

Unlocking the Value of Scotland's Data

**Public Engagement around the Access
of Public Sector Data with or by Private
Sector Organisations**

August 2023

Table of Contents

Executive Summary	2
Introduction.....	3
Methodology.....	5
Search strategy and inclusion criteria	5
Data extraction and analysis.....	5
Flow diagram of the selection process	6
Inclusion and exclusion criteria	6
Limitations	7
Results	8
Included studies	8
Public engagement and research methods used	8
Findings from public engagement and studies on public views: themes	10
Discussion	21
Conclusions and Recommendations	23
References.....	26

Executive Summary

This report highlights the findings from a literature review commissioned by the Scottish Government on public engagement regarding the use of public sector data by or with the private sector over the last 10 years both in the UK and internationally. Unlocking the full value of Scotland's public sector data for public benefit is conditional upon public sector organisations' ability to make their data accessible and useable by third parties, including private sector organisations, in ways that are publicly acceptable and retain public trust. The aim of this review is to inform the design of a Scottish Government programme of public engagement in this area. We used an adapted systematic review and thematic synthesis approach of academic and grey literature to answer the following questions:

- What public engagement has been undertaken regarding the use of public sector data by or with the private sector?
 - What methods were used?
 - What were the results?

Forty-four papers were included in the review, published after 2010 and a thematic synthesis was conducted of the reported findings. We identified the kinds of public engagement and research methods that have been used in this area; these were mostly deliberative, dialogic and qualitative. We generated eight key themes from across the results documented in the literature.

Our key findings are:

- Deliberative and dialogue based qualitative public engagement and research methods are effective in identifying informed and considered public views on private sector use of public sector data, and they can enable the construction of a public consensus that can be used to inform decision making.
- There is a low level of public awareness and understanding of private sector access to public sector data and how this data is used. Publics tend to express negative 'gut reactions' towards the topic, but when provided with more information and opportunities to reflect on or deliberate it, they often change their minds.
- There is widespread conditional acceptance of private sector use of public sector data especially among informed publics. Acceptability is most conditioned by the rationales for the data use, but also by the type of data being used and the type of the private sector organisation using it. Public benefit is the primary driver of acceptability and commercial gain or private profit the primary driver of unacceptability.
- Demonstratable public benefit is the most prevalent consideration that publics have around private sector access to and use of public sector data. While the definition and scope of 'public benefit' is open and contested, publics want to see evidence that public benefit of some kind is the primary driver of public sector data access, that it can actually be achieved, and that it outweighs any possible private benefits.
- Publics want to see the development of equitable benefit-sharing models for collaborations or partnerships between private and public sector organisations, as

they expect benefits – including profits – to be returned to publics and reinvested into the public sector.

- Public trust and distrust are key factors around private sector access to public sector data. While publics tend to be relatively distrustful of private sector organisations, they generally have a high level of trust in the public sector, and this is shaped by perceptions that the public sector is acting for public benefit whereas the private sector is motivated by private interests. Publics are more trusting of private sector uses of public sector data when public sector organisations retain control over the data during collaborations with the private sector.
- Publics expect to see stringent oversight, governance, and safeguard arrangements around private sector use of public sector data, especially concerning an oversight or governance body, transparency and accountability processes, and arrangements for data security and safety, consent, and confidentiality. However, the precise nature of what the safeguards should be is contested, and it may be that the nature of the safeguards is less important than the fact that effective safeguards exist.
- Publics want there to be public involvement or engagement processes and activities around private sector use of public sector data, but the precise nature of what this should look like, who should be involved and in what ways is contested – while some want to be actively involved in decision making, others prefer more passive forms of communication and information distribution, and proportionality matters.
- There is no singular ‘public perspective’ on private sector use of public sector data, but rather, while overarching patterns can be identified, publics are plural, and individuals’ views are shaped by a diverse range of intersecting demographic and attitudinal variables.

Introduction

Public sector organisations in Scotland hold rich datasets that have high value. Unlocking this value and realising the full potential of Scotland’s data to create public benefit requires collaboration and partnerships between public sector organisations and third parties, including private sector organisations that have skills and knowledge which can enable public sector organisations to more effectively make use of and harness the benefits of the data that they hold. There are significant incentives to capitalise on this data, including to improve public services, processes and outcomes and foster data driven innovation in Scotland and beyond, but this is all conditional upon the ability of public sector organisations to make their data accessible and useable by third parties in ways that are acceptable to publics and retain public trust. This creates challenges because the use of public sector data by or with the private sector is a contentious area that has incited public debate and concern, including due to a stream of previous public controversies around data sharing initiatives that did not gain a sufficient social licence to operate.

Prominent examples include the failures of the care.data programme, launched in 2013 with the aim of enabling access to and use of routinely collected health and social care data for commissioning and other purposes beyond direct patient care. It has been argued that the public controversy generated around care.data was due to its failure to gain a social licence, including because communication about the initiative was not appropriately transparent to enable publics to trust that their data was sufficiently protected against misuse and because

publics were not convinced that care.data served public benefit (see e.g. Carter, Laurie, and Dixon-Woods, 2015). Similarly, the recent controversy over the sharing of the Royal Free NHS Trust identifiable patient data with the Google Deep Mind subsidiary to create a diagnostic application generated public concerns including around the potential commercialisation of patient data, insufficient transparency around the initiative, and lack of consent from patients for this data use (see Ballantyne & Stewart, 2019). Further, these challenges are not limited to the UK. For example, in 2019, Ascension, a large US non-profit healthcare system, announced a data partnership with Google called “Nightingale,” which involved the moving of identifiable patient records onto a Google cloud server for data analysis and processing. This incited public alarm as well as an investigation by the US Department of Health and Human Services Office for Civil Rights due to concerns over a healthcare system sharing patient data with commercial companies in general and distrust in Google’s promises to keep this data private, secure and separate from their other data in particular (see Wachter & Cassel, 2020).

Beyond the public controversies in this area, as we will show below, publics generally have negative spontaneous reactions and relatively low levels of trust in private sector organisations, combined with a strong expectation that any use of public sector data will be transparent, accountable, serve public interests and, most importantly, result in public benefit over and above any private or commercial benefits that may be realised from it. This brings to the fore that the success of data sharing ventures and partnerships between private and public sector organisations is dependent on understanding how to secure public support and trust for such ventures, and on ensuring that each initiative is publicly acceptable and geared to deliver public benefit.

In what follows, we report the findings from a literature review of previous public engagement regarding the use of public sector data by or with the private sector over the last 10 years both in the UK and internationally. The aim of this is to inform the design of a programme of public engagement for the Scottish Government in this area including appropriate further engagement activities. Accordingly, the review was driven by the question, what public engagement has been undertaken regarding the use of public sector data by or with the private sector? We undertook a rapid review of the existing literature on public engagement and public views on private sector use of public sector data using an adapted systematic review and thematic synthesis approach, through which we identified, analysed, and synthesised academic and grey literature. We found that deliberative and dialogue based qualitative public engagement and research methods are effective in identifying informed and considered public views on private sector use of public sector data, and there were eight key themes across the findings from the included literature: low public awareness, ‘gut reactions,’ and changing perceptions around data use; acceptability of private sector data uses; the centrality of public benefit; importance of benefit-sharing and distribution; trust and distrust; oversight, governance, and safeguards; public involvement and engagement; and the impact of demographic differences of people’s views. We conclude that these themes should be accounted for in the design of public engagement in this area so that engagements can move informed debate towards acceptable solutions, and we provide some corresponding recommendations.

Methodology

Search strategy and inclusion criteria

This was a rapid review, using an adapted systematic review approach. We conducted a systematic literature search of two electronic academic journal databases (Scopus and Web of Science) on 6th of June 2021 using the following key word chain, tailored for each database:

```
((public* OR patient* OR citizen* OR community*) W/2 (engagement OR involvement OR participat* OR perception* OR attitude* OR view* OR perspective* OR opinion* OR concern*)) AND ((public OR government* OR state OR administrat*) W/5 (data* OR record*)) AND (share* OR link* OR (private W/2 sector) OR commercial* OR (third W/2 part*) OR ((public-private OR private-public) AND partner)))
```

The key word chain was reviewed by and constructed with input from a University of Edinburgh academic support librarian. In addition, we citation traced the included articles and two topic relevant systematic reviews, scanning the references to identify further articles, and we conducted Google searches and grey literature database searches (Open Grey and gov.uk) to identify relevant grey literature. The searches were implemented on the 23rd of June 2021, and the scope was limited to documents published after 2010 in the English language, but there were no geographical restrictions or restrictions on publication type.

After the exclusion of duplicates, the academic journal database search resulted in a total of 1463 articles. SE screened the titles, abstracts and, in some cases, full texts of the articles, of which 21 met the inclusion criteria and were included into the study. Additionally, 23 documents identified through citation tracing and grey literature searches were included after title, abstract and full text screening undertaken by SE, resulting in a final set of 44 included documents. The selection process is depicted in figure 1. Due to time restrictions, we did not conduct quality appraisal of the included studies. Documents that met the inclusion criteria, listed in table 1, were included. Notably, documents were included even if private sector use of public sector data was a secondary rather than a primary topic, as long as this topic was an intentional part of the design of the public engagement activity or research that was reported. Studies where participants spontaneously expressed views on private sector use of public sector data, even though this topic was not part of the study design, were excluded for two reasons: firstly, to manage the number of documents considering time limitations, and secondly, these studies did not report in depth findings on this topic and generally only included a mention of it.

Data extraction and analysis

Full texts for all included documents were downloaded and reviewed by SE. For each included document, main characteristics of the study of public engagement activity were extracted and entered into an Excel sheet, depicted in appendix 1. We applied a modified thematic synthesis approach to analyse the documents to enable rapid review. This, builds on but adapts the method described by Thomas and Harden (2008), which involves three stages: coding; developing descriptive themes; and, generating analytic themes. We applied

a combination of (inductive) line-by-line and paragraph level coding, where SE coded lines or paragraphs of text from each included document according to their meaning and content, using the NVivo qualitative data analysis software. Documents where the primary focus or topic was private sector use of public sector data were coded in full, line-by-line. However, for documents where private sector use of public sector data was a secondary topic, coding was only undertaken at a paragraph level and for the sections of text where private sector use of public sector data was discussed.¹ This generated a list of initial codes. SE then searched for similarities and differences between the codes, identified higher order descriptive themes, and organised these into analytical themes that relate to and address the research questions of this study.

Flow diagram of the selection process

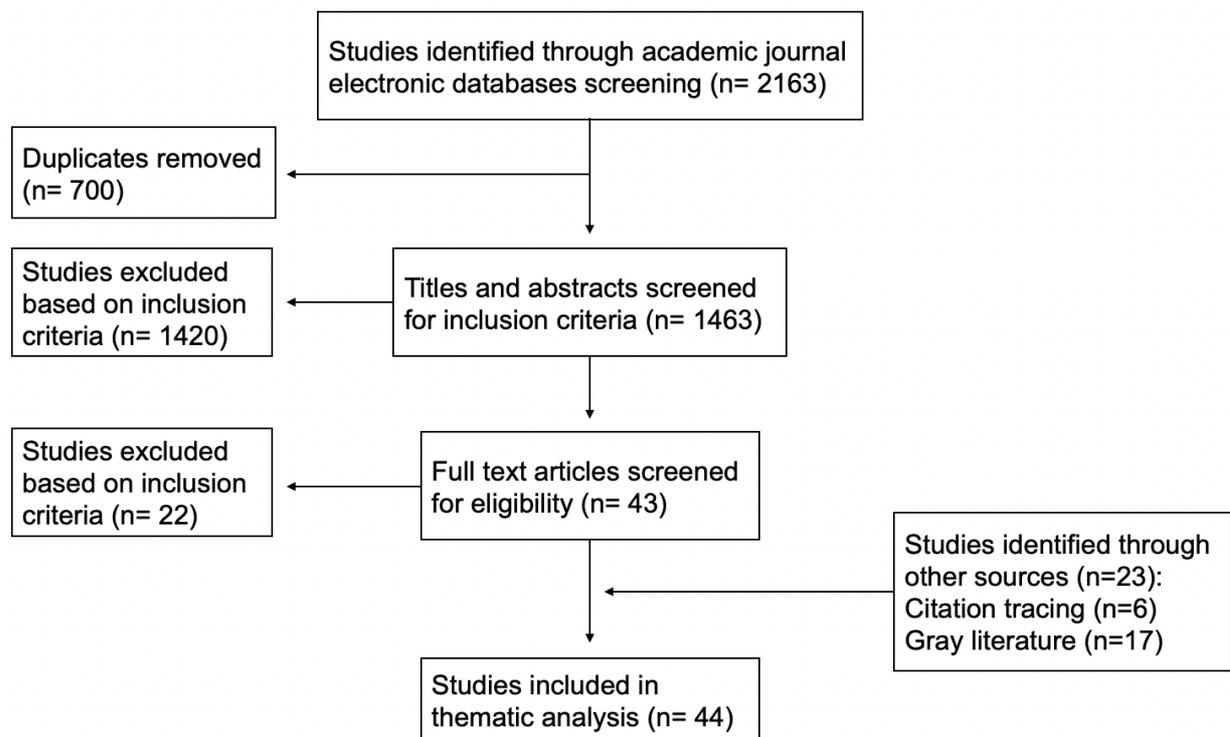


Fig. 1. Selection process based on PRISMA flow diagram

Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
All articles / documents reporting public engagement / research on public sector data use by or with the private sector as a	All articles / documents not reporting public engagement or research on the topic of public sector data use by or with the private sector

¹Some of the documents included findings from reviews undertaken alongside primary data collection, but where this was the case, findings from reviews were not included in the analysis, and thus not coded.

primary topic or focus of the public engagement / research	
All articles / documents reporting public engagement / research on public sector data use by or with the private sector as a secondary topic of the public engagement / research, where this was topic was explicitly designed into the engagement / study	All articles / documents reporting public engagement / research on public sector data use by or with the private sector as a secondary topic of the public engagement / research, but where this topic arouse spontaneously and was not explicitly designed into the engagement / study
All articles / documents reporting primary evidence from public engagement activity or research on the topic	All articles / documents based on secondary evidence (e.g. literature reviews, opinion or commentary articles)

Table 1. Inclusion and exclusion criteria

Limitations

This was a rapid review, which presents some notable limitations. Firstly, the literature search, data extraction, and analysis were performed by a single researcher, SE, which means that there is a higher likelihood of the researcher’s personal and professional biases having influenced the inclusion and exclusion of documents as well as the analysis and interpretation. To mitigate this, early decisions were informed by consultation with an academic librarian and SCB. Secondly, this study was conducted in a restricted timeline, which entailed some compromises on rigour. Most notably, the fact that we did not conduct quality appraisal of the included studies means that it is possible that studies of low quality have been included into the analysis. Also, notably, with the exception of the 9 documents where private sector use of public sector data was the primary topic and line-by-line coding was applied to the whole document, only paragraph level coding of topic relevant sections of the text was undertaken for most documents. This may have influenced the construction of analytic themes. Additionally, the literature search was restricted to two academic journal databases, two grey literature databases, although these cover a very wide range of academic journals, and a limited number of Google key word searches. While citation tracing was used to identify further relevant documents, it is possible that more relevant documents would have been identified if a wider range of academic journal databases and grey literature databases were used to search for documents. Despite these limitations, this rapid review provides a transparent and robust analysis of the current state of the art with respect to public engagement on private sector use of public sector data.

Results

Included studies

Forty-four documents were included into the analysis, of which 9 were focused on private sector use of public sector data as a primary topic. Most of the documents retrieved reported studies conducted in the UK. Out of the 44 documents, 26 studies involved UK populations only (6 in Scotland, 12 in England, 1 in Northern Ireland, and 8 across the UK), and 3 were comparative studies involving populations in the UK and other countries (UK and Sweden; US, Canada; and Australia). The rest involved populations in other high-income countries (5 in the US, 2 in Canada, 2 in Australia, 1 in Singapore, 1 across Europe, and 1 comparative study between the US and Japan). None involved populations in low- and middle-income countries. Out of the 9 primary topic documents, 7 studies involved UK populations only (3 in England, in 1 Scotland, and 2 across the UK), and the other two involved US and Australian populations. The documents also mainly reported studies focused on health or medical (including genomics) data. 36 of all the included documents and 8 of the primary topic documents were exclusively focused on health or medical data. The 9 documents that covered other kinds of data all also included health and medical data. The research or engagement activity participants included representatives of general publics and members of specific communities, patients and people with long term health conditions, and medical or health research and databank participants, with a diverse range of demographic characteristics and backgrounds.

Eight analytical themes were identified across the included documents, and we outline these below. While data across all documents was included in the analysis, in discussing the themes here, we place more emphasis on the 9 documents for which private sector use of public sector data was a primary topic. This is because these documents considered the related issues in more depth and their findings cover a more comprehensive range of considerations relevant to answering our research questions.

Public engagement and research methods used

The included studies used a range of different qualitative public engagement and research methods, mixed qualitative and quantitative methods, and quantitative survey, questionnaire, and, in one case, polling methods.

Fourteen used quantitative survey and questionnaire methods. Most were surveys or questionnaires designed to assess public attitudes, beliefs, knowledge, or awareness of data related topics, generally based on random sampling across a national population or from specific groups such as regional or patient populations, but 4 used stated preference, discrete and scenario based choice experiments. For example, Aitken et al (2018) and Tully et al (2020) used discrete choice experiment studies that enabled participants to select their preferred choice from a set of fixed options around questions concerning data sharing and linkage, and to establish the relative importance of different considerations around each option. The survey and questionnaire based studies generally did not, however, provide participants with information about the topic beforehand, with the exception of one study where a short informational video was provided for survey participants to watch before filling the survey. In addition, 3 of the included documents were based on mixed methods

that combined survey or questionnaire methods (in 2 cases) or polling methods (in 1 case) with qualitative workshop or interview methods.²

The majority of the included documents were based on qualitative research and public engagement methods. This included focus groups, but many studies and engagement activities used deliberative and group dialogue methods (4 used Citizen Juries, 1 used a Citizen Summit, 6 used workshop methods, 3 used other deliberative methods, and 5 used a combination of qualitative methods including one or more of the above deliberative and dialogue methods as well as round tables combined with focus groups or interviews). Most of these were based on the formation of 'mini publics'(c.f. Setälä & Smith, 2018), which bring together randomly selected members of publics who are broadly representative of the population in terms key demographic and, often, attitudinal characteristics in a deliberative environment. These methods generally involved a range of deliberation activities and discussion, and often including prompts, scenarios or case studies of actual or proposed uses of public sector data to scope and assess public perspectives around different examples of possible data uses, which allowed the identification of nuances, complexities and contextual considerations that members of publics had. They generally included facilitators and expert witnesses who provided background information on the topic from different perspectives. Deliberative methods including citizen juries and the citizen summit also enabled the participants to collectively develop recommendations that could be used to inform decision making.

Nearly all studies using these kinds of methods highlighted that they enable informed and in depth discussion, reflection and engagement from publics, resulting in more considered and consequently more nuanced views, and that they enabled consensus of recommendations. For example, Tully et al (2019) highlighted that the deliberative process, like citizen juries, enable publics to learn more about an issue, be exposed to a range of evidence and perspectives, and arrive at informed conclusion about them. Street et al (2021) similarly emphasised the informed nature of the deliberations that these methods enable, and the ability of citizen juries to rapidly increase participants' knowledge and understanding, which allows participants to effectively engage with policy relevant challenges and generate policy-guiding recommendations that are collectively produced and agreed. Understanding Patient Data and Ada Lovelace Institute (2020) noted specifically that public deliberation is a valuable method for public involvement in good data governance and can be purposely designed to ensure public views feed into the development of policy frameworks, while Teng et al (2019) among others emphasised that deliberative methods can enable people with diverse views, opinions, and interests to reach a consensus and contribute to policy through the diverse life experiences they bring to bear on a topic.

Further, Tully et al (2018, 2019) among others highlighted that publics are often unaware how data is used and why, which can mean that responses to quantitative surveys and questionnaires are given from a position of relative ignorance on the topic, which can have an impact on the nature of those responses. They also noted that qualitative methods including deliberative and dialogue methods, but also focus groups that involve the

² Many studies that used deliberative methods included an initial questionnaire and/or an exist questionnaire, but they are here nonetheless counted as a qualitative only rather than as mixed methods. Where a questionnaire was a separate part of a wider study that also included qualitative methods, the study is counted as having used mixed quantitative and qualitative methods.

provision of information and reflection on the topic, enable publics to arrive at and provide more informed views. Indeed, this is supported by our review, which found that the reported findings from quantitative studies tended to report more negative reactions towards private sector access to public sector data than studies and engagement activities that provided participants with information on the topic beforehand and opportunities for discussion and reflection (see the low public awareness, 'gut reactions,' and changing perceptions theme below). Tully et al (2018) also noted that the citizen jury process in particular has been specifically designed to enable decision makers to hear publics' views on complex topics, making it ideal for situations where decision makers wish to learn more about what informed publics think.

Combining quantitative methods with qualitative deliberative and dialogue methods also allowed some studies and engagement programmes to specifically identify how and to what extent the deliberative and dialogue activities influenced or shifted public's views. For example, the Royal Academy of Engineering (2010) combined electronic polling with focus groups and deliberative activities for their public engagement programme around the Breathing Country performance, which was designed to explore young people's views on the development of electronic patient records. This allowed the identification of how publics' perceptions changed from before to after the performance.

Overall, across the included documents, qualitative deliberative and group dialogue methods enabled the emergence of more nuanced and considered public views, likely because they enable participants to become more informed about the topic and reflect on and discuss it with others, whereas quantitative methods generally uncovered views that were based on publics' pre-existing knowledge and spontaneous attitudes towards the topic.

Findings from public engagement and studies on public views: themes

1. Low public awareness, 'gut reactions,' and changing perceptions

Many studies found that public awareness and understanding of data in general and private sector use of public sector data particularly is low. For example, the TNS (2012) and Ipsos MORI (2016) found that there is a limited understanding of how data is used in and shared beyond the public sector and what constitutes 'data' in the first place, including confusion between data and concepts like 'information' and 'knowledge.' Several studies found low awareness of the kinds of data sharing that already occur between the public and private sectors and about the role that private sector companies play in public sector service delivery. Hopkins, Kinsella, and Hopkins Van Mil (2020), Castell, Robinson, and Ashford (2018), and Ipsos MORI (2016), for example, found that most members of general publics are unaware of existing partnerships and data sharing between the NHS and private sector organisations, and when they are aware of it, they tend to underestimate the extent to which cross-sector data sharing already occurs.

Several studies also found that when members of publics are provided with more information about and opportunities to reflect more deeply on private sector uses of public sector data, many – albeit not all – change their minds about the topic, generally towards more positive views. Tully et al. (2019) for example found that over half of their study participants changed their views from feeling negative to feeling positive about private sector use of public sector data after deliberating on and becoming more informed about it,

especially about the safeguards that are in place, while Street et al (2021) found that many of their study participants shifted from complete opposition to recognition of the value of cross-sectoral data sharing. It should be noted, however, that while the shift in perspectives was generally found to be from negative to positive, some, including Chico, Hunn, and Taylor (2019), also found that a small minority change their minds in the opposite direction.

Several studies also suggest that there is a link between low awareness and understanding of private sector use of public sector data on the one hand and negative perceptions of private sector access to public sector data on the other. Ipsos MORI (2016), for example, found that lack of awareness of the current state of affairs led to concerns over private sector access to public sector data and to negative 'knee jerk reactions' towards it, while participants who changed their minds over the course of Chico, Hunn, and Taylor's (2019) study from positive to negative reported that their initial views indicated their spontaneous 'gut reaction' towards the topic. Their views changed to become more accepting after they became more informed. It is important to note that this likely bears relevance to and functions as a caveat when it comes to studies that report findings about the acceptability of private sector access to public sector data, especially when these studies report public views at the 'baseline,' where participants were not provided with further information or opportunities to reflect on the topic before their views were collected.

2. Acceptability

Several studies found that the spontaneous views of members of the public are generally negative towards private sector access to public sector data, and significantly higher numbers of people find it either unacceptable or undesirable than acceptable when their initial views are collected. This is linked to public perceptions about the rationales or motives that drive private sector organisations, which tend to be perceived as connected to profit making. For example, Rand Europe (2015) found that an overall pattern in their large cross-European survey was that Europeans are averse to the idea of private sector actors accessing their health data, while Davidson et al (2013) found that in the UK, there is a strong spontaneous opposition to public data being used by the private sector, for the purpose of profit maximisation. Hill et al (2013), among others, similarly found that when it comes to using data for research, there is a dichotomy between uses that are seen as acceptable versus uses that are seen as unacceptable, and this is based on perceptions concerning whether the research is for public benefit or for profit, respectively; these are seen as clashing motives.

All studies that considered public perceptions of acceptability beyond spontaneous reactions found that when publics' views are unpacked in more depth, they become more nuanced than simple aversion or opposition. Publics actually tend to find private sector use of public sector data acceptable under certain conditions, depending on the rationales for the use, the type of the data used, and the type of the private sector organisation using it. Illustrative of this is Ipsos MORI (2016) finding that participants in their study applied four key tests to decide how acceptable any given case of health data use is, with acceptability depending on why the data is being used, who is using it, what kind of data it is, and how it is used. They found that acceptability is assessed on a scale where uses of aggregate data by public health providers for public benefit in secure and regulated environments are most

acceptable, while uses of identifiable personal data by private organisations with no link to improving public health for purely private benefit in environments with poor security or regulatory oversight are least acceptable.

Relatedly, several studies found that the 'why' question – the rationales and motives for private sector use of data – was central. Generally, if the rationales and motives were directed towards public benefit, private sector use of public sector data was considered acceptable overall, and if the rationales and motives were perceived to be profit driven, it was considered less acceptable or unacceptable. Indeed, while Ipsos MORI (2016) found that their participants applied the above mentioned four key tests to decide how acceptable any case of data use is, they also found that the overall purpose was the key consideration: if the overall purpose for why data is used are considered acceptable, concerns over the involvement of private sector companies fade. This was supported by many other studies, and Tully et al (2018) and Grande et al (2013) for example found that the reasons why data is being used matter more than who is using it. Publics care about the user, but less than they care about the use and rationales for use.

Nonetheless, some studies found hierarchies of acceptability around the types of data that were being accessed and the type of organisation accessing it. When it comes to data type, anonymised and passively collected aggregate data is generally seen as most acceptable, and identifiable, personal, and sensitive data being least acceptable. For example, Davidson et al (2013) found that postcode data, sexual orientation data, and commercial data were particularly sensitive, while Hopkins Van Mil (2021) found that genomics and mental health data as well as qualitative data were seen as especially sensitive because they were considered personal in nature and because more care is needed to interpret qualitative data. When it comes to the acceptability of different private sector organisations accessing public data, the hierarchy of acceptability is shaped by perceptions of the organisations' motives and ability to deliver public benefits. Studies that found a hierarchy of organisations consistently found that at the bottom were insurance companies, with marketing companies following as a close second bottom. Hill et al (2013), for example, found that their study participants had concerns that insurance companies using public data would affect premiums or cover. Ipsos MORI (2016, 2021) similarly found that their participants opposed access being granted to insurance companies in particular, and that insurance and marketing companies were perceived as detrimental to individual and public interests. Because of this, they were seen as a red line, although some considered targeted marketing of health products to be potentially acceptable in cases where it might serve public interests in some way. Internet service, social media and retail companies using public sector data was found less acceptable than pharmaceutical companies. Davidson et al (2013) for example found that while internet service providers and social media companies were associated with online fraud, the majority of their participants felt that pharmaceutical companies contribute to understanding of health and disease and the development of drugs and treatments, and thus should be able to access data, even though there were concerns over profit making. Atkin et al (2021) additionally found that across different kinds of private organisations, participants in their study were likely to see healthcare focused companies' uses of data as more acceptable than non-healthcare focused companies.

Overall, there were different levels of acceptability for private sector use to public sector data that were generally determined by perceptions of what the rationale for data use is and who profits or benefits from the use, where public benefit was the primary driver of acceptability and commercial gain or private profit was the driver of unacceptability. Uses that were perceived oppositional or harmful to public benefit (such as insurance uses) were consistently found to be a red line that should not be crossed. It should be noted, however, that even if publics generally support conditional private sector access to public sector data when it is perceived to carry public benefit, Ipsos Mori (2016), among others, found that there remains a segment of the population that will find it hard to accept whatever the conditions or circumstances are.

3. Public benefit

Public benefit (for which terms like public interest and social benefit were also sometimes used) was the most prominent theme across the included documents and, as highlighted above, a key condition of acceptability for private sector uses of public sector data. Overall, there is widespread public support for private sector uses of the data if these uses can be convincingly demonstrated to deliver or have high potential to deliver public benefit, and this was generally the case even if these uses also carry private or commercial benefit as long as public benefit outweighs private benefit.

The TNS (2012), for example, found that public interest arguments generally trumped all other concerns, while Tully et al (2018, 2019) found not only that public benefit is a key condition of acceptability for private sector uses of public sector data but also that most of their participants accepted commercial gain if it accrued secondary to public benefit. Indeed, across different studies, a key underlying reason for opposition to private sector use of public sector data was the concern that it would be motivated by private or commercial rather than public interest: these were seen as incompatible. Davidson et al (2013), among others, found that their participants opposed private sector involvement when they perceived it as profit driven or did not trust private sector organisations to act in the public interest, while Tully et al (2018) found that the reasons their participants gave for why organisations should not have access centrally included failure to clearly indicate that data use would principally be for public benefit rather than private gain. Some studies also found that the potential for public benefit from any proposed private sector use of public sector data must be convincingly demonstrated rather than just speculated. Connected Health Cities (2017), for example, found that participants in their study required that public benefit as opposed to commercial gain was satisfactorily demonstrated, and generally did not support proposals of private sector use when there remained doubts about whether public benefit would be realised, while Chico, Hunn, and Taylor (2019) found that a major consideration was whether there was evidence that any alleged public benefit would materialise.

What, precisely, 'public benefit' is taken to entail is, however, variable and contested. Several definitions of the concept were proposed by participants across the different studies, and different groups of participants placed emphasis on different kinds of benefit. In some studies, such as that undertaken by Hopkins Van Mil (2021), public benefit was defined expansively to include both direct and indirect benefits, such as improving health

and social care outcomes and knowledge development. Participants in the study undertaken by Aitken et al (2018) used wide ranging conceptualisations of public benefit but considered that it should be as broad as possible, including benefit to individuals, specific groups, and society as a whole. Ipsos MORI (2016) similarly found examples of public benefit to include a wide range of areas and activities from developing new drugs and treatments, to improvements in care, to higher level of service and greater access to services for vulnerable groups.

Several studies conducted in the UK specifically identified that private sector access to or use of NHS data should bring both outcome based and financial benefits directly to the NHS. Tully et al (2019) and Connected Health Cities (2017) both found that their participants wanted to see demonstrable potential for improvements in areas such as cost savings and targeted use of resources as well as treatments, while participants in the study undertaken by Hopkins, Kinsella, and Hopkins Van Mil (2020) wanted the NHS to reap the benefits from any private sector partnership through improvements in patient care as well as resource and administrative efficiency. Chico, Hunn, and Taylor's (2019) participants, on the other hand, identified two main ways in which the NHS should benefit: any products or services developed using NHS data should be made available to the NHS at a preferential rate, and the NHS should have unlimited access to any knowledge or insight arising out of the use of NHS data.

There was some ambivalence around the relationship between public benefit and individual benefit, with some studies subsuming individual benefit under public benefit while others considered it separate, and some finding that public benefit outweighs individual benefit and others finding that it may not do so. Chico, Hunn, and Taylor's (2019) participants could perceive data use to be of public benefit even where they could not see clear and direct benefits to individuals, while Hopkins, Kinsella, and Van Mil Hopkins' (2020) participants saw a clear link between benefits for individual patients and wider societal benefit, and Ipsos MORI (2016) found that individual benefit was seen as valuable but less so than wider societal benefit. Participants in Castell, Robinson, and Ashford's (2018) study found it difficult to assess the idea that uses of personally identifiable patient data may create wider public benefits without benefitting the individuals whose data is used, but Davidson et al (2013) found that whether benefits should accrue to individuals or wider publics depended on the type of data. They found that uses of data that require individuals' proactive participation (e.g. research involving genetic data) should directly benefit individuals whose data is used, whereas uses of routinely collected aggregate data need not benefit individuals directly but should be of wider public benefit.

While the conceptualisations of 'public benefit' and what it should encompass were variable, the general lesson across the included studies is that demonstrable public benefit broadly conceived is either one of the most important considerations or the most important consideration that publics have around private sector use of public sector data. This theme is overlapping with but can be distinguished from another prevalent theme, which was that of benefit-sharing and distribution.

4. Benefit-sharing and distribution

Beyond the general demand that private sector use of public sector data should deliver or have demonstratable high potential to deliver public benefit, a related but separate theme across the included studies pertained to the question of how benefits should be shared or distributed to enable public benefit to be realised most appropriately or equitably. Considerations included fair distribution of benefits, profit sharing and reinvestment, and, for private sector uses of NHS data in the UK, the question of whether the NHS should charge for data access.

Participants in several studies expressed the view that benefits should be distributed fairly across the population. For example, studies undertaken by Understanding Patient Data and Ada Lovelace Institute (2020) and by Hopkins, Kinsella, and Van Mil Hopkins (2020) found that participants generally felt that fair distribution entailed an even distribution of outcomes, benefits and rewards across the country, at least in the longer term, even if short term benefits would first be realised at local or regional levels. Hopkins, Kinsella, and Hopkins Van Mil's (2020) participants also considered that any benefits that accrue from the use of NHS data should be fairly distributed across the NHS, but also that in case of partnerships between private and public organisations, there should be mutual benefit for all parties, including for the private sector organisations involved.

Davidson et al (2013) additionally found that their participants considered benefit-sharing to be more relevant and important for private sector access to public sector data than for other kinds of public data use because of the prevailing perception that private sector organisations were especially motivated by profit, and thus benefit-sharing models are particularly needed to ensure public benefit. Some studies also specifically considered questions around profit sharing and reinvestment, including whether public sector organisations should charge for access to their data. Davidson et al (2013) and Ipsos MORI (2021) found support for the idea that any products or services developed using public sector data and especially NHS data should be provided at a lower cost for the NHS. Ipsos MORI's (2021) participants generally felt that it was important for the NHS to receive a return investment and recover costs of providing and maintaining data, while participants in the Davidson et al (2013) study considered that since the public sector invests considerable funds to data collection, it should recoup some of these funds from those wishing to use the data. They also considered that the cost could vary depending on the type of the organisation, and third sector and other public organisations could be given data for free or at a lower cost than private sector companies, who were presumed to generally make more profit out of data use.

Members of publics generally felt that benefits from private sector use of public sector data should be equitably shared and returned to publics and patients.

5. Trust and distrust

The extent to which publics trust and distrust different types of organisations across the private, public and third sectors, and why, were prevalent questions across the included studies, and many found that trust in and the perceived trustworthiness of the organisations involved with the data are key considerations for publics.

As noted above, several studies found that the spontaneous views of members of publics towards private sector companies accessing public sector data tend to be negative. This is likely connected with many studies finding that publics tend to express spontaneous distrust towards private sector companies especially because of worries over commercial motives but also because of concerns over security and unethical practices. Ipsos MORI (2016), for example, found that the private sector in general was distrusted and participants were concerned about conflicting commercial priorities and about the ability of private organisations to store data safely. Tully et al (2019) similarly found that many of their participants were uncomfortable with commercial access to public sector data because they did not trust private companies' motives or 'agenda,' because the companies did not have a track record of protecting data, or because they worried that companies might use data to manipulate or exploit individuals or populations in ways that support their own agenda. The question of trust was also connected with the above discussed hierarchy of acceptability around different private sector organisations accessing public data, with Krahea, Milligan, and Reilly (2019) and Robinson and Dolk (2016) finding that organisations whose access to data was considered least acceptable were the same organisations that were least trusted and perceived to be the most untrustworthy, namely, insurance companies. Chico, Hunn, and Taylor (2019) found that whether private sector organisations were perceived as trustworthy had two elements: members of publics needed to be able to trust that the organisation was committed to public benefit and that the organisation was able to protect the data they were using.

However, several studies also found that members of publics became more positive about initiatives involving private sector uses of public sector data (regardless of the type of the private organisation) if public sector organisations were involved as key partners in these initiatives and especially if public sector organisations retained control over the partnership and data use. This is linked to higher levels public trust in public sector organisations, which were generally perceived to be working for public benefit and were consequently perceived as more trustworthy than private sector organisations. In the UK, publics placed an especially high level of trust in the NHS, and if the NHS was involved as a partner and retained control over the data and its uses, private sector involvement was perceived more positively. For example, Atkin et al (2021) found that the NHS was consistently identified as the most trusted entity to make decisions about data use, while Tully et al (2019) and Chico, Hunn, and Taylor (2019) found that when the NHS was closely involved, publics had greater trust in the ability of an initiative to deliver public benefit despite private sector involvement, including because the private sector companies would need to accept and adopt NHS principles. Castell, Robinson, and Ashford (2018), among others, also found that participants believed that the NHS owns health data and is ultimately responsible for both the data itself and how it is used, and while trust in the different organisations accessing the data was important, participants believed that the NHS would protect public interest and

ensure no harm comes to individuals. They noted that publics generally trust the NHS unequivocally to protect their data, and because of this, the NHS bears the weight of expectations when it comes to private sector uses of it.

While similar patterns of trust and distrust were evident in other countries as well, notable exceptions were some studies conducted in the US where participants were generally found to have lower levels of trust in the government. Brown Trinidad et al (2010), for example, found that their US based participants expressed widespread distrust in federal agencies and the government in general, while a comparative study undertaken by Ghafur et al (2020) found that distrust in US agencies was significantly higher than in UK agencies, and that this is likely linked with concerns that data might not be protected from commercial end-use within the largely privatised US health system. In ways that may be related to this, Chico, Hunn, and Taylor (2019) found that their participants in the UK expressed lower levels of trust in US based private companies than in private companies based in the UK or Europe, because they did not feel confident that their data would be given the same level of protection in the US.

Overall, the questions around trust, distrust, and trustworthiness were key considerations across different studies, and whether an organisation is trusted or perceived as trustworthy was centrally linked to perceptions of whether the organisations is acting in the public interest or for private benefit. In the UK, there is a high level of trust in public sector organisations, and in the NHS in particular, and private sector access to data is perceived more positively, especially if the NHS oversees and retains control over the process, largely because the NHS is seen as driven by public benefit and because it is trusted to act in public interest.

6. Oversight, governance, and safeguards

Across the included studies, oversight, governance and safeguards around public sector data in general and private sector access to and use of public sector data in particular were a key reoccurring consideration. This encompassed questions around different mechanisms that could or should be implemented to govern data use and who does or should have control over data sharing and partnerships, questions around transparency, accountability, confidentiality, consent processes and questions around safety and security arrangements.

Participants across different studies wanted there to be a governing body or an oversight committee responsible for overseeing how public sector data is shared and used outside the public sector. Understanding Patient Data and Ada Lovelace Institute's (2020) and Hopkins, Kinsella, and Van Mil Hopkins's (2020) participants wanted a single governing body to be held accountable for NHS data sharing, which would undertake proactive, reactive, and monitoring activities including establishing a governance framework, taking regulatory action when needed, and auditing and reporting. Hopkins, Kinsella, and Van Mil Hopkins's (2020) also found that participants wanted this body to be governed by both legal requirements and guiding principles that are binding and subject to ongoing review.

There were, however, some differences in views on what the oversight or governing body should look like and who it should encompass, with some studies finding support for an independent oversight body with diverse representatives and others finding support for public sector bodies to take responsibility for oversight. For example, Hopkins, Kinsella, and Van Mil Hopkins' (2020) participants wanted non-hierarchical representation from the NHS, academic, charity and industry, and members of communities and publics, and Davidson et al (2013) found strong support for the establishment of an oversight body that would comprise of a range of different stakeholders. Aitken et al (2018), however, found that their participants were divided on whether data linkage processes should be overseen by the relevant public service or by another (non-governmental) independent body, while some studies in the UK found that publics want the NHS to retain oversight and control over health data. For example, Castell, Robinson, and Ashford (2018) found that their participants were relatively uninterested in questions around safeguards and governance around health data, but this was because they devolved this responsibility to the NHS, reflecting their high level of trust in the NHS as an organisation presumed to act in the public interest. Their participants presumed that the NHS should retain control and ownership of the data throughout the process as well as decide and ensure adherence to rules of data use.

Several studies found that participants considered transparency and accountability around private sector use of public sector data important. For example, Aitken et al (2016) found that publics' perceptions of different organisations' level of transparency and accountability was crucial for shaping their views, while Hopkins, Kinsella, and Van Mil Hopkins' (2020) participants considered that transparency and accountability to publics should be prerequisites for all NHS data sharing partnership. They saw this to include a public register of organisations using NHS data and transparency around the aims, process, results, and benefits (financial and otherwise) that all partners would receive. Similarly, Davidson (2013) found that transparency was seen as key to mitigating against the possibility of data misuses and instilling public confidence, but also that their participants wanted mechanisms in place that would enable publics to hold data users accountable and monitor or challenge instances of data use, including the extent to which they are contributing to public benefit, the level of profit to private companies, and whether they are achieving their intended outcomes. More generally, Hopkins Van Mil (2021) found that transparency is not an 'add-on' to but rather a part of realising public benefit, including because it is required to demonstrate how public benefit is or will be achieved.

Some studies also raised issues around safety and security arrangements as well as around consent, confidentiality or privacy. Davidson et al (2013) and Hopkins, Kinsella, and Van Mil Hopkins (2020) found that concerns about data security and privacy featured prominently in their studies: participants worried over the risk of data falling in the 'wrong hands,' and generally wanted reassurance that security and privacy would be safeguarded. Ipsos MORI's (2016) participants, on the other hand, generally felt that no amount of security could fully remove the risks involved in data sharing (such as leaks and hacking) and the involvement of private sector organisations intensified this concern because it was perceived to entail that public sector data is moving beyond the public systems that are or should be accountable for it. Ipsos MORI (2016), among others, also found nuances and complexities around publics' views on the issue of consent: while a majority of their

participants would prefer that consent was sought whenever their NHS data is shared with private sector organisations, they also found that participants shifted their views when they had more information, especially when they were presented with projects that they wanted to happen, but it would be impractical to seek consent from everyone whose data is used. Taylor and Talyor (2014) relatedly found that their participants were willing to accept consent models other than those which they may ideally prefer, including consent models with a lower level of individual control over how their data is used, if there are other safeguard arrangements in place to govern how the data can be used. These included provisions such as security, controls over detrimental use and commercialisation, transparency, the existence of an independent ‘watchdog’ body, and confidentiality protections. Indeed, confidentiality was identified as particularly significant in the context of private sector access in some studies, including by Aitken et al (2016) who found that confidentiality and anonymity were seen as having a higher level of importance when data might be accessed by private companies.

Overall, Ipsos MORI (2016) found that having safeguards in place makes a difference to publics’ views on private sector use of public sector data, but there is no clear preference on the precise nature of what the safeguards should be – rather, the nature of the safeguards is less important than knowing that safeguards exist in the first place. This is to some extent reflected across the other studies as well, which generally found that publics want to see stringent oversight, governance, and safeguard arrangements, especially concerning an oversight or governance body, transparency and accountability processes, and arrangements for data security and safety, consent, and confidentiality. However, the precise nature of what these should look like varied across the different studies.

7. Public involvement and engagement

Many studies identified a need for further public engagement and involvement around data use in general and private sector access to public sector data in particular. This partially overlaps with but can distinguished from the theme of oversight, governance and safeguards, as public involvement and engagement was often identified as a component of oversight or transparency, but it was also raised as a separate element that should be considered.

Several studies identified a need to involve publics in decision making about how their data is used, but there were different ideas about the nature and extent of this involvement and what it should look like in practice. For example, Davidson et al (2013) found unanimous agreement that there should be public involvement in decision making on data sharing, including how benefit-sharing models should be developed and in setting the rules for how data is governed. They found, however, that participants favoured feedback and consultation models over more active forms of engagement, such as participation in agenda-setting and representation, partly because the participants felt that publics may not have the specific skills required to be involved in an oversight body. Yet, they also found that views on this varied and participants often struggled to articulate who should be involved and what involvement should look like. Hopkins, Kinsella, and Van Mil Hopkins (2020) and Understanding Patient Data and Ada Lovelace Institute (2020) on the other hand found that their participants strongly supported active forms of public involvement in

decision making about data use. Hopkins, Kinsella, and Van Mil Hopkins (2020) reported a desire for public involvement at various levels, including policy and practice and the establishment and management of partnerships, via forms such as Citizen Juries and deliberation activities for key decisions, public votes for partnerships, and representation in governance boards. Understanding Patient Data and Ada Lovelace Institute (2020) highlighted however that some participants also felt that proportionality matters, and they wanted to ensure that public involvement did not result in overly complex bureaucracy that might hinder positive developments. There was a suggestion that public involvement could focus on 'grey area' cases of data use that attract a diversity of views and perspectives, whereas involvement may not be needed for broadly acceptable cases, where there is clear public benefit and low risks, or for broadly unacceptable 'red line' cases, such as cases involving insurance or marketing uses or clear commercial exploitation of data.

Generally, while several studies identified a need for public involvement or engagement, the precise nature of what this should look like, who should be involved and in what ways was contested.

8. Impact of demographic differences

Finally, many studies found that demographic differences had an impact on public perceptions and views on private sector use of public sector data, and several different demographic variables were found to be significant. Ipsos MORI (2016), among others, found that many interconnected factors including educational attainment, age, socioeconomic status as well as existing knowledge and awareness of data uses are linked to the extent to which individuals accept private sector access to public sector data. Other studies found links especially pertaining to age, socioeconomic and employment status, education and work sector, gender and ethnicity, and the status of having a long term health condition.

For example, Atkin et al (2021), Castell, Robinson, and Ashford (2018) and Chico, Hunn, and Taylor (2019) found differences linked to age and life stage. They found that generally younger people tended to be more accepting of data sharing and use by private sector organisations than older adults, while Aitken et al (2018) found that older participants were most concerned with how profits are managed. Ipsos MORI (2016), however, also found that the relationship between age and acceptance of private sector use of data is complex and non-linear, and acceptance could be seen as being most shaped by individuals' mindset concerning issues such as risks and benefits of data use. Castell, Robinson, and Ashford (2018), among others, found that socioeconomic and employment status and work sector impacted participants' views, while Aitken et al (2018) found that those who are not in full time work were more concerned with oversight and type of data used whereas those in full time work were less concerned with oversight and more with the purpose of data use. Aitken et al (2018) also found that men were more concerned about oversight arrangements in place than women, while a comparative study across the UK, US, and Australia undertaken by Milne et al (2019) on genomic data sharing arrangements found that men were more likely to have high overall trust across different kinds of organisations involved in genomic data sharing, including private sector organisations, than women. In the US, Grande et al (2013) found that African American and Hispanic participants were less

willing to share their health information for private sector (especially marketing) uses than white participants, while Castell, Robinson, and Ashford (2018), among others, found that people with long term health conditions and a patient status tended to be more positive and open to data sharing than general publics. Chico, Hunn, and Taylor (2019), on the other hand, found that patients were less likely to support scenarios involving targeted marketing of health products and pharmaceutical companies accessing public sector health data than members of general publics.

Overall, these differences show that individuals' social positions and life experiences shape their perceptions of and attitudes towards private sector uses of public sector data, but the ways in which they do so are complex and not always consistent across different studies. This highlights that publics do not have a homogenous perspective or shared understanding of this topic and while patterns can be identified across several demographic variables, there are likely to be complicated intersections across the different variables.

Discussion

The studies included in this review suggest that deliberative and dialogue based qualitative public engagement methods enable the identification of more nuanced public perspectives around private sector use of public sector data than (especially quantitative) methods that generally collect publics' spontaneous attitudes or 'gut reactions' towards the topic. Deliberative methods can also enable a consensus to be built among people with diverse backgrounds and perspectives. This is connected with the finding that there is a low level of public awareness of private sector use of public sector data and this low awareness has an impact on the spontaneous reactions and views that publics have towards the topic. Publics tend to express negative 'gut reactions' towards private sector organisations in general and towards them accessing public sector data in particular, and these negative reactions are influenced by presumptions about the motives that drive private sector interest in public sector data: publics express concern that private sector organisations are motivated by commercial interests and profiteering rather than by public benefit. However, when publics are given more information about the topic and opportunities to discuss and reflect on it in more depth over the course of deliberative and dialogue based public engagement activities, their views tend to become more nuanced, and many change their minds towards more positive views.

The above is also connected to the finding, reported in many of the included studies, that publics tend to spontaneously express the view that private sector use of public sector data is undesirable or unacceptable. However, when people's views are unpacked and probed in more depth, and when more information about the topic and opportunities to reflect on it are provided, publics actually tend to be supportive of private sector access to public sector data, subject to certain conditions and caveats. These include, more notably, the rationale and motives for private sector use of the data, which can matter more than who is using it. When private sector involvement can be shown to have high potential to deliver public benefit, publics tend to find it acceptable, but when publics suspect that the rationale is mostly private benefit or profits, they tend to find it undesirable or unacceptable. There are, however, related hierarchies of acceptability depending on the kind of private sector organisation and type of data involved. When companies are seen to be working (at least to a large extent) towards public benefit, their involvement is perceived as more acceptable,

but when companies are seen to be working in ways that are perceived as contrary to public benefit, their involvement is seen as unacceptable. Publics also generally perceive the use anonymised and aggregate data as most acceptable, but they are more concerned about identifiable, personal, and sensitive data.

Across the included documents, public benefit was the most prevalent theme raised by publics, and indeed, the acceptability of private sector use of public sector data was largely conditional upon the extent to which public benefit could be realised from it. The ability of private sector use of public sector data not only generally trumped all other concerns, but publics were also willing to allow some private benefit or profits for private sector organisations as long as private benefits are not driving motive, and so long as public benefit outweighs private benefits or profit. They also wanted to see convincing evidence and justifications that public benefit will or could in fact be realised. The exact meaning or definitions of public benefit were variable and contested across the different studies, including benefits to individuals, specific groups, society as a whole, and, in the UK, to the NHS specifically via avenues such as cost savings and outcome and service improvements. Despite this, it is clear that the ability to deliver demonstratable public benefit, broadly conceived, is a central consideration for private sector use of public sector data. Relatedly, the included studies suggest that benefit-sharing is a key consideration for publics, and people want to see a fair distribution of benefits from private sector access to public sector data that are realised across the population, including in terms of profit sharing and reinvestment. Especially with private sector use of NHS data in UK, publics wish to see a return of investment and cost recovery for the NHS.

The issue of public trust and distrust in different private (and public) sector organisations was also a key issue across the different studies, and connected to publics' views on acceptability, public benefit and benefit-sharing. Publics' negative 'gut reactions', as well as many of their key enduring concerns, were linked to the level of trust and distrust they had towards different organisations, and many studies found both spontaneous and enduring distrust in private sector organisations' motives, which were perceived by many as profit driven. However, when public sector organisations were centrally involved in directing and controlling private sector uses and access to the data, public trust in the whole process was much higher, because publics generally expressed significantly higher levels of trust in public sector organisations. They generally presumed that public sector organisations' motives are principally driven by public benefit, and that public sector organisations would consequently work to protect public interest against any possible competing private interests. In the UK, this was especially the case when the NHS was involved: publics generally expressed very high levels of trust in the NHS and believed that if the NHS retains control of the data in the context of private sector partnerships, it will protect public interest and ensure public benefit. The few included studies focussing on the US showed less willingness to trust state/government agencies than is the case in the UK.

Oversight, governance, and safeguards were widely reported as key issues of public interest and concern in private sector use of public sector data, and publics generally wanted and expected to see mechanisms in place that would govern how data is used. While the exact nature of what oversight, governance and safeguards should look like was contested and varied across different studies, generally, publics wanted either a governing body or an

oversight committee consisting of different stakeholders to take responsibility for overseeing how public sector data is shared and used, or they wanted public sector organisations, such as the NHS, to hold this responsibility. They also expected there to be transparency and accountability over private sector uses of public sector data, including ways for publics to access and monitor how and why public sector data is used by private companies, as well as rigorous safety and security arrangements, including to avoid data falling in the 'wrong hands.' While consent and confidentiality were also raised, public views on this tended to be conditioned by the nature of other safeguards that were in place to govern how data is used and protected. Overall, there was a suggestion that having oversight, governance, and safeguard mechanisms in place to begin with may be more important than the exact nature of these safeguards when it comes to public trust in and acceptability of private sector use of public sector data.

Many of the included studies identified that publics wish there to be public engagement and involvement around private sector use of public sector data, but there were differing views and perspectives on what this should look like, at what level it should be undertaken, and what publics should be involved in exactly. Some wanted active public involvement in decision making about data use across different levels while others preferred involvement only in some areas like agenda setting and development or governance arrangements, or only in making decisions about contested 'grey area' cases. Despite these differences, overall there was a general desire for public engagement and involvement around this policy area.

Importantly, across the included studies, public views and perspectives were not homogenous, rather the views and perspectives of different people were shaped and influenced by a range of intersecting demographic and attitudinal attributes that had a bearing on how individuals perceive and feel about private sector use of public sector data. This serves as a reminder that while one can identify trends and patterns across publics, 'the public' is not a homogenous entity but rather composed of multiple publics, which in turn are composed of multiple groups and individuals who may have different perspectives, concerns, or areas of emphasis. Dialogic or deliberative public engagement is able to bring such difference together to reflect and debate, enabling some kind of consensus or emergent set of views or positions.

Conclusions and Recommendations

We identified that deliberative and dialogue based qualitative public engagement and research methods have the ability to shift public views, and they tend to enable the development and identification of more nuanced and considered public perspectives around private sector use of public sector data than quantitative methods that generally collect publics' spontaneous attitudes or 'gut reactions.' Deliberative methods also enable the identification of collectively developed and agreed recommendations from members of publics. This suggests that deliberative and dialogue based public engagement methods are likely to be most effective in identifying informed and considered public views on private sector use of public sector data, whereas quantitative methods can be used to assess publics' spontaneous reactions and levels of existing awareness and knowledge.

We identified eight overarching key themes across the findings from the included documents that should be emphasised and taken into account when it comes to private sector use of public sector data in general and in designing public engagement activities around this area in particular: low public awareness, 'gut reactions,' and changing perceptions around data use; acceptability of private sector data uses; the centrality of public benefit; importance of benefit-sharing and distribution; trust and distrust; oversight, governance, and safeguards; public involvement and engagement; and the impact of demographic differences of people's views.

The impact of low levels of public awareness on publics' views concerning private sector use of public sector data suggests that there is a pressing need to better inform and engage publics about how and why private sector organisations are accessing and using public sector data, and providing information should be a part of related public engagement activities. The central emphasis on public benefit across the included studies, including as a key condition of public acceptability for private sector use of public sector data, suggests that any proposed or actual case of private sector involvement with public sector data should be able to deliver or have high potential to deliver public benefit in order for it to be acceptable to publics and have public support. This ability should be convincingly evidenced and justified, and concerns over excessive private benefits and profit motives mitigated. However, there is no consensus on the definition of 'public benefit' in the existing literature, which suggests that public engagement activities should be designed to clarify the meaning and scope of this concept. Relatedly, publics wish to see a fair distribution of benefits from private sector uses of public sector data, suggesting that the development of benefit and profit sharing models is of central importance and should involve publics. While publics tend to be relatively distrustful of private sector organisations, the included studies show that publics generally have a high level of trust in public sector organisations and especially the NHS in the UK. Publics are generally more trusting of partnerships between the private and public sector if public sector organisations, and especially the NHS, retain control of the data and its uses during such partnerships, because publics presume that public sector organisations will protect public interest and ensure public benefit is realised. This suggests that to gain and retain public trust in and support for private sector uses of public sector data, public sector organisations should be centrally involved as key partners, and the data, as well as how it is used, should remain under the control of public sector organisations. The included studies also suggest that publics expect there to be both stringent oversight, governance, and safeguards mechanisms, and public involvement and engagement mechanisms in place. However, there is no consensus in existing literature on the precise nature of these oversight, governance, and safeguards, or public involvement and engagement mechanisms, and there is some ambivalence around these areas. This suggests that public engagement activities should be designed to identify, clarify and define what kinds of oversight, governance, and safeguards should be in place. It also suggests that more work is needed to identify precisely where, to what extent, and at what levels publics wish to be involved when it comes to decision making about private sector use of public sector data.

Based on the above, we recommend a focus on the following:

- Use deliberative and dialogue based public engagement methods to involve publics in decision making about private sector use of public sector data
- Develop public engagement activities to identify where, to what extent, and at what levels, publics wish to be involved in decision making about private sector use of public sector data
- Develop public engagement activities to identify and derive an agreed definition of the meaning and scope of 'public benefit,' to develop a useable definition that is informed and supported by a public consensus
- Ensure that all cases of private sector use of public sector data are centrally motivated by and have a demonstratable potential to deliver public benefit, and provide convincing evidence and justifications for how public benefit will be realised
- Involve publics in the development of effective benefit-sharing models for private sector partnerships, including profit sharing and reinvestment of profits into the public sector
- Ensure that all cases of private sector use of public sector data have an in-built benefit-sharing system based on these benefit-sharing models
- Develop public engagement activities to identify the best and most acceptable oversight, governance, and safeguard mechanisms that should be implemented to govern private sector uses of public sector data in ways that ensure that data is protected, and that public benefit is realised
- Ensure that all cases of private sector use of public sector data have stringent oversight, governance, and safeguard mechanisms that publics find acceptable and trustworthy
- Ensure that all cases of private sector use of public sector data are transparent and clearly communicated to publics, and implement educational campaigns that inform publics about private sector use of public sector data more generally

References

- Aitken, M., McAteer, G., Davidson, S., Frostick, C., & Cunningham-Burley, S. (2018) Public Preferences regarding Data Linkage for Health Research: A Discrete Choice Experiment. *International Journal of Population Data Science*, 3: 11
- Aitken, M., Cunningham-Burley, S., & Pagliari, C. (2016) Moving from trust to trustworthiness: Experiences of public engagement in the Scottish Health Informatics Programme. *Science and Public Policy*, 43(5): 713-723
- Atkin, C., Crosby, B., Dunn, K., Price, G., Marston, E., Crawford, C., O'Hara, M., et al. (2021) Perceptions of anonymised data use and awareness of the NHS data opt-out amongst patients, carers and healthcare staff. *Research Involvement and Engagement*, 7: 40
- Ballantyne, A., & Stewart, C. (2019) Big Data and Public-Private Partnerships in Healthcare and Research. *Asian Bioethics Review*, 11: 315-326
- Brown Trinidad, S., Fullerton, S. M., Bares, J. M., Jarvik, G.P., Larson, E. B., & Burke, W. (2010) Genomic research and wide data sharing: Views of prospective participants. *Genetics in Medicine*, 12(8): 486-495
- Carter, P, Laurie, G., & Dixon-Wood, M. (2015) The social licence for research: why care.data ran into trouble. *Journal of Medical Ethics*, 41: 404-409
- Castell, S., Robinson, L., & Ashford, H. (2018) Future data-driven technologies and the implications for use of patient data Dialogue with public, patients and healthcare professionals. Ipsos MORI.
- Chico, V., Hunn, A., & Taylor, M. (2019) Public views on sharing anonymised patient-level data where there is a mixed public and private benefit. NHS Health Research Authority & University of Sheffield
- Connected Health Cities. (2017) Connected Health Cities Citizens' Juries Report. A report of two citizens' juries designed to explore whether the planned and potential uses of health data by Connected Health Cities are acceptable to the public.
- Davidson, S., McLean, C., Treanor, S., Ipsos MORI, Aitken, M., Cunningham-Burley, S., Laurie, G., Pagliari, C., & Sethi, N. (2013) Public Acceptability of Data Sharing Between the Public, Private and Third Sectors for Research Purposes. The Scottish Government
- Ghafur, S., Van Dael, J., Leis, M., Darzi, A., & Sheikh, A. (2020) Public perceptions on data sharing: key insights from the UK and the USA. *The Lancet*, 2: e444-e446
- Grande, D., Mitra, N., Shah, A., Wan, F., & Asch, D. A. (2013) Public Preferences About Secondary Uses of Electronic Health Information. *JAMA Internal Medicine*, 173(19): 1798-1806

Hill, E. M., Turner, E. L., Martin, R. M., & Donovan, J., L. (2013) "Let's get the best quality research we can": public awareness and acceptance of consent to use existing data in health research: a systematic review and qualitative study. *BMC Medical Research Methodology*, 13: 72

Hopkins Van Mil (2021) Putting Good into Practice: A public dialogue on making public benefit assessments when using health and care data.

Hopkins, H., Kinsella, S., & Van Mil Hopkins, A. (2020) Foundations of Fairness: view on uses of NHS patients' data and NHS operational data. A mixed methods public engagement programme with integrated Citizens' Juries. *Creating connections Hopkins van Mil*.

Ipsos MORI (2021) Public deliberation in the use of health and care data.

Ipsos MORI. (2016) The one-way mirror: public attitudes to commercial access to health data. Report prepared for the Wellcome Trust.

Krahea, M., Milligan, E., & Reilly, S. (2019) Personal health information in research: Perceived risk, trustworthiness and opinions from patients attending a tertiary healthcare facility. *Journal of Biomedical Informatics*, 95: 103222

Milne, R., Morley, K. I., Howard, H., Niemiec, E., Nicol, D., Critchley, C., Prainsack, B. et al. (2019) Trust in genomic data sharing among members of the general public in the UK, USA, Canada and Australia. *Human Genetics*, 138: 1237–1246

Setälä, M., & Smith, G. (2018). Mini-Publics and deliberative democracy. In A. Bächtiger, J. S. Dryzek, J. Mansbridge, & M. E. Warren (Eds.), *The Oxford handbook of deliberative democracy* (pp. 300–314). Oxford: Oxford University Press.

Street, J., Fabrianesi, B., Adams, C., Flack, F., Smith, M., Carter, S. M., Lybrand, S., et al. (2021) Sharing administrative health data with private industry: A report on two citizens' juries. *Health Expectations*, 00: 1-12

Taylor, M. & Talyor, N. (2014) Health research access to personal confidential data in England and Wales: assessing any gap in public attitude between preferable and acceptable models of consent. *Life Science, Society and Policy*, 10: 15

Teng, J., Bentley, C., Burgess, M. M., O'Doherty, K.C., & McGrail, K. M. (2019) Sharing linked data sets for research: results from a deliberative public engagement event in British Columbia, Canada. *International Journal of Population Data Science*, 4(1): 13

Thomas, J. & Harden, A. (2008) Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8: 45

TNS (2012) Open data dialogue Final Report.

Tully, M. P., Bernsten, C., Aitken, M., & Vass, C. (2020) Public preferences regarding data linkage for research: a discrete choice experiment comparing Scotland and Sweden. *BMC Medical Informatics and Decision Making*, 20: 109

Tully, M. P., Hassan, L., Oswald, M., & Ainsworth, J. (2019) Commercial use of health data - a public "trial" by citizen jury. *Learning Health Systems*, 3: e10200

Tully, M., Bozentko, K., Clement, S., Hunn, A., Hassan, L., Norris, R., Oswald, M., & Peek, N. (2018) Investigating the Extent to Which Patients Should Control Access to Patient Records for Research: A Deliberative Process Using Citizens' Juries. *Journal of Medical Internet Research*, 20(3): e122.

Rand Europe (2015) Privacy of health records Europeans' preferences on electronic health data storage and sharing.

Royal Academy of Engineering. (2010) Privacy and prejudice: Young people's views on the development and use of Electronic Patient Records.

Understanding Patient Data & Ada Lovelace Institute. (2020) Foundations of Fairness: Where next for NHS health data partnerships?

Wachter, R. M. & Cassel, C. K. (2020) Sharing Healthcare Data with Digital Giants: Overcoming Obstacles and Reaping Benefits While Protecting Patients. *JAMA*, 323 (6): 507-508

**Unlocking the Value of Scotland's Data:
Public Engagement around the Access of Public Sector Data with or by Private Sector Organisations**

Dr Sonja Erikainen, Professor Sarah Cunningham-Burley,
Centre for Biomedicine, Self and Society,
University of Edinburgh
2021



© Crown copyright 2023

OGL

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3 or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

This publication is available at www.gov.scot

Any enquiries regarding this publication should be sent to us at
The Scottish Government
St Andrew's House
Edinburgh
EH1 3DG

ISBN: 978-1-80525-775-2 (web only)

Published by The Scottish Government, August 2023

Produced for The Scottish Government by APS Group Scotland, 21 Tennant Street, Edinburgh EH6 5NA
PPDAS1238683 (08/23)