National Centre for Social Research

Llywodraeth Cymru
Welsh Government

Using health data to identify and approach people about health and care research: A Public Dialogue

Authors: Harry Pearse, John Evans, Alessandra Radicati, and Eleanor O'Keeffe,

Prepared for: Welsh Government

JULY 2024

We believe that social research has the power to make life better.

By really understanding the complexity of people's lives and what they think about the issues that affect them, we give the public a powerful and influential role in shaping decisions and services that can make a difference to everyone. And as an independent, not for profit organisation we're able to put all our time and energy into delivering social research that works for society.

National Centre
for Social Research
35 Northampton Square
London EC1V 0AX
020 7250 1866

www.natcen.ac.uk

A Company Limited by Guarantee

Registered in England No.4392418

A Charity registered in England and Wales (1091768) and Scotland (SC038454)

This project was carried out in compliance with ISO20252

designbysoapbox.com

Contents

Executive Summary

In November 2023, the Welsh Government commissioned the Centre for Deliberation at the National Centre for Social Research (NatCen) to run a Public Dialogue with members of the Welsh public. The dialogue sought to understand participants' views on processes that use health data that identify and approach people about taking part in research studies hosted by NHS Wales. The aim of the dialogue was to inform policymaking in Wales and the UK on the question of whether such processes should be allowed to access patient healthcare data without prior consent.

Drawing from the NatCen panel – a random probability sample of over 22,000 people from across the UK – 35 Welsh participants were purposively recruited for the dialogue, reflecting a range of demographic characteristics and backgrounds. Participants were given briefing materials to ensure everyone was deliberating from a shared evidentiary basis, and throughout the dialogue, participants were briefed by, and had opportunities to put questions to, a range of expert speakers.

The dialogue was split into four sessions. The first three, which were online, each lasting 2.5hrs, were designed to introduce participants to key elements of the processes used by UK health services to identify and approach prospective research study participants, and the role that health data has played, and could play, in those processes. The fourth, deliberative session was an all-day face-to-face event, in which participants were split between two locations: Bangor and Cardiff. As the culmination of the three preceding exercises, the final session generated the most informed and reflexive data, and, as such, is the focus of this report. During this final session, participants engaged in a series of scenario exercises.

- The first scenario encouraged them to think about the relative importance of diversity, equity, and inclusion in processes that rely on patient health data to identify and approach people about research studies.
- The second explored participant views on the role and status of personal autonomy in such processes, and how patient privacy and confidentiality ought to be balanced against patient and public benefits.
- The third asked participants to assess a particular digital process for identifying and inviting people to take part in health-related research, and to decide whether they would be comfortable with a process that relied on accessing health data without prior consent.
- In the final exercise, participants developed and voted on a set of expectations regarding future processes for identifying and approaching prospective research study participants in Wales.

The dialogue's key question – posed in different forms throughout the in-person session – was: **how comfortable are participants with processes that identify and approach people about healthcare studies by accessing their NHS health data without prior consent**. On this, and every other question posed in the dialogue, there were no unanimous decisions. Some participants remained concerned about data-led processes, pointing to the distress people might feel if they were contacted about what might be considered a sensitive health condition – for example, a mental health condition. However, other participants expressed support, arguing, among other things, that participation in research studies would likely be advantageous for the invitee, and/or have a positive bearing on the public interest – improving research efficiency or outcomes and possibly public health benefits.

Nevertheless, on balance, our analysis shows that a sizeable constituency *would* be comfortable with processes that access un-consented patient health data, provided certain conditions are met. The participants comfortable with such processes usually based their assessment on **conditions** and **safeguarding principles**. And it was often the absence of these same conditions that caused other participants to withhold their support. Thus, if the following conditions are met, a broad constituency would likely be satisfied:

- **Research quality:** accessing and using patient health data must have a material impact on health and care research – improving recruitment efficiency or facilitating more precise or better-quality research outcomes.

- **Equity:** participants held that using health data in processes that identify and approach participants should contribute to efforts to make health and care research more inclusive, diverse, and accessible.

- **Public benefit:** uses of health data for such processes should prioritise public needs and generate benefits for the entire Welsh population – public benefits must have primacy, particularly over commercial or profit-driven considerations.

- **Data safeguarding:** processes that rely on patient health data must be carefully regulated and subject to oversight procedures.

- **Public involvement:** the Welsh public should be kept up to date on future initiatives via public engagement campaigns; provided with relevant information and consulted about potential plans.

- **Individual autonomy and privacy:** must be respected. For a minority of participants, personal autonomy and privacy outweighed all other considerations – in some cases, ruling out the possibility of permissive data-driven processes. However, for others, a commitment to protecting individual privacy was important, though not more important than the potential benefits of developing more effective processes to identify and approach people about research studies.

1. Introduction

1.1 Context

In recent years, in Wales and across the UK, recruitment processes that access routinely collected NHS health data to identify and approach individuals about research studies have advanced in various ways. Broadly, there are two models for these data-based processes: opt-in registers or platforms ('consent to contact'), and digital intermediaries (operated via local care providers or at a national level). Furthermore, during the Covid-19 pandemic, exemptions to normal information governance processes established an alternative model for using data to identify and approach individuals for research (in this case, for Covid-19-related studies). Although each of these models has, at times, been used effectively, they also have drawbacks or limitations.[1] (Dialogue participants were briefed on, and discussed these uses and drawbacks in the dialogue's early information-sharing sessions).

Opt-in platforms are registers or websites that allow members of the public to sign up to be notified about research opportunities. These platforms operate in different ways. For example, they might invite people to express an interest in a particular medical area or type of research, or to declare their patient information, or to consent to having their care record linked to their identifiable confidential patient information. Alternatively, people can sign up to be contacted about research opportunities without first submitting their medical data.

Several such platforms/registers are currently operating in the UK. Some are managed by NHS organisations or medical charities, and are operated locally, regionally, or nationally. Usually, they are publicly advertised – on posters in GP surgeries or relevant websites – and are available to anyone. However, because they are voluntary – and rely on people being aware of the platform/register and consenting to be contacted about participation – opt-in platforms tend to produce uneven recruitment pools; skewed towards certain demographics and less representative of under-served communities.

Digital intermediaries, on the other hand, are organisations (commercial or non-commercial) that match potential research participants to research studies. Often, they work by presenting de-identified patient information (data that cannot be linked to a specific person because identifying information, like a name or address, has been removed) to researchers outside the patient's care team. This allows researchers to search existing electronic health records without accessing individual records or identifying individual patients. Alternatively, a digital intermediary might move de-identified patient information from electronic health records to a separate database. Such databases might collate information from various sources – e.g. GP and hospital records – and be searchable to identify eligible candidates for different studies.

To stay within legislative boundaries and information governance rules, some digital intermediaries work with direct care professionals. For example, they might supply GP Practices with an electronic health system that crosschecks patient records against research study inclusion criteria and flags eligible individuals to their GP. The success of such approaches relies on the availability and willingness of direct care professionals to initiate recruitment, many of whom are overworked and busy. In other cases, organisations extract de-identified data from multiple sources and combine them into searchable databases that match studies to potential participants. NHS Digitrials is one such example.[2] This approach relies on the Confidentiality Advisory Committee (CAG) – empowered by the Health and Care Act section 251 and Control of Patient Information (COPI) Regulation – deciding that using confidential patient data without consent is acceptable. The CAG takes these decisions on a case by case basis.

During the Covid-19 pandemic, the standards for determining when identifiable health data could be shared (beyond one's direct care professional) were temporarily relaxed to ensure health research processes could

[1] Health and Care Research Wales (2023), *Research matters: our plan for improving health and care research in Wales 2022-2025*. Available at: Research matters: our plan for improving health and care research in Wales 2022 - 2025 (healthandcareresearchwales.org). (Accessed 26/07/24).
[2] NHS England (2019), *NHS DigiTrials*. Available at: NHS DigiTrials - NHS England Digital. (Accessed 26/07/24).

respond to the demands of a public health crisis. This was done through legal notices issued under The Health Service (Control of Patient Information) Regulations 2002 by the Secretary of State for Health and Social Care and Welsh Government, instructing NHS organisations in England and Wales to share identifiable confidential patient information for research and other purposes.[3] The exemptions have since lapsed. However, while in effect, they demonstrated the viability of an alternative, more rapid, process for identifying and approaching people about research studies, leading to calls from some within the research community to make such governance arrangements permanent.

Policy discussions are ongoing in Wales and across the UK about how to most appropriately develop and deploy processes to support the use of data to identify and approach individuals – particularly digital approaches. To inform the discussions, Welsh policymakers want to understand public views and preferences. And consequently, the Welsh government commissioned the Centre for Deliberation – part of the National Centre for Social Research (NatCen) – to design and conduct a Public Dialogue to capture and map the perspectives of Welsh citizens. Its objectives were:

- To understand the spread and weight of participant **attitudes towards NHS Wales' use of health data for contacting people about research opportunities**.
    - Specifically, to gauge whether the Welsh public would **accept processes that identify and contact people about research studies by accessing their health data without explicit prior consent**.
    - To explore the balance between ideas of **individual privacy** and **confidentiality**, and **the public good**.
- To uncover the Welsh public's **expectations about future uses of health data in processes that identify and approach people about research**.
- To **inform Welsh government policymaking** on processes that use health data, as well as broader UK-wide discussions about similar issues.

1.2 Method

To fulfil these research objectives, NatCen delivered a four-part Public Dialogue in Wales in April 2024.

A Public Dialogue is a form of deliberative research in which a small group of citizens – broadly reflective of a larger population – talk with one another and experts about particular policy questions. To prepare for a dialogue, participants are given briefing materials illuminating the issues to be discussed, and during the events they watch presentations by, and interact with, experts. Throughout, they are encouraged to engage with the views, values, and preferences of other participants. A dialogue runs for several hours, usually over a number of days, which gives participants the time and space required to reach considered, and sometimes consensual, decisions – on policy recommendations, principles, or a set of public expectations.

To deepen participants' engagement with processes that use health data to identify and approach people about research studies, this Public Dialogue took a scenario methods approach to public deliberation.

Scenario methodologies are discursive tools that encourage people to think through the experiences, preferences, and choices of other people. At root, they are a form of storytelling; exposing people to hypothetical scenarios involving hypothetical people, and asking them to navigate a fictional situation through the eyes of the imagined person. The aim is to give people insights into points of view that might otherwise be unknown or occluded to them. Scenarios encourage participants to consider other people's experiences when evaluating public policy, and to begin working with concepts of the public good. In this project, this method was used to introduce participants to a variety of individual or constituency perspectives on the use of health data, as well as a more distant, systemic

[3] NHS Health Research Authority (2019), The Health Service (Control of Patient Information) Regulations 2002: regulation 5 decision procedure for research applications. Available at: The Health Service (Control of Patient Information) Regulations 2002: regulation 5 decision procedure for research applications - Health Research Authority (hra.nhs.uk). (Accessed 26/07/24).

viewpoint. As detailed below, all major project decisions – regarding key policy questions and scenario design – were taken in consultation with the Welsh Government and its Data for Research Working Group.

Design: The dialogue was comprised of three 2.5 hours online sessions (hosted between 9 and 22 April 2024), and an all-day face-to-face session on 27 April 2024. Where necessary, participants were given technical support and advice – loaned computer hardware and offered Zoom tutorials. And to make attendance of the fourth session as easy and convenient as possible, participants were split across two locations – half in Bangor, the other half in Cardiff.

The first three sessions were introductory and covered the following ideas and policy areas:

- Session 1: Health data, health and care research, and the importance of recruitment. The benefits of using health data in processes that identify and approach potential participants (generating larger, more diverse, and more efficiently managed, research samples); as well as the challenges (questions of privacy, confidentiality, and institutional trustworthiness).

- Session 2: Current models for using health data in this way – opt-in and data intermediaries (direct care and digital). The benefits, drawbacks, and trade-offs of each.

- Session 3: Data and recruitment safeguards. The role of the HRA Confidentiality Advisory Group (CAG), the time-limited exemptions to particular safeguards introduced during the Covid-19 pandemic, and whether the rules developed in a period of public emergency can or should be utilised in ordinary circumstances.

However, the analytic focus of the dialogue centred on the fourth, day-long session. And therefore, although we collected data on participant views in the three prior sessions, the objective of these sessions was to ensure participants were familiar with the relevant subject matter and prepared to participate in the final, deliberative session on 27 April.

The fourth session was structured differently, revolving around scenario exercises, and encouraging participants to think through the experiences of different constituencies, as well as the broader public benefit. At the end, they engaged in a group decision-making process about the future of processes that contact people about research opportunities. As intended, this day-long session yielded the most reflexive, deliberative data. And it is this data that forms the basis for this report.

The themes, scenarios, and research objectives explored in the fourth session were as follows. (Full details of the scenarios can be found in the Digital Appendix).

Scenario 1	**Theme:** Diversity, Equity, Inclusion. **Scenario:** 'Kath' is contacted and approached – based on her ethnicity, gender, and mental health background – to take part in a study about a new form of therapy. The process of identifying and approaching Kath is one that is not currently common practice. (It operated for several Covid studies during the pandemic, though only through exemptions provided by legal notices). **Objective:** To understand whether, or to what degree, processes that identify people according to demographic identifiers or personal health conditions, affect their willingness to take part in research studies. To gauge whether being approached about a 'sensitive' health condition impacts someone's view on that process. Likewise, to understand whether being approached about more generic or population-wide studies influences people's openness to participating.
Scenario 2	**Theme:** Personal Autonomy. **Scenario:** 'John' – who is overweight and has hypertension and does not like interacting with his GP – is contacted and approached about taking part in a study regarding early prostate cancer detecting and treatment. Again, the process of identifying and approaching John is not common practice at present. However, a variable to the scenario – introduced

	after participants' initial discussions – changed the process to one that is commonly accepted as legitimate within current information governance rules. **Objective:** To explore how comfortable participants are with processes that identify and approach people regarding conditions they do not currently or knowingly suffer from (though may do in the future) – people who might therefore be reluctant to engage with healthcare professionals and infrastructure. Whether the public benefit, or the potential upshots for invitees themselves, offsets people's reluctance to be contacted about research. And whether the source of an invitation to participate – be it a commonly accepted method (via a GP) or an alternative (via NHS staff not directly related to an individuals' care) – affects people's willingness to participate.
Scenario 3	**Theme:** Permissiveness of data practices. **Scenario:** 'Jenny' must evaluate a research application to use a digital intermediary tool – weighing up the risks, benefits, stakeholder perspectives etc. She must advise the Welsh Information Governance Board to approve or reject the application. **Objective:** To gauge people's comfort levels with the prospect of processes that identify and approach potential participants based on controlled but un-consented access to their health data. What would such processes need to look like to assuage public concerns about privacy and confidentiality.
Future Expectations	**Theme:** Public expectations for future processes. **Activity:** Draft principles and expectations on flipcharts. Then – as a group – vote for and against the most and least supported statements. **Objective:** Develop and winnow down a set of expectations about future processes for identify and approaching people about research studies in Wales

Sampling and recruitment: Participants were purposively recruited from the NatCen panel – a 22,000-person random probability sample frame. Of the panel's Welsh cohort, a total of 35 people – stratified by age, sex, education, socio-economic background, ethnicity, rurality or urbanity – completed the dialogue (see Appendix for full breakdown of participant demographics). During the recruitment process, to ensure participants' informed consent, NatCen circulated a privacy notice and a dialogue briefing document. Participants who completed the dialogue were offered a £250 incentive.

Delivery: Before the first session, participants completed a survey assessing their general views and understanding of data issues and policymaking (see Appendix). The survey revealed that most participants held data to be a socially important topic; that many of them were concerned about their data being stolen or used for profit; and that most of them felt they knew at least a little about how their data was collected and used.

Before each of the three learning sessions, participants received briefing notes and explainer videos – drafted and compiled by NatCen researchers and our academic partner, Dr Rachel Thompson (Swansea University / University of Oxford). The format of each session alternated between plenary (whole group) exercises and breakout (small group) discussions. During the plenary sessions, participants were briefed by representatives from Understanding Patient Data, MedConfidential, use My data, the Centre for Trials Research, the University of Oxford, and the Ada Lovelace Institute (for the full list of speakers, their titles and institutional affiliations, see Appendix). The same speakers then circulated among the breakout groups, fielding questions, and engaging in discussion with participants. In their small groups, participants discussed what they had heard and conducted short elicitation exercises – on interactive whiteboards (Miro) or Google Forms. The sessions were audio recorded and facilitators took notes throughout.

At the final face-to-face session, participants worked primarily in small groups, recording their responses to, and assessments of, scenario exercises on handwritten worksheets. At the end of the session, they drafted statements

outlining their expectations for the future – displayed on flip charts. The entire group (split across Bangor and Cardiff) then voted on their favourite and least favourite statements. All sessions were overseen by a lead NatCen facilitator, and small group discussions were moderated by NatCen facilitators, or, in the case of the Welsh language groups, Datblygiadau Egni Gwledig (DEG).

Once the deliberations were over, participants completed a final survey, confirming the most significant public benefits or concerns about data-driven recruitment processes (see Appendix).

1.3 Approach to Analysis

Data was captured prior to, during, and after the workshops. The evidence generated from the fourth session and used in this report derives from participants' worksheets and flip charts, and the audio recordings of their exercises. In the previous recorded sessions, participants used Miro and Google Forms. Participants also completed pre- and post-deliberative surveys.

NatCen researchers collated and inductively coded the data from the fourth session – reviewing and discussing the findings and identifying emerging codes and themes. The scenario worksheets completed by participants asked particular questions – for example, about how people might feel or what choices they might make in the given scenario. Our inductive codes were informed by these framing questions. The final exercise in the fourth session, in which participants set out their preferences for how processes for identifying and approaching research study participants in the future, was framed around the three categories: expectations, principles, and recommendations. These categories formed the foundation for our inductive coding process.

To reach our final conclusion about participants comfort, or otherwise, with processes that identify and approach people about research studies based on un-consented access to their health data, we compared and integrated their responses to the range of exercises conducted in the fourth deliberative session.

1.4 Reading this Report

Chapters 2, 3, and 4 summarise participant responses to, respectively, the first, second, and third scenarios discussed in the fourth dialogue session. They also set out our analysis of these responses. Chapter 5 outlines participants' expectations for the development and use and data-based processes in the future.

1.5 Working with the Welsh Government

This dialogue could not interrogate every aspect of the different possible processes for identifying and approaching people about participating in health and care studies. Instead, it prioritised themes and inquiries of interest to the Welsh government and achievable within deliberative parameters. The process of selecting what to include in the dialogue was collaborative and iterative, undertaken by the Welsh Government and NatCen researchers, drawing on advice from the Welsh government's Data for Research Working Group (DRWG) – a group of health data experts from NHS Wales and Welsh academic institutions, NHS researchers and research delivery staff, Welsh Government digital policy officials, and lay members. (The groups and organisations comprising the DRWG can be found in the Appendix).

The Welsh government began by conducting informal conversations with data experts and researchers across the UK and reviewing relevant academic and grey literature.[4] From this, it concluded that certain policy inquiries had not received adequate attention – including how the public felt about patient data being used to identify and approach people about research studies without their prior consent. The Welsh government and NatCen then convened a series of DRWG meetings (January – March 2024) to identify the issues due to be explored in the dialogue and identify the variables likely to affect public expectations about data-driven processes (see Appendix for longlist of variables discussed by the DRWG). The group also agreed on the contextual information necessary for participants to engage in the dialogue, and how this information could be conveyed in an accessible way.

Ultimately, the Welsh government and NatCen resolved that the dialogue should explore participants attitudes to NHS data-based processes that identify and approach eligible individuals about research, both though 'expected'

[4] Understanding Patient Data (2021), *How do people feel about the use of data?* Available at: How do people feel about the use of data? | Understanding patient data. (Accessed 26/07/24); Understanding Patient Data (2024). New research: Public attitudes towards patient data for planning and population health. Available at: New research: Public attitudes towards patient data for planning and population health | Understanding patient data. (Accessed 26/07/24); and NHS England (2024), *Public attitudes to data in the NHS and social care*. Available at: Public attitudes to data in the NHS and social care - NHS England Digital. (Accessed 26/07/24).

routes (i.e. via a GP or clinician), and through centralised processes that do not rely on the explicit involvement of invitees' direct care professionals. Themes of diversity and inclusion, personal autonomy, the perceived sensitivity of a condition and/or type of data, and the permissibility and safeguards of any potential processes were chosen as key issues. The final exercise of the deliberation focused on public expectations about data-based processes in the future.

2. Diversity, Equity, and Inclusion

2.1 Introduction

This chapter will set out the first scenario covered during the in-person, all day dialogues which ran in Bangor and Cardiff. The scenario introduced participants to a hypothetical person – Kath – who had been approached to participate in a research study due to her ethnicity, gender and condition. Participants were taken through three exercises: First, they were asked to discuss how Kath would feel being identified and approached about research in this way. They were then prompted to discuss whether the sensitivity of Kath's condition would influence her response. Finally, participants were asked how Kath would feel about being asked to participate in research concerning a less sensitive condition, such as the flu. The aim of this scenario was to prompt participants to think about how sensitive, identity-defining characteristics, like ethnicity, gender, and particular medical experiences, might influence the reaction of someone who had been identified to participate in a health and care research study on this basis. Underlying these inquiries were two broader questions about: whether participants thought data practices affect different individuals and communities differently; and whether participants thought sensitivities regarding identity/demographics/health conditions should be accommodated in practices that use healthcare data to identify potentially eligible individuals for research. The scenario described a method of identifying and approaching individuals by NHS staff not directly involved in the participants care. This method is not currently common practice, though it operated for several Covid studies during the pandemic, and only on an exceptional basis.

This chapter draws on the worksheets that participants completed during the exercise, as well as the transcripts of their discussions. The details of the Scenario are set out in the box below.

Scenario 1: Kath

Kath is of Bangladeshi heritage and has suffered from anxiety and post-natal depression for the past three years. She found it tough to talk about her feelings at first and does not trust healthcare professionals due to a negative experience in the past. However, she has found a sympathetic GP who has helped her manage her condition. She is currently on a course of anti-depressants and is starting to feel better and more optimistic.

One day, Kath receives an email, addressed to her, and signed by an NHS researcher, informing her of an NHS research project that could be relevant to her. The email says that Kath's health data suggests she is eligible for the study due to her gender, ethnicity, and her mental health condition. The project is trialling a new therapy to help people from Kath's background and area, who have similar mental health conditions. This includes a new online CBT programme, alongside access to a community support hub. There is a number for Kath to call if she's interested in learning more. Kath closes the email. She can't think of how they got her email address or knew about her eligibility.

2.2 Key Findings

Participants were asked ***how Kath would or should feel about being identified and approached on the basis of her gender, ethnicity, and mental health condition to participate in a research study about a new form of therapy***. Many participants' responses were **negative**, suggesting that Kath would feel targeted, anxious, confused, angry, and betrayed that her health data had been shared. However, there were some responses which were more positive about the situation. These generally surrounded the potential benefits to either Kath herself (through enhanced treatment opportunities) or the wider public (through providing data which could help others in similar contexts). The following section will unpack each of these reactions, starting with the predominant feeling that contact of this nature would provoke a negative response.

Targeted

The fact that Kath was identified due to her ethnicity – or status in a marginalised group – was liable to make her feel 'singled out' and 'vulnerable', according to some participants. Being identified on the basis of her gender and condition (post-natal depression) could also be unsettling to her, provoking concerns about how researchers knew this information, and whether the study was real and legitimate. As it was known that Kath had experienced anxiety and low mood in the past, some participants felt the researchers should have anticipated that this form of contact – referencing personal medical experience and sensitive characteristics – would evoke a negative emotional reaction.

> *'She'd be upset and worried especially as she's already taking medication for depression…I think mental health issues are somewhat taboo and private, so she'd probably a bit sensitive about that. Also, it could have been that the email alienated her, targeted her, that she felt discriminated against because of her race.'*

However, **most of the negative responses referred to more general (i.e. not diversity/inclusion-specific) concerns** about Kath's likely reaction:

Anxious

Participants were concerned about how and where researchers had accessed Kath's health data, and where else that same information might have ended up. Some assumed it had been shared by Kath's General Practitioner (GP). However, others raised concerns about who else might have access to it, and the amount of data now potentially available to others. This concern about data security extended to worries about how information about Kath's condition would be managed. For example, there were discussions about how stigma surrounding mental health would make Kath concerned about this information being available to current or future employers. These findings echo concerns expressed by participants in the surveys conducted before and after the dialogue, where data being used for purposes beyond the intention it was collected for (being stolen or shared illegally, for example), topped participants' list of concerns about the operation of data-driven processes for identifying and approaching people about research studies.

Confused

Participants said Kath would be confused about why her GP had not contacted her about the study, arguing that communication with a direct care professional would have felt less 'abrupt' and 'out of the blue' than contact from an NHS researcher. The mode of delivery was also seen as a problem – participants said emails could be perceived as impersonal, and an inappropriate channel for raising highly personal and sensitive topics. It was similarly argued that some people may not be comfortable conducting important and personal matters via email as it can be hard to know what is authentic or spam. Emails also provide no immediate recourse for questions or discussion. Participants described how Kath would likely want to talk to someone about being approached to participate, to find out, among other things, how researchers had accessed her data. For some participants, the ability to respond to the email by telephone would make Kath more likely to participate in the study, as she would be able to access guidance and understand the next steps in the process. Nevertheless, for others, even if the original email approach included a phone number, the invitation would be viewed with suspicion.

> *'I think she thought: is this genuine or a scam … because I would wonder where it's come from, why, how do they know about me, where did they get everything from? I would worry.'*

Angry

Without knowing how her data was accessed, Kath may assume her privacy had been breached or violated. One participant described how being approached about research without any information about *how* she was identified would feel like an 'invasion of privacy'.

> *'Depending on what's included, if you suddenly out of the blue received an email with, "we know that you've got post-natal depression after your second kid"…you're more likely to feel really angry and not want to do it because of how it's been presented to you. It really does have to be very sensitive, not just [in regards to] the sensitivity of the data but in the sensitive way it's presented to her to get the outcome that the researcher wants.'*

Betrayed and Doubtful

Given the above misgivings, Kath's already-low confidence in healthcare professionals was seen as likely to wane further. Participants went as far as to suggest that her faith in the legitimacy of the healthcare system might also decline.

> *'I think it would affect her a lot because of the mental health and gender issues. It boils down to a lack of trust and betrayal, feeling betrayed. I'd feel so betrayed by the person [who had shared my information] - it's obviously come from the doctor.'*

Nevertheless, as well as these negative reactions, several participants saw reason to be **more optimistic** about Kath's position, arguing that she may feel:

Intrigued

The prospect of acquiring better and more specific treatment for her condition was seen as a positive opportunity for Kath. Participants argued, for example, that she would welcome the option to explore treatment options, as well as the opportunity to discuss her condition with others. They also said that Kath's surprise and worry at being contacted unexpectedly would be offset by the assumption that her data was obtained legally by NHS researchers, who would presumably have Kath's treatment at the forefront of their considerations.

Potential Public Good

Participants believed Kath would also be cognisant of the potential public good of participating in health & care research and that this might offset or attenuate her anxieties regarding the availability of her health data:

> *'If you shop in a supermarket, you've already handed your data over to so many different people. If you've got a loyalty card, if you click "accept cookies". So, when we live in that world where our data is being given [away] for nothing…we have to redress the balance a little bit and say, "Okay, we can look at a different approach…for recruiting to research that is for the public benefit"…[And so] if the public benefit is high, it's worth a little bit of leeway in terms of the risk.'*

Some participants described how Kath's experience of anxiety and depression in the past, and her empathy for others in similar situations, may motivate her to participate in the study. Although, for a very small minority of participants, such an outcome would amount to guilt-based coercion – compelling Kath to participate even if she did not really want to, and in the face of the negative emotions described in this chapter.

Language and Cultural Differences

Kath's demographic identity caused some participants to reflect on whether language barriers or different cultural expectations might affect how the approach was perceived or understood. Participants discussed citizens' variable familiarity with NHS processes, and how they might be misinterpreted by people or communities who have minimal contact with, or trust in, NHS infrastructure. Consequently, it was argued, individuals' first languages, as well as the norms and beliefs associated with different cultural groups, should be considered when deciding how people are contacted and the content of that communication.

> *'How far is she integrated? Is she [Kath] first, or third, or fourth generation Bangladeshi?...She might have experienced, if she wasn't second or third generation Bangladeshi, less benign forms of government. She may well have experience of her society or community being negatively impacted by either inefficient or corrupt government bodies.'*

Participants were then asked ***whether the sensitivity of the condition Kath was contacted about was likely to affect her response***. Most participants thought this factor would have **a critical, determining effect** on how one would respond to being approached about participating in a research study.

Mental Health

Participants agreed mental health was seen as a more sensitive and stigmatising than other, physical conditions; and something that people would be anxious to keep from their friends, family, and work colleagues. For example, participants spoke about how Kath's job security might be unfairly thrown into question if her condition was known by her employer, due to misconceptions about how it would affect her ability at work. Across the workshop, participants felt that the mention of mental health conditions would substantially influence how Kath would react to being approached about research.

> *'I think it's the mental health condition which is one of the principal problems.'*

Some participants also highlighted the intersection between cultural perspectives and mental health, explaining how mental health could be viewed differently in different cultures, and correspondingly be subject to different stigmas. In this vein, participants urged caution in the recruitment approach on the grounds that Kath's family – with their potentially specific, culturally informed perspectives on mental health – may not be aware of her health conditions and may somehow see the email detailing her conditions.

Participants also remarked on the stigma attached to post-natal depression – a mental health condition stigmatised across different cultures due to norms and expectations surrounding motherhood. The sensitivity of mental health conditions also emerged as a finding in the pre-dialogue survey, where it was identified as more sensitive than chronic conditions (like asthma or diabetes), prescriptions or medicine, or demographic characteristics.

Variable discussion: the relationship between the study and the condition:

Finally, participants were asked ***if the relevance of the research study to Kath's particular condition was likely to affect her response***. To explore this question, participants were given a revised version of the scenario, in which Kath was identified as a potential participant, not for a new therapy programme, but for a new flu treatment.

Scenario variation: One day, Kath gets an email from an NHS researcher, asking if she wants to participate in a flu drug trial. The trial could lead to better flu treatments for the Welsh population. The email explains that Kath has been identified from her health data because of her age and ethnicity. The research team have attached information about the medicine on trial and how the study works.

Kath follows up with an email to the project team, asking how they accessed her health data and identified her for contact. She gets a reply quickly, saying that, about four months ago, Kath signed up to be part of a platform called HealthWise Wales. She had given her permission for researchers to access her health data and to contact her about studies that were relevant to specific conditions or health issues she had.
Kath had been busy with her kids that day and had forgotten that she had signed up.

In this revised scenario, most participants felt that Kath would feel more comfortable being contacted about a less personally relevant or triggering condition, like the flu. Participants thought Kath would feel less singled out or targeted (despite still being identified through her age and ethnicity) because flu-related issues tend to be less personally and culturally sensitive. They also thought that widespread familiarity with flu conditions ('Flu inquiry is fine – vaccines regularly advertised on TV') would detoxify requests to participate in studies exploring related, or similarly population-wide, issues – such as those related to Covid-19. Indeed, inquiries about population-wide studies were thought to better invoke the idea of the public good as they are less personal and more public oriented with clear cut social goods. Participants spoke about how those approached would be more likely to feel part of a community which is doing something to benefit society: 'Flu is [a] more prevalent [condition], so more people [will be] approached' for these sorts of studies; further, when they are approached, people are more likely to consider 'the community' than 'the individual'. And when the public benefit was more evident, people had fewer concerns about data privacy or personal risk. That said, some participants felt Kath would be less willing to participate in studies which were not linked to her personal condition. However, this resistance might well be offset if – as would likely be the case – Kath's family members or friends had suffered from the condition.

2.3 Research Questions and Conclusions

Throughout their discussions, participants were asked direct questions, to which they gave direct answers. At the same time, the scenarios were designed to elicit broader, and in many cases implicit, reflections on two additional questions. Although these questions were not put directly to the participants, based on our analysis of their direct responses, we can derive the following conclusions or reflections:

Do the Welsh public think data practices affect different individuals and communities differently?

To some extent, yes: Participants felt that some people will feel anxious and singled out if they are identified and approached to participate in research studies based on their demographic information – like ethnicity or gender. However, **more weight was placed on a few cross-cutting concerns** (i.e. concerns that are not specific to a particular demographic, and which apply to anyone being contacted about research). For example, being contacted about a study concerning a personally sensitive condition, such as a mental health condition; and worrying about how, when, and where this health data was accessed and by whom, and where the information might end up?

Consequently, in many cases, participants were more open to being identified and approached to participate in studies focused on population-wide health conditions, rather than personally relevant ones. This is because they had fewer concerns about their health data being accessed and used in, as well as it being easier to grasp the public benefit of, population-wide studies.

How could sensitivities regarding identity/demographics/health conditions be accommodated in research practices?

Participants highlighted mental health conditions as particularly sensitive and potentially stigmatising – in different ways and in different communities. Their exposure through potentially insecure modes of contact should therefore be considered thoroughly before contact is made.

> *'There's a concern that she [Kath] is in a shared household…[so] how private is her email? Could anybody else look at it?...She may feel put out…that personal data was [mentioned] in…the very first email. I think you shouldn't [mention that]…on the very first email, it should [solely] be about participating.'*

Recruiting for population-level studies on generic conditions, like the flu, was seen as less problematic. Common conditions carry less social stigma, which reduces people's concerns about their health data being accessed or known by others.

Participants made it clear that processes for accessing and managing their health data should be open and transparent about where data was gathered from, how it was processed, and who – if anyone else – has access to it. They also said there should be clarity about the broad privacy safeguards in place to reassure those contacted that their data has only been provided to researchers, and that the researchers themselves are subject to strict controls surrounding how they use that data.

Participants were clear that, when contacting people on the basis of their specific demographics or conditions, the approach should be explicit about why an individual has been selected, as well as frame their participation in terms of the public benefit. The scenario exercises primarily explored how Kath would personally feel about being identified and approached to take part in a healthcare study, inviting participants to focus on questions of personal privacy over questions of the public good. Nevertheless, they also noted that outlining a research study's impact on public benefit might offset people's initial concerns about data privacy. We can infer from participants other remarks that this would be easier to achieve if the study was about a common and much talked about conditions (like flu). The post-workshop survey also revealed that participants think more information on how and why NHS Wales might use data to identify potential participants, would help build public trust in NHS Wales.

3. Personal Autonomy

3.1 Introduction

This chapter will address participant responses to a scenario exploring the importance of personal autonomy. The scenario introduced participants to John, who has several health conditions but is managing them adequately. John generally avoids going to the GP and is reluctant to engage with health services. Participants were asked to consider how John would feel about being contacted by his GP surgery via a phone call, followed by an email, to inform him that some of his health conditions make him eligible to participate in an NHS research study.

The scenario aims to understand two key components of using health data to identify and approach individuals about research studies. First, it investigates whether the dialogue participants are comfortable with data-led processes that contact people, like John – who may be generally averse to seeking medical care – about a medical condition they do not currently have. Second, it seeks to understand if participants' concerns about such processes change when the person approaching people about research studies changes. The initial scenario describes a commonly used process for contacting people that is in keeping with current information governance practices (i.e. via GP with direct care relationship with individual). Later, we varied the scenario so the contacting process differed from what is commonly considered acceptable within current information governance practices (i.e. contact by an NHS staff member not directly involved in an individual's care). Overall, the scenario seeks to illuminate how, or the degree to which, questions of personal autonomy shape or restrict people's views on processes which use health data to identify and approach individuals about research studies, and the degree to which the person making contact with an individual is part of the direct care team, or wider NHS structures, is important to shaping views on acceptability of these methods . Evidence was gathered from participants' written responses to the scenario and its variables, as well as transcripts of their group discussions. The full scenario is presented in detail below.

Scenario: John

John is 53. He has hypertension and diabetes but manages these conditions well. He is otherwise healthy and enjoys life. He lives with his family – his wife and two children – in Brecon. He is active and plays sports with the kids every weekend.

John doesn't like going to the GP much – he says he has got better things to do and going to the GP surgery makes him feel like he's a sick person. He knows he is overweight and is gradually reducing his weight by making small adjustments to his diet. He does not like hearing from the GP about it.

One day, John gets a call from his GP surgery receptionist informing him about an NHS research study that he might be eligible for. John's health record, which documents his weight issues, shows he has a higher-than-average risk of developing prostate cancer.

The research project will trial a new test that could help identify and treat prostate cancer earlier, which would ultimately lead to better health outcomes for men with the disease.

The GP surgery receptionist sends John an email containing more information about the study, as well as a link to register his interest in participating in the study.

3.2 Key Findings

Participants were first asked ***how John would or should feel about being contacted on the basis of his weight and disease-susceptibility to participate in a research study about prostate cancer?*** Responses revealed a range of positive interpretations including: sympathy for the communication protocol described; predictions that John would feel enlightened about his health conditions as a result of being contacted; and discussions of both personal and public benefits arising from this scenario. More negative responses focused on John potentially feeling worried, irritated, stigmatised, angry, or rushed into making decisions about his health.

Overall, participants' responses to the scenario and its variation indicated that most were comfortable with the idea of a person being contacted for healthcare research about a condition they do not currently have, but which they could develop in the future based on certain risk factors. Participants' responses suggest this can override the individual's dislike of being in contact with healthcare professionals or systems.

Communication Protocol

Participants noted several positive aspects of the mode of communication described in the scenario – a phone call followed by an email. Most participants who mentioned the form of communication noted the upsides of being contacted by a GP's office – that it would feel familiar and trustworthy. Others pointed to the personal touch of being contacted by phone, and the benefit of follow up communications being conducted by email – ensuring John was sufficiently informed and had time and space to make a sensible choice. One participant noted that: 'a call is more personal. He has a choice whether to go ahead with it [the research study] or not. [The] follow up email also provides more info'.

Personal Benefits and Learning about Health Conditions

Some participants suggested that John may end up feeling enlightened about aspects of his own health. Though John is aware of his conditions, it is plausible he learned something new about them or their implications through being contacted to participate in research. For example, one participant remarked that maybe he 'didn't realise that [being] overweight could cause prostate problems'. Being contacted about the study could thus help John realise the connection between aspects of his current health and his risk factors for other conditions. Becoming more interested in or concerned about his condition might make John feel 'lucky,' as well as being a source of relief. One participant observed that:

> *'Despite his [John's] reluctance to admit and/or discuss his health problems and risk factors, I think that he would realise how lucky he was that he was being offered this new test that could hopefully show that he is OK but would otherwise give him early warning of one of the biggest killers of men.'*

Several participants thought it likely John would acknowledge the personal benefits and health advantages to being contacted about this particular study. Whether in receipt of new, or already or partly known, information, John would feel 'grateful' his health needs were being catered to. With this knowledge, John's outlook or behaviour might change for the better. He might feel encouraged to help himself, either for his own sake or for that of his family.

Public Benefit

Some participants thought that John might also feel inclined to participate to serve the public benefit**.** Because prostate cancer is a widespread health condition with serious implications, participating in the study might help researchers better understand the disease and its associated risk factors:

> *'[John] may be more likely to give it [the study] more consideration as it is about improving outcomes for others and [he] will have time to look at this when conveniently accompanied by a follow-up email.'*

Echoing an ambivalence first surfaced in relation to the previous scenario (about Kath), some participants worried that John would feel obligated or pressured to participate, while others noted that he could still decide whether or not to get involved in the study. One participant observed that while '[John] probably feels obliged to partake [in the research study] for the public benefit…he's not being forced to'. Participants' attitudes towards the scenario echoed post-workshop survey findings in which many people said a key public benefits of using health data to identify people to take part in research was that this would create *a larger potential pool of people invited to health*

and care research studies, which could increase the numbers of people participating in research studies and the potential benefits of the research for the population.

Negative Responses

While most participant responses reflected positive sentiments and interpretations of the scenario, some participants highlighted concerns, specifically negative emotions John may experience as a result of being contacted by his GPs office. While many thought the communication protocol was positive, some participants argued that phone communication was too 'intrusive' for an interaction or request of this nature.

In addition, participants identified other negative emotions John may experience. Some felt that being identified and contacted was likely to make John feel 'annoyed' 'embarrassed', 'targeted', and 'irritated', while others speculated that he might feel 'singled-out' or like 'a failure – he's [already] trying to improve his health'. By contrast, one participant concluded that 'ignorance was bliss'. Another suggested that given John's predisposition towards the healthcare system, the attempt at recruiting him would be especially unwelcome:

> *'He already feels as though contact with healthcare is a waste of time. [So, he] may be a little irritated about it as he doesn't feel he needs healthcare.'*

Some noted that he may feel offended especially given the sensitive nature of the health conditions being discussed (e.g., his weight and the link to prostate cancer): "he may feel offended they mention his weight. He might feel he's healthy and this would not be something that would affect him". This was echoed by another participant who observed that John may feel 'insulted' by this approach as 'weight is a delicate issue'. By contrast, some participants noted that prostate cancer, while sensitive, is also a widely discussed health issue. John would be 'interested – but not overly concerned. Prostate [cancer] and testing is currently on the TV daily.' While the negative emotions participants identified were significant, many participants who named these potential negative reactions were also clear that these did not outweigh the potential benefits of contacting John.

Gender and Age

Finally, conversations touched on John's age and gender and how these factors may impact his response. Some participants noted that 'men in particular' are less likely to want to discuss sensitive medical issues, and thought that prostate cancer screening could be a delicate subject. Others suggested that 'Welsh men' want to be 'he-man,' referencing specific ideas of masculinity and how this informs men's attitudes towards healthcare services more broadly. When discussing the communication used in this scenario – a phone call issued by the GP surgery – some participants observed that this would be more welcome for a person of John's age, as opposed to "Gen Z" or younger people who may strongly dislike receiving phone calls.

Variable Discussion: point of contact

Scenario variation

John gets a text from an NHS researcher working for the local health board informing him about the same research study on prostate cancer. The text says John's health records show he has a higher-than-average risk of developing prostate cancer in the future. The text says that this doesn't mean that John has cancer now or will develop it.

The text includes a link to a website containing more information about the study, and a number to call if John wants to ask questions.

Participants were asked **if John was likely to feel differently if he was contacted by an NHS researcher instead of the GP's surgery?** Many participants assumed that John's feelings would be similar to the previous exercise, and very little was made of the fact that the person contacting John had changed. Although not the primary focus of the scenario variation, participants spent a good deal of time discussing the mode of communication, which suggests that members of the public find this to be a salient aspect of processes for identifying and approaching people about research studies. Some participants felt a text message was an entirely appropriate way to make contact, whereas others suggested it was too impersonal or might be seen as less trustworthy by recipients who would misclassify it as spam.

Overall, participants assumed that many of John's feelings would be similar to the previous scenario. That is, he might feel concerned (about his condition) or angry (about being contacted); but also possibly relieved or grateful (that his health concerns were being tended to and taken seriously).

As noted in the previous chapter, some participants had expressed misgivings about 'Kath' not being contacted by her GP. By contrast, in the John scenario, participants were less concerned about whether he had been contacted by a primary care professional. However, some participants noted that an approach from an NHS Researcher rather than his GP might put John more at ease. These participants felt that when one hears from their GP, it is likely to be bad news: "your stomach turns over, because you don't know why [they are contacting you]". Being contacted by someone else in the NHS gives rise to less anxiety because it is less likely to concern one's specific health conditions. As one participant noted, '[being contacted by someone in the NHS other than your GP] immediately puts [the conversation] into a different area and [makes it about] something else other than your immediate health'.

Other changes to the process of contacting John, specifically the use of a text message, were thought likely to have significant implications. Some participants thought the new approach was entirely appropriate. More of the responses, however, thought text was an inappropriate mode of communication as it was too impersonal and likely to be considered suspicious (especially by the elderly or people who are not phone literate) or simply not taken seriously. While the scenario was not designed to test participants' views on the use of text messages to contact individuals about research studies – and, though the process described is currently common practice – participants' responses to this aspect of the scenario highlight a wider issue of effective communication practices and mistrust of specific forms of communication.

3.3 Research Questions and Conclusions

At its core, the John scenario asked participants to consider the likely reactions of a person being contacted for participation in healthcare research who is generally reluctant to interact with the NHS and healthcare infrastructure. Among the range of views, participants' responses were weighted more towards the positive aspects of this type of approach, highlighting both the personal and public benefits to participation in the hypothetical study (e.g., John learning more about his condition and potential vulnerabilities, as well as having the option to participate in research that may impact others). While participants seemed generally positive about the variation to the scenario, in which John is contacted by an NHS Researcher rather than his GP surgery, their responses highlighted that the format in which he is contacted mattered to them. Their comments suggested a high level of scepticism about the use of text message, and a preference for a phone call followed by an email.

The scenario exercise asked participants direct questions about a specific, though hypothetical person, John. However, to a significant degree, participants' responses can stand in as answers to a related, but broader policy question about processes which use health data to identify and approach individuals about research studies. Although participants did not reflect on this explicitly, our analysis suggests that based on the responses in this group, **the dialogue participants would be comfortable with people being contacted to take part in healthcare research who may be generally reluctant to engage with healthcare infrastructure**. While participants noted in detail the negative personal emotions this could elicit (anger, irritation, feeling stigmatised), these were outweighed by the potential personal benefits and public benefits of the research being undertaken (e.g., taking part in a study that addresses an important health issue affecting a significant number of men). Responses about the mode of communication suggest some ambiguity around how people should be approached to take part in healthcare research, with different opinions on the extent to which phone calls, emails and text messages may appear trustworthy to people being contacted. This suggests that further work could be done to design processes to identify and contact people in order to mitigate concerns about privacy. It also highlights the importance of any form of communication used to contact members of the public being executed with a high degree of professionalism and clarity.

4. Permissiveness of Data Processes

4.1 Introduction

This chapter will cover participant responses to a scenario concerning the permissiveness of data processes. In the final scenario, 'Jenny & the Welsh Information Governance Board,' participants were asked to consider whether to approve the utilisation of a tool that relies on widespread access to patient health data (through a digital intermediary). We used this scenario as a lens to examine how participants think decisions ought to be taken about the adoption and operation of digital intermediary services; how they balance and prioritise different claims or arguments in relation to health data use; and their comfort levels towards processes that rely on data systems to identify and approach people about research studies. Ultimately, participants explored the trade-offs between individual privacy and the benefits of using a large-scale automated system to identify and contact potential participants for research studies. This scenario thus gives us a good understanding of public attitudes towards the acceptability of digital tools for these purposes, as well as an indication of the public's 'red lines' vis-à-vis data-based services to identify and approach individuals about research studies that do *not* ask for opt-in consent.

This scenario intentionally included some of the topics participants had encountered during previous sessions of the public dialogue: the tool under review closely resembled a blend of the 'digital intermediaries' systems they had learned about in Session 2; the information governance and safeguards processes recalled their discussions in Session 3. When Jenny opened the application out to public consultation, she received three submissions, reflecting stakeholder perspectives participants had encountered, in person, in Sessions 2–3. Participants had expressed largely cautious attitudes in relation to 'digital intermediaries' in Session 2. One of the aims of this scenario was to revisit the use of digital intermediaries while asking participants to explicitly consider the public benefit.

Most people in the dialogue recommended that Jenny should approve the application, with only a small minority rejecting the application. In some cases, those who supported the application did so with conditions – expressing similar concerns to people who voted 'no' to the application. This suggests that, despite disagreeing over whether or not to approve, there was a shared understanding of the major risks associated with this type of research tool. However, it further suggests that, provided certain conditions and safeguards are in place, using health data to identify and contact people about health and care research studies without their specific consent would be largely acceptable.

Rationales for approval of the hypothetical research tool centred around:

- its public health benefits;
- its research benefits;
- its governance and safeguarding provisions;
- a sense that the tool conformed with participants' privacy and confidentiality expectations.

Participants also specified conditions they believed should accompany approval of the research tool, regarding:

- oversight and governance;
- transparency and communications;
- management and infrastructure of its database;
- legitimacy of the research process overall.

Rationales for rejection of the research tool focused on:

- doubts over public trust and legitimacy;
- concerns about privacy and risk to personal data;
- uncertainty about data security and management of the database;

- lack of clarity overall and lingering uncertainties.

Jenny is a member of the Welsh Information Governance Board that oversees applications to access and use health data for research purposes that go beyond the direct care of individuals. When reviewing applications, Jenny must balance concerns about individuals' confidentiality and privacy with considerations of the public benefits of the research.

Jenny is deciding whether to approve this application:

A team of central NHS staff at Health and Care Research Wales have developed a proposal to use patient heath data from across all the health boards and GP practices in Wales to match with research study eligibility criteria. The team has been working with a company that specialises in privacy protecting technology to develop software that de-identifies patient electronic records at source (i.e. before they leave the health board or GP system) and then collates them in a searchable database.

When the NHS research team receives research study criteria from research groups looking to open their studies in Wales, they can then match the study eligibility criteria to the de-identified data (housed in the searchable database), creating a list of eligible individuals, displayed as unique codes. These codes can then be linked back to the individuals' contact details located in the source data systems (health boards or GP systems) and emails sent to potentially eligible participants to inform them of the study option.

Jenny also learns the following:

- The tool only displays deidentified data (i.e. researchers have no access information that might identify individual patients).
- People are included in the database without their consent or an opt-in process; however, they can opt-out at a later stage if they want.
- As a large dataset (all Welsh patient data), the database is more likely to be targeted by hackers. However, the company who developed the deidentification software meet robust security and safeguarding standards.
- Only approved studies will be granted access to the database.

Jenny runs a series of consultations to gauge public perceptions of the tool. She speaks to privacy campaigners (who worry about the NHS' capacity to manage such a large data set, and the lack of patient control over data); health and care professionals (who say the tool will make identification and contacting of potential participants more efficient and spread its benefits more evenly across the Welsh population); and patient groups (who welcome tools that broaden access to health and care studies but remain concerned that some constituencies will still be excluded). Jenny takes all these perspectives under advisement.

As part of the exercise, participants were asked to vote and explain their rationales in a free text box on their worksheets, which also invited them to outline queries or conditions that would need to be addressed if they were to approve the software tool. They also discussed their responses with one another. We undertook an inductive thematic analysis of the three free text boxes (rationales for approval, conditions, rationales for rejection) as well as the recordings of participants' discussions to understand the attitudes relating to each of these decisions.

4.2 Key Findings

Rationales for Approval

A primary consideration for participants who voted to approve the tool was the idea that health data use could increase the public benefits gained from research. One participant suggested that 'the risk to individual rights [is less than the] risk of not recruiting for public benefits.' The scenario did not give an example of a specific application of the tool, so the discussion of public benefits remained general. However, it was clear that, for most participants who approved of the tool, its potential for large-scale (as opposed to individual) benefits was a motivating factor. As another participant remarked: 'research should benefit the larger group…everyone…that's a hope, that's what we want from having this huge database.'

The second most common rationale listed by those who approved the tool related to research benefits, such as being able to identify and contact large samples of people at speed (thus facilitating population-wide studies) or generating more diverse samples (ensuring more effective and useable research findings). As one participant who voted yes noted, '[the tool offers] the big benefit of quickly recruiting the most suitable candidates…and enough of them for the research'. Another pointed out that 'gaining more access to health records means more diversity and [makes research] more relevant', while another described it as 'achieving better health outcomes through more effective and efficient research…making research faster and cheaper'. Some also believed that this type of research tool may have financial benefits for the NHS by streamlining processes, making them more efficient and potentially generating new streams of income: 'the thing is the research itself brings money into the NHS…[we can] sell it to the world and make lots of money. [It] comes back to us'.

Another set of rationales for approving the tool focused on the governance and safeguards in Jenny's investigation. For example, one participant who voted to approve the tool noted that it would be 'governed by a robust ethics committee' as per the scenario description. Another approved of the steps taken by Jenny as laid out in the scenario: 'it seems that Jenny thoroughly looks into aspects of privacy and data sharing', suggesting that the process described for evaluating the tool were sufficient to reassure this participant. Another noted that they voted yes 'because there are checks and balances [and] data security regulation'. Some responses from participants who approved the tool referenced the use of de-identified data as a motivating factor for their vote, for example one participant stated that 'the data is suitably de-identified by using codes so the data is kept secure'.

Conditions for Approval

Participants were also asked to share what conditions, if any, they would attach to approval. Levels of detail given for these conditions ranged from more to less specific. Conditions clustered around four key themes, some of which overlapped with participants' rationales for full approval: oversight and governance, transparency and communications, management of the database in question, and legitimacy of the research process.

Some participants touched on the oversight and governance arrangements of the research tool, though there was not always a great deal of detail in participants' responses as to which aspects of governance they were most concerned about. Some referenced the importance of ensuring the tool remained under the governance of the NHS: '[Yes, Jenny should approve], on condition the information held on the database is held and operated within the NHS'. Others suggested that governance should balance public interest with privacy and confidentiality when making decisions – and that only 'necessary' or 'high public interest' projects should be approved, for instance.

Although the scenario specified that Jenny had sought public perspectives in the process of making her decisions, there had not been any reference to how information about the tool should be communicated to the public if it were implemented. One participant suggested that '[data security] policy should be widespread and accessible and known to the public well in advance.' Participants talked about wanting 'transparency about the process' as a condition for approval, although they did not explore this in depth beyond the idea that information should be widely available. One participant stated that 'it would be nice to know that patients know what their details are going to be used for in terms of the particular research dynamics'.

As noted above, some participants felt that the data used by the digital tool should stay within the NHS. Elaborating on this, some talked about data security, accuracy, and quality, specifying that the system should be 'audited' for biases and data accuracy checked. For instance, one participant suggested that any 'inbuilt bias towards older people and other groups should be audited to prevent unconscious bias/discrimination'. Another participant suggested that '[it is important] that data is not sold for [the] highest price and that controls are in place to identify and take action when breached'.

Finally, many participants focused on the consent protocol that would take place during a subsequent recruitment process. Though this was not in the scope of the dialogue, which focused on the process of identifying and approaching people about research studies, but not yet recruiting them, it remained an ongoing area of concern. For these participants, it was important that research invitations are clearly communicated, that information about the project and research aims are adequately described, and that 'sign up and consent processes are clear'.

Rationales for Rejection

There were a small number of rejections, but a relatively high number of rationales for why participants had ongoing concerns with this form of digital intermediary system. Rationales were not always presented in huge detail, but pointed to serious concerns on the part of some participants who voted to reject the tool. A few participants expressed doubt that the wider public would support this intervention or be able to trust these processes,

speculating speculated that the database 'could prevent people accessing NHS care if they knew this database existed'. Other participants situated this within the wider political landscape:

> *'Trust between society and government [is] at an all-time low. Having a system that does not obtain consent before adding [people] to it is a breach of privacy and could further destroy human rights and trust.'*

For a few participants, the fact that the tool employed de-identified data was not enough to assuage concerns about privacy and risks to personal data, and some specifically objected to the opt-out, rather than opt-in, nature of the database: 'opt-out policy is not good enough'. Another noted, 'this is a form of no-choice opt-in'. As one participant stated, 'the data seems traceable even if it's not identifiable', which, although nothing in the scenario suggested the data would be traceable, suggests they did not find the scenario explanation of how data would be handled reassuring. Another remarked that 'data is automatically given – even if de-identified'.

Those who rejected the application expressed ongoing uncertainty about the technology described, as well as the NHS' capacity and infrastructure to handle such a tool and any potential data breaches. In the words of one participant, '[the] future is unpredictable…[and the NHS could] lessen [my] concern about privacy by taking action on data breaches seriously', which suggests they did not think the NHS's current protocols take data security sufficiently seriously. Another participant remarked that '[the] NHS is in crisis [and it is] not equipped to ensure adequate safeguards', demonstrating some doubts that the database envisaged in the scenario could work well. One participant expressed concern that the 'NHS doesn't have data apparatus or capacity' to complete this database.

Many of the rationales listed for rejections were framed as questions asking for more clarification, rather than statements. Some participants stated that they agreed with the tool in principle but there was not enough information relating to the use or processing of 'deidentified' data, which meant they were not comfortable recommending approval. This suggests that with further reassurances and additional clarifications, some participants may find their concerns assuaged and would vote to accept the use of the tool. For others, their 'no' vote was rooted in a deeper sense of scepticism and unease with the NHS and with the research tool itself.

4.3 Conclusion

Though grounded in the specifics of the 'Jenny' scenario, this exercise offers some generalisable insights into public attitudes towards processes that identify and approach people about research studies, including the use of de-identified data, processing large volumes of data, and the use of a digital intermediary. The exercise asked participants to weigh up and prioritise evidence and perspectives in the public interest. Overall, it suggests a broad level of assent to the use of digital intermediaries. But only in the context of some significant conditions/caveats.

In the end, the rationales and conditions for approval, and the reasons to reject, express many of the same arguments and concerns about the scenario and processes to identify and contact individuals about research studies. This suggests that even between those who voted to reject the application and those who voted to approve it, certain key principles (the importance of public trust, appropriate governance and safeguards, and appropriate levels of transparency and communication) are held in common, with participants disagreeing on whether the tool described in the scenario sufficiently upheld these shared principles.

5. Future Expectations

5.1 Introduction

This chapter sets out the findings from the final exercise conducted during the dialogue's in-person workshops. In this exercise, participants were asked to discuss, and try to agree upon, a set of rules or principles they would like to see applied in Wales in the future to processes that identify and approach people to participate in health and care research studies by accessing their health data without prior consent. Initially working in small groups, participants distilled their hopes and requirements into agreed statements which were reviewed and voted on by the wider group – participants voting 'for' or 'against' each other's suggestions.

The votes against some statements make sense in the context of participants' other, positive voting decisions. For example, a statement saying AI should be used to speed up data processes received many negative votes reflecting participant concerns about a lack of human oversight and control. This outcome is broadly congruent with the group's decision to award positive votes to several statements advocating *for* human oversight. However, other statements that received 'against' votes are harder to explain. For instance, one negatively received statement proposed the public be included in forthcoming decisions around health data. The rejection of this view contradicts the support given elsewhere to statements in favour of public consultation. This exercise was less discursive than the previous scenario exercises, which means we have limited evidence to explore these tensions or to gauge whether they reflect substantive disagreements or if participants simply misunderstood some of the statements. Moreover, the aim of the exercise was to produce a set of broadly shared expectations – positive statements. And thus, this chapter will not discuss the negatively received statements and will instead focus on the statements that received broad support.

The statements that received the highest number of positive votes (and no or minimal rejection) were thematised into overarching, and in some cases overlapping, participant expectations – set out below. Although participants were asked to group their statements into different categories (expectations, principles, and recommendations), in the end, their responses did not reflect these nuances. Consequently, the findings of this exercise are presented under the single heading, 'expectations', highlighting areas of broad agreement on what the Welsh government should prioritise in the future.

This process was done simultaneously in Cardiff and Bangor, and our findings bring together and integrate the results from both locations. The evidence set out in this chapter draws from the statements drafted by participants, as well as the transcriptions of their table conversations which took place prior to the voting exercise. Although the following themes are presented in order, starting with the category that received the highest number of positive votes, the difference in vote share – given the overall number of participants – is not hugely significant.

5.2 Key Findings – Public Expectations

Research Quality

Across the workshops, participants said that processes that identify and approach people to participate in health and care research studies must lead to demonstrable research benefits. For instance, there was strong agreement that using health data for identifying and approaching individuals should be contingent on making the recruitment process more efficient (at achieving demographic diversity etc.). Participants also agreed that such processes should have a clearly identifiable impact on health outcomes in Wales. One participant said: 'My hope would be [to see] the evidence of impacts and benefits from the data collected.' The same sentiment was captured in the statement: 'We can demonstrate how it [a research study supported by a data-based process for identifying and approaching leads to better health outcomes.'

Oversight Ensuring Standards and Security

Participants widely supported the notion that processes for identifying and approaching people about health and care studies should be subject to scrupulous regulatory and research ethics oversight – primarily for the purposes of proper data use and security – and it was widely held that misuses of data should be swiftly detected and prosecuted. Participants expected a robust regulatory framework, agile in its response to evolving technologies that could be used for data processing. One statement read: 'Regulations to be flexible and future-proofed to evolving technological advancements.' In addition, participants supported the establishment of demographically diverse human oversight panels to determine who may be allowed access to health data, and to ensure that data is handled correctly and securely. 'Fully vet who has access to my personal information' read one statement. The oversight process should be regularly reviewed to ensure it continues to meet these aims.

> *'We need assurances that data is protected…committees that are formed around how the data will be used and explanation of who were in those committees and how those are decisions made.'*

Equity

Another core expectation was that different communities and regions across Wales should have equitable access to healthcare infrastructure – both the processes for identifying and approaching people to take part in research, and the outcomes and benefits of that research. One statement read: 'as many people as possible [should be] able to take part [in research] and benefit.' Indeed, various participants held that healthcare policy and outcomes should benefit the full spectrum of Welsh citizens. Another statement called for: 'fully inclusive data, including all areas of…[society] and [the] economy and diverse statuses [within the Welsh population]'. Participants said that data-based processes for identifying and approaching people to participate in research studies ought to produce more representative samples, ultimately facilitating more targeted health interventions. We must 'Make sure the research is diverse…[and] try and get the hard-to-reach groups [and] the most vulnerable'. This – it was argued – would minimise the discrepancies in healthcare access between rural and urban regions and different demographic communities.

Public Consultation

Many participants thought the transparency and trustworthiness of processes for identifying and approaching people about research studies could be achieved if those people responsible for data (e.g., researchers or research commissioners) communicated effectively with the public. One statement illustrating this read: 'There should be a mechanism for public debate'. People referenced large scale media campaigns (a measure which was also supported in the post-workshop survey), using simple, direct, and jargon-free language to ensure widespread comprehension. Participants advocated for an 'increase [in] public awareness of what health data is and how it is used'. Moreover, participants supported the idea of venues in which the public could hear about and debate the use of health data for identifying potential research participants. One participant explained how greater public awareness could influence how such data-driven processes are received:

> *'Could there be better public awareness of health data in general? Because I didn't really know much about it, and if people generally knew a lot more about public health data and how it's used and the benefits of it, when they were approached, they'd be like, "oh I know what this is, this is a good thing!"'*

Public Benefit

Participants said the public benefits of health and care research were central to the legitimacy of using health data to identify and approach individuals about taking part in research. Public benefit was construed as an improvement to the wellbeing of ordinary Welsh citizens. Participants also stressed the importance of maintaining a 'balance between public benefit and the risks' of health data being more widely available to researchers (such as, data breaches and accidental misuse). This mandate was particularly important in regard to profit-making: some participants maintained that profit should not be part of the decision to use people's health data ('Health data for recruitment should not be exploited for profit'). In doing so, participants explained that using health data to help identify and contact potential research participants could maximise public benefit while minimising risks of exploitation.

'A big hope is that it benefits the Welsh population…And I suppose that also says something about potentially there's researchers from outside Wales coming in. And one of the hopes, I guess, might be if data is being made accessible, that there will be some kind of benefit for the Welsh populations.'

Transparency and Trustworthiness

For the Welsh government to secure the 'buy in', participants called for transparency regarding the use of their health in these processes. This included clear communication about who is granted access to their data, what these people or organisations are accessing it for, and where it is being stored. This was particularly important for establishing the trustworthiness of these processes with communities who are typically more disconnected from the healthcare system than other constituencies: 'there is mistrust, especially in some communities, regarding NHS data'. Greater transparency was seen as remedying mistrust and encouraging participation.

'Transparency should be exercised as a priority, regardless of what decisions are made... Everything should be explained to the public.'

5.3 Conclusion

This exercise surfaced a range of participant expectations about the future of Wales' data-driven processes for identifying and approaching people about research studies. These included: implementing measures to ensure data security; justifying data use on the basis of public interest or benefit; and establishing robust processes of regulatory and public oversight – insuring against risk and mitigating against harms (like data breaches) and unwanted outcomes (like selling data to private companies). Participants advocated for continued public engagement, providing space for communication between data users and the public as well as increasing general awareness around data practices. Not only are these protections deemed necessary to offset the risks associated with data use and privacy; they are also – for many participants – a precondition for increasing research equity and quality, another expectation for future data practices. Many of these expectations are familiar, echoing the conditions discussed in previous sessions by participants broadly comfortable with data-driven processes. And without these conditions and safeguards, many participants would be hesitant and/or sceptical about processes that relied on unconsented access to health data to identify and approach people about health and care studies.

6. Conclusion

To help the Welsh government streamline policy on processes that identify and approach people about research studies, NatCen convened 35 Welsh citizens from across the nation to participate in a Public Dialogue. At root, the dialogue was concerned with participants' views on the role and availability of health data. And throughout, participants were asked to weigh the relative importance of patient privacy and confidentiality, and the public benefits of improved data-driven processes.

Participants voiced various policy and political priorities, and there were enduring disagreements about issues ranging from the salience of patient data privacy to the suitability of different modes of approaching people to participate in research studies – phone calls, texts, emails etc.

However, the underlying question of the dialogue – asked in different forms across all three scenario exercises and the discussion of future expectations – was whether participants were **comfortable with processes that approach people about research studies having identified them as eligible by accessing their patient data without prior consent**. Some participants were not comfortable, citing an individual's right to privacy and the idea that personal data should not be available to non-healthcare professionals. However, many others were comfortable with such processes, especially if certain conditions were met.

- The findings from the first scenario exercise show participants were only minimally concerned by the idea of people being approached about research studies on the basis of their identity. More emphasis was placed on generic reservations about how data is processed and who might have access to it. These concerns were much greater in situations when people were contacted about research studies concerning sensitive health issues, and diminished when studies were about a population-wide condition. Many of these concerns were addressed and resolved in the future expectation exercise.

- The findings from the second scenario question reveal that participants put great store in the potential for research studies to benefit the general population, or even (in rarer cases) the individuals who sign up to them or people like them or people they know. These considerations often trump concerns about processes that approach people based on un-consented access to their healthcare record data.

- In the third scenario exercise, a sizeable majority of participants said they were content with a process that identifies and approaches people about research based on un-consented access to their health data. There were a small number of dissenting participants. However, their reasons for rejecting the process under discussion were, in several cases, addressed and resolved in the future expectations exercise.

- The future expectation exercise distilled the conditions and safeguarding principles required for participants to feel comfortable with processes that rely on un-consented access to patient health data. The absence of these conditions was what caused some participants – in the previous exercises – to express doubts about data-driven processes of this sort.

Taken together – that is, by assessing the full range of participant responses, balancing the contributions made in one scenario with the contributions made in the others, and in particular the futures exercise – our analysis suggests that **processes reliant on un-consented health data would satisfy a broad constituency within the dialogue, provided the following conditions were met**:

- Accessing health data must be done in the service of making processes for identifying potential participants more efficient, and research outcomes more robust.

- Integral to this, data-based processes for identifying and approaching individuals must operate to make health and care research more inclusive, diverse, and accessible.

- Processes that use unconsented data should serve the public good – over and above commercial interests.

- All processes for using data to identify and approach individuals for research studies must be carefully regulated and operate with continuous oversight.

- The Welsh public should be consulted about forthcoming innovations, and efforts made to ensure broad public understanding of how data is used.

- Within and across these conditions, individual privacy and confidentiality must be respected and preserved (within the boundaries of a model that identifies and approaches individuals based on unconsented access to health data).

7. Appendix

7.1 Sample

The sample was recruited to broadly reflect the Welsh public in key demographics like gender and ethnicity, and strike a balance in demographics like age, household income, and indices of multiple deprivation. We also recruited according to whether participants believed they were directly affected by the issues discussed in the dialogue, aiming for around 50% of participants to identify themselves as such. In total, 35 participants attended the fourth and final session. The following table breaks down that sample.

Note: Numbers do not always add up to 35 due to missing data or respondent refusal.

Characteristic	**Measure**	**Participants (*n*)**
Ethnicity	Asian or Asian British	1
	Black, Black African, Black Caribbean or Black British	3
	Mixed or multiple ethnic groups	1
	White or White British	30
Age Group	16-25	2
	26-45	10
	46-65	12
	66+	11
Gender	Female	22
	Male	12
	Other	1
Education Level	Below A-Level or equivalent	5
	A-Levels or equivalent	7
	Foundation degree or equivalent	3
	Undergraduate Degree or equivalent and above	20
Monthly Household Income	Above £5k	3
	£2k to 5k	13
	Up to £2k	11
Directly Affected by Issues Discussed in this Project	No	16
	Yes	16
WIMD Decile Rank *(2019 data linked to participants via postcode; 1 = most deprived)*	1	1
	2	3
	3	4
	4	5
	5	7
	6	5
	7	3
	8	1
	9	3
	10	2

North/South Wales	North Wales South Wales	19 16

7.2 Expert Speakers/Stakeholders

- Dr Rachel Thompson (University of Oxford) – presented at Session 1
- Dr Fiona Lugg-Widger (Centre for Trials Research, University of Cardiff) – presented at Session 1
- Nicola Hamilton (Understanding Patient Data) – presented at Sessions 1, 2, and 3
- Prof Mike Robling (Centre for Trials Research) – presented at Session 2
- Sam Smith (MedConfidential) – presented at Sessions 2 and 3
- David Chuter (Use My Data) – presented at Sessions 2 and 3
- Prof Andrew Carson Stevens – presented at Session 3
- Anna Studman (The Ada Lovelace Institute) – presented at Session 3

7.3 Welsh Government Data for Research Working Group

The Data for Research Working Group convenes research data experts and key stakeholders in Wales to provide advice to the Welsh Government Health and Social Services Research & Development team on policies that use health data to support health and social care research. Membership includes:

- Welsh Government policy officials from the Health and Social Services R&D team, the Chief Digital Officer for Health team, and the Chief Digital Officer team;
- Digital Health and Care Wales officials responsible for data management and information governance;
- Data research academics from Cardiff University and Swansea University;
- Clinical Researchers;
- NHS Research delivery staff;
- Primary Care representatives; and
- Public members, drawn from the Health and Care Research Wales public involvement community

7.4 Variables that might Affect Public Views on Research Recruitment – Discussed by the Data for Research Working Group

- The scale of data being accessed – whether pre-screening should take place (narrowing the list of contactable individuals to those most likely to be eligible), or if a shallower view of the data (and a wider set of possible contacts) was preferable.
- The sensitivity of the condition under study.
- Whether people are contacted about a condition they have a known diagnosis vs. a prospective condition based on demographic or other characteristic predictors.
- The trial phase participants are invited into (to explore perception of riskiness of the study).
- Who is sponsoring the research (commercial vs non-commercial).
- Whether there are individual or population-wide benefits of the study.
- Whether the NHS staff involved in the recruitment process are direct care professionals or people 'further' away from the day-to-day care of potential invitees.
- The role of third parties in the data recruitment process (including anonymisation/automation of data processing).
- The method of contact (letter, email, phone-call).
- An individual's choice about whether their data should be available to data-led recruitment processes.

7.5 Pre-dialogue Survey Results (ascertain participants general views on data):

Prior to and after the dialogue, participants completed surveys which broadly aimed to, respectively:

a) Ascertain their level of knowledge about general data practices
b) Collect their views on using health data to identify and invite participants to research practices, specifically.

The following sections present visualisations of the responses to each survey.

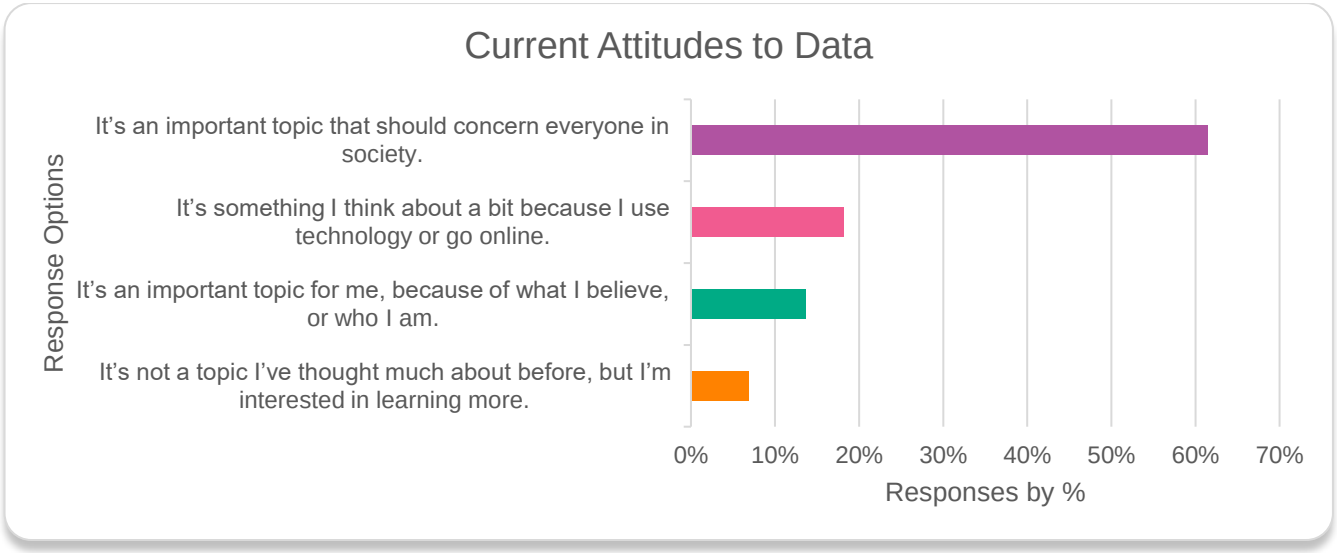

Figure 1. Percentage of votes received for each of the above statements representing participants' attitudes towards data.

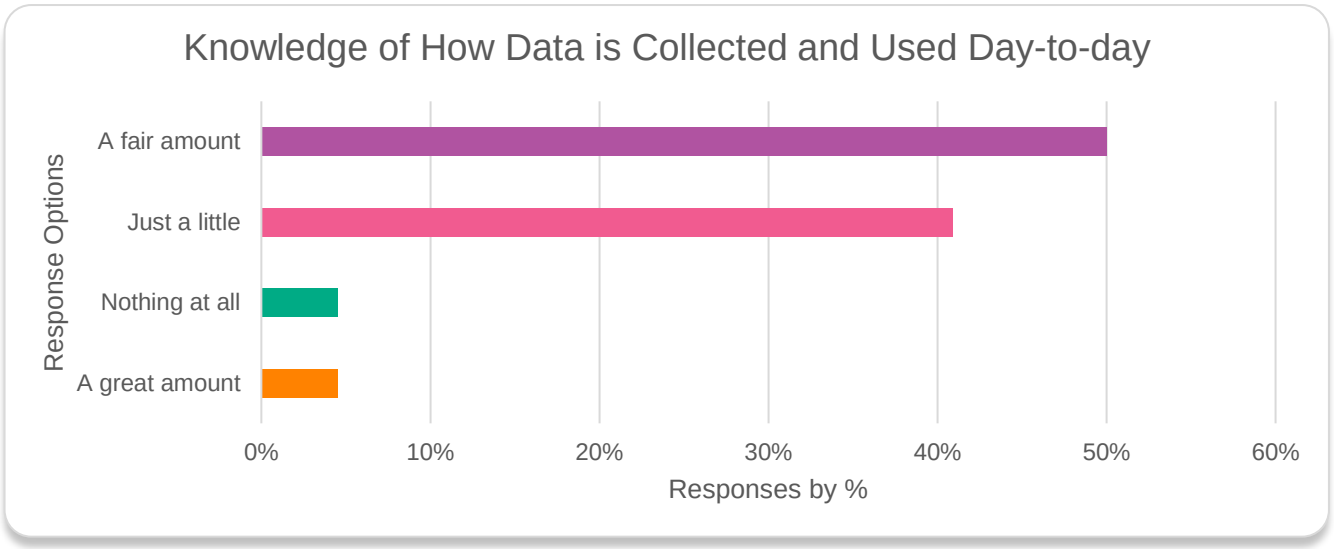

Figure 2. Percentage of participants who felt they knew how data about them was collected and used throughout their daily lives.

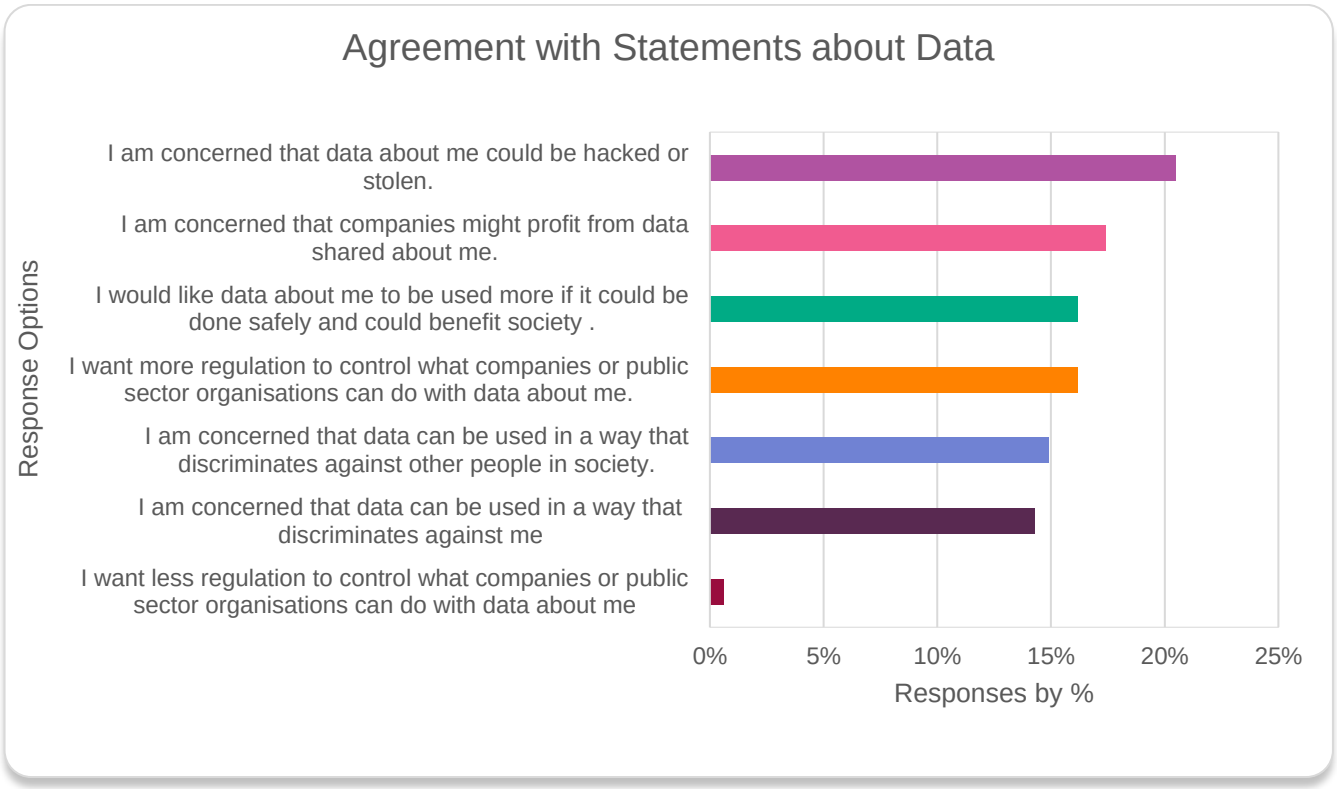

Figure 3. Percentage of votes received for each of the above statements representing participants' agreement.

7.6 Post-dialogue Survey Results (confirming participant views on benefits and risks):

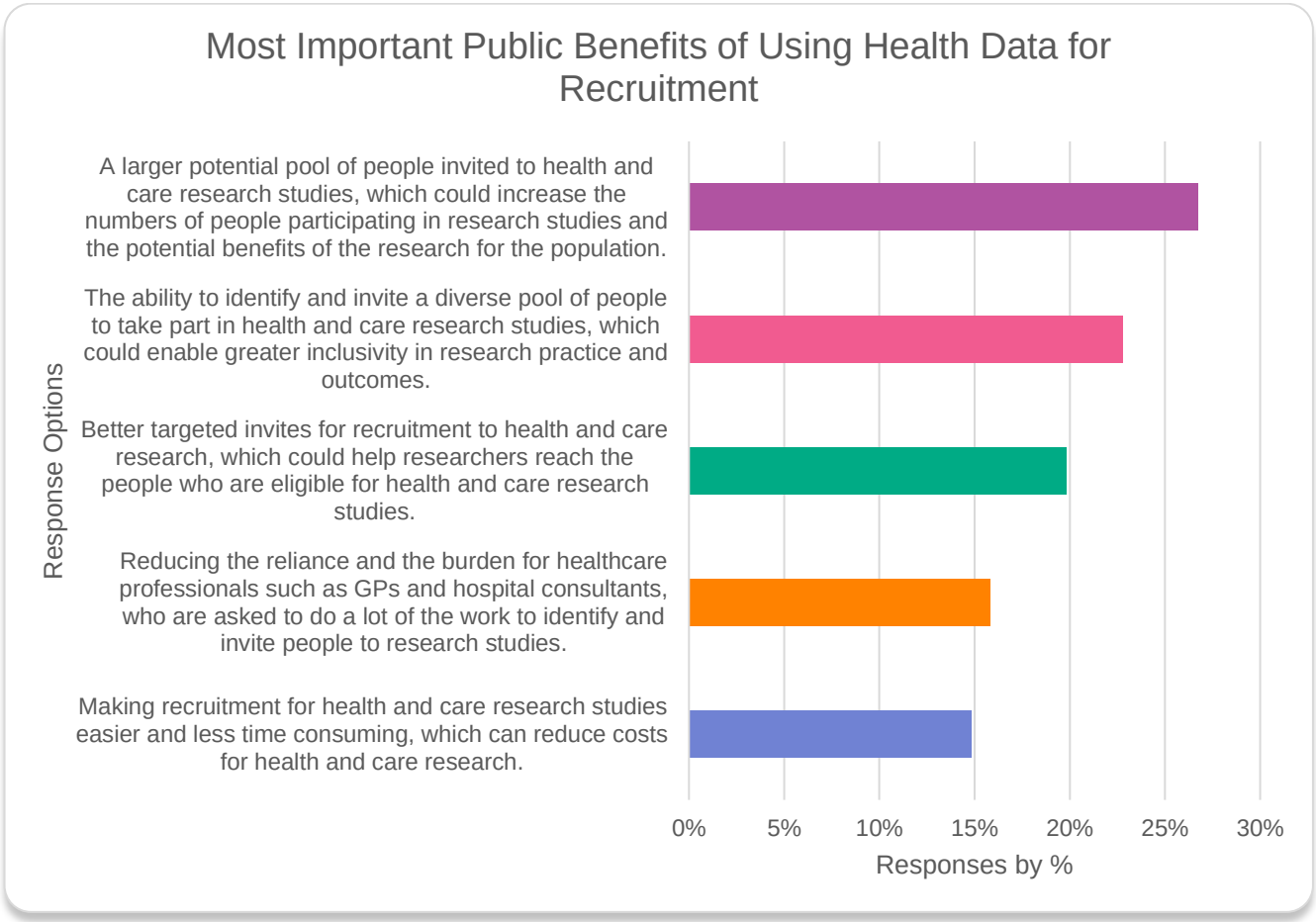

Figure 4. Percentage of votes received for each of the above statements, representing the potential public benefits of using health data to identify and invite people to take part in health and care research, without their consent.

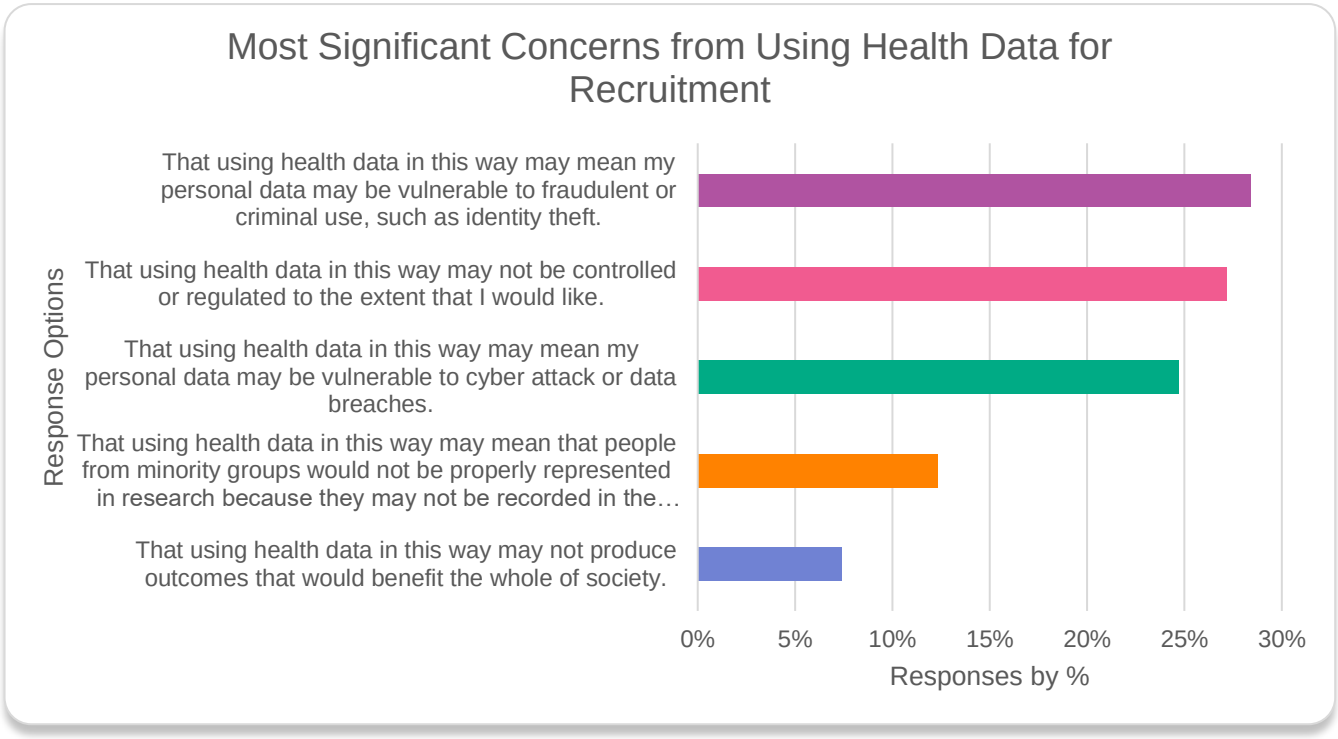

Figure 5. Percentage of votes received for each of the above statements, representing participants' concerns of using health data to identify and invite people to take part in health and care research, without their consent.

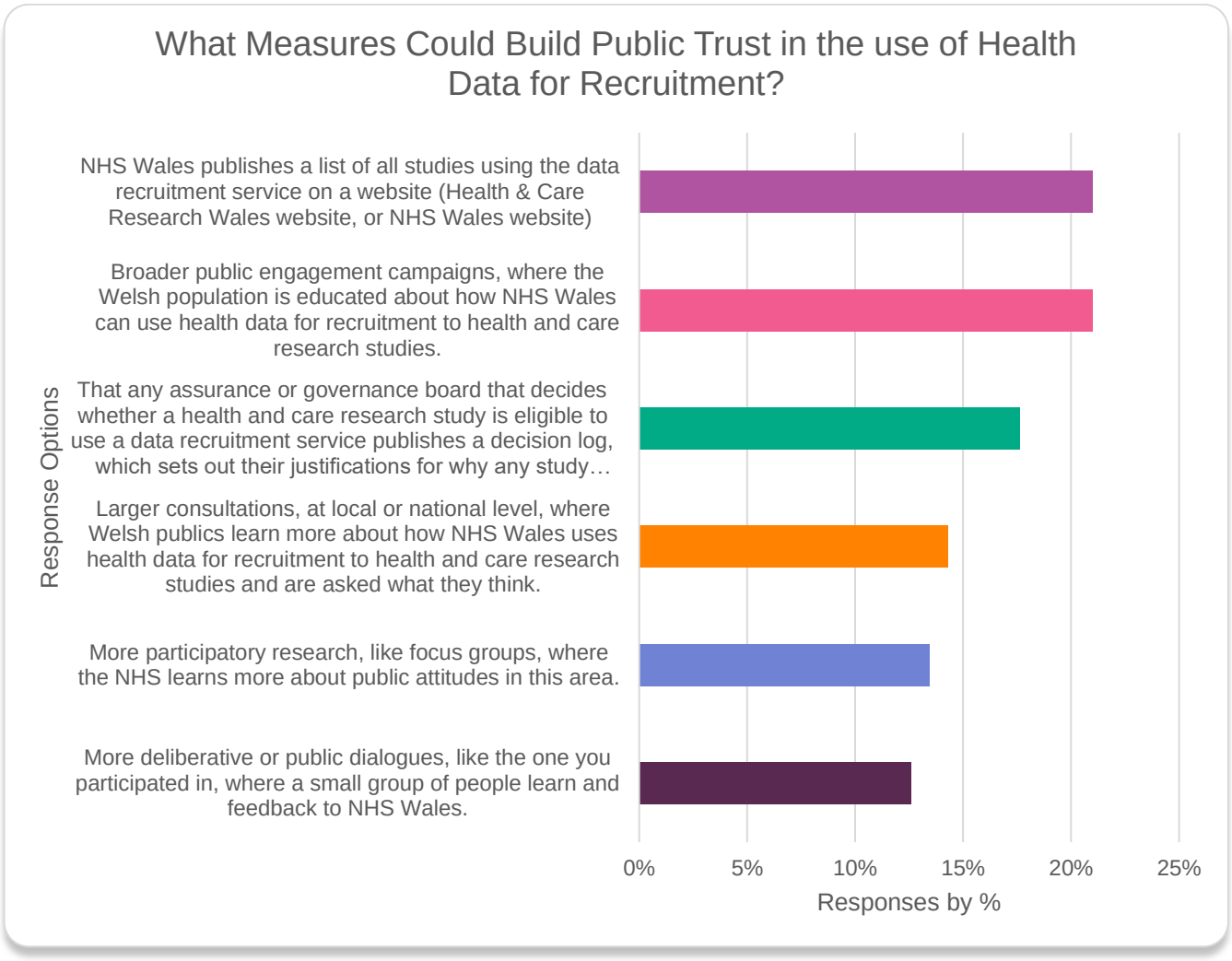

Figure 6. Percentage of votes received for each of the above statements, representing measures which could better build public trust in NHS Wales using health data to identify and invite people to take part in health and care research, without their consent.

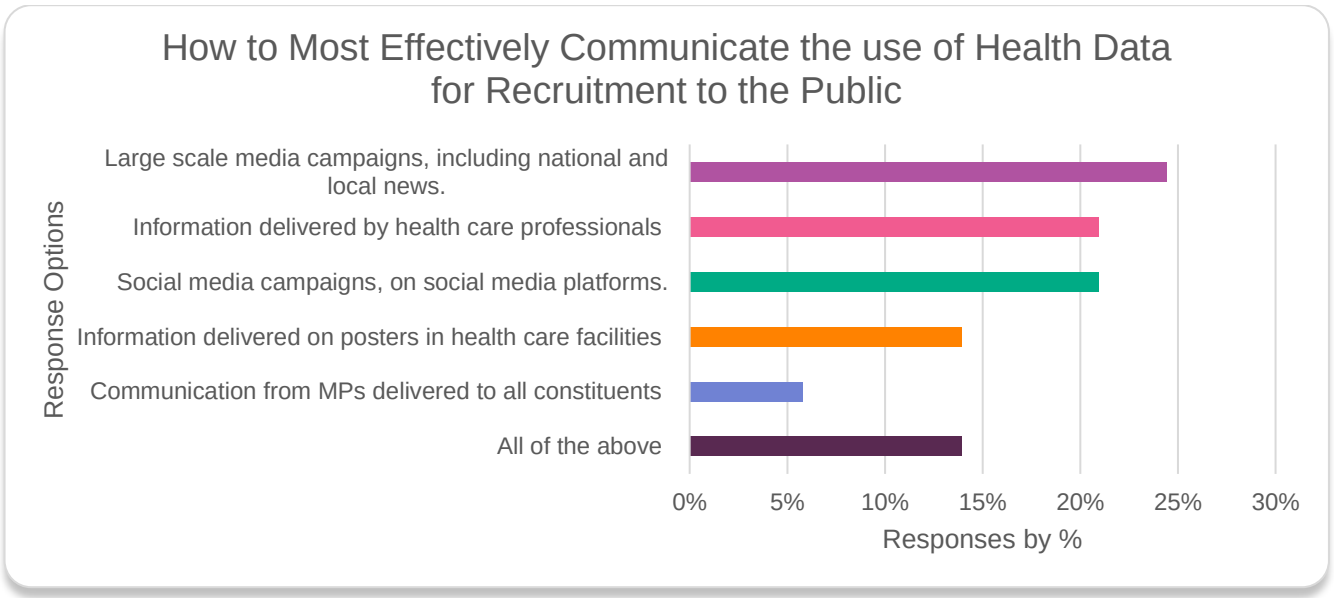

Figure 7. Percentage of votes received for each of the above statements, representing ways to communicate to the public about using health data to identify and invite people to take part in health and care research, without their consent.


National Centre
for Social Research