civic data coop

Data Into Action

Cheshire and Merseyside
Secure Data Environment
Public and Patient Engagement:

The Use of Administrative Health Data for Research

Prepared on behalf of the
Data into Action Programme
by the Liverpool City Region
Civic Data Cooperative and
Civic Health Innovation Labs
May 2024

Contents

Executive Summary

Three key messages:

1 **Six public focus groups were conducted on the use of secondary health data in the context of a Secure Data Environment (SDE) from November 2023 to February 2024 in Cheshire and Merseyside in England.**

2 **Participants were broadly supportive of health data sharing for public benefit and the SDE plans, and desired strong public communication to ensure the success of the plan.**

3 **Participant support was nuanced by concerns on the use of data for artificial intelligence, potential commercial profit-making, and how data systems may exacerbate existing inequalities and stigma in healthcare.**

The Cheshire and Merseyside Secure Data Environment and Data into Action teams worked with a range of partners to conduct a programme of public engagement on the design and policies of a proposed secure data environment (SDE) and specifically perspectives on sharing de-identified health data for secondary uses without explicit consent. The work described in this report forms part of a larger initiative of public and patient participation and engagement planned for the Data into Action programme, of which the SDE is part.

A series of six focus groups were run with 46 residents from November 2023 to February 2024. Residents were asked their perspectives on the use of health data for risk stratification, population health management, research, and public health planning in the context of an SDE. Partners included the Liverpool City Region Civic Data Co-operative; Health Innovation North West Coast; Lancashire and South Cumbria NHS Foundation Trust; NHS Cheshire and Merseyside; and Arden & GEM.

Focus group participants discussed four main themes regarding their perspectives, hopes, and fears around the secondary use of health data:

1. Participants expressed broad support for health data sharing for clear public benefit even without consent. Participants were clear on the benefits of both population health and research data sharing. This support was nuanced by concerns around the use of data for risk stratification, automated algorithmic decision-making, and artificial intelligence, which participants felt may contribute to increased diagnostic errors and decreased autonomy in their interactions with the healthcare system.
2. Some of the participants felt they were more familiar with data and research than the average public or patient group. Therefore, participants were strongly supportive of active, direct, and understandable communication on the SDE to a range of communities across Cheshire and Merseyside. Effective communication both on the SDE and the benefits of health data projects was described as essential to data sharing acceptability.
3. Participants described deidentification as creating a sense of safety in data sharing. They felt that deidentification layered with clear public benefit would influence most communities and individuals in Cheshire and Merseyside to support the SDE plan. The reidentification of certain kinds of data like gender or rare conditions was perceived to have potential negative impacts on health and wellbeing and warranted further public deliberation.
4. Participants wanted to see recognition and communication on general data concerns. General data concerns included things not explicitly related to the SDE but relevant to the data ecosystem as a whole. For example, risks of data gaps and missing information, the use of artificial intelligence, and how the SDE fits in to wider national initiatives for health data sharing. Participants were critical of data sharing for commercial profit-making.
5. Finally, participants discussed the importance of opt-out mechanisms and choice in assuring the acceptability of not using explicit consent. To prevent harm and maintain a sense of autonomy, participants desired transparent and simple opt-out mechanisms.

As the Data into Action programme expands its plans for the SDE and public participation, the findings from these focus groups will form the key base for topics for further debate and direction-setting the design of the programme. Recommendations for further public deliberation include:

1. Cheshire and Merseyside resident perspectives on sufficient public benefit for commercial access to health data.
2. The regulation and governance of AI technologies that are built from health data in the Secure Data Environment.
3. Additional regulation and requirements around communication and consent for data access to sensitive data; both the value of access to this data for community benefit and prevention of harm.
4. Public preferences on identifying and addressing data gaps in the Secure Data Environment considering both data accuracy and potential 'missing' datasets.

Introduction

This report presents a summary of public engagement for the Cheshire and Merseyside Secure Data Environment (SDE) commissioned by the Data into Action programme. The Cheshire and Merseyside Secure Data Environment team worked with a range of partners to plan a series of focus groups to seek feedback on the design and policies of a proposed secure data environment and specifically perspectives on sharing de-identified health data for secondary uses without explicit consent.

The work described in this report forms part of a larger initiative of public and patient participation and engagement planned for the Data into Action programme, of which the SDE is part. Partners included the Liverpool City Region Civic Data Co-operative; Health Innovation North West Coast; Lancashire and South Cumbria NHS Foundation Trust; NHS Cheshire and Merseyside and NHS Arden & GEM Commissioning Support Unit.

The introduction below provides an overview of the Data into Action programme, a summary of existing knowledge on public preferences for health data sharing, and the aims and objectives of this public engagement.

The Cheshire and Merseyside Secure Data Environment and the Data into Action programme

Data into Action uses health and social care information to improve services across Cheshire and Merseyside.

It is a programme of work which sees our NHS, councils and the University of Liverpool working together to make things better for all 2.6 million people in our region.

When people see a GP, go to hospital, or receive care in a care home, information is recorded to help them receive the best possible care.

Data into Action uses that data to see what our health and social care services are doing well, where they can improve, and how changes can be made.

About Data into Action

Data into Action is a population health management platform, which was established in three months across Cheshire and Merseyside to help the health and care system manage the Coronavirus crisis and drive its recovery. It is a collaboration involving the NHS, local government and the University of Liverpool. It was central to the region's response and recovery, including virtual ward programmes. It was also key to mass testing, vaccine delivery and national investigations on re-opening events.

The Data into Action expansion is co-ordinating the roll out of the programme across a population in excess of 16 million residents in the North West of England. Data into Action is acting as an enabler, supporting access to data sets and the technical capabilities required for the mature population analytics now required in the NHS 20/21 planning and operational guidance.

The programme brings together the activities and projects that access and use the Cheshire and Merseyside data asset – the CM Secure Data Environment (SDE), that was previously called CIPHA - with the aim of delivering data into action through a unified programme.

About the Cheshire and Merseyside Secure Data Environment

The Cheshire and Merseyside SDE is an online platform that securely stores health and care data.

It brings together information from several important areas of health and care within Cheshire and Merseyside. This includes GP practices, community and mental health services, hospital services, and social care.

The Cheshire and Merseyside SDE will be used by health, care and research professionals to plan health and care services, carry out research, and make sure that patients receive the best possible care. This includes academic partners from universities. In the future, we may work with other industry or non-NHS partners. An example of this would be working with a pharmaceutical organisation to explore the potential of new treatments for health conditions. Another example would be working with social services to understand how we can help people with complex health and social care needs, such as elderly individuals with multiple chronic conditions or individuals with disabilities requiring both medical and social support.

Alongside the technical work, the programme includes a series of Patient and Public Involvement and Engagement activities to ensure that patient and public perspectives are included in decision-making processes. Patient and public voice will be an essential component in the operation of both the Data into Action programme and the SDE.

The Cheshire and Merseyside SDE is part of a wider regional and national NHS Secure Data Environment network, and it will support the new North West Secure Data Environment.

You can learn more about both the SDE and Data into Action programme on the Data into Action webpages: datainto action.cheshireandmerseyside.nhs.uk

Public and patient perceptions of data sharing in the UK

Extensive academic research and public consultation has been done on patient and public perceptions on the use of health data for research and other secondary data sharing over the past 25 years (Aitken et al., 2016, 2018; Cascini et al., 2024; Howe et al., 2018; Ipsos MORI et al., 2016; Jones et al., 2022; Kalkman et al., 2022; Rempel et al., 2018; Yates et al., n.d.). This includes work on a variety of health data sharing topics ranging from the use of biobank data, to commercial access to health data, to concepts of public benefit, and variations in preferences by data type (Aitken et al., 2018; Cascini et al., 2024; Ipsos MORI et al., 2016; Kirkham et al., 2022). This focus on public preferences is driven both by the dramatic expansion in the amount, use, and access to secondary health data in the past decades but also examples of public disquiet towards data sharing initiatives like Care.Data that have resulted in those initiatives' failure (Ford et al., 2020; Jones et al., 2022). Therefore, questions on public preferences towards the use of health data are often driven by both a normative and substantive aim, i.e., public opinion is sought both to ensure the success of data sharing initiatives and to improve ethical data sharing practices (Rempel, 2018).

Systematic reviews have highlighted broad support both in the UK and globally for health data sharing underpinned by clear public benefit and transparency (Aitken et al., 2016; Jones et al., 2022; Kalkman et al., 2022; National Data Guardian, 2024). In a deliberative exercise on defining public benefit in Scotland, participants highlighted the desire to keep the conception of public and societal benefit broad but to prioritise benefit to vulnerable communities (Aitken et al., 2018). Consistent key concerns by community across quantitative and qualitative feedback mechanisms on health data sharing include privacy, security, data leaks, and who has access to data (Kalkman et al., 2022). While privacy, security, and data leaks are generally attended to through communication on safe data sharing practices, data access is more complex. A common finding is a preference for academic and government research above commercial research, as well as a strong disinclination for commercial access to data, again, without clear public benefit. Who has access to data is highly significant on influencing whether individuals want to share health data in the UK, where profit-making by commercial organisations is seen as having a high risk of potential harm (Jones et al., 2022).

Despite a general finding of support, important variations in preferences for health data sharing exist. Researchers have examined how ethnic background, perceptions of risk, age, and experiences of service use may influence the inclination for an individual to want to share their data (Jones et al., 2022; Kalkman et al., 2022; Kirkham et al., 2022). While findings on individual variation are inconsistent and further research is warranted, it is important to note that trust in data sharing is inextricably linked to trust in the health system. There is, however, a consistent finding that individuals and communities are not aware that data is not shared (Jones et al., 2022). As Aitken (2016) phrases it participants often ask: 'doesn't this happen already'? This has led researchers to call for not only a focus on aligning public preferences for data sharing to practice but communicating widely and clearly on data literacy and on how data is already being used (Paprica et al., 2019; Yates et al., n.d.).

While systematic reviews demonstrate a wide literature on the subject of public preferences for health data sharing, there is also a significant focus on the public role in data governance within health and civic structures (Patel, 2021; Rempel et al., 2018). Namely, what should the ongoing role of the public be in the governance of their own data? Drawing on political theory, public involvement ranges in intensity from informing public on how their data is governed to fully empowering public to be the main decision-maker in how data governance is designed and operationalised (Patel, 2021). Several authors have proposed the concept of a 'social license' to ensure that health data sharing practice includes public preferences above and beyond legal frameworks and requirements (Aitken et al., 2018). While the debate on public governance and stewardship of data is beyond the scope of this report, it is important to note activities like these focus groups sit within these literatures and practices as well.

The debate on preferences for health data sharing can at times feel broadly settled; however new political movements, the lack of data on specific communities' preferences and constantly evolving data technologies demonstrate the need for ongoing public dialogue. In particular, renewed calls to commercialise health data within the United Kingdom as a political talking point are re-evidencing the need for a public dialogue on the boundaries of health data sharing for commercial and profit-making uses (BBC, 2023; Donnelly, 2024). These not uncontroversial calls underline a proposed moral imperative that sharing data would be a panacea to the underfunding of the health and care systems. While innumerable dialogues have been held on this very topic (Ipsos MORI et al., 2016), their lack of translation in to practice puts pain to re-open the debate. Adding to these renewed political and moral arguments around data sharing, rapid technological advances in Artificial Intelligence that rely on extensive and robust datasets add further pressure for immediate direction setting on the next decade of health and care policy on health data sharing (Ada Lovelace Institute & The Alan Turing Institute, 2023). While public preferences research demonstrates strong support for the concept of data sharing for health improvement, the operation and impact of data sharing needs further deliberation.

Aim and Objectives

The aim of these focus groups is to explore the use of de-identified personal data for research and population health support without consent in the context of a secure data environment amongst Cheshire and Merseyside residents. The objectives were to:

1. Explore residents' qualitative familiarity and understanding of health data through a data literacy discussion exercise.
2. Explore residents' perspective on the Secure Data Environment policies on deidentification, consent, and the Five Safes data access model through presentation and focus group discussion.
3. Explore residents' perspective on health data sharing for population health/planning, research, and risk stratification through case study exploration and focus group discussion.
4. Explore residents' preferences for communication of the SDE plan and health data sharing through focus group discussion.

Methods

Six online focus groups were held with 46 residents from November 2023 to February 2024. An initial two pilot focus groups, made up of residents from the North West of England, including Cheshire and Merseyside, were held in November 2023. The following main four workshops included 34 residents from Cheshire and Merseyside only and ran from January to February 2024. A table of demographic characteristics for participants in the following main four focus groups only is listed in Appendix A.

Recruitment

Participants were recruited online through existing patient and public engagement networks. Participants were remunerated £50 by the North West Secure Data Environment programme for their time preparing for and taking part in the focus group following NIHR recommended rates. Arden & GEM recruited and managed the participants for all four of the main focus groups. Participants were recruited online via email, newsletter, and email lists. Recruitment materials are available in Appendix B. Key contacts were identified through the Health Innovation North West Coast (NWC) and the Cheshire and Merseyside Integrated Care Board (ICB), and cascaded to their existing patient groups and contacts. Representatives from the Liverpool City Region VS6 Partnership, the Cheshire and Warrington Infrastructure Partnership, NHS England's Lived Experience team, and Wirral Hospital NHS Trust also cascaded the call for participants.

Arden & GEM compiled a database of contacts for underserved groups, and reached out to these directly. The key categories were:

- Ethnic minority groups
- Youth audiences
- Healthwatch
- Men
- Expectant parents and new parents
- Armed forces
- LGBTQIA+
- Mental health
- Disability and long-term conditions
- Other community links

Participants were invited to express their interest via email and were provided with details on the focus groups and pre-reading outlined below. A table of demographic characteristics for participants is listed in Appendix A.

Focus Group Guide Development

The focus group guide was designed in collaboration with the University of Liverpool's Liverpool City Region Civic Data Cooperative, NW Secure Data Environment, Arden & GEM, and NHS Lancashire and South Cumbria Foundation Trust. The focus group covered three main concepts: understanding of health data, perspectives on secondary uses of health data, and the use of health data without explicit consent. This included a specific question driven by the NHS Confidentiality Advisory Group on secondary uses of data for planning, population health, risk stratification, and research. Discussion questions and activities were designed in a collaborative manner drawing on established focus group guides from Our Data Bodies' Digital Defence Playbook (Lewis et al., 2018) and the Round 'Ere Community Wellbeing Data Hub project (Capacity & LCR Civic Data Cooperative, 2023). The focus group guide was refined after the first two online focus groups and will be used as an asset moving forward for future public participation. The refined focus group included the same initial sections and questions from the draft focus group guide but added additional information on health data sharing examples.

Focus Group Content

Each focus group lasted approximately two hours and was held online using Microsoft Teams software. Each focus group included a facilitator, notetaker, and SDE representative to answer questions about the SDE. The focus groups were set up and managed by Arden & GEM and were facilitated and moderated by the LCR Civic Data Cooperative with support from the NW Secure Data Environment team.

Participants were provided with pre-reading materials to familiarise themselves with both health data and the rationale for health data sharing prior to the focus group.

Participants were sent one PDF for pre-reading that outlined the content of the focus groups and provided links to two videos:

1. Patient data saves lives available at: https://www.youtube.com/watch?v=fJ2hyXCOOyQ&t=152s
2. Why health and care data matters to me available at: https://www.youtube.com/watch?v=3QptmDnBoFE

The final focus group guide is attached in Appendix B and included the following:

1. Introductions and reminder of health data sharing pre-reading materials
2. Health Data Literacy Activity 1 – word association
3. Health Data Literacy Activity 2 – definitions of data
4. Presentation and group discussion on data anonymisation and deidentification
5. Presentation and group discussion on the '5 Safes' of Data Sharing in the UK
6. Group discussion: Health data sharing for research
7. Presentation and group discussion on health data consent.
8. Case study discussion on data sharing for secondary uses: planning/population health, risk stratification, and research.
9. Group discussion on SDE communication needs

Participants were also asked to fill in a post-focus group evaluation survey, see Evaluation below for full details.

Analysis

All six focus groups were held, recorded, and transcribed using Microsoft Teams Software. The transcripts and notes were thematically analysed for key concerns and hopes related to the secondary use of health data without explicit consent in the context of the Secure Data Environment proposal (Braun & Clarke, 2006). Lumivero NVivo14 Qualitative software was used to conduct the thematic analysis. All analysis was completed by ER, see Authorship and Acknowledgements.

Findings

Five key themes were identified from the six focus groups. They are listed below alongside indicative quotes. Quotes are listed verbatim and not edited for grammar.

Broad support for health data sharing for clear public benefit even without consent

Participants across all focus groups were highly supportive of the use of health data in research for improving health outcomes. In fact, several participants felt that the benefits of sharing data for population health research outweighed individual risks of reidentification. Participants were clear on the benefits of both population health and research data sharing. This was led by strong descriptions of altruistic motivations for the use of health data to support healthier communities. Participants described public benefit as data being used in research and planning to fix issues and improve health for individuals, communities, and the wider population. They often described a strong hope for improved treatments and service provision for themselves, evidencing personal motivations in addition to altruistic hopes. As the quotes describe below, participants are sharing not necessarily their hopes for health data sharing but their hopes for health improvement and that data will play a role in this. This was phrased at times as a concern that not using the data, or fears that others not understanding the benefit of using the data, may prevent positive outcomes.

> **"My hope is that data will help to develop new approaches to health interventions that reflect current conditions without old assumptions."**
>
> *Participant in Focus Group 3*

> **"I think I come from a obviously a bit of a point of bias because I I want to make sure that the public health and NHS are prepared for whatever and equalities that affecting our Community and they can make sure that the address those inequalities now without the data, without the consent of Community members. We wouldn't be able to collect that data to make sure that we can address those inequalities and see if what we do has an impact and on the outcomes of the health outcomes of our Community."**
>
> *Participant in Focus Group 1*

> **"What's my hopes for using this data? I'd wanted to kind of fix a lot more issues. I'd wanted to kind of look at a lot of things that are local."**
>
> *Participant in Focus Group 3*

> **"I'm sure there will be some people who you know if if somebody was doing this sort of research and then approach somebody who was just like, how did you get my information? I don't want your help, you know, stay out of my life. There are people who are like that. You know, you can't please everybody. But I think in general. If this is if this helps society be healthier, how can it be a bad thing?"**
>
> *Participant in Focus Group 6*

> **"It's benefit should outweigh personal risk".**
>
> *Participant in Focus Group 2*

> **"So I think it's if we didn't have that data, it would be very difficult to actually try and predict ahead to see you know what we're going to need in the future."**
>
> *Participant in Focus Group 4*

Participants were not entirely hopeful and expressed nuanced concerns about the risks of how data would be used and who had access to it. They expressed concern on the potential negative impacts of data sharing for risk stratification, participants were concerned that risk stratification without health professional review may mis-categorise patients. Similarly, they described several personal examples and fears around how mis-categorisation, even by health professionals, could reinforce stigma or prevent needed care. They connected risk stratification to broader concerns on new data technologies like Artificial Intelligence systems and how they may increase the potential for mis-categorisation errors and a lack of autonomy in patient interactions with the health care system. This theme around autonomy was expanded in several participant's concern that commercial organisations may use data full stop and in particular to create technologies that do not hold public good and public benefit at their core. Returning to preferences around consent, there was greater concern for data sharing and therefore a desire for more direct consent processes where there was perceived to be a lack of transparency around the use of data for public benefit.

> **"And I think the third one [risk stratification] raised a few issues with me in terms of whether actually, how do we know they're looking at the right data and what data are they bringing in and you know, and I and I suppose and it also links back to a question in the last one, which is, you know who's making these decisions about what is should be considered OK and what isn't."**
>
> *Participant in Focus Group 2*

> **"Yes, I know we are all discussing this data in a good intention. [Risk] stratification was that the one patient is developing medical disease (referring to example given). That's the one, bit worried. Most of the people I think are worried about that."**
>
> *Participant in Focus Group 3*

> **"I think the potential for research potential for finding answers fantastic. In terms of concerns. Who actually at the end of the day owns the data? And if the funding runs out, if things change in, say, 2050 years, how can you guarantee that that won't ever be sold on to anyone else?"**
>
> *Participant in Focus Group 6*

> **"And certainly I would be quite concerned if there is invasion from, say, pharmaceutical company to use it for their own profit."**
>
> *Participant in Focus Group 6*

> **"But if if things are accessible to people and or to organisations or whoever, and then the interpretations can be wrong. Did -do you see? Do you see what I'm getting at? So that it can be used to stigmatize things, to stigmatize people and their conditions."**
>
> *Participant in Focus Group 1*

> **"What immediately came to my mind was Big Brother is watching us. I'm sorry, 1984 and all that. You know what I mean? I'm an old person. I see things differently. I worry about AI and how it's used in an algorithms like this, things like that and what it learns. If you take something I know you shouldn't say this, but people, if they want to smoke, then that's their problem. OK, we will try and everything we can to help them to get over that and realign their lives, but we cannot be Big Brother and say you've got to stop smoking and things like this, you know, I mean, we're taking your cigarettes away and we're doing this and the other and that's just an example of what I'm trying to explain."**
>
> *Participant in Focus Group 5*

Communication and understandability are essential to data sharing acceptability

Some of the participants, in general, felt they were more familiar with data and research than the average public or patient group. For that reason, they strongly supported improved communication on how data is shared and with whom to ensure data sharing without consent was successful, i.e., important health research was not stalled by significant patient opt-out. This was at times described as getting the public 'on side' with the SDE plan. Participants felt other groups would need to better understand the benefits of data sharing, including contextualising those benefits in personal healthcare improvements, to prevent opt-outs due to confusion on how their data was being used. Several participants felt the focus group itself was an effective tool for reassuring them on the safety and security of the SDE plan. The desire for active and repeated communication on the SDE and research projects that had access to health data within it, was described as essential to the trustworthiness of the SDE plan.

> **"I think getting the public on your side is so important because I appreciate how important this project is." – Participant in Focus Group 3.**
>
> **"it's just about raising awareness of the benefits of using that as long as you can't be identified and and these are the things that researchers need to do their research."**
>
> *Participant in Focus Group 2.*

> **"You know, you've got your chosen few who go to certain meetings. But but how do you get that to the general public? Sort of, you know, be 2 1/2 million people who live in Cheshire and Merseyside."**
>
> *Participant in Focus Group 6*

> **"The message that's been that already and one or two other points about the communication, it's the NHSE produce a lot of information to do a lot of good work on comms...But generally speaking, I don't think they do a great piece of work on actually making it available in a way that the general public understands... And I think, uh, it's taking the public with you."**
>
> *Participant in Focus Group 5.*

> **"And now that we've understood it, as long as people understand personally, me, yes, if it means my health is gonna be better. And you know my friend's kids in the future is is, you know, that they're gonna find, you know, more treatments, better treatments, more efficient treatments. Yes, by all means, use it. I have no problem."**
>
> *Participant in Focus Group 6*

> **"I'd be concerned in the future about sharing my data with consent in situations where I know it would help because as a person I'm very consent heavy even to the point of people posting photographs of me and I can feel quite upset if things are done without my consent. So I think I would definitely be really a lot more cautious in the future if I felt like something wasn't. If I felt like my my information was being shared without my knowledge"**
>
> *Participant in Focus Group 5*

Clear communication was described as accessible and understandable to all and jargon-free, taking in to account patients and publics who may have very low digital literacy or access to digital devices like mobile phones. In addition, participants requested information in multiple languages, particularly for new immigrant groups who may be less familiar with the NHS. This also included a strong desire for clear information on what exact data was being shared, where it came from, and who had access to it. Participants specifically described active and clear communication on what organisations, i.e., private and commercial companies, had access to data as being essential to the acceptability of the SDE.

> **"Who has access to the data to private companies? That's the that's the big worry for lots of people."**
>
> *Participant in Focus Group 6.*

> **"So sometimes it's it can be easy to not communicate about data because if you don't communicate about it, nobody asks any questions about it. Whereas if you're being honest, you should be communicating about it regularly"**
>
> *Participant in Focus Group 3.*

> **"I've got friends who secure data environment would be a foreign language to them. So, if the actual consent form was in this foreign language, they wouldn't know where to start. I've got friends who never go on the Internet. Never picked up a mouse, you know? Then no idea. So, you've gotta put it in really simple language to understand basically."**
>
> *Participant in Focus Group 4.*

> **"And also and I don't even know where to look and I know a lot of people of my age because I'm in my 70s who don't have access to IT and you know, you say what's on the website and everything they say...Well, I haven't got a computer so I'm very, very worried about also about people who are technology poor who cannot get to the to the things that we can get to."**
>
> *Participant in Focus Group 5.*

> **"So you know most of the people [referring to ethnic minority communities] they are not aware of what is research, what is data, do you understand? So it's a kind of yeah. According to me my experience what I feel that more and more kind of what you say campaign or maybe awareness among the community it should be done by the NHS or like the charity organisation like us you know."**
>
> *Participant in Focus Group 6.*

Additionally, taking in to account digital literacy was also described as providing information on the more complex technical aspects of the SDE. Participants wanted to understand both the safety measures put in place to protect the security of data and the technical specifications like the use of cloud computing services. Participants described familiarity with news stories on data sharing harms and science fiction stories as needing to be overcome with clear information on the security features of the SDE. For example, what technical infrastructure and policy measures prevent data leaks, inappropriate data sharing, and bad actors from gaining access to the SDE.

> "However, whatever way you put it, pitch it, you still gotta answer questions from people who will be more curious like me, you would say, well, what's standards are you talking about when it comes to security? Are you using the cloud, for instance, were physically on what servers is the data held?"
>
> *Participant in Focus Group 4*

> "I mean, I saw to know what obviously about what the you know, the gist of health data, but it's the negative connotations to me that springs to mind thinking of things obviously like Black Mirror and you know, Cambridge Analytica. There's a big concern for me that if I am selling this, it's all gonna end up either gonna get leaked somehow or it even as I've watched the videos before and I know that it's, you know, it's meant to be all secure and and anonymized. But my question is first of all, how secure is that really?"
>
> *Participant in Focus Group 5*

Deidentification creates a sense of safety in data sharing

Deidentification of data was seen as a key safeguard in making data sharing without consent acceptable. Deidentification was often described colloquially by participants as ensuring people outside of the health system were unable to connect personal information to individuals. In this sense deidentification was described as protecting privacy and preventing harm, particularly from the commercialisation of healthcare data, e.g., an oft-repeated concern was the use of private health data used to increase travel insurance rates. These concepts were also described by participants as needing to connect to public benefit. Deidentification was not sufficient for trustworthy data sharing practice in and of itself but a core component of building towards trustworthy practice.

> **“And I’m just thinking that if-if information is depersonalized and then I-I suppose my, you know, sitting in a another place is as a as a researcher, I’d kind of say, well, does it matter what data I use as long as it’s in that cohort of depersonalized information?”**
>
> *Participant in Focus Group 2*

> **“And no problem in sharing diagnosis information to help the the wider good, as it were, as long as it wasn’t linked into more specific personal information, that would I, you know, narrow it down and identify and identify me or someone like me, I guess.”**
>
> *Participant in Focus Group 5*

The reidentification of certain kinds of data like sex, gender, mental health, or rare conditions was described as having more potential for negative impact due to societal stigma. In terms of gender and sex data, participants described two opposing concerns. On one hand they worried about sex data not being used, and therefore reinforcing historical bias in health care research towards excluding people who are female. Conversely, participants highlighted the risks of non-binary and trans individuals in their data being used to further stigmatise and marginalise that group. In this example and others, concerns were focussed for some participants specifically on the risk of sensitive information being reidentified by health care professionals. Thus, participants supported further public participation and debate on opt-out, minimisation, and aggregation of data in the SDE in these contexts.

> **“Gender identity as a form of data can be quite an issue. We’ve had someone who is trans at present that to their GP and then the issues that followed from them declaring that it’s like steam rolled into more problems than solutions and and even then many people trans people don’t necessarily want the whole world to know their gender identity is changed. So then that can steamroll into mental health issues.”**
>
> *Participant in Focus Group 4*

> **“And when you’re looking at any kind of data, you’ve got to look at it by sex, because sex will affect every part, every part of of how we respond to disease and to drugs.”**
>
> *Participant in Focus Group 3*

> **“Things like rare diseases, and particularly if you’re like I am one in 51 in the world. So obviously it would be quite easy to identify and somebody from a so called rare disease if there if it is extremely rare.”**
>
> *Participant in Focus Group 2*

> **“I know people who have had mental health problems. I know people, even my family. You’ve got certain other health issues that have, you know, big stigmas attached to them. So I would imagine that for them to to to disclose that that should be at the discretion and the choice and all times really.”**
>
> *Participant in Focus Group 5*

> **“On those specific things, you know, having those included in research is. It’s kind of scary, you know? What are they going to do if they know that I only make this much or that, you know, my sexuality is this or my ethnic background is this? Is that going to, you know, cloud someone’s judgement on, you know, with research, you know, I’m sure that you try to. Be very unbiased, but because people come up against bias every single day, that makes everything so much more sensitive.”**
>
> *Participant in Focus Group 6*

Public and patients want to see recognition and communication on general data concerns

Participants discussed key concerns around data gaps, future technologies, inappropriate data sharing, and abuse of data by non-health government authorities. Although this falls outside of both the data held by the SDE and the explicit purpose of the SDE, participants felt others in their communities would have questions about these concerns regardless of the SDE parameters. This included discussion on how the SDE fits within the larger initiatives for data sharing across England and the UK.

> **“I wouldn’t expect you to comment on this, but obviously the the Federated data platform is is being introduced. Palantir have won the contract. That’s quite controversial. And and and things like the contract with Palantir, where a lot of the information that’s been redacted. So it generates a lot of questions and it’s about transparency. People understanding what the data is being used for, where it’s going, how it’s gonna be communicated back to patients on an ongoing basis about the outcomes and what third parties can can access that information. So that’s my I think it people’s general perception probably if you ask the man on the street, the information goes into a big black hole and they’re not really clear about what happens to it.”**
>
> *Participant in Focus Group 5*

Concerns around data gaps included worries that secondary health data may not fully capture people’s lived experiences and many personal characteristics like language or disability may prevent people from sharing data or accessing health services. Participants hoped that researchers accessing data through the SDE would need to consider the context of how the data was collected and if it accurately represented the population being researched.

> **“Using data does have it’s limits which I find sometimes is forgotten. Ie the context of HOW and WHY the research data was collected - we’ve seen in the past how numbers can been played with to fit agendas…I also wonder about how people can draw attention to where there is a lack of data, being trans I hit so many health roadblocks and gatekeepers which can be mentally and physically detrimental..”**
>
> *Participant in Focus Group 3*

> **“the other side is that a worry would be that the incomplete data capture cause obviously the data that is captured and doesn’t always in my mind include the patient journey, you know, and obviously the the qualitative subnet which is very, very valid.”**
>
> *Participant in Focus Group 3*

Regarding future technologies, participants expressed hope that the SDE would consider and create proactive measures around ensuring new data technologies were appropriately evaluated and deidentified, where relevant. Technologies discussed included both alternative kinds of data like retina scans as well as the use of artificial intelligence (AI). AI, when mentioned, was described as 'scary' and 'frightening'. Participants wanted to understand if and for which AI technologies the SDE data would be used.

> **"I think it'll become more important with artificial intelligence and I'm not sure whether any humans can actually be involved and whether that's whether that's a risk to us all or whether it's actually. Again, I don't know really. That's what I'm thinking about for the future."**
>
> *Participant in Focus Group 4*

> **"I don't like the idea of and too much control being in the hands of computer algorithms or AI or something like that. It worries me. Where does it stop? Or we going into something in the future where it takes over completely, or are we just using it at the moment? What regulations are we got round it?"**
>
> *Participant in Focus Group 5*

> **"I accept the prevention of abuse of data is not easy but with AI etc this does need to be a priority for consideration."**
>
> *Participant in Focus Group 4.*

> **"I'm sure when when people came up with social media, you know, with like Facebook, I don't think it was, you know, the originally planned for it to be used for, you know, you know, messing around with elections and votes and stuff like that. So for me, looking at like the wider implications and sort of like safeguards again in place and to make sure that this sort of thing isn't done and it it's just again sorted like the road to hell is paved with good intentions. So making sure it is good and just for the purpose it says it is really."**
>
> *Participant in Focus Group 5.*

Inappropriate data sharing was discussed as both sharing data outside of the health system and specifically for commercial purposes. Participants often discussed fears of data being sold by the NHS for profit. Similarly, participants also described fears around specifically non-health government use of data. These concerns were not necessarily based on concrete examples but revolved more broadly around their perception of other people's dystopic fears of being controlled by government authorities.

> **"It's going out, but obviously my main interest is the abuse of data and how that is going to be and, you know, hope making sure that the the the data is not abused and it in the right hands."**
>
> *Participant in Focus Group 4*

> **"And a good way, but the use of data in controlling populations, I mean long. I'm talking big, big picture now. At the end of the day, we some some people have a a real fear that data can be used as a control of of population. Yeah, politically as well."**
>
> *Participant in Focus Group 1*

> **"I think things like people's mental health, sexual health and medication are important because the people who want that information to monetize it all your pharmaceutical companies, your insurance companies, and that would affect everyone around you."**
>
> *Participant in Focus Group 4.*

Importance of opt-out mechanisms and choice in assuring the acceptability of not using explicit consent

The clear option to opt-out was described as essential to trustworthy data sharing. Participants emphasised the need for clear opt-out instructions to be central to transparent communication. Participants hoped that the option to opt-out would be sufficient to protect people who may have particularly high risks due to data sharing leaks, e.g. people with trans experience. The discussion was tempered by a recognition that opting out held risks for the success of the SDE. Participants hoped the clear option to opt-out would demonstrate trustworthy behaviour so that individuals would choose not to opt-out.

Some participants described their preferences and perspectives on withdrawing consent as valuing choice and autonomy in how they lived their lives. They saw withdrawing consent as a method of 'voting with their feet'. For example, if data was shared in unexpected ways like for profit or commercialisation, they would withdraw their consent to protest that use. This was not opposition to sharing their data in general but instead was a desire to exercise their right to choose who had access over the data and to ensure the use of data aligned with their personal values.

> **"I think that it kind of needs to be actively promoted that people are aware that they can opt out. And I honestly think if people find out inadvertently that they could have opted out, they will feel more likely to opt out because they're kind of feel that it's been because it they weren't aware of it"**
>
> *Participant in Focus Group 2.*

> **"So I'm if it's just for sort of like random sales purposes or something like that? Absolutely not. If it's something for like as we mentioned, like the greater good or you know, I'm happy to share more, but it's gotta be my my discretion. It's gonna be my choice. And who I choose to share it with myself? Not a decision made by someone else."**
>
> *Participant in Focus Group 5.*

> **"Obviously your health data, but the thing that links to me is like the NHS is managing it, but the patients actually own it. NHS don't own it. It's the patients that take ownership of it."**
>
> *Participant in Focus Group 3.*

Discussion and Recommendations

The focus groups demonstrably evidenced that Cheshire & Merseyside residents see strong value in the SDE and its associated plan for sharing health data.

Consistent with existing research and engagement, health data sharing without consent for research is broadly supported by a range of communities and individuals. This report adds to this literature by expanding this to the sharing of health data within a Secure Data Environment context (Cascini et al., 2024). The use of the 'Five Safes' and deidentification was positioned and described by participants as helping to address common concerns on data leaks, privacy, and security in health data sharing. Similarly, the support for data sharing within the context of altruistic motivations, i.e. 'the public good', and the lack of support for profit-focused data sharing is consistent with other reviews and research on this topic (Aitken et al., 2016; Jones et al., 2022). It is important to note that while renewed calls for the commercialisation of health data draw narratives to public benefit, engagement exercises like this continue to demonstrate that communities have strong reservations. We echo Aitken's (2018) finding that public benefit and profit are not connected in the minds of publics and communities. Commercial sharing of data must evidence public benefit to a range of communities to be seen as acceptable.

These findings add to a growing literature on the importance of considering public preferences towards data technologies like artificial intelligence (AI) and automated decision-making (Ada Lovelace Institute & The Alan Turing Institute, 2023). While health data use for risk stratification is a specific application of data that can use automated decision-making techniques for defining risk categories for patient care, it's spontaneous connection to concepts of AI by participants demonstrates the high level to which AI narratives and news stories are in the public consciousness. A recent survey of the UK population demonstrated that healthcare was the highest preferred priority for the application of AI over applications like marketing and policing (Ada Lovelace Institute & The Alan Turing Institute, 2023). Conversely in our findings, AI is positioned as a concern and not a value, consistent with dystopic narratives of AI in popular media (Sartori & Bocca, 2023). This dichotomy warrants significant room for future debate at the national and local level on public preferences for AI application and regulation in the context of a Secure Data Environment.

The discussions around AI and national SDE initiatives also warrant further attentiveness to communications on this SDE project and how to address wider concerns on the data ecosystem. Communication and transparency are important to any complex initiative like an SDE, but even more so when there are competing narratives around data sharing concerns and harms with associated National programmes. Active communication to a wide range of communities will be imperative to avoid opt-outs that undermine the benefit of the SDE for research. Confusion and concern will fill any gaps left by communication initiatives. As evidenced in the focus groups, residents are highly supportive of the SDE plans if they can demonstrate and communicate public benefit. It is the responsibility of the programme to ensure that communities are clear on how this initiative differs to wider programmes and how public benefit fits into all data projects.

Recommendations for further deliberation

In line with the preferences of the focus group participants, we recommend the following topics for engagement and deliberation with communities in the context of the NHS Cheshire & Merseyside Secure Data Environment.

1. **Cheshire & Merseyside resident perspectives on sufficient public benefit for commercial access to health data**
2. **The regulation and governance of AI technologies that are built from health data in the Secure Data Environment**
3. **Additional regulation and requirements around communication and consent for data access to sensitive data; both the value of access to this data for community benefit and prevention of harm**
4. **Public preferences on identifying and addressing data gaps in the Secure Data Environment considering both data accuracy and potential 'missing' datasets.**

Next steps

This work represents the first step in the Data into Action programme's Patient and Public Involvement and Engagement work. These focus groups have supported the approval of the Cheshire & Merseyside Secure Data Environment by the NHS Confidentiality Advisory Group. Next steps include hosting in-person events to include the perspectives of individuals who may not have the ability or access to take part in an online focus group. The programme is also setting up a Patient and Public Advisory Group that will continue to advise on and provide feedback to a range of core issues related to the Data into Action programme and its partners.

Evaluation

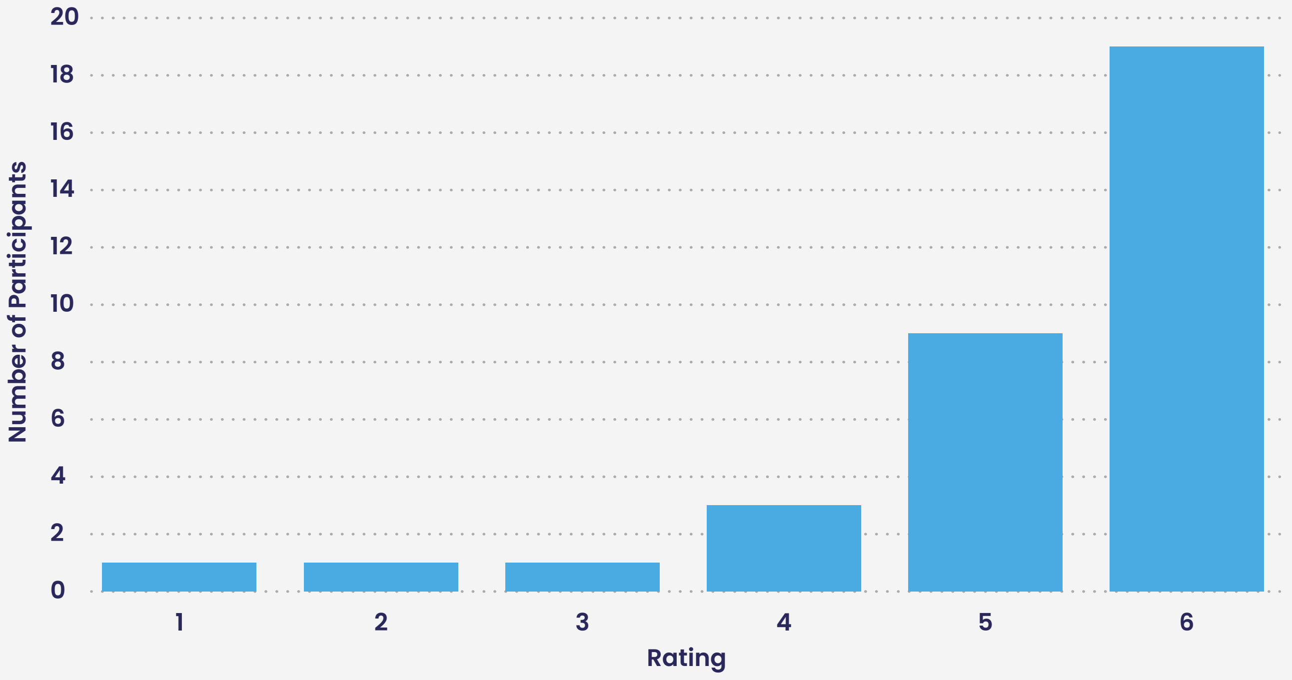
Following each focus group taking place, a survey was sent to each participant to collect their views and opinions on several factors including:

- Satisfaction with how the session was run
- Understanding of the concepts presented in the focus group
- Suggestions for improvements on how the focus group could run
- Thoughts around the idea of health data for research
- Ideas around what the SDE must include and do well to improve public trust and acceptability of using health data for research, planning and risk

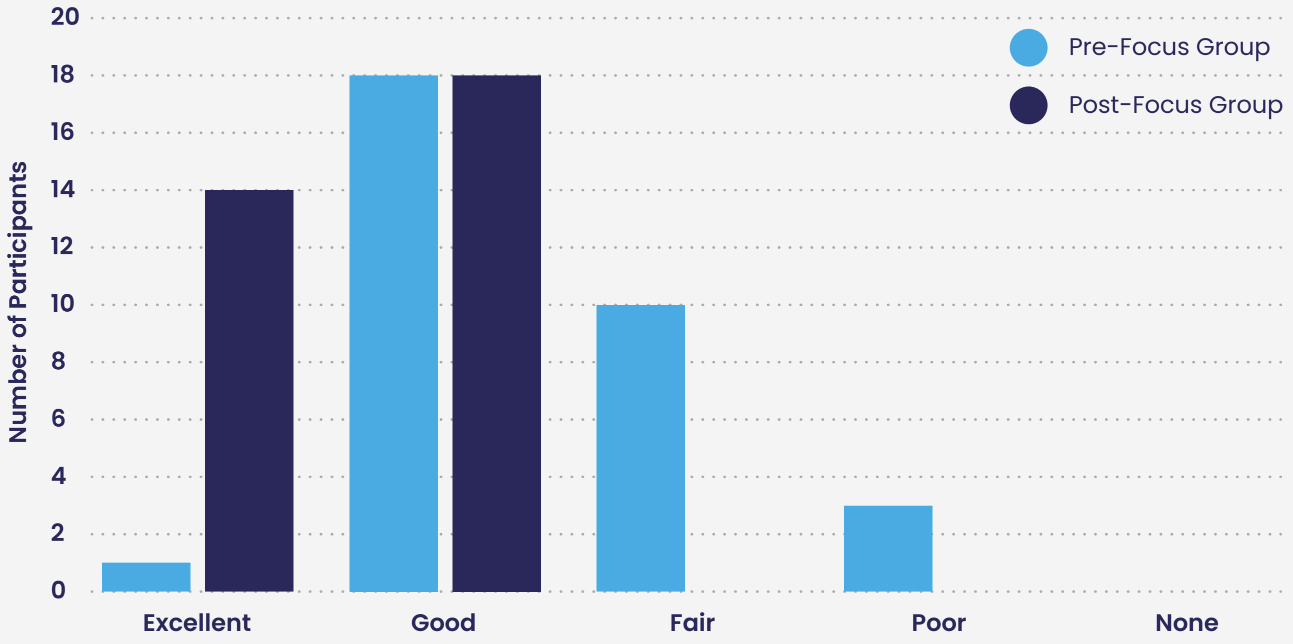
A snapshot of three questions are listed below. The full evaluation slides and figures are on the Data into Action website at: datraintoaction.cheshireandmerseyside.nhs.uk

How would you rate the session?

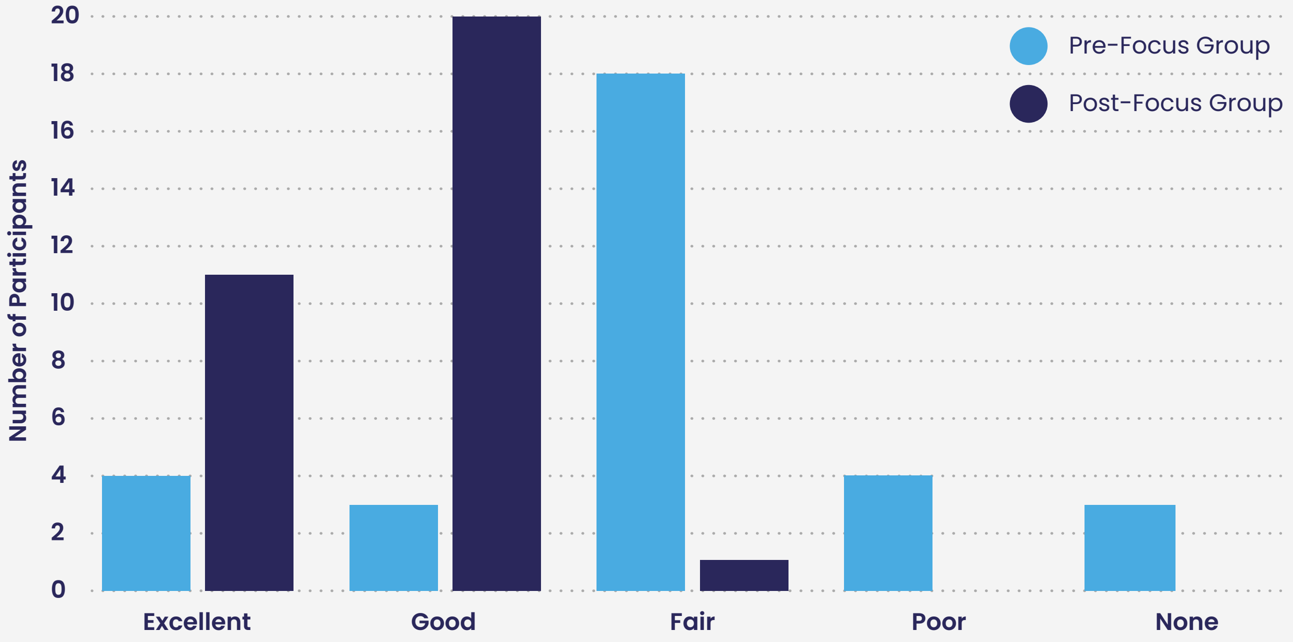
Participants were asked to rate the session, with 6 stars meaning ‘excellent’ and 1 star meaning ‘terrible’



How would you rate your understanding of Health Data pre- and post-focus group?



How would you rate your understanding of a Secure Data Environment Data pre- and post-focus group?



References

Ada Lovelace Institute, & The Alan Turing Institute. (2023). How do people feel about AI? A nationally representative survey of public attitudes to artificial intelligence in Britain. https://adalovelaceinstitute.org/report/public-attitudes-ai

Aitken, M., de St. Jorre, J., Pagliari, C., Jepson, R., & Cunningham-Burley, S. (2016). Public responses to the sharing and linkage of health data for research purposes: A systematic review and thematic synthesis of qualitative studies. BMC Medical Ethics, 17(1), 73. https://doi.org/10.1186/s12910-016-0153-x

Aitken, M., Porteous, C., Creamer, E., & Cunningham-Burley, S. (2018). Who benefits and how? Public expectations of public benefits from data-intensive health research. Big Data & Society, 5(2), 2053951718816724. https://doi.org/10.1177/2053951718816724

BBC. (2023, October 29). Sale of NHS data up to government, says tech boss. BBC. https://www.bbc.com/news/technology-67254010

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Capacity, & LCR Civic Data Cooperative. (2023). Round 'Ere Widnes: Final Project Report. https://civicdatacooperative.com/app/uploads/2023/11/Round-Ere-Report.pdf

Cascini, F., Pantovic, A., Al-Ajlouni, Y. A., Puleo, V., De Maio, L., & Ricciardi, W. (2024). Health data sharing attitudes towards primary and secondary use of data: A systematic review. EClinicalMedicine, 71, 102551. https://doi.org/10.1016/j.eclinm.2024.102551

Donnelly, L. (2024, March 19). Labour vows to save NHS by sharing data from GP records. The Telegraph. https://www.telegraph.co.uk/politics/2024/03/19/labour-vows-to-save-nhs-by-sharing-data-from-gp-records/

Ford, E., Kazempour, Y., Cooper, M. J., Vittal Katikireddi, S., & Boyd, A. (2020). Media content analysis of general practitioners' reactions to care. data expressed in the media: What lessons can be learned for future NHS data- sharing initiatives? BMJ Open, 10(e038006), 1–9. https://doi.org/10.1136/bmjopen-2020-038006

Howe, N., Giles, E., Newbury-Birch, D., & McColl, E. (2018). Systematic review of participants' attitudes towards data sharing: A thematic synthesis. Journal of Health Services Research & Policy, 23(2), 123–133. https://doi.org/10.1177/1355819617751555

Ipsos MORI, Castell, S., & Evans, H. (2016). The One-Way Mirror: Public attitudes to commercial access to health data.

Jones, L. A., Nelder, J. R., Fryer, J. M., Alsop, P. H., Geary, M. R., Prince, M., & Cardinal, R. N. (2022). Public opinion on sharing data from health services for clinical and research purposes without explicit consent: An anonymous online survey in the UK. BMJ Open, 12(4), e057579. https://doi.org/10.1136/bmjopen-2021-057579

Kalkman, S., van Delden, J., Banerjee, A., Tyl, B., Mostert, M., & van Thiel, G. (2022). Patients' and public views and attitudes towards the sharing of health data for research: A narrative review of the empirical evidence. Journal of Medical Ethics, 48(1), 3–13. https://doi.org/10.1136/medethics-2019-105651

Kirkham, E. J., Lawrie, S. M., Crompton, C. J., Iveson, M. H., Jenkins, N. D., Goerdten, J., Beange, I., Chan, S. W. Y., McIntosh, A., & Fletcher-Watson, S. (2022). Experience of clinical services shapes attitudes to mental health data sharing: Findings from a UK-wide survey. BMC Public Health, 22(1), 357. https://doi.org/10.1186/s12889-022-12694-z

Lewis, T., Gangadharan, S. P., Saba, M., & Petty, T. (2018). Digital defense playbook: Community power tools for reclaiming data. Detroit: Our Data Bodies. https://www.odbproject.org/wp-content/uploads/2019/03/ODB_DDP_HighRes_Single.pdf

National Data Guardian. (2024). National Data Guardian 2022-2023 report. https://www.gov.uk/government/publications/national-data-guardian-2022-2023-report/national-data-guardian-2022-2023-report

Paprica, P. A., McGrail, K. M., & Schull, M. J. (2019). Notches on the dial: A call to action to develop plain language communication with the public about users and uses of health data. International Journal of Population Data Science, 4(1). https://doi.org/10.23889/ijpds.v4i1.1106

Patel, R. (2021). Participatory data stewardship: A framework for involving people in the use of data. Ada Lovelace Institute. https://www.adalovelaceinstitute.org/wp-content/uploads/2021/11/ADA_Participatory-Data-Stewardship.pdf

Rempel, E. S., Barnett, J., & Durrant, H. (2018). Public engagement with UK government data science: Propositions from a literature review of public engagement on new technologies. Government Information Quarterly, 35(4), 569–578. https://doi.org/10.1016/j.giq.2018.08.002

Sartori, L., & Bocca, G. (2023). Minding the gap(s): Public perceptions of AI and socio-technical imaginaries. AI & SOCIETY, 38(2), 443–458. https://doi.org/10.1007/s00146-022-01422-1

Yates, P. S. J., Carmi, D. E., Lockley, D. E., Wessels, P. B., & Pawluczuk, D. A. (n.d.). Me and my big data: Understanding citizens data literacies research report

Appendices

Appendix A: Participant Demographic Characteristics

Table 1: Self-Reported Participant Demographic Characteristics of Four Cheshire and Merseyside Focus Groups, n=34

Demographic Category	Count (n)	Percentage (%)
Ethnicity		
White: English / Welsh / Scottish / Northern Irish / British	26	76.5
Asian / Asian British: Indian	2	5.9
Asian / Asian British: Pakistani	2	5.9
Black / African / Caribbean / Black British: African	2	5.9
Asian / Asian British: Chinese	1	2.9
Mixed/multiple ethnic groups: any other Mixed/Multiple ethnic background	1	2.9
English as a First Language		
Yes	31	91.2
No	3	8.8
Relationship Status		
Married	14	41.2
Single	10	29.4
Living with partner	5	14.7
Divorced	4	11.8
Widowed	1	2.9
Health Problem or Disability that limits day-to-day activities lasting or expected to last at least 12 months		
Yes	10	29.4
No	18	52.9
Somewhat	6	17.6
Self-reported Disability (multiple selections)		
Physical disability	10	29.4
Mental health condition	7	20.6
Long-term illness	6	17.6
Neurodivergence	4	11.8
Sensory disability (deaf, hard of hearing, blind, visually impaired)	1	2.9
Prefer not to say	1	2.9
No/None	2	5.9
Carer		
Yes	15	44.1
No	18	52.9
Not answered	1	2.9
Sex and/or Trans Identification		
Female	22	64.7
Male	10	29.4
Trans man	2	5.9
Trans woman	0	0
Age		
16-19	2	5.9
20-24	1	2.9
25-29	2	5.9
30-34	1	2.9
35-39	5	14.7
40-44	1	2.9
45-49	2	5.9
50-54	4	11.8
55-59	3	8.8
60-64	5	14.7
65-69	3	8.8
70-74	4	11.8
75-79	1	2.9
80+	0	0.0

Appendix B: Recruitment Materials and Focus Group Guide Links

Recruitment materials and a copy of the focus group guide are available on the NHS Cheshire and Merseyside Data into Action website: www.cipha.nhs.uk/about-us

A written script of the slides and focus group content is listed below.

Secure Data Environment Focus Group Script

Focus Group Introduction

Today, we will be exploring your thoughts about the use of personal health data for:

- research
- population health management
- risk stratification

We will explain these terms as we go through the session.

We will do three interactive activities over the course of two hours, with a 5-10 minute break between activities.

Before we start, if you haven't done so already, please use the link in the chat to fill in the consent form.

Participant Introductions

Why is health data so important?

Information collected by health and care organisations helps to:

- improve individual care
- speed up diagnosis
- plan local services
- research new treatments.
- Data saves lives. We all benefit from how it's been used by researchers over many years
- It also saves everybody time, and the NHS and care services money that can be put back into patient care.

Remind participants of the video links provided in the pre-reading.

Example of a health data projects

An example of a project that uses health data would be linking GP mental health data regarding low-level depression with anti-depressant use to explore the effectiveness of current drugs and suggest new treatments to patients.

It might show how talking therapies boosted their recovery, or how starting to exercise improved their mental wellbeing.

Activity 1

In the chat:

What's the first thing that comes to mind when you hear the term 'health data'?

What words come to mind?

What images do you see?

What experiences does it make you think about?

Activity 2

Health data = facts, details, statistics, or any information about health pulled together for reference or analysis.

What does or doesn't make sense about this definition?

What's missing from this definition?

Using NHS health data for "primary uses"

Using health data for "primary use" means using data for self-care and direct care, like when you see your GP or go to the pharmacist.

The data is being used for its primary reason – to provide you with direct care.

Using NHS health data for "secondary uses"

Using health data for "secondary uses" means using health data for purposes secondary to why it was originally collected

Secondary uses include:

- carrying out research
- assessing a patient's risk of developing certain diseases or medical conditions, also known as risk stratification
- planning and improving health and care services, also known as population health management.

Secondary uses help **support** direct care. They help not just you, but also others who use health and care services.

Introduce main question of session:

What are your **hopes, expectations, or worries** with **using de-identified personal data** for secondary use cases in health **without explicit individual consent**

Break

Presentation and Discussion: What does "de-identification" look like?

Data unique to you like NHS number. These are removed for secondary uses, which helps to protect privacy.

Identifiable data like your post code is not unique but could identify you. This data is aggregated. Data aggregation combines and summarises information to analyse trends and patterns. Identifiable data is made less detailed for research to protect privacy.

Data can also be minimised. Data minimisation is a principle that emphasises collecting and processing only the essential information necessary for the specific purpose. Identifiable data is **removed** to protect privacy.

But personal or sensitive information can sometimes help us to answer important questions, like "Why are some people more likely to get cancer or die young?" De-identification protects privacy while still allowing research to be carried out.

Not all information is personal. **Working with the public will be important to determine what data needs extra safeguards and protections due to sensitivities.**

Introducing the Five Safes in the Secure Data Environment: Safe People, Projects, Settings, Data, and Outputs.

- Only approved researchers can access
- Only projects that show public benefit can use data
- Data stays in the SDE like a digital laboratory
- Data is de-identified
- Only anonymous data is used in publications and reports. This means data is grouped together and not individual.

Discussing consent

- In 'traditional' research, like a trial, you are approached because you have a condition that is being researched at your hospital.
- They may do additional tests or try new medications with you to see how this impacts your condition and health over time.
- This would require 'explicit consent' – signed forms, with specific explanations.
- Healthcare providers are allowed to collect data under the legal framework of 'public task'.
- This means it is essential that certain information, including identifiable information, is collected and stored to deliver a public service – without your consent.
- This is because it's necessary for them to provide you with the best care as part of their duty to serve the public.
- Secondary use cases, such as research and service planning, do not require explicit consent.
- This is because de-identified health data used in research and planning poses very low personal risk compared to clinical trials, as it doesn't involve direct interventions like new medications or additional tests.
- De-identified data is primarily used to the greater benefit of the public. It helps researchers analyse trends, patterns and outcomes without directly impacting individuals.
- De-identification also ensures that individual identities are removed from the data, safeguarding privacy and reducing the need for consent.

Activity 3

When you think about using health data for secondary uses without direct consent...

What is the first concern you think of when you hear about health data and research?

What is the first hope you think of when you hear about health data and research?

What kinds of concerns do you think your family or friends might have about de-identified health data for research?

Case Studies

Example 1: Connecting child asthma, air pollution and home sensor data to evaluate what is impacting increased asthma rates to create policy recommendations.

Example 2: Your family member has noticed that their COPD isn't as well controlled. They have needed a lot more steroids and antibiotics recently.

A smart health system would also notice this. It could send a notification or prompt to their GP or a specialist to ask if a review to improve preventative measures could be carried out. This could include adding a new regular medication, checking for damp and mould in the house or encouraging the individual to stop smoking.

Activity 4

When you think about research, what kinds of things come to mind?

Do the research examples match what you thought health data research is?

This research does not use individual consent. How does that make you feel?

How important or not would you rate using health data for research like this?

Activity 5

Let's talk more specifically about communication around the secure data environment and what you think other people would want to know about it.

What kinds of questions come to mind when you hear the phrase "secure data environment"?

What is the main question you think people will have about the secure data environment?

What and how should we be telling people about it?

Close

Authorship and Acknowledgements

Dr Emily Rempel from the University of Liverpool co-designed the draft focus group guide, facilitated the focus groups, completed the analysis, and wrote the report. Sophie Powers, Steven Birch and Christine Elliott from Arden & GEM supported the design of the focus group, completed the focus group process plan, recruited participants, and conducted the evaluation of the focus groups. Dale Kirkwood from NHS Lancashire and South Cumbria Foundation Trust co-designed the focus group and supported the facilitation of the first two pilot focus groups.

The project team would like to thank the focus group participants for their time and thoughts. We'd also like to acknowledge the support of Health Innovation NWC in recruitment, design, and notetaking for the focus groups. Finally, a thank you to Leadership within the Data into Action programme and Secure Data Environment teams for devoting their time and authenticity in hosting these important discussions.