refers to the systems and processes in place to support data research. It includes physical systems, such as the data centres where the data itself is held; the computer software that researchers use to analyse data; the governance that exists around it; and the people who run the systems and do the research. It is everything that makes data research happen.

### Public involvement and public engagement
Public involvement refers to activity which captures and addresses the views and concerns of the public. The primary goal of public involvement is for activities to be carried out 'with' or 'by' the public, rather than 'to', 'about' or 'for' them; it is to seek public input and make sure it is taken on board.[1] Public involvement can exist at different levels, from exploring views regarding a topic or issue, to involving the public in decision-making.

Public engagement is the dissemination of information to the public in a forum in which questions can be asked and views expressed. The primary goal of public engagement is to offer a space for information sharing and dialogue.

### SafePod
A SafePod is a standardised safe setting that provides the security and controls for data that requires secure access for research. A SafePod includes a door control access system, CCTV, a researcher area for dataset analysis, secure IT cupboard and a height adjustable desk.[2] SafePods are able to connect to TREs in different locations via a secure remote connection.

### Secure remote connection
It is possible for researchers to connect to a TRE via a secure remote connection. This allows researchers to securely access data held within a TRE from a convenient location, instead of being physically present within the TRE building. When accessed via a secure remote connection, data can be viewed on the researcher's computer screen but not downloaded to the computer – the data itself remains in the TRE, and the researcher accesses the system via a virtual desktop. Strong security controls protect the data, and access is monitored by TRE staff.

[1] NIHR (2021). 'Briefing notes for researchers – public involvement in NHS, health and social care research'. Accessed 04/04/2022.

[2] Safepod Network homepage, accessed 11/04/2022.

## Sensitive data

Sensitive data includes data which contains personally identifiable information such as names, addresses and identifying numbers. This data can still be sensitive once it has been de-identified (has had all personal identifiable information removed), particularly if there is potential for re-identification when used with other data. Commercial data such as retail information, business details or confidential product details may also be considered sensitive when used for research.

## Trusted research environment (TRE)

A trusted research environment (TRE) is a digital system that securely holds and provides access to sensitive data for approved researchers. The data does not leave the TRE, and strict security measures protect the privacy of the people the data is about. TREs significantly reduce the potential for data misuse or the possibility of re-identification of data which has been de-identified (had all personal identifying information like names and addresses removed).

# 2 / Objective

**This dialogue sought to explore the views of the UK public regarding the design and delivery of a more joined-up, efficient and trustworthy national data research infrastructure. It aimed to deepen public conversation around data research practices on a national scale, and, by taking a deliberative approach, sought to capture tangible actions that could be taken forward by those holding and using sensitive data for research, in order to address public views.**

# 3 / Methodology

## Deliberative approach

Deliberation was considered the best methodology to equip the public with the necessary understanding to explore this topic, as per recent reviews of public involvement and engagement best practice in data research (Aitken et al. 2019; Jones et al. 2020). Deliberation emphasises logically building people's understanding about a topic, drawing out multiple perspectives and trade-offs rather than driving at a consensus and allowing time for views to be expressed and developed.

An initial deliberative workshop was planned to repeat five times in January 2022: once in a location in each of the four nations of the UK, and once online for those unable to attend in person. Due to public health restrictions linked to COVID-19, however, the workshops were all moved online and held over two days across Thursday 13 January and Friday 14 January.

Of the original participants who were recruited, only one could not take part on the new dates and was replaced from a reserve list. Minor adaptations to the original workshop design – such as incorporating the use of a digital whiteboarding tool – were necessary to enable discussions to as far as possible be as effective in an online forum as they would have been in the planned in-person format.

A single follow-up workshop was then held online on Tuesday 22 February with a cross-section of 10 participants from the two initial workshops. The purpose was to check that analysis of the initial workshops had accurately captured participants' views and expectations, and to bring those expectations to life through discussion of tangible actions that could be taken forward by the data research community to address them.

## Oversight Group

The deliberative workshop design was informed by an independent Oversight Group made up of 13 professionals working within the field of public involvement and engagement with research and data, plus one public contributor. The purpose of this group – which met three times over the course of the Dialogue process – was to provide input and advice on the workshop design and on the interpretation, contextualisation and implementation of findings. For a list of members, see Appendix 1.

# Recruitment

50 members of the public were recruited to take part in the deliberative workshops. On the day of the workshops, six people dropped out or did not attend, resulting in a final total of 44 participants.

To be as inclusive as possible and bring together participants from a diversity of backgrounds and identities from across the different countries and regions of the UK, a local, community-based recruitment approach was adopted. This involved advertising the workshops using mixed methods to suit a broad range of people. The recruitment methods used can be seen opposite.

In addition, to capture a range of different views on the subject – including both spontaneous/uninformed views and, as the workshops progressed, informed views – it was important that participation focussed on people who had not previously taken part in involvement, engagement or professional activities related to the use of sensitive data in research.

At least 10 participants were recruited from each of the four nations of the UK. Participants were asked to confirm that they were over 18 and had not previously been involved in public involvement and engagement with data research; other than these criteria, participants were accepted as they signed up.

**'Community Researchers'**
assisted with workshop recruitment in the four cities the face-to-face workshops would have been held. These researchers – who were themselves recruited through local charities – were lay people embedded in their local communities. They tailored their recruitment strategies to their localities and to a diversity of demographics, speaking face-to-face with people on the street.

**Fliers were distributed**
in local libraries, community housing noticeboards, GP surgeries, transport hubs, grocery shops, barbers and hair salons, restaurants and faith organisations.

**Fliers were posted**
to community groups and public services outside of the major cities to be distributed to local people.

**Social media platforms**
Facebook, Instagram and Twitter were used to reach rural, coastal, and further afield communities through community and common interest groups.

In recognition of their valuable time and input, participants were offered a £150 digital voucher as a thank you for the initial, full-day workshops, and those who took part in the half-day follow-up workshop were offered a further £75.

All 50 members of the public who originally signed up to participate were asked prior to the initial workshops to complete a voluntary, anonymous demographic information form. Not all participants completed the form and/or each question, and we cannot be sure which demographic backgrounds the six people who ultimately did not attend belonged to. The complete demographic characteristics of those who responded can be seen in Appendix 2.

## Initial workshops

In each of the two initial workshops, participants were grouped across five virtual 'breakout rooms' for discussions, each with four to five participants. On the Thursday, participants were grouped by resident nation, whereas on the Friday they were grouped into UK-wide breakout rooms with at least one participant from each of the four nations. This was based on the original workshop structure, which would have involved one face-to-face workshop in each of the four nations and one UK-wide online workshop. Each group had a facilitator and a note taker and was able to ask a scientist to join the room for more information or to ask questions if needed. The workshop was devised in a series of cycles designed to 'build understanding' of a number of topics, followed by interactive activities and group reflection time to enable exploration of a set of key questions related to each topic. The rationale for these questions was based upon:

1. **DARE UK's mission** to design and deliver a joined-up, efficient and trustworthy national data research infrastructure, and the need to understand public views regarding how this should be done.

2. **Existing gaps in knowledge of public views.** As discussed, we know from previous research that the public broadly supports data research if certain conditions are met. However, there is less evidence regarding views towards methods of data access for approved researchers, and there is also a need to explore whether views may have changed over time, particularly since the beginning of the COVID-19 pandemic and associated developments in how sensitive data is accessed and used for research.

The topics and questions explored are outlined on the next page. Please note that each of these questions was not directly asked of participants; rather, the workshop activities were designed to explore the overarching topics and questions in a deliberative fashion.

## Key questions explored in the deliberative workshops:

| Topic | Questions explored |
|---|---|
| **Trustworthiness** | What would enable you to feel trust in the UK's national data research infrastructure?<br>How can trustworthiness be demonstrated in the way sensitive data is managed and used in research?<br>What does your trust most rely on? |
| **Access and accreditation** | Who should be granted access to sensitive data and what accreditation processes should they have to go through?<br>What uses/projects should be granted access to sensitive data?<br>What methods are you most comfortable with for accessing sensitive data? How do you feel about remote access to a secure environment?<br>Have your views towards sensitive data access been affected by the pandemic?<br>How do you feel about data being shared across the four nations of the UK? |
| **Balancing risks and benefits** | How should the risks of data use be minimised?<br>How can the benefits of data research be effectively articulated to the public?<br>How can we ensure that data research directly and positively impacts the people the data represents?<br>How do you perceive the risk-benefit tension? Are there tiers of risk, for example for different types of data or research project? |
| **Public involvement and communications** | How do you think the public should be involved in data research processes?<br>How should we keep the public informed about data research and the infrastructure that supports it?<br>What sort of language should we use/avoid when describing data research and its infrastructure to a public audience? |

Short presentations from researchers and technologists, along with time for questions and discussion, were designed to build participants' understanding of the key concepts involved in the use of sensitive data in research. A full list of speakers and presentation topics can be seen in Appendix 3. Interactive activities were designed to explore the questions above, draw out the potential trade-offs each issue brings about, and explore multiple perspectives. Participants were given time and encouragement to talk to each other and help each other reflect about each issue, as well as being prompted by the facilitator.

The workshop activities were produced by Kohlrabi Consulting and the DARE UK team, before being refined in response to feedback from the team of community researchers. Facilitators followed a script to ensure the topics were covered consistently across all breakout rooms on both days. We have made the workshop scripts and presentations available on the DARE UK website, in case they are of use to other researchers exploring public attitudes towards sensitive data use.

## Follow-up workshop

A single follow-up workshop, which took place five weeks after the initial workshops, was designed to check that analysis of the initial workshops had accurately captured participants' expectations for a national research infrastructure for sensitive data, and to bring those expectations further to life. 10 participants from across the four nations were invited back to attend this workshop, with one participant invited from each of the initial workshop breakout rooms.

The aim of the final workshop was to help shape a set of tangible actions – or recommendations – to be taken forward based on public expectations regarding the implementation of a more joined-up, efficient and trustworthy national data research infrastructure.

## Analysis

Following the initial workshops, the note takers' documents were coded and themes were identified in a separate document in the form of code clusters. Relevant quotes and excerpts from the notes were grouped under these code clusters, along with a brief summary of the themes and sub-themes that highlighted any tensions and contradictions.

Validity of the extracted themes was checked by a second researcher using workshop transcripts and recordings, and validation was then triangulated with feedback from each breakout room facilitator.

At the follow-up workshop, the extracted themes were presented to participants to ensure they accurately represented their understanding of the initial workshops. Participants were given time to discuss whether they felt the overview was accurate, or to further deepen conversations. The facilitators then prompted participants on actions that could be taken forward in response to their views.

# Strengths and limitations of the approach

Despite starting the day with little knowledge of the use of sensitive data in research, participants appeared engaged during the deliberative activities and many expressed enjoyment of the experience. The findings from their conversations build on previous published literature to allow DARE UK to better understand how to build and maintain a national research data infrastructure in a way which aligns with public views.

Several aspects of the approach taken should be kept in mind. The mixed recruitment methods achieved a broad sample of members of the public who acknowledged that they would not ordinarily have seen the invitation to be part of public involvement and engagement. However, the views expressed in this report may not be reflective of those of the wider UK public – a sample of 44 people cannot represent the diversity of views across England, Northern Ireland, Scotland and Wales. Moreover, the demographic information of the sample is incomplete, as many participants did not contribute their personal characteristics.

Deliberative activities helped participants gain an understanding of the relevant processes and trade-offs over the course of the workshops, thereby enabling them to give views. However, by the end of the process this also made the sample slightly less representative of the general public, who may not have the same level of awareness.

There is always the chance in group discussions for a tendency towards agreement rather than disagreement, which might obscure the variety of views in a given group. Having several different sessions and breakout groups was intended to mitigate the impact of this on the findings. For those invited to participate in the follow-up workshop, this was also on the basis that they had taken part in a variety of different breakout rooms during the initial workshops, to make sure the spread of viewpoints expressed and discussions had during the initial workhops was covered.

Concerns about tokenism in public involvement were expressed by workshop participants, and although these were not directed at DARE UK per se, we try to address these concerns by keeping participants and the wider community informed at all stages about how their involvement is affecting DARE UK Phase 1 outcomes. Specifically, how their input is feeding into wider recommendations to UKRI regarding mechanisms that could be taken forward in the development of a more joined-up, efficient and trustworthy national data research infrastructure.

# 4 / Findings

**Participants appeared to have several expectations relating to the use of sensitive data in research which they returned to repeatedly and confirmed in the follow-up workshop. These key expectations have been grouped into the six themes discussed below.**

## Transparency

The large majority of participants started the initial workshops with low understanding that research takes place using sensitive data for research, and were largely unaware of the procedures governing that data use. As participants' understanding grew throughout the workshops, they demonstrated an acceptance of sensitive data use when this was in the 'public benefit', and were positive when hearing examples of research findings and their potential implications for society. The message from participants was that they wanted what they had learnt during the workshop to be common knowledge.

Based on their own understanding, participants felt society in general had low awareness of the processes surrounding the use of sensitive data in research. Most participants recommended ongoing generalised awareness campaigns (including covering the existence of trusted research environments (TREs) and other secure data infrastructure) through health service professionals, the education system, social media, print media, television and radio.

Participants emphasised that everyone needs access to this information, so both the messaging itself and the mode of access to that messaging must be tailored to include different people. Common suggestions for such efforts were to incorporate translations in engagement materials, and for the materials to "not overload people", be too complicated or "need a law degree to read".

In addition, multiple participants voiced concern that some people might not be included or able to access a general awareness campaign. People agreed that those handling and using sensitive data for research need to make proactive efforts to reach specific publics. Their examples were people without access to the internet, people who don't have much interaction with or trust of public services, and those who are geographically isolated.

The reasons cited for greater efforts in transparency were largely to do with gaining public trust and helping people feel in control of what their data is used for.

> *"There's a gap between the great work being done and what people actually know. People are badly informed."*
>
> Workshop participant

Some participants were explicit that if people don't know sensitive data is used and how it is used then they can't help to make decisions governing it. They wanted to know what sensitive data is collected from the public, how and where it is stored, the technologies involved in data privacy (such as de-identification), and how researchers access the data, right up to being informed of the intended findings and societal implications of the research.

Making sure that the findings and implications of data research reach the public seemed particularly important. There was a perception that the people whose data is used in research are forgotten by researchers once the data has been studied. Several participants re-told experiences where they had been assured that they would be fed back to after interactions with research teams or public services, but had not received anything. Some wanted the findings of research brought to them so they could use them in work or charity endeavours; some just wanted to feel that the research was being used to improve society at large.

A few participants seemed keen to have more transparency about sensitive data use on an individual level, such as being notified each time their data is used for research, or where it is going and being held. Some wondered whether the public understand that their data will be used for research when it is collected by public services – this was linked to a lingering generalised sense of low control over their data, largely based on negative experiences with marketing companies and the perception that online advertising was targeting them. Some participants suggested that more information about sensitive data use in research should be given at the point at which people 'give' that data – in other words, when they connect with a public service. A few people suggested the public should be able to opt out of providing their data for research.

In the follow-up workshop, participants returned to the need for data custodians and researchers to raise public understanding of sensitive data use.

Participants spoke a lot about their desire for public benefit to be brought about through research using their de-identified data (data which has had all personal identifying elements such as names and addresses removed); they did not want to opt out of their data being used. However, they keenly wanted the research and data practices to be brought to their attention more frequently, and with more effort to reach them and be understood by them.

As well as confirming the need for general awareness campaigns to make sensitive data use visible and understandable, participants re-iterated that they would like someone trusted, like GPs, to let them know that their information can be used for research to benefit society.

# Public involvement and engagement

Participants stressed the importance of data custodians and researchers making more efforts to come to them with information and ways for them to be involved in research using sensitive data. Before the workshop, almost all were unaware that public involvement and engagement exists in research. There was a sense of amusement verging on exasperation that public involvement opportunities are advertised on organisations' websites or Twitter, which members of the public may not naturally see. The sense from participants was that public involvement and engagement need to be far more proactive, that data custodians and researchers need to come to where the public is, rather than expect the public to come to them.

As to what form this involvement and engagement might take, participants had many suggestions. In theory, they felt it was important that members of the public represent the public voice on panels for decision-making. They wondered if those members of the public could also canvas opinion from communities they have ties to. However, they were not certain that they themselves would want to join a panel. They had the sense that the application process would be challenging and asked for the process to be simple, the opportunity itself to be shorter than they presumed (imagining having to sit for 4-5 hours), and their time remunerated.

## Diversity and inclusion

The desire for diversity and inclusion in public involvement and engagement – in which a variety of backgrounds, identities and viewpoints are included in outreach and activities – came up frequently.

Participants had a sense that there is not representation of all people living in the UK in engagement and involvement activities; for example, of all nationalities, ethnicities, ages, socio-economic positions and interests. They did not want anyone to miss out on potential efforts that might be made to increase public understanding of the use of sensitive data in research, but also felt that some people's views on it had not yet been heard.

Several participants were concerned that people who don't have access to the internet, or who are geographically or societally isolated, will be left behind. The repeated insistence for young people to be included in giving their views, or to learn about the use of sensitive data in research at school, was almost an ask for the next generations to grow up with clarity and involvement.

Multiple participants spoke about how their views on the sensitivity of data have been shaped by their experiences as a minority or disenfranchised group in the UK, such as Black, Asian, Muslim, working class, migrants or refugees, or politically opposed to their government. Most breakout groups mentioned different levels of trust regarding sensitive data collection across the communities individual participants had ties to, and suggested there was a need for better efforts to engage those people and incorporate their views.

> *"Some ethnic minorities are too scared to give information. This is something I've noticed in our community* – 'don't write that, don't give that information'."
>
> Workshop participant


Differences in the perception of the sensitivity of data were explored. Several participants either worked with refugees or had lived experience of being a refugee or migrant. The consensus in their groups was that if you are living in fear that your data may be used for life-changing decisions against you, then that data is far more sensitive than it is to people who are living in stability. They wondered if that data should not be used, or if it should be subject to more assurances, as more trust is needed. There wasn't an answer, but they thought that it should be explored further with those groups themselves.

Participants expressed a feeling that the public's views might not be used in a meaningful or genuine way. For example, one or two participants had experienced some form of public consultation where they felt they had been included as a 'tick box' exercise. This conclusion was largely based on the frustration of not getting fed back to about the findings or downstream societal implications of the consultation. Annoyance at inclusion not feeling genuine reflected some participants' feeling that the public was 'talked down to' by researchers.

'Classism' came up as a perceived issue along with the suggestion that this leads to research not serving working class people. A sense of separation between the general public and academia was not disagreed with by other participants.

One participant expressed dislike for researchers referring to ethnic minorities as "hard to reach communities". They suggested that the problem is that researchers come in and "take what they want", and the community involved never sees the benefit of being involved. People felt that more genuine dialogue could happen if researchers are more transparent and welcoming, and if participants in public involvement and engagement activities are given the requisite understanding to take part and make decisions.

## Recruitment

In terms of reaching members of the public for public involvement and engagement, there was reiteration of people being unique within their diversity. They therefore felt that researchers need to use mixed methods of communication within their target groups.

Offline communication was highly regarded. Many participants wanted researchers to come directly into people's communities – for example, talking to people on the street, distributing fliers and using public noticeboards where the public can put their research concerns and priorities. They knew it would take effort, but they felt it was needed. Several participants felt that local councils who understand and work with their local communities could help to meaningfully increase inclusion.

Some participants favoured social media, or a database they could sign up for to receive newsletters and offers of participation. Participants in the follow-up workshop agreed that they would be happy to be part of a central database – independent from research bodies – where people willing to participate in public involvement and engagement can sign up and specify what areas of involvement they are interested in. They felt that the portal could be used to keep participants updated about research progress and outcomes, perhaps in a newsletter, which would again address concerns around tokenism. There was agreement across the board that communication should be 'quick and snappy'. Participants in the follow-up workshop wanted to know about involvement opportunities, but felt they did not have much time to read about them.

Workshop participants desired a targeted approach to rectify the gaps in inclusion that many of them saw. Many, but particularly those themselves from ethnic minority backgrounds, spoke about the need for much more time and effort from researchers to reach and involve older members of their communities, as well as those with negative experiences with, or lack of connections to, the government and public services.

A frequent suggestion was to bridge gaps between researchers and members of the public by building relationships with trusted community members, visiting physical locations (faith organisations were mentioned) and conducting public involvement using translators and trusted community members. It was repeatedly acknowledged that these efforts would take time, but there was consensus that time was needed to build trust. In the follow-up workshop, one breakout group recommended targeted advertising on protected characteristics to rectify perceived historical lack of inclusivity for some groups.

At the same time as asking for targeted efforts to rectify imbalances, participants agreed that no individual is representative of a whole group and should not be treated as such if public involvement and engagement is to be meaningful. Multiple participants said that as Black or South Asian people (in their particular cases) they had the perception of being "put in a box" by researchers – they did not want to just be sought out through community or faith groups, and be used to satisfy criteria. Participants of the follow-up workshop who wanted inclusion improved with targeted advertising cautioned against only using stereotyped channels; they emphasised that people have lives and interests away from their demographic information.

# Data security and access

Conversations about data security highlighted confusion and feelings of a low sense of control surrounding sensitive data more generally. Participants expressed not knowing where their data was or who was using it, with almost a sense that they had lost it and could not get it back. Use of sensitive data for research was initially conflated with practices like market research, targeted online marketing and selling people's data to industry.

The majority of participants expressed positive surprise on learning about how sensitive data was securely stored and handled for use in research. As understanding of the layers of security to protect the data used in research grew, many participants who initially expressed little trust in data security acknowledged that they were reassured. Sharing of sensitive data was considered acceptable as long as it was de-identified. Participants who remained unsure acknowledged that their concerns had been shaped by a perception that governments might distort statistics based on sensitive data collection for potentially harmful political gain.

## Data security and governance

Despite most participants expressing reassurance with safety procedures, the group frequently returned to mention of hacking and data loss.

There was a sense of mystery around data breaches or misuse, whether they occur, how frequently, and what the consequences of those incidents are for the public and the people involved. Participants said they wanted to have the same information about these issues as those handling and using sensitive data.

At least one participant per breakout group brought up a suggestion of how data could be misused by individual researchers. Some breakout groups said the suggestions were unlikely, calling them "James Bond plots". However, amidst amusement and verbalised reassurance, the conversations kept circling back to a low sense of control about data in general. The lingering concern appeared to be about individual researchers and what they might do with the data after they have been granted access.

To manage their concerns, participants wanted rigorous vetting of researchers requesting access to data, and close monitoring of what the researchers are doing with the data. They also placed trust in 'systems' over individual people and asked for regular reviews of governance structures and security updates, which they felt would restrict misuse.

> *"What happens when researchers breach the rules around sensitive data? Feels very cloak and dagger."*
>
> Workshop participants

More general participant suggestions of how to improve the public's sense of control over their data, including an independent monitoring body for those handling and using sensitive data and a data misuse helpline, were met with approval by other participants. Participants also felt that transparency could play a role in alleviating this sense of low control, particularly for future generations, by teaching in school. In the follow-up workshop, the hope was verbalised that young people might translate their understanding to their parents and grandparents.

## Data access

The different methods of data access for approved researchers, from access in a secure room in TRE buildings, to access via a secure remote connection to the TRE, were presented to participants. There was consistent support across the breakout groups of the different ways in which sensitive data is securely accessed by researchers.

Participants felt it is safest when researchers access data in a secure room at the TRE where the data is held. They felt access in a safe room at an approved university was the second best option, followed by access via a SafePod based at a university. Researchers accessing data on their laptop over a secure remote connection – either in their office or at home, as opposed to in a secure room – were seen as the most vulnerable to accidental or intentional data loss or misuse. While some participants were deeply resistant to this sort of remote computing, others felt it is secure enough and makes sense in the context of the new working from home culture.

In the follow-up workshop, the majority of participants were supportive of approved individual researchers accessing data in their homes via a secure remote connection – as long as security features, researcher vetting and monitoring of activity are in place. There was by no means automatic trust of individual researchers, with lingering worries such as researchers being able to photograph data on their screen at home. One or two members were adamant that they would still prefer data to be accessed in a secure physical environment, such as a SafePod.

Again, participants stressed that public acceptance relies on transparency. They agreed that in general they had faith in the data protection safeguards and regulations they had learnt about, but that for trust, they needed transparency about who has access to what data and for what reasons. They felt that public understanding has a long way to go, and that public awareness campaigns and demonstrations of how the security works – perhaps with tours of TREs – would be reassuring to people.

## One nation approach

Participants largely wanted the UK to be unified in its approaches to the use of sensitive data in research, while wishing to be mindful of some unique, country-specific needs and issues.

Some participants from Scotland, Wales and Northern Ireland suggested the need for governance to be consistent and timely across the UK so that their countries didn't get left behind or missed out of activity happening in England. However, there was acknowledgement that the four countries have differences, so some governance processes might need a different approach. For example, several participants from Northern Ireland pointed out that in small communities, people may be more identifiable from their personal characteristics, even without their names, so data breaches may have more serious ramifications than in the rest of the UK.

There was agreement that data sharing across countries would be publicly beneficial, for example to create a greater sample size and improve study quality. Two participants not living in England mentioned that they were unsure about researchers in England "looking at their data". These feelings were based on perceived imbalances between England and the other UK nations as to the benefit of research using sensitive data. However, participants generally approved of sharing data across the UK, and even Europe and internationally if relevant and in the interest of the public good and to reduce redundancy.

Participants in the follow-up workshop appeared satisfied with a national approach to the use of and protection of sensitive data, for example by setting national core values or standards. This came with a caveat for participants from Scotland, who reminded the group of the need to consider the possibility of independence. There was the suggestion to build-in a safeguard to ensure that Scottish political independence would not affect any standardisation which is implemented.

In terms of how these standards should be set up, participants agreed that a panel or governing body with at least one member of the public from each nation should be involved. There was acknowledgement that this may be not as straightforward in Northern Ireland, given strong and differing political and religious views and identities across the country. As for the physical location of a governing body, the participants' assumption was that it would be based in England, which they felt unfair.

Within national standards, participants in the follow-up workshop affirmed that they were also keen to see a localised approach to some aspects of data use and public awareness raising in different nations. They wanted data research tailored to country-specific health and social needs. Several participants also suggested that, if there were a generalised awareness campaign about sensitive data use in research, each country and potentially each local authority should personalise how they take that message to the population.

## Standardisation and centralisation

A desire for centralised procedures around sensitive data use in research came up repeatedly. This feeling appeared to come from two places: wanting to improve the sense of public control over data use; and to speed up research benefit for the public.

In terms of sense of control, many participants expressed a lingering feeling of not being able to visualise where their data is being kept or who is using it and for what purposes. There was acknowledgement that greater transparency on these matters, with active efforts to reach the public with information, would lessen the low sense of control felt around the use of sensitive data in general. Many felt that if procedures were centralised and streamlined then they as the public could understand processes better and therefore feel more confident about data use.

A few participants felt that sensitive data storage facilities should be merged and centralised to hold all data from across the public sector, so they could easily know where their data is being kept and what security is protecting it. Some felt that security would be greater if more data was in one location, whereas others were concerned that this might put it more at risk than if it were held separately. There was interest in adding an independent central regulatory body for the governance of TREs and matters of data security more generally, particularly amongst participants with little trust in the government.

When described the processes for applying for and accessing sensitive data for research, participants were generally very surprised that it takes so long end-to-end. Several participants worried that existing bureaucracy ("red tape") would compromise the value of the research by delaying findings or discouraging research entirely. Many felt that a centralised approach to applying for access to data, as well as for researcher training and approval, would streamline the process to allow public benefit to be realised more quickly.


Some did not want safety restrictions lessened or streamlined, but suggested access requirements could vary depending on the sensitivity of the data and associated risk of data misuse if data were identified, and/or the urgency of the research. This would enable some research projects to happen more quickly than perhaps they do now, but not all.

In the follow-up workshop, there was a move away from suggesting different tiers of access for data of differing levels of sensitivity.

There was stronger agreement with research for public benefit being more easily facilitated through a standardised and streamlined access procedure. Several participants voiced that, as long as the public are aware that data is being kept safe and secure, different types of data should not be subject to differing access requirements.

Participants were prompted to discuss who would be deciding which access measures to streamline and which research would be prioritised in access processes. There was a high level of trust in one breakout group, with agreement that professionals handling and using sensitive data could guide the public on these matters. The other breakout group wanted to include members of the public on a data access panel. However, both groups felt that the public at large did not need or indeed want to be asked each time a study was granted access. What came out as more important to participants was ensuring that the access process was made very clear to the public from the beginning; that at some point public research priorities were listened to; and that there was communication of the risks and potential benefits of each piece of research taking place.

## Who uses the data

Workshop participants agreed that they were comfortable with academics at universities studying sensitive data for research in the public benefit – subject to robust security. There were wide ranging views about trusting private companies and government researchers with sensitive data, from acceptance to deep mistrust of both political and commercial agendas.

There was disagreement over whether people should be able to gain commercially from data. Several participants argued that commercial gain is acceptable if the research is valuable to the public. They reasoned that private access to data maximises the use of sensitive data and can be a win-win situation, and emphasised there is nothing wrong with benefiting financially. They did not see any good reason for differentiating between public and private actors, as long as the data is handled in a transparent, safe and secure manner.

Others worried that public interests will in practice be secondary to commercial – or "vested" – interests, such that research that is seemingly conducted to advance the public good ultimately serves the company's interests. Relatedly, there was suspicion about how accurately commercial companies may present the findings of research using sensitive data when commercial gain is involved.

Many participants contrasted their apprehensions about private use of sensitive data with their acceptance of government use. A sharp distinction was drawn between the private and public sector, with the former inspiring suspicion and the latter trust. Others differentiated between political parties and public institutions, such as the NHS. In general, there was almost universal agreement about the value of research that benefitted the NHS.

> *"Profit should not be a barrier to research which is valuable to the public."*
>
> Workshop participant

At the same time, there was concern about government researchers studying sensitive data and using the findings to progress their agendas. This concern was strongly voiced by participants with lived experience linked to political agendas at home and abroad. Several participants living in both England and Wales felt that people's data had been used by the government to further policies which they felt had increased their experiences of racism and personal distress. Some participants living in Northern Ireland felt suspicious of governments there presenting biased data to society and basing policies on research with that data. Participants as a whole recognised that potential harmful uses of data affect certain groups more negatively than others. There was the suggestion that these angles should be considered when access panels are deciding who gets access to what data and for what purpose.

In the initial workshops, while some participants were opposed to private and government access to sensitive data in principle, others proposed more stringent security requirements for these groups. By the end of the follow-up workshop, however, participants demonstrated trust in the security procedures they had learnt about previously and agreed that there shouldn't be different criteria for different types of organisations accessing sensitive data, as long as it is assessed as being in the public benefit above all. There was a strong sense of wanting sensitive data to be used toward the greater public good.

> *"I don't trust the government. It's not necessarily that the people are not qualified, but I think the agenda that they push doesn't serve me. If it was a researcher or academic, I'd give them what they want."*
>
> Workshop participant

The caveat for commercial or government researchers using data was an insistence on increased transparency from these users – for the public to be told who is using what data, what for and how they are ensuring it is unbiased. One participant pointed out that each organisation involved in collecting, granting access to or studying data are having conversations behind the scenes before they make decisions regarding the data. This reiteration of the need for transparency may link to several participants in the initial workshops expressing the wish to see research outputs from governments and commercial companies being checked for bias.

In the follow-up workshop, one breakout group agreed that the public would like to hear the 'behind-the scenes' conversations so they can be reassured that whoever is using the data is doing so safely and carefully. They felt that they as the public couldn't make decisions or present their opinion about data use and data users without full transparency from data custodians and researchers, including those from within government and industry.

Transparency from the users of data extended to discussions around the 'public good' or 'public benefit' of research. Many of the breakout groups brought attention to the subjectiveness of this phrase – they wanted the public to be deciding what is in the public benefit, and hearing about it, particularly to avoid suspicions that companies or the government might 'spin' that something is in the public interest. By pointing out that research outcomes may not be in the community's interest, these participants raised an important question: who is the imagined 'public' when we speak of the 'public good'? Some participants, whether because of class or race, did not feel that the public researchers were aiming to benefit included them. Those participants did not hear about research findings in their daily lives, and therefore expressed that they did not see or feel their benefit.

> *"Who decides what is in the public benefit anyway? The public should be able to decide if something is in the public benefit or not."*
>
> Workshop participant

# 5 / Recommendations

**The following recommendations focus on tangible actions that could be taken forward to address the views and feelings of dialogue participants regarding the implementation of a more joined-up, efficient and trustworthy national data research infrastructure.**

In general, workshop participants felt both an overarching low sense of control over how their sensitive data is used in research, and a disconnect with the research community. Many expressed feelings of being 'talked down to' when they hear or read about research practices, leading to a sense that greater efforts need to be made to speak to and involve the public in a more meaningful way. The actions suggested here may therefore also help to alleviate these overarching feelings, to enable the public to feel confident in how their data is used for research and avoid a sense of being excluded from data research practices.

## 1 Proactive transparency should be practiced by those handling and using sensitive data for research

Transparency around all processes relating to the use of sensitive data in research – including security and access processes and the goals, outcomes and impacts of data research projects – was seen as essential to demonstrating trustworthiness to the public.

Approaches to transparency should not focus on merely making information publicly available on websites or via existing networks for people to seek out themselves.

Information should be proactively brought into people's lives through mixed methods: for example, via trusted community leaders such as GPs, local councils, and faith organisations, and via social media campaigns and advertising, teaching in schools, on community noticeboards, and by talking to people and distributing information in public areas.

## 2 Public involvement and engagement should be meaningful and inclusive

At all levels, public involvement and engagement with research using sensitive data should be inclusive, from the way it is designed to the way it is recruited and reported.

A diverse and inclusive public should be included in initiatives, from raising awareness of how sensitive data is used in research, to involvement in decision-making. Proactive and targeted outreach should be used to increase inclusion of neglected groups

and make involvement and engagement more representative of the UK, or if particular groups are implicated in the data use or research project.

Public involvement and engagement activities should equip members of the public with sufficient understanding to take part in genuine conversations about data research processes and offer informed input. In addition, researchers should be equipped with the necessary skills to engage with the public and address their concerns in a meaningful way.

Facilitators and researchers using sensitive data should feed back to the public about their findings and the societal implications of their involvement or the inclusion of their data in research.

To ensure inclusion and accessibility, fresh, public-informed methods and efforts to reach people should be used, and there should be appropriate remuneration for participants' time; a simple application process; and opportunities to get involved that do not require extensive time commitments.

## 3 Efforts should be made to raise awareness of security processes to protect data, and make sure those processes remain fit for purpose

There should be greater efforts to increase public understanding of data security processes, such as via ongoing public awareness campaigns and demonstrations of how the security works, perhaps with tours of TREs.

Proactive transparency about data misuse or breaches should be practiced to remove mystery and help the public feel that they are on the 'same side' as data custodians and researchers.

To increase confidence in data security, there could be an independent monitoring body and efforts to increase awareness around who can be contacted in case of individual concerns around data use.

Currently, there is enough faith in the security procedures for researchers to access data at home via a secure remote connection to a TRE – as long as there is strong technological security and governance, vetting of researchers and monitoring of activity.

Security processes should be regularly reviewed to ensure they continue to be fit for purpose as technology advances.

## 4 The processes and systems supporting data research across the UK should be unified in their approaches

The UK should be unified in its use of and approach to data where possible – for example, in its approaches to access, accreditation and data security and governance – while mindful of the unique needs and circumstances of each individual nation.

Data should be shared across the UK, Europe and internationally if secure, relevant and in the interest of the public good, and to reduce redundancy.

Governance should be consistent and timely across the UK where possible, so that countries don't get left behind or missed out of activity happening in other areas.

A panel or governing body agreeing standards across the UK should meaningfully involve at least one member of the public from each nation, and preferably multiple people representing different groups and interests.

There should also be a localised approach to some aspects of data use – for example, regionally or nationally – including identifying research priorities and perceptions of public benefit, as well as when building public understanding of data research.

## 5 Where feasible, processes enabling access to sensitive data for research should be standardised and centralised

To improve public benefit from sensitive data research, many data processes could be centralised and standardised. A streamlined application process could be created to avoid delaying or discouraging important research.

As long as the public are aware that data is being kept safe and secure, different types of sensitive data should not be subject to differing access requirements.

Public representatives from different groups and communities across the UK could be included on access panels.

## 6 Sensitive data should be made available for research when it is in the public benefit

It should not matter who is using the data (for example, people working on behalf of academia, industry or government), as long as there are assurances that the research is for the public good, and the end-to-end process is transparent. The public should be told not only who is using the data, why, and how; research findings should also be openly and proactively communicated so the public can judge the quality of conclusions.

With regards to establishing what data uses are in the 'public good', members of the public should be involved in decision-making. Ensuring that the general public – and particularly the groups of people whose data is being studied – are involved in decision-making and hear about research findings will also improve the potential for benefit to be achieved.

Organisations applying to use sensitive data for research in the public benefit should ensure their conversations "behind the scenes" are also transparent and reach the public. Public support is not static, and without transparency, the public cannot give their views or make informed decisions.

There should be strong governance processes and rigorous vetting and monitoring of individual researchers accessing sensitive data for research.

# 6 / Conclusion and next steps

**This public dialogue has highlighted tangible actions that could be taken forward by those handling and using sensitive data for research, to ensure processes are aligned to public views. Above all else, the public clearly want more information about how and why their data is being used in research, and want to be meaningfully involved in data research processes.**

**This work has also highlighted some areas that could benefit from further research with the public:**

- As discussed, some workshop participants expressed an overarching feeling of a low sense of control over what happens to their data after it is collected, with suggestions made about how this could be alleviated. Some discussion centred around the possibility of individuals being able to opt out of their data being used in research. Future public dialogue could explore how this might work in practice, and the potential impacts upon research findings, to gather informed views on the issue.

- Participants discussed the concept of 'public benefit' or 'public good' when considering issues of sensitive data use in research, and stressed that this might mean different things for different people. They questioned who the imagined 'public' is when we speak about public good, who makes decisions about what is and is not in the public benefit, and which groups are benefitting in each case. Further research could explore the concept of public good in more depth, and what it might mean for different groups in society in the context of data research, as this may have an impact on how decisions are made about data access for research.

- Workshop participants also spoke at length about the need to use different methods to reach different groups in society, both with information about data research and to recruit for involvement and engagement activities. Further research could explore with different groups what the most effective methods and messaging are for reaching them and their communities, to help inform more inclusive public involvement and communications practices.

The findings of this public dialogue are invaluable for informing DARE UK's ongoing work to design and deliver a more joined-up, efficient and trustworthy national data research infrastructure, as well as being a useful addition to the wider evidence base on public attitudes towards the use of sensitive data for research. The DARE UK team will report back on how these recommendations have fed into and affected the programme's wider work and impact. We will also maintain an ongoing dialogue with the public throughout future Phases of the programme to remain informed of public views and expectations.

# 7 / References


**Aitken, M., K. H. Jones, M. Kaarakainen, F. Lugg-Widger, K. McGrail, A. McKenzie, R. Moran, M. J. Murtagh, M. Oswald and A. Paprica (2019)**
Consensus statement on public involvement and engagement with data-intensive health research
International Journal of Population Data Science 4(1).

**Atkinson, S., S. Badger, R. Milne and C. Brayne (2017)**
Ethical, legal and social issues in dementia research
Dementias Platform UK.

**Elias, P. (2021)**
Promoting public engagement with longitudinal research: A report to the Economic and Social Research Council
Warwick Institute for Employment Research.

**Jones, K. H., S. Heys, R. Thompson, L. Cross and D. Ford (2020)**
Public involvement and engagement in the work of a data safe haven: a case study of the SAIL Databank (Including Jones, K. (2020)
SAIL Databank: Public Involvement and Engagement Policy).
International Journal of Population Data Science 5(3).

**Kispeter, E. (2019)**
Public support for accessing and linking data about people from various sources: Literature review
Warwick Institute for Employment Research,
University of Warwick.

**Scott, K. (2018)**
Data for Public Benefit: Balancing the risks and benefits of data sharing
Involve, Carnegie UK Trust and Understanding Patient Data.

**Waind, E. (2020)**
Trust, Security and Public Interest: Striking the Balance – A review of previous literature on public attitudes towards the sharing and linking of administrative data for research
Administrative Data Research UK.

# 8 / Appendices

## Appendix 1: Oversight Group members

**Harriet Baird**
Impact and Knowledge Exchange Manager, Scottish Centre for Administrative Data Research

**Graham Bukowski**
Public Engagement Lead, Public Dialogue and Public Opinion, UK Research and Innovation

**Chris Carrigan**
Chief Operating Officer and Patient and Public Involvement Lead, DATA-CAN: The Health Data Research Hub for Cancer

**Dr Victoria Chico**
Senior Privacy Specialist, Office of the National Data Guardian

**Joyce Fox**
Phase 1 Delivery Team Public Advisor, DARE UK

**Niccola Hutchinson-Pascal**
Head of Co-Production Collective, Co-Production Collective, University College London

**Emily Jarratt**
Senior Policy Advisor, Centre for Data Ethics and Innovation (CDEI)

**Shayda Kashef**
Public Engagement Manager, ADR UK (Administrative Data Research UK)

**Professor Kerina Jones**
Associate Director of Information Governance and Public Engagement, Population Data Science Dept., Swansea University Medical School

**Wiktoria Kulik**
Policy Advisor, Centre for Data Ethics and Innovation (CDEI)

**William Lammons**
Parent, Patient, and Public Involvement Research Lead, Section of Neonatal Medicine, Imperial College London, and Qualitative Researcher, Our Future Health

**Sinduja Manohar**
Public Engagement and Involvement Manager, Health Data Research UK (HDR UK)

**Dr Richard Milne**
Head of Research and Dialogue, Wellcome Connecting Science

**Elizabeth Nelson**
Public Engagement, Communications and Impact Manager, Administrative Data Research Centre Northern Ireland

# **Appendix 2:** Demographic characteristics of registered workshop participants

| Ethnicity | No. | Country of residence | No. | Age | No. |
|---|---|---|---|---|---|
| Arab | 2 | England | 13 | 18-24 | 4 |
| Bangladeshi | 5 | Wales | 12 | 25-34 | 2 |
| Black African | 3 | Scotland | 11 | 35-44 | 12 |
| Black Caribbean | 2 | Northern Ireland | 10 | 45-54 | 7 |
| Black Other | 1 | Missing data | 4 | 55-64 | 6 |
| Chinese | 1 | | | 65-74 | 3 |
| Pakistani | 6 | **Gender** | **No.** | 75+ | 3 |
| Other mixed background | 1 | Female | 29 | Missing data | 13 |
| White British | 12 | Male | 13 | | |
| White non-British | 6 | Non-binary | 1 | | |
| Missing data | 11 | Missing data | 7 | | |

NB: All 50 participants who signed up to attend the workshops were asked prior to the initial workshops to complete a voluntary, anonymous demographic information form. Not all participants completed the form and/or each question – for each question, the number of participants who did not respond is denoted as ‘missing data’. On the day of the workshop, six people dropped out or did not attend, resulting in a final total of 44 workshop participants, but we cannot be sure which demographic backgrounds those people belong to.

# **Appendix 3:** Workshop speakers

## Initial workshop 1 (Thursday 13 January 2022)

| Name | Affiliation | Presentation topic |
|---|---|---|
| **Gerry Reilly** | DARE UK Phase 1 Delivery Team | **'What is a trusted research environment (TRE)?'**<br>*Presented to all workshop participants* |
| **Dr Angela Sorsby** | University of Sheffield | **Data research case study: 'An investigation into racial bias in court case outcomes in England and Wales'**<br>*Presented to the 2x England breakout rooms* |
| **Dr Babak Jahanshahi and Dr Neil Rowland** | Queen's University Belfast | **Data research case study: 'Air pollution and health in Northern Ireland'**<br>*Presented to the Northern Ireland breakout room* |
| **Jan Savinc** | Edinburgh Napier University | **Data research case study: 'Increased deaths at home in Scotland during COVID-19 pandemic'**<br>*Presented to the Scotland breakout room* |
| **Dr Michaela James** | University of Swansea | **Data research case study: 'HAPPEN Wales: The health and attainment of pupils in a primary evaluation network'**<br>*Presented to the Wales breakout room* |

## Initial workshop 2 (Friday 14 January 2022)

| Name | Affiliation | Presentation topic |
|---|---|---|
| **Dr Susheel Varma** | DARE UK Phase 1 Delivery Team | **'What is a trusted research environment (TRE)?'**<br>*Presented to all workshop participants* |
| **Dr Robert French** | Cardiff University | **Data research case study: 'Children and young people with type 1 diabetes: data linkage beyond health'**<br>*Presented to all workshop participants* |

# DARE UK

## Get in touch


enquiries@dareuk.org.uk

www.dareuk.org.uk

@DARE_UK1


ADR UK

HDR UK Health Data Research UK

UKRI UK Research and Innovation

Kohlrabi Consulting

DARE UK 2022. DOI: 10.5281/zenodo.6451935